SPARSE RECOVERY ASSISTED DOA ESTIMATION UTILIZING SPARSE BAYESIAN LEARNING

Min Huang, Lei Huang

College of Information Engineering, Shenzhen University, China

ABSTRACT

This paper proposes a novel approach to sparse recovery assisted direction-of-arrival (SR-DOA) estimation. By exploiting the sparsity inherent in the spatial spectrum, the DOA estimation is formulated as a sparse nonnegative least squares problem. Meanwhile, in order to enhance the estimation accuracy, the devised method is able to suppress the additive Gaussian noise but at the expense of a few degrees-of-freedom, and mitigate the sampling errors by exploiting its asymptotic distribution. Subsequently, the sparse Bayesian learning with nonnegative Laplace prior is utilized to yield the DOA estimation. The performances of the proposed SR-DOA estimator along with other two existing approaches are investigated and compared. Numerical results show that the proposed SR-DOA algorithm is superior to the state-of-the-art methods in terms of the estimation accuracy.

***Index Terms*—** DOA, sparse recovery, sparse nonnegative least squares problem, sparse Bayesian learning, nonnegative Laplace prior

1. INTRODUCTION

Since direction-of-arrival (DOA) estimation can be widely used in many areas, such as radar, sonar and wireless communications, it has received considerable attention in literature. However, the estimation performances of the existing algorithms usually degrade seriously in the situations of low signal-to-noise ratio (SNR) or small number of snapshots.

Because the number of source signals is usually limited, the spatial spectrum observed is sparse. Thus, by properly utilizing the compressive sensing techniques, the sparse property inherent in the array signal model can be exploited to improve the DOA estimation performance. In particular, DOA is estimated by minimizing the data fitting error as well as the sparsity of solution. In [1], the sparsity of solution is generated by forming an ℓ_1-norm penalty function. In order to enforce the sparsity of solution, a weighted ℓ_1-norm penalty function which utilizes the property of noise subspace is proposed by [2]. In addition, the regularization parameter, which is used to control the trade-off between the data fitting error and the sparsity of solution, is obtained with the aid of the Lagrangian duality in [3, 4]. Although, the regularization parameter derived in [4] does not rely on any $a\ priori$ knowledge, it is suboptimal because it only satisfies the sufficient conditions of optimal solution. Furthermore, when the grid is dense, the computational complexities of the above algorithms are unaffordable.

Alternatively, the sparse Bayesian learning (SBL) [5], which avoids the regularization parameter selection, can be used to solve the sparse recovery problem. Specifically, the DOA is determined by maximizing its posterior probability, namely, the product of the likelihood probability and the prior probability. In order to obtain more degrees-of-freedom (DOFs) to estimate DOA, the DOA estimation problem is firstly converted to a sparse nonnegative least squares (S-NNLS) problem [6]. Then, the nonnegative Gaussian probability density function (PDF) is considered as the prior probability to solve the S-NNLS problem [7]. To enforce the sparsity of solution, the nonnegative Gaussian PDF is replaced by the Laplace PDF [8]. Although it is is not conjugate to the Gaussian likelihood function, it can be implemented by a hierarchical way. In other words, the Laplace PDF can be constructed by Gaussian PDF and Exponential PDF. Nevertheless, according to [9, 10], the power of Gaussian noise estimated by SBL may not be accurate, leading performance degradation.

In this paper, a sparse recovery assisted DOA (SR-DOA) estimator is proposed. First, we show that the high-resolution DOA estimation can be formulated as a sparse optimization problem. Second, a selection matrix is designed for mitigating the effect of additive Gaussian noise but at the expense of a small amount of DOFs. Third, a whitening filter is introduced for coping with the sampling errors. Last, SBL with nonnegative Laplace is used to determined the DOAs. Our simulation results show that the proposed SR-DOA estimator outperforms all the benchmark DOA estimators when SNR is in low region.

This work was supported in part by the Natural Science Foundation of China under Grants U1501253, 61601300, 6160010632, 61501300, 61601304 and 61501485; in part by the China Postdoctoral Science Foundation under Grant 2016M592535; in part by the Natural Science Funding of Guangdong Province under Grant 2017A030313336; in part by the Foundation of Shenzhen under Grants JCYJ20170302142545828 and JCYJ20160520165659418; in part by the Foundation of Shenzhen University under Grants 2016057 and 201557; and in part by the China Scholarship Council.

