GRIDLESS DOA ESTIMATION VIA. ALTERNATING PROJECTIONS

Mark Wagner, Peter Gerstoft

University of California San Diego
Electrical and Computer Engineering
8820 Shellback Way, La Jolla, CA, 92037

Yongsung Park

Seoul National University
Naval Architecture and Ocean Engineering
Seoul, 151-744, South Korea

ABSTRACT

An alternative method for solving the gridless direction-of-arrival (DOA) estimation problem is presented. Gridless DOA estimation involves solving the semi-definite characterization of a rank minimization problem. We show that the original non-convex formulation of the gridless DOA estimation problem can be solved efficiently using the method of alternating projections (AP). We deem our solution 'alternating projections based gridless DOA estimation,' or APG. Using insight from the derivation of APG we present a reduced dimension variation of APG, (RD-APG). The presented algorithms are compared in speed and accuracy to gridless DOA estimation solved using the current state of the art SDP solver.

Index Terms— Gridless DOA estimation, Alternating Projections, Optimization, DOA estimation

1. INTRODUCTION

The process of estimating direction of arrival (DOA) of one or more signals impinging upon an array of sensors is a well studied topic in array signal processing. Because each sensor is located at a slightly different spatial position the phase of each arriving signal will vary across the sensors in a well defined pattern depending on the DOA of the incoming signal. Given knowledge of the sensor positions, the DOA of each signal can be inferred.

Array processing has recently seen advances with the introduction of compressive sensing (CS) [1], which has demonstrated improved resolution is achievable if the solution is known to be sparse. This is ideal for DOA estimation where the number of DOAs is typically small. Many CS based DOA estimation techniques [2, 3] have recently emerged which show DOA estimation can be achieved even from one snapshot of data with correlated sources and noise.

An issue with CS based DOA estimation techniques is that they (generally) require a dictionary of possible phase patterns experienced by the array from impinging signals from a grid of possible directions of arrival. This can result in basis mismatch [4] which occurs when the sources do not fall on a discrete search grid. To overcome basis mismatch, recent works [5–8] propose gridless CS based techniques utilizing a continuous sparsity measure. Gridless methods have shown superior performance to their gridded counterparts [9, 10].

Gridless DOA estimation is achieved by solving a semidefinite program (SDP) which is arrived at as a convex relaxation of a rank minimization problem. Alternating direction method of multipliers (ADMM) [11] is the current state of the art SDP solver and consists of a reasonably fast iteratively updating algorithm based on the augmented Lagrange function. In this paper we propose an alternate method of solving gridless DOA estimation for uniform linear array (ULA) structures using an optimization algorithm known as alternating projections (AP). The proposed method leads to two new algorithms for gridless DOA estimation with improved speed and comparable performance to ADMM.

2. OVERVIEW OF GRIDLESS DOA ESTIMATION

2.1. Signal Model

Assume K narrowband sources are located in the far-field from the array and arrive at the array from directions $\theta_k, k = 1, \ldots, K$. The sensors form a ULA with M sensors. Measurements $\mathbf{Y}$ are considered a weighted superposition of plane waves,

$$\mathbf{Y} = \mathbf{A}_{\mathrm{s}}\mathbf{X} + \mathbf{N} \tag{1}$$

where $\mathbf{Y} \in \mathbb{C}^{M\times L}$, L is the number of snapshots, $\mathbf{A}_{\mathrm{s}} = [\mathbf{a}_{\mathrm{s}}(\theta_1), \ldots, \mathbf{a}_{\mathrm{s}}(\theta_K)]$, $\mathbf{a}_{\mathrm{s}}(\theta) \in \mathbb{C}^{M\times 1}$ is the steering vector which characterizes the phase pattern of the associated plane wave for a certain DOA θ, i.e. $\mathbf{a}_{\mathrm{s}}(\theta) = [1, e^{j(2\pi d/\lambda \sin\theta}, \ldots, e^{j2\pi d/\lambda (M-1)\sin\theta}]^{\mathsf{T}}$, λ is the wavelength, d is the inter-sensor spacing, $\mathbf{X} \in \mathbb{C}^{K\times L}$ is the amplitude of the source signal, $\mathbf{N} \in \mathbb{C}^{M\times L}$ is the measurement noise, and $\theta \in [-90°, 90°]$.

