ARRAY COVARIANCE MATRIX-BASED ATOMIC NORM MINIMIZATION FOR OFF-GRID COHERENT DIRECTION-OF-ARRIVAL ESTIMATION

Yu Zhang, Gong Zhang, Xinhai Wang

Key Lab of Radar Imaging and Microwave Photonics, Ministry of Education,
Nanjing University of Aeronautics and Astronautics, Nanjing, China

ABSTRACT

A two-stage method for off-grid coherent direction-of-arrival (DOA) estimation using atomic norm minimization based on the covariance matrix is proposed in this paper. In the first stage, by vectorizing the covariance matrix, a new off-grid model matched as a linear combination of two dimensional harmonic is presented, where the proposed denoising covariance matrix-based atomic norm minimization (DCMANM) is applied for the vectorized covariance matrix denoising. Then the simplified dual polynomial method (SDPM) is used for DOA estimation. Unlike most of existing methods, the proposed method requires knowing neither the number of signals nor the statistics of noise. Numerical simulations demonstrate the outperformance of the proposed method in both the angle resolution and DOA estimation precision compared to the state-of-the-art approaches.

***Index Terms*—** DOA estimation, covariance matrix, atomic norm, dual polynomial

1. INTRODUCTION

Direction-of-arrival (DOA) estimation as an important subject in array signal processing has attracted much interest, and many classical methods sprung up over last few decades [1–4]. MUSIC [2], as a representative of the classical methods, has been applied widely. However, it suffers from certain limitations such as it requires the prior information about the number of signals and the uncorrelated property among signals. Recently, inspired by compressed sensing (CS) and sparse representation theorem, [5] proposes L1-SVD method to estimate DOAs. Compared with MUSIC, L1-SVD is more robust to the numbers of signals and correlated signals. We often refer to L1-SVD as a on-grid method since it formulates signal sparse representation on the predefined discrete dictionary or equally assumes the true DOAs exactly lie on the predefined fixed grid. However, in reality the assumption rarely holds since the DOAs are continuous-valued and the unavoidable basis mismatch may result in the performance degradation. More recently, the method named atomic norm soft thresholding for multiple measurement vectors (AST for MMV) using atomic norm technology for line spectrum estimation is proposed in [6,7] to alleviate the effect of basis mismatch. In contrast to L1-SVD, AST for MMV is referred to as an off-grid method. It's worth noting that, in both L1-SVD and AST for MMV, the prior information about the statistics of noise is required to determine a desired regularization parameter. [8] proposes the covariance matrix reconstruction approach (CMRA), another off-grid method, for DOA estimation, which is based on the covariance matrix of observed signals and can be carried out without knowing the number of signals and the statistics of noise, while it can only handle uncorrelated signals.

In this paper, firstly, by vectorization operator on the covariance matrix of observed signals, an off-grid model matched as a linear combination of two dimensional harmonic is presented, where an off-grid method named denoising covariance matrix-based atomic norm minimization (DCMANM) is proposed for the vectorized covariance matrix denoising. Secondly, simplified dual polynomial method (SDPM) is proposed for DOA estimation, of which both peak searching and rooting formulations are presented. It's the first time the atomic norm for two dimensional harmonic is utilized into uniform linear array (ULA) DOA estimation. Moreover, The proposed two-stage method (DCMANM+SDPM) not only can be carried out without the prior information of the number of signals and the statistics of noise, but is robust to correlated or even coherent signals.

Throughout this paper, matrices and vectors are denoted by bold letters and scalars are denoted by unbolded letters. For a matrix $\boldsymbol{X}$, $\mathrm{vec}(\boldsymbol{X})$ denotes a vector whose elements are taken column-wise from $\boldsymbol{X}$. We use $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ to denote the transpose, the conjugate and the conjugate transpose operation, respectively.