2. DOA ESTIMATION PROBLEM FORMULATION

Consider K uncorrelated narrowband far-field signals, $s_k(t)$, $k = 1, 2, \cdots, K$, impinging on a linear sparse array which consists of M omnidirectional sensors located at $[0,\ d_1,\ \cdots,\ d_{M-1}]$, where d_m represents the distance between the $(m+1)$-th sensor and the first sensor. Then, the array output vector $\boldsymbol{x}(t)$ of T snapshots can be expressed as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t),\ t = 1, 2, \cdots, T \tag{1}$$

where $\boldsymbol{s}(t) = [s_1(t)], s_2(t), \cdots, s_K(t)]^T$ and $\boldsymbol{n}(t)$ denote the source signal and additive Gaussian noise, respectively. $\boldsymbol{A}$ consists of K steering vectors, i.e., $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_K)]$, with

$$\boldsymbol{a}(\theta) = \left[1, e^{-\frac{j2\pi d_1 \sin(\theta)}{\tau}}, \cdots, e^{-\frac{j2\pi d_{M-1} \sin(\theta)}{\tau}}\right]^T$$

where τ is the wavelength.

Note that the DOA of the k-th source signal is distributed in the range of $(-90°, 90)$. Thus, by invoking all the possible DOAs, $\mathbf{x}(t)$ in (1) can be written in a high-resolution and sparse representation as

$$\boldsymbol{x}(t) = \bar{\boldsymbol{A}}\bar{\boldsymbol{s}}(t) + \boldsymbol{n}(t),\ t = 1, 2, \cdots, T \tag{2}$$

where $\boldsymbol{A} = [\boldsymbol{a}(\bar{\theta}_1), \boldsymbol{a}(\bar{\theta}_2), \cdots, \boldsymbol{a}(\bar{\theta}_{\bar{K}})]$ and the set of $\bar{\boldsymbol{\theta}} = \{\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_{\bar{K}}\}$ gives a sampling grid of all possible DOAs, while $\bar{\boldsymbol{s}}(t) = [\bar{s}_1(t)], \bar{s}_2(t), \cdots, \bar{s}_K(t)]^T$ with $\bar{s}_k(t)$ being the possible source signal. In general, we have $\bar{K} \gg K$. Therefore, $\bar{\boldsymbol{s}}(t)$ is a sparse vector, whose k-th row is nonzero and equals to the corresponding row of $\boldsymbol{s}(t)$ in (1). Consequently, the problem of DOA estimation based on (1) is equivalent to identifying the positions of the nonzero rows of $\boldsymbol{x}(t)$ in (2).

3. SPARSE RECOVERY ASSISTED DOA ESTIMATION ALGORITHM

3.1. S-NNLS modeling

To begin with, the sample covariance matrix of $\boldsymbol{x}(t)$ of (2) can be derived as

$$\hat{\boldsymbol{R}} = \bar{\boldsymbol{A}}\boldsymbol{R}_s\bar{\boldsymbol{A}}^H + \boldsymbol{R}_n + \boldsymbol{E} \tag{3}$$

where $\boldsymbol{R}_s = \mathbb{E}[\bar{\boldsymbol{s}}(t)\bar{\boldsymbol{s}}(t)^H] = \mathrm{diag}\{\sigma_1^2, \cdots, \sigma_{\bar{K}}^2\}$ with $\sigma_k^2 = \mathbb{E}[\bar{s}_k(t)\bar{s}_k(t)^H]$ being the power received from the k-th source signal, $\boldsymbol{R}_n = \mathrm{diag}\{\sigma^2, \cdots, \sigma^2\}$ with σ^2 being the variance of noise, while $\boldsymbol{E}$ reflects the error between the covariance matrix of $\boldsymbol{x}(t)$ given in (2), which is $\bar{\boldsymbol{A}}\boldsymbol{R}_s\bar{\boldsymbol{A}}^H + \boldsymbol{R}_n$, and its sample covariance matrix $\hat{\boldsymbol{R}}$ of (3). Let us vectorize (3), yielding an M^2-length vector, which is