2.2. Gridless DOA Estimation and Rank Minimization

Gridless DOA estimation is performed by exploiting sparsity in the atomic norm of a set of measurements. Atomic norm minimization (ANM) [5, 7, 12] has been widely applied to a variety of estimation problems including DOA estimation [9, 10], and can be applied in both the single and multiple snapshot cases [13]. Gridless DOA estimation, which is based on ANM, uses steering vectors $\mathbf{a}_s(t)$ characterized by continuous variable t as the set of atoms

$$\mathcal{A} = \{\mathbf{A}(t, \mathbf{b}) = \mathbf{a}_{\mathrm{s}}(t)\mathbf{b}^{\mathsf{H}} : t \in [-1, 1), \|\mathbf{b}\|_2 = 1\}. \quad (2)$$

The atomic set defined in (2) is the set of all rank 1 matrices of constrained norm which can be created from an array steering vector and arbitrary measurement. The atomic norm of noiseless measurements $\mathbf{Y}$ is,

$$\begin{aligned} \|\mathbf{Y}\|_{\mathcal{A}} &= \inf\left\{ \sum_{k=1}^{K} \tilde{x}_k : \mathbf{Y} = \sum_{k=1}^{K} \tilde{x}_k \mathbf{a}_{\mathrm{s}}(t_k)\mathbf{b}_k^{\mathsf{H}}, \tilde{x}_k > 0 \right\} \\ &= \inf\left\{ \sum_{k=1}^{K} \|\mathbf{x}_k\|_2 : \mathbf{Y} = \sum_{k=1}^{K} \mathbf{a}_{\mathrm{s}}(t_k)\mathbf{x}_k^{\mathsf{H}} \right\}. \quad (3) \end{aligned}$$

which produces an exact decomposition if the measurements are noise free, and an approximate decomposition otherwise.

We solve the original non-convex l_0 norm minimization formulation of gridless DOA estimation [10], which is defined as

$$\|\mathbf{Y}\|_{\mathcal{A},0} = \inf\left\{ K : \mathbf{Y} = \sum_{k=1}^{K} \mathbf{a}_{\mathrm{s}}(t_k)\mathbf{x}_k^{\mathsf{H}} \right\}. \quad (4)$$

By Theorem 6.9 of [10], $\|\mathbf{Y}\|_{\mathcal{A},0}$ can be found from the optimal solution to the following rank minimization problem,

$$\begin{aligned} &\underset{\mathbf{u},\mathbf{Z}}{\text{minimize}} \quad \operatorname{rank}(\mathbf{T}(\mathbf{u})) \\ &\text{subject to} \quad \mathbf{S} \succeq 0, \end{aligned} \quad (5)$$

where

$$\mathbf{S} = \begin{bmatrix} \mathbf{T}(\mathbf{u}) & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \mathbf{Z} \end{bmatrix}, \quad (6)$$

$\mathbf{Z} \in \mathbb{C}^{L\times L}$ is a free variable which is a conjugate symmetric matrix and $\mathbf{T}(\mathbf{u}) \in \mathbb{C}^{M\times M}$ is a rank $K \leq M$ Hermitian Toeplitz matrix whose first column is $\mathbf{u}$.

Once $\mathbf{T}(\mathbf{u})$ is estimated the angles t_k are recovered through Vandermonde decomposition of $\mathbf{T}(\mathbf{u})$ [14]. The Caratheodory theorem states that any Toeplitz matrix $\mathbf{T}(\mathbf{u})$ can be represented as follows,

$$\mathbf{T}(\mathbf{u}) = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathsf{H}}, \quad (7)$$

where $\mathbf{V} = [\mathbf{a}(t_1), \ldots, \mathbf{a}(t_K)]$ is a Vandermonde matrix, $\mathbf{D} = \operatorname{diag}([\tilde{x}_1, \ldots, \tilde{x}_K])$, and $\tilde{x}_k$ are real positive numbers. The decomposition of (7) is the Vandermonde decomposition. Moreover, from theorem 6.1 of [10], the Vandermonde decomposition of a Toeplitz matrix $\mathbf{T}(\mathbf{u}) \in \mathbb{C}^{M\times M}$ is unique if $\mathbf{T}$ is positive semi-definite (PSD) and $\operatorname{rank}(\mathbf{T}(\mathbf{u})) \leq M$. The Vandermonde decomposition is computed efficiently via several different methods [15, 16], and once t_k is recovered the amplitudes of the sources $\mathbf{x}_k$ are readily inferred.