2. DCMANM WITH ULA

Consider K far-field narrowband signals $s_k(t)$, $k = 1, \ldots, K$, impinging on a ULA of M omnidirectional sensors which are uniformly spaced with a spacing of d from distinct directions $\theta_k \in [-90^\circ, 90^\circ)$, $k = 1, \ldots, K$. The observation model can

This work was supported by the Chinese National Natural Science Foundation under contract nos. 61471191, 61071163, 61071164 and 61501233.

be represented as

$$\boldsymbol{y}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t) + \boldsymbol{n}(t), t = 1, 2, \ldots, L \tag{1}$$

where t indexes the snapshot and L denotes the snapshot number, $\boldsymbol{y}(t) = [y_1(t), \ldots, y_M(t)]^T$, $\boldsymbol{s}(t) = [s_1(t), \ldots, s_K(t)]^T$ and $\boldsymbol{n}(t) = [n_1(t), \ldots, n_M(t)]^T$ denote the observed snapshot, the vector of source signals and the vector of noise at the snapshot t, respectively. $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_K)]$ is the array mainfold matrix and $\boldsymbol{a}(\theta_k)$ is the steering vector of the k-th source which satisfies

$$\boldsymbol{a}(\theta_k) = [1, e^{j2\pi \frac{d}{\lambda}\sin(\theta_k)}, \ldots, e^{j2\pi \frac{d}{\lambda}(M-1)\sin(\theta_k)}]^T, \tag{2}$$

where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_K]$ is the vector of directions, and in this paper, the spaced distance d is assumed to be equal to half of the wavelength λ, i.e., $d = \lambda/2$.

Moreover, the source signals $\boldsymbol{s}(t)$ can be correlated and $\boldsymbol{n}(t)$ are additive Gaussian white noise with same distribution which satisfy

$$\begin{aligned} &E[\boldsymbol{s}(t_1)\boldsymbol{s}^H(t_2)] = \boldsymbol{P}\delta_{t_1,t_2}, \\ &E[\boldsymbol{n}(t_1)\boldsymbol{n}^H(t_2)] = \mathrm{diag}(\boldsymbol{\sigma})\delta_{t_1,t_2}, \\ &E[\boldsymbol{s}(t_1)\boldsymbol{n}^H(t_2)] = E[\boldsymbol{n}(t_1)\boldsymbol{s}^H(t_2)] = \boldsymbol{0}, \end{aligned} \tag{3}$$

where $\boldsymbol{P} = \{P_{i,j}\} \in \mathbb{C}^{K\times K}$ is the source correlation matrix, $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_M]^T \in \mathbb{R}_+^M$ denotes the noise variance parameter, and δ_{t_1,t_2} denotes the Kronecker delta function which equals 1 if $t_1 = t_2$ or 0 otherwise. It's worthy noting that in amount of covariance matrix-based algorithms such as MUSIC and CMRA, the source correlation matrix $\boldsymbol{P}$ is assumed as a diagonal matrix, while it may not be hold in practice. In this paper, we assume $\boldsymbol{P}$ is a non-diagonal matrix to exhibit the robustness of the proposed method to correlated signals in theory. Under the above assumptions, the covariance matrix $\boldsymbol{R}$ of observed signals has the following decomposition

$$\boldsymbol{R} = E[\boldsymbol{y}(t)\boldsymbol{y}^H(t)] = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{P}\boldsymbol{A}^H(\boldsymbol{\theta}) + \mathrm{diag}(\boldsymbol{\sigma}). \tag{4}$$

Let $f_k = \frac{1}{2}\sin(\theta_k) \in [-\frac{1}{2}, \frac{1}{2})$ and $\Omega = \{f_1, \ldots, f_K\}$ denote the frequency that corresponds to the direction θ_k and the set of corresponding frequencies, respectively. Then (2) can be written equivalently as

$$\boldsymbol{a}(f_k) = [1, e^{j2\pi f_k}, \ldots, e^{j2\pi(M-1)f_k}]^T, \tag{5}$$

and similarly, $\boldsymbol{A}(\boldsymbol{f}) = [\boldsymbol{a}(f_1), \ldots, \boldsymbol{a}(f_K)]$, where $\boldsymbol{f} = [f_1, \ldots, f_K]$ is the vector of corresponding frequencies. It's easy to find that $\boldsymbol{A}(\boldsymbol{f})$ satisfies

$$\boldsymbol{A}^H(\boldsymbol{f}) = \boldsymbol{A}^T(-\boldsymbol{f}). \tag{6}$$

