Gridless DOA Estimation and Root-MUSIC for Non-Uniform Linear Arrays

Mark Wagner, *Member, IEEE*, Yongsung Park, *Member, IEEE*, and Peter Gerstoft, *Senior Member, IEEE*

***Abstract*—Gridless direction of arrival (DOA) estimation is addressed for a 1-D non-uniform array (NUA) of arbitrary geometry. Currently, gridless DOA estimation is solved via convex relaxation, and is applicable only to uniform linear arrays (ULA). The ULA sample covariance matrix has Toeplitz structure, and gridless DOA is based on the Vandermonde decomposition of this matrix. The Vandermonde decomposition breaks a Toeplitz matrix into its harmonic components, from which the DOAs are estimated. First, we present the irregular Toeplitz matrix and irregular Vandermonde decomposition (IVD), which generalizes the Vandermonde decomposition. It is shown that IVD is related to the MUSIC and root-MUSIC algorithms. Next, atomic norm minimization (ANM) for gridless DOA is generalized to NUA using the IVD. The resulting non-convex optimization is solved using alternating projections (AP). Simulations show the AP based solution for NUA/ULA has similar accuracy as convex relaxation of gridless DOA for ULA.**

***Index Terms*—Continuous compressed sensing, DOA estimation, gridless DOA estimation, matrix decomposition, optimization, root-MUSIC.**

I. Introduction

ESTIMATING the direction of arrival (DOA) of one or more signals arriving at an array of sensors is an important topic in array signal processing and has a wide range of applications in radar, sonar, wireless communications, etc. Recently, the focus of DOA estimation has turned from classical subspace based DOA algorithms including MUSIC, root-MUSIC, and ESPRIT [1]–[4] to newer high resolution [5], [6] and compressive sensing based methods such as compressive DOA [7]–[9]. Compressive methods have the advantage that they are high resolution (resolve nearby DOAs), and require only a single measurement snapshot.

Early compressive DOA techniques approximated the measurements as a linear combination of a few DOA array patterns picked from a finite grid of possible DOA array patterns. However, this technique suffers from errors due to grid mismatch because true DOAs are not on a grid [10]. As a response came the family of off-grid, or *gridless* methods, which exploit sparsity in the *atomic norm* of the measurements [11]–[17]. The atomic norm is the minimum number of "atoms" from a manifold required to reconstruct a vector, thus gridless DOA is the continuous analog to gridded compressive DOA.

Gridless DOA is an application of continuous compressed sensing (CCS) spectral estimation, which was introduced in [11]. There have been many adaptations of CCS to related problems [18], [19], however, the atomic norm minimization (ANM) formulation of gridless DOA for 1-D arrays is the focus of this work [20], [21]. There are other similar gridless DOA algorithms such as the enhanced matrix completion (EMaC) method [22] and gridless SPICE [14] based on the principals of CSS, but formulated differently from ANM. See [23] for a comprehensive review of CCS for DOA.

CCS solves a semi-definite programming problem (SDP) whose objective is to find the lowest rank Toeplitz matrix which can explain the measurements. The frequencies composing the signal are then recovered through Vandermonde decomposition of the optimal Toeplitz matrix [24]. This decomposition is known to be unique when the Toeplitz matrix is rank deficient [4]. In the context of DOA estimation, the parameters of the Vandermonde matrices (known as harmonics) indicate the DOAs. The weakness of gridless methods is that they are limited to regularly sampled measurements that can only be taken from a uniform linear array (ULA).

Early work towards extending gridless DOA to non-uniform arrays (NUAs) was focused around special array geometries such as minimum redundancy array (MRA) [16], co-prime arrays [25], [26], and sparse arrays [27]. While gridless DOA in general does not extend to arbitrary array geometries, these special arrays form a small subset of NUAs for which gridless DOA is still applicable (as shown in Section III-B). The novelty of such arrays was never their NUA geometry, but their ability to resolve more sources than sensors composing the array. This work is concerned towards extending the usefulness of gridless DOA past ULAs, to arrays of arbitrary 1D geometry. The term 'NUA' in this work refers to any array of arbitrary 1D geometry.

There has been some recent effort towards extending gridless DOA to NUAs [28]–[30], which is achieved by approximating the NUA array manifold by a Fourier series. This idea traces back to "Fourier domain root-MUSIC" [31]. There also exist older array interpolation techniques which transform the ULA manifold to a NUA manifold [32]–[38], however, this technique is limited by a built in tradeoff between accuracy and computational complexity. The proposed algorithm is based on a new matrix decomposition. Other recent works involve generalization of gridless DOA to higher dimensional array

Manuscript received February 27, 2020; revised October 13, 2020 and February 26, 2021; accepted March 17, 2021. Date of publication March 23, 2021; date of current version April 16, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xue Jiang. This work was supported by the Office of Naval Research (ONR) under Grant N00014-18-1-2400. *(Corresponding author: Mark Wagner.)*

The authors are with the Scripps Institute of Oceanography, The University of California San Diego, La Jolla, CA 92093 USA (e-mail: m2wagner@eng.ucsd.edu; yop001@ucsd.edu; gerstoft@ucsd.edu).

Digital Object Identifier 10.1109/TSP.2021.3068353

geometries [39], [40], though the work presented in this paper is limited to 1D arrays.