$$\boldsymbol{y} \triangleq \mathrm{vec}\{\hat{\boldsymbol{R}}\} = \boldsymbol{V}\boldsymbol{\varsigma} + \boldsymbol{\rho} + \boldsymbol{\xi} \tag{4}$$

where $\boldsymbol{V} \triangleq \bar{\boldsymbol{A}}^* \odot \bar{\boldsymbol{A}}$, $\boldsymbol{\varsigma} \triangleq [\sigma_1^2, \cdots \sigma_{\bar{K}}^2]^T$, $\boldsymbol{\rho} \triangleq \mathrm{vec}(\boldsymbol{R}_n) = [\sigma^2\boldsymbol{e}_1^T, \cdots, \sigma^2\boldsymbol{e}_M^T]^T$ and $\boldsymbol{\xi} \triangleq \mathrm{vec}(\boldsymbol{E})$. Here, $(\cdot)^*$, $\odot$ and $\boldsymbol{e}_i$ denote, respectively, the complex conjugate, Khatri-Rao product [11], and the i-th column of the identity matrix $\boldsymbol{I}_M$. Based on (4), our DOA estimation problem is converted to a problem of identifying the locations of nonzero elements in $\boldsymbol{\varsigma}$.

Firstly, we convert (4) into its real form, which can be expressed as

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{V}}\boldsymbol{\varsigma} + \hat{\boldsymbol{\rho}} + \hat{\boldsymbol{\xi}} \tag{5}$$

where $\hat{\boldsymbol{y}} = [\Re\{\boldsymbol{y}\}^T, \Im\{\boldsymbol{y}\}^T]^T$, $\hat{\boldsymbol{V}} = [\Re\{\boldsymbol{V}\}^T, \Im\{\boldsymbol{V}\}^T]^T$, $\hat{\boldsymbol{\rho}} = [\boldsymbol{\rho}^T, \boldsymbol{0}^T]^T$ and $\hat{\boldsymbol{\xi}} = [\Re\{\boldsymbol{\xi}\}^T, \Im\{\boldsymbol{\xi}\}^T]^T$. Here, $\boldsymbol{0}$ is an $M^2 \times 1$ zero vector.

From (5) we know that there are M nonzero elements in the noise resulted component of $\hat{\boldsymbol{\rho}}$, whose positions are known. Therefore, they can be canceled by deleting the corresponding elements in $\hat{\boldsymbol{y}}$ but at the cost of the loss of M DOFs. In mathematical form, the cancellation of the noise resultant components in (5) can be implemented by pre-multiplying a selection matrix $\boldsymbol{J}$ satisfying $\boldsymbol{J}\hat{\boldsymbol{\rho}} = \boldsymbol{0}$ on $\hat{\boldsymbol{y}}$, yielding

$$\boldsymbol{u} \triangleq \boldsymbol{J}\hat{\boldsymbol{y}} = \boldsymbol{J}\hat{\boldsymbol{V}}\boldsymbol{\varsigma} + \boldsymbol{J}\hat{\boldsymbol{\xi}}. \tag{6}$$

Note that, according to the structure of $\boldsymbol{e}_i$, $\boldsymbol{J}$ is constructed from the identity matrix $\boldsymbol{I}_{2M^2}$ by removing its $\{0 \times M + 1, 1 \times M + 2, \cdots, (M-1) \times M + M\}$ rows.

According to [12], $\boldsymbol{\xi}$ of (4) is asymptotically complex normal ($\mathcal{CN}$) distributed, if T is sufficiently large. In this case, the distribution of $\hat{\boldsymbol{\xi}}$ in (5) can be approximated as [13]

$$\hat{\boldsymbol{\xi}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\hat{\boldsymbol{\xi}}}) \tag{7}$$

where

$$\boldsymbol{R}_{\hat{\boldsymbol{\xi}}} = \frac{1}{2}\left[\Re\{\boldsymbol{R}_{\boldsymbol{\xi}}\}\ -\Im\{\boldsymbol{R}_{\boldsymbol{\xi}}\};\ \Im\{\boldsymbol{R}_{\boldsymbol{\xi}}\}\ \Re\{\boldsymbol{R}_{\boldsymbol{\xi}}\}\right]$$

with

$$\boldsymbol{R}_{\boldsymbol{\xi}} \sim \mathcal{CN}(\boldsymbol{0}, (\hat{\boldsymbol{R}}^T \otimes \hat{\boldsymbol{R}})/P).$$