Thereby, the covariance matrix $\boldsymbol{R}$ can be rewritten as

$$\begin{aligned} \boldsymbol{R} =& \boldsymbol{A}(\boldsymbol{f})\boldsymbol{P}\boldsymbol{A}^T(-\boldsymbol{f}) + \mathrm{diag}(\boldsymbol{\sigma}) \\ =& \boldsymbol{A}_{[K]}^K \mathrm{diag}(\mathrm{vec}(\boldsymbol{P}^T))(\boldsymbol{A}^K(-\boldsymbol{f}))^T \\ &+ \mathrm{diag}(\boldsymbol{\sigma}), \end{aligned} \tag{7}$$

where $\boldsymbol{A}_{[K]}^K = [\boldsymbol{A}_1^K, \ldots, \boldsymbol{A}_K^K]$ with $\boldsymbol{A}_i^K = [\boldsymbol{a}(f_i), \ldots, \boldsymbol{a}(f_i)] \in \mathbb{C}^{M\times K}$ and $\boldsymbol{A}^K(-\boldsymbol{f}) = [\boldsymbol{A}(-\boldsymbol{f}), \ldots, \boldsymbol{A}(-\boldsymbol{f})] \in \mathbb{C}^{M\times K^2}$. Denote by $\boldsymbol{r}^\star = \mathrm{vec}(\boldsymbol{R}^T) \in \mathbb{C}^{M^2}$ the vectorized covariance matrix $\boldsymbol{R}$, then one has

$$\begin{aligned} \boldsymbol{r}^\star =& (\boldsymbol{A}_{[K]}^K \otimes \boldsymbol{A}^K(-\boldsymbol{f}))\mathrm{vec}(\mathrm{diag}(\mathrm{vec}(\boldsymbol{P}^T))) \\ &+ \mathrm{vec}(\mathrm{diag}^T(\boldsymbol{\sigma})) \\ =& \sum_{j=1}^K \sum_{i=1}^K P_{i,j}\boldsymbol{a}(f_i) \otimes \boldsymbol{a}(-f_j) + \mathrm{vec}(\mathrm{diag}(\boldsymbol{\sigma})) \\ =& \sum_{j=1}^K \sum_{i=1}^K P_{i,j}\boldsymbol{c}(f_i, f_j) + \mathrm{vec}(\mathrm{diag}(\boldsymbol{\sigma})), \end{aligned} \tag{8}$$

where $\boldsymbol{c}(f_i, f_j) := \boldsymbol{a}(f_i) \otimes \boldsymbol{a}(-f_j)$, $f_i, f_j \in \Omega$. It's straightforward to find that the first item of (8) is linear combination of elements from the atomic set defined in [9, 10] and the atomic norm minimization (ANM) approach can be applied to obtain the decomposition of the first item of (8) with denoising the second item from $\boldsymbol{r}^\star$.

Remark 1 We refer to (8) as a complete formulation since herein we consider overall $P_{i,j} \neq 0$. In fact, $P_{i,j} (i \neq j)$ will be equal to zero when the two source signals are uncorrelated, resulting in the sum items in (8) less than K^2. We denote the set as the collection of all feasible $[f_i, f_j]$ by Λ, i.e., $\Lambda = \{[f_i, f_j] | f_i, f_j \in \Omega, P_{i,j} \neq 0\}$, where $|\Lambda| = U$ denotes the cardinal number of Λ.

Remark 2 Note that the feasible region of atomic set defined in [9] as $[0, 1) \times [0, 1)$ is different from the feasible region of $\boldsymbol{c}(f_i, f_j)$ which is $[-\frac{1}{2}, \frac{1}{2}) \times [-\frac{1}{2}, \frac{1}{2})$. The equivalence between them can be easily obtained since $\boldsymbol{a}(f)$ is periodic with period one.