In general, the optimization of (5) is thought to be NP-hard and is replaced with its convex relaxation which takes the form of an SDP. The contribution of this paper is to show that (5) can be directly solved with speed and accuracy comparable to that of the current state of the art SDP solver applied to the SDP relaxation of (5). The insight gained from our direct optimization of (5) leads us to a high speed reduced dimension variant of gridless DOA estimation which we term 'RD-APG'.

2.3. Alternating Projections

Alternating Projections (AP) is a basic algorithm used to find a point of intersection between two or more convex sets. AP is guaranteed to converge to a solution when each of the sets are closed and convex [17]. When one or more sets are not closed and convex there is no theoretical guarantee on convergence.

2.4. Important Sets and Their Projections

2.4.1. The Rank Constrained Set

Let C be the rank constrained set, narrowly defined here as

$$C = \{\mathbf{A} \in \mathbb{C}^{N\times N} : \operatorname{rank}(\mathbf{A}) \leq K,\ K \leq N\}. \quad (8)$$

The projection operator onto C is

$$P_C(\mathbf{A}) = \sum_{k=1}^{K} \sigma_k \mathbf{u}_k \mathbf{v}_k^{\mathsf{H}} \quad (9)$$

where $\sigma_k, k = [1, \ldots, K]$ are the K singular values of $\mathbf{A}$ with largest magnitude and $\mathbf{u}_k$, $\mathbf{v}_k$ are the corresponding left and right singular vectors.

Set C is non-convex. Regardless, there are some reported instances [18] of AP converging when applied between set C and other convex sets.

2.4.2. The Hermitian Toeplitz Set

Let D be the set of Hermitian Toeplitz matrices. The projection of a matrix onto D is the Toeplitz matrix parameterized by the mean of the elements along the diagonals [19]

$$P_D(\mathbf{A}) = \mathbf{T}(\mathbf{u}), \quad (10)$$

$$u_i = \frac{1}{2(N-i)} \sum_{j=1}^{(N-i)} \mathbf{A}_{j,j+i-1} + \mathbf{A}^*_{j+i-1,j} \quad (11)$$

where $\mathbf{A}_{i,j}$ is the (i, j)th element of $\mathbf{A}$ and u_i is the ith element of $\mathbf{u}$. Set D is convex [20].

2.4.3. The Positive Semi-Definite Set

Let E be the set of complex $N \times N$ PSD matrices

$$E = \{\mathbf{A} \in \mathbb{C}^{N\times N} : \lambda_i \geq 0, \quad \forall\, i\}. \quad (12)$$

Given: $\mathbf{Y} \in \mathbb{C}^{M\times L}, K, \epsilon$
1. Initialize randomly: $\mathbf{T}^{(0)} \in \mathbb{C}^{M\times M}$, $\mathbf{Z}^{(0)} \in \mathbb{C}^{L\times L}$
2. for $i = 0$: max iterations a: $\tilde{\mathbf{T}}^{(i)} = P_D(P_C(\mathbf{T}^{(i)}))$ b: $\mathbf{S}^{(i)} = P_E\left(\begin{bmatrix} \tilde{\mathbf{T}}^{(i)} & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \tilde{\mathbf{Z}}^{(i)} \end{bmatrix}\right)$ c: if $\frac{1}{M+L}\|\mathbf{S}^{(i)} - \mathbf{S}^{(i-1)}\|_{\mathcal{F}} \leq \epsilon$, break d: $\mathbf{T}^{(i+1)} = \mathbf{S}^{(i)}(1:M, 1:M)$ e: $\mathbf{Z}^{(i+1)} = \mathbf{S}^{(i)}(M+1:M+L, M+1:M+L)$
3. Decompose: $\mathbf{T} = \mathbf{V}\mathbf{D}\mathbf{V}^H$
The array steering vectors are the columns of $\mathbf{V}$

Table 1. The APG algorithm

where λ_i is the ith largest eigenvalue of $\mathbf{A}$. The set E, forms a cone and is convex [21]. The projection operator into E is

$$P_E(\mathbf{A}) = \sum_{i=1}^{N} \max(0, \lambda_i)\mathbf{e}_i\mathbf{e}_i^{\mathsf{H}} \tag{13}$$

where $\mathbf{e}_i$ is the ith eigenvector of $\mathbf{A}$.