Here, $\otimes$ denotes the Kronecker product [11]. Therefore, $\boldsymbol{J}\hat{\boldsymbol{\xi}}$ in (6) follows the distribution of $\mathcal{N}(\boldsymbol{0}, \boldsymbol{G})$ with $\boldsymbol{G} \triangleq \boldsymbol{J}\boldsymbol{R}_{\hat{\boldsymbol{\xi}}}\boldsymbol{J}^T$. Explicitly, the transformed noise $\boldsymbol{J}\hat{\boldsymbol{\xi}}$ is colored, which may lead to severe performance degradation. In order to alleviate its effect, we may whiten $\boldsymbol{J}\hat{\boldsymbol{\xi}}$ through multiplying $\boldsymbol{u}$ of (6) by $\boldsymbol{G}^{-\frac{1}{2}}$, yielding an S-NNLS model, i.e.,

$$\hat{\boldsymbol{u}} \triangleq \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{\varsigma} + \boldsymbol{\nu} \tag{8}$$

where $\boldsymbol{\Psi} \triangleq \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{J}\hat{\boldsymbol{V}}$ and $\boldsymbol{\nu} \triangleq \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{J}\hat{\boldsymbol{\xi}} \sim \mathcal{N}(0, \boldsymbol{I}_{2M^2-M})$ is now a white Gaussian noise vector.

3.2. Sparse Bayesian learning with nonnegative Laplace prior

For the model (8), we have the Gaussian likelihood function as

$$p(\hat{\boldsymbol{u}}|\boldsymbol{\varsigma}) \sim \mathcal{N}(\boldsymbol{\Psi}\boldsymbol{\varsigma}, \boldsymbol{I}_{2M^2-M}). \tag{9}$$

In addition, according to [8], the prior for $\boldsymbol{\varsigma}$ can be considered as a nonnegative Laplace distribution. However, since the nonnegative Laplace prior is not conjugate to the Gaussian likelihood function, it cannot be directly applied to the sparse Bayesian learning. In order to solve this issue, we can model it in a hierarchical way [5, 8], which is described as

$$\begin{aligned} p(\boldsymbol{\varsigma}|\lambda) &= \int p(\boldsymbol{\varsigma}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}|\lambda)d\boldsymbol{\gamma} \\ &= \lambda^{\bar{K}} e^{-\lambda \sum\limits_{k=1}^{\bar{K}} \boldsymbol{\varsigma}_k} \end{aligned} \tag{10}$$

where

$$p(\boldsymbol{\varsigma}|\boldsymbol{\gamma}) = \prod_{k=1}^{\bar{K}} \mathcal{N}_+(\boldsymbol{\varsigma}_k|0, \boldsymbol{\gamma}_k) \tag{11}$$

with

$$\mathcal{N}_+(\boldsymbol{\varsigma}_k|0, \boldsymbol{\gamma}_k) = 2\mathcal{N}(\boldsymbol{\varsigma}_k|0, \boldsymbol{\gamma}_k) \tag{12}$$

being the zero-mean nonnegative Gaussian probability density function (PDF), while

$$p(\boldsymbol{\gamma}|\lambda) = \prod_{k=1}^{\bar{K}} p(\boldsymbol{\gamma}_k|\lambda)$$

with

$$p(\boldsymbol{\gamma}_k|\lambda) = \frac{\lambda}{2} e^{-\frac{\lambda_\gamma}{2}}$$

being the Exponential PDF. Furthermore, the hyperprior of λ in (10) is assumed to follow Gamma distribution, i.e.,

$$p(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \tag{13}$$

where $\Gamma(a)$ is a Gamma function, while the a and b is set to a very small constant c for simplification.