According to the atomic norm minimization theory [9,10], and mitigating the effect of noise by eliminating the M entries of $\boldsymbol{r}^\star$ which contain the noisy element σ_i, we have the following atomic norm minimization problem

$$\min_{\boldsymbol{r}} \|\boldsymbol{r}\|_A \ \text{ s.t. } \boldsymbol{\Gamma}_T\boldsymbol{r} = \boldsymbol{\Gamma}_T\boldsymbol{r}^\star, \tag{9}$$

where $\boldsymbol{\Gamma}_T = \mathrm{blkdiag}\{\boldsymbol{E}_1, \ldots, \boldsymbol{E}_M\}$ is the noise elimination operator with $\boldsymbol{E}_m = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{m-1}, \boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_M]^T$ and $\boldsymbol{e}_m$ being the standard orthogonal basis of 1 at the m-th position. $\mathrm{blkdiag}\{\cdot\}$ denotes the block diagonal operator.

In practical applications, $\boldsymbol{R}$ is estimated from the L snapshots as follows

$$\hat{\boldsymbol{R}} = \frac{1}{L}\sum_{l=1}^L \boldsymbol{y}(t_l)\boldsymbol{y}^H(t_l), \tag{10}$$

which contains estimation error since finite snapshots effect. Denote $\triangle\boldsymbol{r} = \boldsymbol{\Gamma}_T(\mathrm{vec}(\hat{\boldsymbol{R}}^T) - \boldsymbol{r}^\star)$ as the partial estimation error, then following from [11] and [12], $\triangle\boldsymbol{r}$ satisfies the asymptotic normal distribution as

$$\triangle\boldsymbol{r} \sim \mathrm{AsN}(0, \boldsymbol{W}^*), \tag{11}$$

where $\boldsymbol{W} = \frac{1}{L}\boldsymbol{\Gamma}_T(\boldsymbol{R}^T \otimes \boldsymbol{R})(\boldsymbol{\Gamma}_T)^T$ and can be approximately estimated as $\hat{\boldsymbol{W}} = \frac{1}{L}\boldsymbol{\Gamma}_T(\hat{\boldsymbol{R}}^T \otimes \hat{\boldsymbol{R}})(\boldsymbol{\Gamma}_T)^T$. It's straightforward to deduce that

$$\|(\hat{\boldsymbol{W}}^*)^{-\frac{1}{2}}\triangle \boldsymbol{r}\|_2^2 \sim \mathrm{As}\chi^2(M^2 - M), \tag{12}$$

where $\mathrm{As}\chi^2(M^2 - M)$ denotes the asymptotic Chi-square distribution with $M^2 - M$ degrees of freedom. A alternative constraint is introduced as

$$\|(\hat{\boldsymbol{W}}^*)^{-\frac{1}{2}}\triangle \boldsymbol{r}\|_2^2 \leq \eta \tag{13}$$

with η making (13) hold with a high probability $1-p$, where p is small value. Imposing the constraint (13) instead of the constraint in (9), we have the following denoising covariance matrix-based atomic norm minimization (DCMANM) method for DOA estimation

$$\hat{\boldsymbol{r}} = \arg\min_{\boldsymbol{r}} \|\boldsymbol{r}\|_A \ \text{ s.t. } \ \|(\hat{\boldsymbol{W}}^*)^{-\frac{1}{2}}\triangle \boldsymbol{r}\|_2^2 \leq \eta. \tag{14}$$

As [9] claims, $\|\boldsymbol{r}\|_A$ has an approximate semidefinite program (SDP) characterization and hence (14) can be solved efficiently using off-the-shelf SDP solver. In this paper, to enhance sparsity and resolution, the reweighted trace minimization (RWTM) proposed in [10] is used to solve (14). We refer interested readers for respective papers for details.