2.4.4. The Set of Matrices with Given Columnspace

Let F be the set of complex $N \times M$ matrices spanning a given K dimensional columnspace. All matrices spanning linear subspaces form convex sets. For any matrix $\mathbf{B} \in F$ and $\mathbf{A} \in \mathbb{C}^{M\times P}$, $P \geq K$, the projection onto F is

$$P_F(\mathbf{A}) = \mathbf{B}\mathbf{B}^{\dagger}\mathbf{A} \tag{14}$$

where † is the Moore-Penrose Pseudo-inverse. It is well known that $\mathrm{rank}(P_F(\mathbf{A})) \leq K$ [22].

3. GRIDLESS DOA ESTIMATION USING ALTERNATING PROJECTIONS (APG)

The gridless DOA estimation problem of (5) is a search for the rank K Toeplitz matrix which completes the PSD matrix defined in (6). Because all principal submatrices of a PSD matrix are also PSD (and symmetric) we know that $\mathbf{S}, \mathbf{T}(\mathbf{u}), \mathbf{Z} \in E, F$. We also know

$$\mathrm{rank}\Big(\mathbf{T}(\mathbf{u})\Big) = \mathrm{rank}\Big(\mathbf{A}_s\mathbf{D}\mathbf{A}_s^{\mathsf{H}}\Big) = \mathrm{rank}\Big(\mathbf{A}_s\Big) = K. \tag{15}$$

From this trivial analysis we can classify each matrix from (5) into one or more of the sets defined in section 2.4.

$$\begin{aligned} \mathbf{T}(\mathbf{u}) &\in C, D, E, \\ \mathbf{Z} &\in E, \\ \mathbf{S} &\in E. \end{aligned} \tag{16}$$

From here (5) is solvable via AP by iteratively projecting each matrix to its proper sets. We note it is unnecessary to project $\mathbf{T}(\mathbf{u})$ and $\mathbf{Z}$ into set E because they are principal submatrices of $\mathbf{S}$ and will become members of E when $\mathbf{S}$ is projected to E. The proposed APG algorithm is given in table 1.

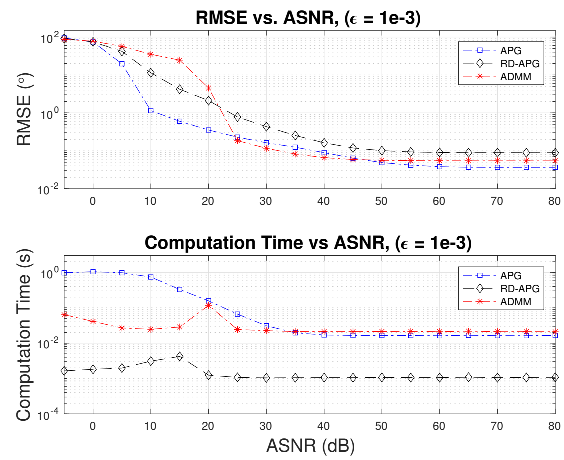

Fig. 1. Top: RMSE vs. ASNR using $\epsilon = .001$, $K = 3$, $M = 20$, $L = 30$. Each point gives RMSE from 250 trials. Bottom: Mean computation time.

APG works as follows: given random initializations of $\mathbf{T}(\mathbf{u})$ and $\mathbf{Z}$, $\mathbf{T}(\mathbf{u})$ is projected to C and D using (9) and (11), $\mathbf{S}$ is then constructed from the resulting projections. $\mathbf{S}$ is projected onto E using (13) and the next estimates of $\mathbf{T}(\mathbf{u})$, $\mathbf{Z}$ are taken from the resulting $\mathbf{S}$ matrix. The process is iterated until convergence. It is assumed that K is known.

3.1. Reduced Dimension APG (RD-APG)

Reduced Dimension APG (RD-APG) is a computationally faster variant of APG which eliminates the SDP formulation of (5). Observe that by substituting (1) and (15) into (5) the optimal $\mathbf{S}$ in the noiseless case is

$$\mathbf{S} = \begin{bmatrix} \mathbf{A}_s\mathbf{D}\mathbf{A}_s^{\mathsf{H}} & \mathbf{A}_s\mathbf{X} \\ (\mathbf{A}_s\mathbf{X})^{\mathsf{H}} & \mathbf{Z} \end{bmatrix}. \tag{17}$$