Based on the Bayes rule, we can estimate $\boldsymbol{\varsigma}$ by maximizing its posterior density, namely,

$$\begin{aligned} \hat{\boldsymbol{\varsigma}} &= \arg\max_{\boldsymbol{\varsigma}} p(\boldsymbol{\varsigma}|\boldsymbol{\gamma}, \lambda, \hat{\boldsymbol{u}}) \\ &\propto \arg\max_{\boldsymbol{\varsigma}} p(\hat{\boldsymbol{u}}|\boldsymbol{\varsigma})p(\boldsymbol{\varsigma}|\boldsymbol{\gamma}) \\ &\propto \arg\max_{\boldsymbol{\varsigma}} \mathcal{N}_+(\boldsymbol{\varsigma}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \boldsymbol{\mu} \end{aligned} \tag{14}$$

where

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{\Psi}^T \hat{\boldsymbol{u}}$$

and

$$\boldsymbol{\Sigma} = \left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \boldsymbol{\Lambda}^{-1}\right)^{-1}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}\{\boldsymbol{\gamma}\}$. From (14), we readily find that $\hat{\boldsymbol{\varsigma}}$ is a function of $\boldsymbol{\gamma}$. Hence, once $\boldsymbol{\gamma}$ is estimated, the Maximum-A-Posteriori (MAP) estimate of $\hat{\boldsymbol{\varsigma}}$ can be determined by (14).

The $\boldsymbol{\gamma}$ and its associated hyperparameter λ can be estimated by maximizing their posterior density, namely, $p(\boldsymbol{\gamma}, \lambda|\hat{\boldsymbol{u}})$. For convenience, considering $\boldsymbol{\varsigma}$ as a hidden variable, $\boldsymbol{\gamma}$ and λ can be readily obtained by expectation-maximization (EM) algorithm [14], which is written as

$$\begin{aligned} \hat{\boldsymbol{\gamma}}, \hat{\lambda} &= \arg\max_{\boldsymbol{\gamma},\boldsymbol{\lambda}} \mathbb{E}[\mathrm{logp}(\boldsymbol{\varsigma}, \boldsymbol{\gamma}, \lambda|\hat{\boldsymbol{u}})] \\ &\propto \arg\max_{\boldsymbol{\gamma},\boldsymbol{\lambda}} \mathbb{E}[\mathrm{logp}(\boldsymbol{\varsigma}, \boldsymbol{\gamma}, \lambda, \hat{\boldsymbol{u}})]. \end{aligned} \tag{15}$$

Furthermore, the $p(\boldsymbol{\varsigma}, \boldsymbol{\gamma}, \lambda, \hat{\boldsymbol{u}})$ of (15) can be expanded as

$$p(\boldsymbol{\varsigma}, \boldsymbol{\gamma}, \lambda, \hat{\boldsymbol{u}}) = p(\hat{\boldsymbol{u}}|\boldsymbol{\varsigma})p(\boldsymbol{\varsigma}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}|\lambda)p(\lambda). \tag{16}$$

Hence, when λ is given, $\hat{\boldsymbol{\gamma}}$ can be computed by

$$\begin{aligned} \hat{\boldsymbol{\gamma}}_k &= \arg\max_{\boldsymbol{\gamma}_k} \mathbb{E}[\mathrm{logp}(\boldsymbol{\varsigma}|\boldsymbol{\gamma}) + \mathrm{logp}(\boldsymbol{\gamma}|\lambda)] \\ &= -\frac{1}{2\lambda} + \sqrt{\frac{1}{4\lambda^2} + \frac{w_k}{\lambda}} \end{aligned} \tag{17}$$

where w_k is the second-order moment of $\boldsymbol{\varsigma}_k$. With the aid of (14), w_k can be obtained by [6]

$$w_k = \boldsymbol{\mu}_k^2 + \boldsymbol{\Sigma}_{k,k} + \boldsymbol{\mu}_k \sqrt{\frac{2\boldsymbol{\Sigma}_{k,k}}{\pi}} \frac{e^{\frac{\boldsymbol{\mu}_k^2}{2\boldsymbol{\Sigma}_{k,k}}}}{\mathrm{erfc}(-\frac{\boldsymbol{\mu}_\mathrm{k}}{\sqrt{2\boldsymbol{\Sigma}_\mathrm{k,k}}})}. \tag{18}$$