3. SDPM FOR DOA ESTIMATION

When the optimal solution $\hat{\boldsymbol{r}}$ of (14) is given, it's of equal importance to exploit its decomposition where the DOAs can be simultaneously retrieved . One computationally effective method is dual polynomial method (DPM) which is proposed in [13]. It can be simply concluded as two stage. In the first stage, the dual optimal solution $\boldsymbol{q}^\star$ of the prime problem

$$\min_{\boldsymbol{r}} \|\boldsymbol{r}\|_A \ \text{ s.t. } \ \boldsymbol{r} = \hat{\boldsymbol{r}} \tag{15}$$

is obtained by solving the precise SDP presented in [13] using SDP solvers. Then in the second stage, the U unknown frequency couples $(f_i, f_j) \in \Lambda$ can be obtained by located where $|\langle \boldsymbol{q}^\star, \boldsymbol{c}(f_i, f_j)\rangle| = 1$. Readers are referred to [13] for more details. However, to retrieve the DOAs, we actually only need to acquire the set of corresponding frequencies Ω. In other words, only all the distinct frequencies contained in Λ is needed to be determined, i.e., the K frequency couples $(f_i, f_i) \in \Lambda$. The other frequency couples belong to Λ are correspond to the correlated source couples and can be omitted. As a simplification of DPM, the expected K frequency couples can be determined by searching $|\langle \boldsymbol{q}^\star, \boldsymbol{c}(f_i, f_j)\rangle|$ whether equal to one with $f_i = f_j$ and we conclude it as the following corollary:

Corollary 1: Define a simplified dual trigonometric polynomial $\hat{Q}(f)$ as $\hat{Q}(f) := \langle \boldsymbol{q}^\star, \boldsymbol{c}(f, f)\rangle$, then the corresponding frequencies can be localized by identifying the locations where $\hat{Q}(f) = \mathrm{sign}(P_{i,i}) = 1$, i.e.,

$$\begin{aligned} &\hat{Q}(f_i) = 1, \ \forall f_i \in \Omega, \\ &|\hat{Q}(f)| < 1, \ \forall f \notin \Omega, \end{aligned} \tag{16}$$

where $\mathrm{sign}(\cdot)$ represents the complex sign.

Following from Corollary 1, we introduce the peak searching function of the proposed simplified dual polynomial method (SDPM) as

$$P_{\text{SDPM}}(f) = \frac{1}{1 - |\hat{Q}(f)|}, f \in [-\frac{1}{2}, \frac{1}{2}). \tag{17}$$

By locating the peak, we can obtain the corresponding frequencies $f_i, i = 1, \ldots, K$.

Next, a closed-form solution of SDPM designated as Root-SDPM is given. Firstly, consider the simplified dual trigonometric polynomial $\hat{Q}(f)$, define $\boldsymbol{Q} = \{Q_{i,j}\} \in \mathbb{C}^{M\times M}$ as a matrix satisfying $\boldsymbol{q}^\star = \mathrm{vec}(\boldsymbol{Q}^T)$, one has

$$\begin{aligned} \hat{Q}(f) &= \langle \boldsymbol{q}^\star, \boldsymbol{c}(f, f)\rangle = \langle \boldsymbol{q}^\star, \boldsymbol{a}(f) \otimes \boldsymbol{a}(-f)\rangle \\ &= \sum_{(m_1, m_2)\in \boldsymbol{M}} q_{(m_1-1)M+m_2} e^{j2\pi[f,-f][m_1-1, m_2-1]^T} \\ &= \sum_{(m_1, m_2)\in \boldsymbol{M}} Q_{m_1,m_2} e^{j2\pi f(m_1-m_2)} \\ &= \sum_{t=-M+1}^{M-1} \left(\sum \mathrm{diag}(\boldsymbol{Q}, t)\right) e^{j2\pi f t} \\ &= \sum_{t=-M+1}^{M-1} u_t e^{j2\pi f t}, \end{aligned} \tag{18}$$