In this form it is easy to verify that

$$\mathrm{colsp}(\mathbf{Y}) = \mathrm{colsp}(\mathbf{A}_s) = \mathrm{colsp}(\mathbf{T}(\mathbf{u})) \tag{18}$$

where colsp() is the columnspace. By uniqueness of the Vandermonde decomposition any Toeplitz PSD matrix with columnspace equal to $\mathbf{A}_s$ must then decompose into $\mathbf{A}_s$ and some diagonal matrix $\mathbf{D}$. Therefore the optimal $\mathbf{T}(\mathbf{u})$ belongs to F and has columnspace equal to the signal subspace of the measurements. The signal subspace is estimated from the sample covariance matrix

$$\mathbf{R}_{yy} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^{\mathsf{H}} = \mathbf{E}_S\mathbf{\Lambda}_S\mathbf{E}_S^{\mathsf{H}} + \mathbf{E}_N\mathbf{\Lambda}_N\mathbf{E}_N^{\mathsf{H}} \tag{19}$$

Given: $\mathbf{Y} \in \mathbb{C}^{M\times L}$, $\mathbf{E}_S \in \mathbb{C}^{M\times K}$, K, ϵ
1. Initialize: Randomly generate $\mathbf{T}^{(0)} \in \mathbb{C}^{M\times M}$
2. for $i = 0$: max iterations a: $\mathbf{T}^{(i)} = \mathbf{E}_S\mathbf{E}_S^{\dagger}(\mathbf{T}^{(i)})$ b: $\mathbf{T}^{(i)} = P_E(\mathbf{T}^{(i)}))$ c: $\mathbf{T}^{(i)} = P_D(\mathbf{T}^{(i)}))$ d: if $\frac{1}{M}\lVert\mathbf{T}^{(i)} - \mathbf{T}^{(i-1)}\rVert_{\mathcal{F}} \leq \epsilon$, break
3. Decompose: $\mathbf{T} = \mathbf{VDV}^H$ The array steering vectors are the columns of $\mathbf{V}$

Table 2. The RD-APG algorithm

where $\mathbf{\Lambda}_S$ is a diagonal matrix containing the 'strong' eigenvalues of $\mathbf{R}_{yy}$, and $\mathbf{E}_S$ is the signal subspace which is composed of the corresponding eigen-vectors [23]. $\mathbf{\Lambda}_N$ and $\mathbf{E}_N$ contain the remaining eigenvalues and eigen-vectors.

The SDP formulation of (5) may therefore be replaced by AP performed between sets D, E, and F. This new formulation solves directly for $\mathbf{T}(\mathbf{u})$ rather than $\mathbf{S}$, reducing the dimension of the problem by L. RD-APG is detailed in table 2.

3.2. Handling Noise

We have observed that the APG and RD-APG algorithms are robust to noise. When noise is introduced into the signal model there will almost certainly be no solution to the optimization of (5). However, this is not an issue because the AP algorithm will converge to points of impasse when no solution exists. The matrix $\mathbf{T}(\mathbf{u})$ after the final projection to sets C and D will be the rank K Toeplitz matrix which most nearly completes the desired PSD $\mathbf{S}$ matrix. Despite the inclusion of noise, this $\mathbf{T}(\mathbf{u})$ matrix can still be decomposed to meaningful estimates of the DOAs.

4. SIMULATION

APG and RD-APG were tested on simulated signals impinging on a ULA of M sensors modeled by (1). An ADMM based algorithm for gridless beamforming detailed in [10] was tested against the proposed AP based algorithms. Each simulation had K randomly generated DOAs chosen such that $|\sin(\theta_i) - \sin(\theta_j)| \leq .1$ for $i \neq j$. Signals contained in the rows of $\mathbf{X}$ and noise contained in the rows of $\mathbf{N}$ were drawn from an L dimensional zero mean complex Gaussian distribution such that $\mathrm{E}[\mathbf{XX}^{\mathsf{H}}] = \sigma_x^2\mathbf{I}$ and $\mathrm{E}[\mathbf{NN}^{\mathsf{H}}] = \sigma_n^2\mathbf{I}$. The signals were generated to meet a specified array SNR (ASNR), where ASNR is

$$\mathrm{ASNR} = 10\log_{10}\left(\frac{M\sigma_x^2}{\sigma_n^2}\right). \quad (20)$$

Vandermonde decomposition was computed via Pisarenko harmonic decomposition. DOA estimation accuracy was calculated as root mean squared error (RMSE) between true and estimated angles,