Similarly, when $\boldsymbol{\gamma}$ is given, λ can be computed by

$$\begin{aligned} \hat{\lambda} &= \arg\max_{\hat{\lambda}} \mathbb{E}[\mathrm{logp}(\boldsymbol{\gamma}|\lambda) + \mathrm{logp}(\lambda)] \\ &= \frac{\bar{K} - 1 + c}{\sum\limits_{k=1}^{\bar{K}} \boldsymbol{\gamma}_k/2 + c}. \end{aligned} \tag{19}$$

From (17) and (19), it is easy to see that $\hat{\boldsymbol{\gamma}}$ and λ are the functions of $\{\hat{\boldsymbol{\varsigma}}, \lambda\}$ and $\boldsymbol{\gamma}$, respectively. Recalling that $\hat{\boldsymbol{\varsigma}}$ is a function of $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\varsigma}}$ can be determined in an iterative way, which is tabulated in Algorithm 1.

4. SIMULATION RESULTS

In this section, we evaluate the performance of our proposed SR-DOA method, and compare it with the state-of-the-art approaches, i.e., the nonnegative SBL (NNSBL) [8] and conventional SBL [5] algorithms. We consider a linear sparse array of 4 sensors with locations $\mathcal{D} = \{0,\ d,\ 4d,\ 6d\}$, where d is the basic element-spacing. Moreover, $(-90°90°)$ space

Algorithm 1 SR-DOA algorithm

Step **1**: Construct $\hat{\boldsymbol{u}}$ and $\boldsymbol{\Psi}$ from (8), and initialize $\boldsymbol{\gamma}$.
Step **2**: Calculate λ by (19).
Step **3**: Determine $\hat{\varsigma}$ by (14).
Step **4**: Update $\boldsymbol{\gamma}$ by (17).
Step **5**: If the norm of the difference betweent the updated $\boldsymbol{\gamma}$ and its last value is small enough, stop the algorithm; Otherwise, jump to Step **2**.

is sampled uniformly with interval $1°$ to obtain the direction set of $\bar{\boldsymbol{\theta}}$. Six source signals, which have the common center frequency of $f = 200$ Hz and common propagation speed of $c = 340$ m/s, are assumed to come from the directions of $\{-54°, -28°, -9°, 10°, 31°, 56°\}$. In addition, the signals are generated by uncorrelated narrowband Gaussian sources.

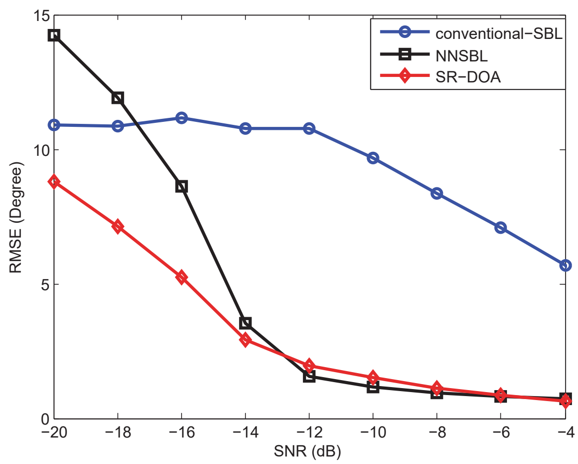

Fig. 1. RMSE versus SNR performance for different DOA estimators at T=200.

Fig. 1 shows the root mean square error (RMSE) versus SNR for different DOA estimators, where the number of snapshots is $T = 200$. We can explicitly observe that our proposed SR-DOA method outperforms the other estimators when the SNR is less than -12.5 dB. This is due to the fact that noise power becomes dominated at low SNRs, which, when properly mitigated, leads to lower RMSE for the proposed approach although it losses some DOFs. Note that although the additive Gaussian noise is estimated and compensated in NNSBL algorithm, the estimate is not accurate as it is interfered by the other hyperparameter, namely, $\boldsymbol{\gamma}$, especially in low SNR region [9, 10]. In addition, when the SNR is larger than -12.5 dB, the estimation performance of our proposed SR-DOA algorithm is slightly worse than that of the NNSBL algorithm. This is because the loss in DOFs becomes dominated at high SNRs and thereby degrades the estimation performance. This is why the degradation may be visible in relatively high SNR