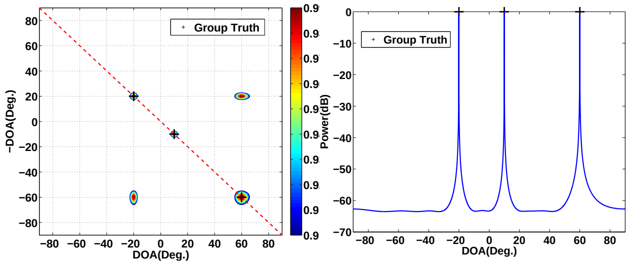
where $\boldsymbol{M} = \{1, \ldots, M\} \times \{1, \ldots, M\}$ denotes the union of the indices, and $u_t = \sum \mathrm{diag}(\boldsymbol{Q}, t)$ is the coefficient corresponding to the degree t. $\mathrm{diag}(\boldsymbol{Q}, t)$ is a vector with the t-th diagonal of $\boldsymbol{Q}$ being its elements. Then, define a trigonometric polynomial of degree $4M - 4$ as

$$\begin{aligned} p_{4M-4}(e^{j2\pi f}) &= e^{j2\pi(2M-2)f}(1 - |\hat{Q}(f)|^2) \\ &= e^{j2\pi(2M-2)f}(1 - \sum_{j=-2M+2}^{2M-2} \mu_j e^{j2\pi f j}), \end{aligned} \tag{19}$$

where $\mu_j = \sum_t u_t u_{t-j}^*$. Note that the corresponding frequencies are the nonzero roots of (19), the phase angles of the corresponding frequencies as $\hat{z}_i$ can be obtained by locating the double roots of $p_{4M-4}(z), z \in \mathbb{C}$ on the unit circle. Finally, with the knowledge of $\hat{z}_i$, $f_i = \frac{1}{2\pi}\arg(\hat{z}_i)$. What's more, as a byproduct, the number of sources can be obtained as the number of double roots of $p_{4M-4}(z)$ on the unit circle.

With the obtained corresponding frequencies, the DOAs can be retrieved by

$$\theta_i = \arcsin(2f_i), \ f_i \in \Omega. \tag{20}$$



(a) DPM in [13]

(b) SDPM

Fig. 1. DOA estimation for correlated sources. the first and the third source are coherent. (a) DPM in [13], (b) Our proposed method (SDPM)

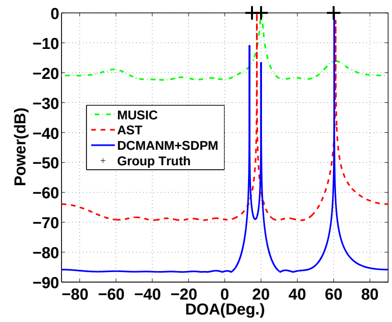


Fig. 2. Spacial Spectra for correlated sources. The first and the third source are coherent.

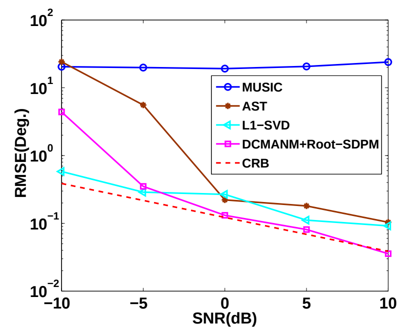


Fig. 3. RMSE comparison of MUSIC, L1-SVD, AST for MMV for two coherent sources with M=10, L=300.

4. NUMERICAL SIMULATIONS

In this section, we present numerical examples to evaluation the performance of the proposed two-stage method (DCMANM+SDPM) for DOA estimation. In simulations, we consider a 10-element ULA and the number of collected snapshots L is set to 300. The parameter η in DCMANM is calculated using the Matlab function chi2inv$(1-p, M^2-M)$, where p is set to 0.1. The proposed DCMANM is implemented using CVX [14].