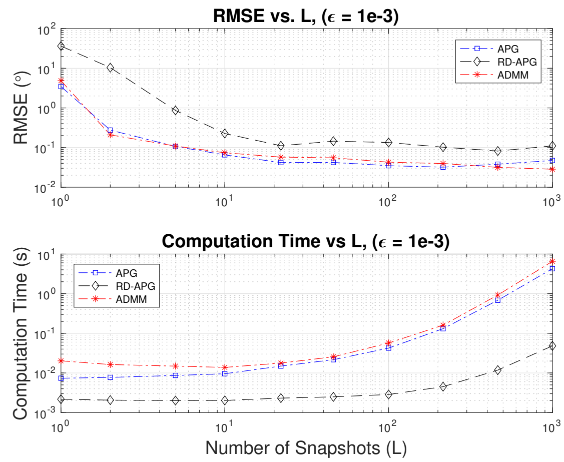

Fig. 2. Top: RMSE vs. L using $\epsilon = .001$, $K = 3$, $M = 20$, $ASNR = 30$ dB. Each point gives RMSE from 250 trials. Bottom: Mean computation time.

$$\mathrm{RMSE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(\theta_k - \hat{\theta}_k)^2}. \quad (21)$$

where $\hat{\theta}_k$ is the estimated DOA of the kth signal.

Figures 1 and 2 compare APG, RD-APG, and ADMM applied to simulated measurements using $\epsilon = .001$ and ADMM τ parameter varied between .01 and 10 to achieve best performance at each ASNR. Results showed APG and ADMM have similar performance. APG is superior at low SNR for many snapshots, RD-APG is best for few snapshots and low SNR, and ADMM is superior for many snapshots. The results suggest that AP based algorithms are competitive with ADMM in most situations, and are better at rejecting noise in low ASNR cases.

5. CONCLUSION

A new method for solving the optimization problem induced by gridless DOA estimation based on the AP algorithm was proposed and compared to the current state of the art optimization algorithm. Insights gained from the AP based algorithm lead to a reduced dimension variant of gridless DOA estimation which was shown to have significantly improved computation time. The two versions of AP based gridless DOA estimation were compared against the current state of the art solution and shown to have similar performance at lower computational cost.

6. REFERENCES

[1] E.J. Candès and M.B. Wakin, "An introduction to compressive sampling," *IEEE Sig. Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[2] G.F. Edelmann and C.F. Gaumond, "Beamforming using compressive sensing," *J. Acoust. Soc. Am.*, vol. 130, no. 4, pp. 232–237, 2011.

[3] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Am.*, vol. 136, no. 1, pp. 260–271, 2014.

[4] Y. Chi, L.L. Scharf, A. Pezeshki, and A.R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal. Process.*, vol. 59, no. 5, pp. 2182–2195, 2011.

[5] G. Tang, B.N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[6] B.N Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal. Process.*, vol. 61, no. 23, pp. 5987–5999, 2013.

[7] E.J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[8] Y. Park, W. Seong, and Y. Choo, "Compressive time delay estimation off the grid," *J. Acoust. Soc. Am.*, vol. 141, no. 6, pp. 585–591, 2017.

[9] A. Xenaki and P. Gerstoft, "Grid-free compressive beamforming," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1923–1935, 2015.

[10] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," *arXiv preprint arXiv:1609.09596*, 2016.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends. Mach. Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[12] R. Heckel, V.I. Morgenshtern, and M. Soltanolkotabi, "Super-resolution radar," *J. Info. and Infer.*, vol. 5, no. 1, pp. 22–75, 2016.

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

[14] V.F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. Int.*, vol. 33, no. 3, pp. 347–366, 1973.

[15] Y. Hua and T.K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 5, pp. 814–824, 1990.

[16] G.H. Golub and C.F. Van Loan, *Matrix computations*, vol. 3, JHU Press, 2012.

[17] H.H. Bauschke and J.M. Borwein, "On the convergence of von-Neumann's alternating projection algorithm for two sets," *Set-Valued Analysis*, vol. 1, no. 2, pp. 185–212, 1993.

[18] T. Cai and A. Zhang, "Rop: Matrix recovery via rank-one projections," *Ann. Stat.*, vol. 43, no. 1, pp. 102–138, 2015.

[19] M.G. Eberle and M.C. Maciel, "Finding the closest Toeplitz matrix," *J. Comp. & Applied Mathe.*, vol. 22, no. 1, pp. 1–18, 2003.

[20] A. Dembo, C.L. Mallows, and L.A. Shepp, "Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1206–1212, 1989.

[21] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[22] G. Strang, *Introduction to linear algebra*, vol. 3, Wellesley-Cambridge Press Wellesley, MA, 1993.

[23] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas. Prop.*, vol. 34, no. 3, pp. 276–280, 1986.