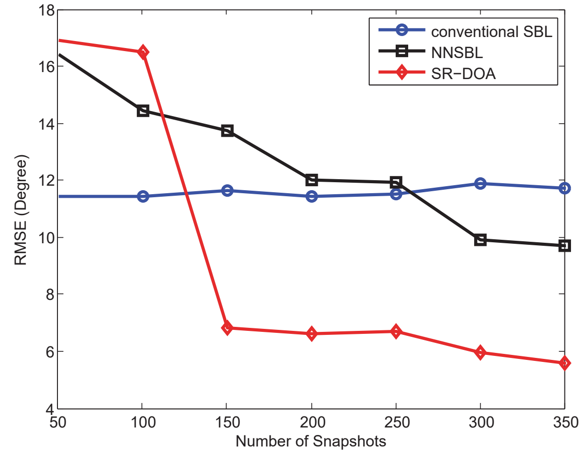

Fig. 2. RMSE versus number of snapshots for different DOA estimators at SNR $= -18$ dB.

region, as indicated in Fig. 1.

Fig. 2 depicts the RMSE versus the number of snapshots for different DOA estimators. As the number of snapshots increases, the RMSE performances of our proposed SR-DOA and NNSBL algorithms improves, while the performance of the conventional SBL algorithm is almost unchanged. This indicates that the conventional SBL algorithm is invalid when the SNR is low. Furthermore, from Fig. 2, we can find that the RMSE of the proposed SR-DOA algorithm is relatively low, as long as the number of snapshots is not less than 150, indicating that the proposed SR-DOA algorithm is superior to other existing schemes in the scenarios of low SNR and small number of snapshots.

5. CONCLUTIONS

A novel SR-DOA estimator has been proposed in this work. Exploiting the sparsity of spatial spectrum, a sparse S-NNLS model is constructed. Therein, the selection matrix and whitening filter have been designed in order to mitigate the effect of noise and sampling error on the DOA estimation. Since the constructed selection matrix is based on the fact that the power of noise is real, the cancellation of noise only at the expense of a few DOFs. Subsequently, the sparse Bayesian learning with nonnegative Laplace prior is used to solve the S-NNLS problem, resulting in the MAP estimate of DOA. Numerical results show that the proposed SR-DOA estimator outperforms the state-of-the-art approaches in terms of estimation accuracy for low SNRs.

6. REFERENCES

[1] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010-3022, Aug. 2005.

[2] C. Zheng, G. Li, H. Zhang, and X. Wang, "An approach of DOA estimation using noise subspace weighted minization," *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2011, pp. 2856-2859.

[3] X. Xu, X. Wei, and Z. Ye, "DOA estimation based on sparse signal recovery utilizing weighted ℓ_1-norm penalty," *IEEE Signal Process. Lett.*, vol. 19, no. 3, pp. 155-158, Mar. 2012.

[4] Z. -Q. He, Z. -P. Shi, L. Hunag and H. C. So, "Underdetermined DOA estimation for wideband signals using robust sparse covariance fitting," *IEEE Signal Process. Lett.*, vol. 24, no. 4, pp. 435-439, Apr. 2015.

[5] S. Babacan, R. Molina, A. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 53-63, Jan. 2010.

[6] A. Nalci, I. Fedorov, B. D. Rao, "Rectified gaussian scale mixtures and the sparse non-negative least squares problem,", arXiv preprint arXiv:1601.06207.

[7] N. Hu, B. Sun, Y. Zhang, J. Dai, J. Wang, and C. Chang, "Underdetermined doa estimation method for wideband signals using joint nonnegative sparse bayesian learning," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 535-539, May 2017.

[8] N. Hu, B. Sun, J. Wang, J. Dai, and C. Chang, "Source localization for sparse array using nonnegative sparse bayesian learning," *Signal Process.*, vol. 127, pp. 37-43, 2016.

[9] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704-3716, Jul. 2007.

[10] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 912-926, Sep. 2011.

[11] P. L. Fackler, *Notes on Matrix Calculus*. Tech. Rep., North Carolina State Univ., 2005.

[12] Z. -M. Liu, Z. -T. Huang and Y. -Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3896-3907, Aug. 2013.

[13] R. G. Gallager, "Circularly-symmetric Gaussian random vectors," preprint, 2008.

[14] M. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, pp. 211C244, 2001.