Firstly, suppose three signals with power [5,5,1] impinge onto the ULA from $[-20°, 10°, 60°]$. Among the signals, the first and the third are coherent and uncorrelated with the second. The SNR is set to 10dB. As shown in Fig.1(a), by treating negative DOAs as another dimension of angle, DPM proposed in [13] not only detects the three true DOAs, but excavates the correlation between the first and the third signals. The proposed peak searching method is shown in Fig.1(b) which can be regarded as a section of Fig.1(a) along the direction of DOA $= -(-$DOA$)$.

We also evaluate the DOA estimation resolution of our proposed method with comparison to MUSIC, L1-SVD and AST for MMV. For computationally effectiveness, AST for MMV is implemented via Alternating Direction Method of Multipliers (ADMM) [6]. We consider three equal-power signals with DOAs $[15°, 20°, 60°]$ impinge onto the ULA. The first signal and the third signal are coherent, but uncorrelated with the second signal. The SNR is set to -3dB. As shown in Fig.2, our proposed method has a higher resolution compared to MUSIC and AST for MMV in low SNR. We omit the result of L1-SVD to enhance the figure visibility where L1-SVD can likewise distinguish these three signals.

Furthermore, we use root mean square error (RMSE) to evaluate DOA estimation precision of our proposed method with the above three algorithm. Assume two coherent signals with equal power impinge onto the ULA from $[-10+\epsilon, 10+\epsilon]$ where ϵ has a uniform distribution in $[-1°, 1°]$. Inhere, Root-SDPM is employed for DOA estimation. The RMSE of DOA estimation against SNR shown in Fig.3 is obtained through a total of 50 trials. From Fig.3, we can see that our proposed method has the best estimation precision than the rest three methods when SNR>-5dB, although our method loses its super-resolution ability when SNR is lower since it detects only one DOA in some runs. Moreover, L1-SVD is comparable with or better than AST for MMV since the iterative grid refinement (IGR) is employed on L1-SVD for accuracy improvement [5] and the ADMM employed converges slowly to an extremely accurate solution [15].

5. CONCLUSIONS

In this paper, a new covariance matrix-based two-staged method for off-grid DOA estimation via atomic norm minimization is proposed. Firstly, in DCMANM, we denoise the vectorized covariance matrix. Then in SDPM, both peak searching method and close-formed root polynomial method are proposed for DOA retrieval. The proposed method is robust to the correlated sources and can be carried out without knowing the source number and the statistics of noise. Numerical simulations demonstrate the outperformance of the proposed method compared to the existing methods.

6. REFERENCES

[1] T. Zisking and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 236, no. 10, pp. 1533-1559, 1988.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol.34, no. 3, pp. 276-280, 1986.

[3] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational in variance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984-995, 1986.

[4] F. Gao and A. B. Gershman, "A generalized esprit approach to direction-of-arrival estimation," *IEEE Signal Processing Letter*, vol. 12, no. 3, pp. 254-257, 2005.

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

[6] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1257-1269, 2016.

[7] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 5145-5157, 2016.

[8] X. Wu, W.-P. Zhu and J. Yan, "Direction-of-arrival estimation based on toeplitz covariance matrix reconstrution," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3236-3240.

[9] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1030-1042, 2015.

[10] Z. Yang, L. Xie and P. Stoica, "Vandermonde decomposition of multilevel toeplitz matrices with application to multidimensional super-resolution," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3685-3701, 2016.

[11] Z. Liu, Z. Huang and Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 1-12, 2013.

[12] C. W. Heltrom, *Probability and stochastic processes for engineers.* New York, NY, USA: Macmillan, 1991.

[13] W. Xu, J.-F. Cai, K. V. Mishra, M. Cho, and A. Kruger, "Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional ($d \leq 2$) off-the-grid frequencies," in *Information Theory and Applications Workshop (ITA)*, 2014, pp. 1-4.

[14] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming," *Online accessible: http://stanford. edu/ boyd/cvx*, 2008.

[15] Z. Yang and L. Xie, "Continuous compressed sensing with a single or multiple measurement vectors," in *IEEE Workshop on Statistical Signal Processing (SSP)*, 2014, pp. 308-311.