The primary contribution of this work is to generalize the ANM formulation of gridless DOA to NUAs based on a new matrix decomposition applicable to NUA measurements. We define two new matrices, the *irregular Toeplitz* and *irregular Vandermonde* matrices. The word 'irregular' refers to the irregular sampling of the wave field by a NUA. The irregular Toeplitz matrix is decomposed into irregular Vandermonde components, similar to Toeplitz and Vandermonde matrices. This allows for gridless DOA (and root-MUSIC) to be extended to NUAs.

The secondary contribution of this work is a computationally efficient solution to the non-convex rank minimization problem resulting from the generalization to NUAs. Non-convex optimization problems are typically substituted for their convex relaxation and solved using a general solver [41], such as the alternating directions method of multipliers (ADMM) [42]. However, the extension to NUAs cannot be easily cast to its convex relaxation. Instead, we propose using the alternating projections (AP) algorithm [43, pg. 606].

The AP algorithm has been used to solve CCS problems [44], and implementation of AP to gridless DOA for NUAs is formulated by projecting onto the irregular Toeplitz set. The use of AP for solving similar non-convex optimization problems has shown promising results [44]–[46]. A comparison is provided which reveals the proposed AP solution has similar or superior performance to competing algorithms, while maintaining the ability to generalize to the NUA case.

The paper is structured as follows: Section II contains an overview of gridless DOA and relevant concepts such as the Vandermonde decomposition and root-MUSIC algorithm. In Section III, the IVD is introduced, which allows any positive semi-definite (PSD) matrix to be decomposed into harmonics sampled at known irregular intervals (NUA sensor locations). The IVD allows for non-Toeplitz matrices to be decomposed into their harmonic components. In Section IV, a modification of the gridless DOA problem is presented such that it can be applied to NUAs. A solution to the reformulated gridless DOA problem based on the AP algorithm is proposed. In Section V, the proposed algorithm is tested on simulated measurements from both ULAs and NUAs. The results are for DOA estimation, but are also applicable to the broader CCS problem to irregularly sampled signals.

In the remainder of this paper, lowercase bold letters represent vectors and uppercase bold letters represent matrices. Below is a short list of notation that will be used.

- $^{\mathsf{T}}$ Vector or matrix transpose.
- * Complex conjugate.
- $^{\mathsf{H}}$ Vector or matrix conjugate transpose.
- † Moore-Penrose pseudo-inverse.
- $\angle$ Phase angle.
- $\perp$ Orthogonal.
- $\succeq$ Positive semi-definite.
- $\|.\|_2$ Two norm.
- $\|.\|_F$ Frobenious norm.
- $\mathbb{E}$ Expected value.
- Tr Matrix trace.
- $\mathrm{diag}(\mathbf{X})$ Diagonal elements of matrix $\mathbf{X}$.
- $\mathrm{diag}(\mathbf{x})$ Diagonal matrix with elements $\mathbf{x}$.
- $\mathrm{span}(\mathbf{X})$ columnspace of matrix $\mathbf{X}$.
- $\mathbf{I}$ Identity matrix.
- $\mathbf{1}$ Ones matrix.
- $\mathbf{0}$ Zeros matrix.
- $\mathbf{x}^y$ Element-wise exponentiation.
- $\mathcal{U}(a,b)$ Uniform distribution from a to b.

II. Background: DOA Estimation for ULAs

A. Model Framework

Consider an array of M sensors receiving uncorrelated signals from K narrowband sources located in the far field of the array. Signals from each source arrive from angles $\boldsymbol{\theta} = [\theta_1 \ \dots \ \theta_K]^{\mathsf{T}}$, given in radians. The sources are assumed to be in the same plane as the array, and the sensors are positioned at points on a line given by sensor position vector $\mathbf{r} = [r_1 \ \dots \ r_M]^{\mathsf{T}}$, where each value r_i denotes the distance of sensor i from an arbitrary origin point in units of half-wavelengths. By casting $\mathbf{r}$ in units of half-wavelengths, many of the equations surrounding DOA estimation are simplified. For L measurement snapshots, the array measurements $\mathbf{Y} \in \mathbb{C}^{M\times L}$ are modeled as

$$\begin{aligned} \mathbf{Y} &= \mathbf{Z} + \mathbf{N}, \\ \mathbf{Z} &= \mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta})\mathbf{X}, \end{aligned} \tag{1}$$

where $\mathbf{X} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_K]^{\mathsf{T}} \in \mathbb{C}^{K\times L}$ contains the signals from each of K sources, Gaussian uncorrelated noise is contained in $\mathbf{N} \in \mathbb{C}^{M\times L}$, and

$$\mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta}) = [\mathbf{a}(\mathbf{r}, \theta_1) \ \dots \ \mathbf{a}(\mathbf{r}, \theta_K)], \tag{2}$$

$$\mathbf{a}(\mathbf{r}, \theta) = [e^{-j\pi r_1 \sin\theta} \ \dots \ e^{-j\pi r_M \sin\theta}]^{\mathsf{T}}, \tag{3}$$

is the array manifold matrix whose columns model the phase pattern across the array from DOA θ. The column vectors are known as *array manifold vectors* and are defined over $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})$. The goal is to recover $\boldsymbol{\theta}$ given knowledge of the sensor positions $\mathbf{r}$ and measurements $\mathbf{Y}$.

B. Vandermonde and Toeplitz Matrices

A Vandermonde matrix $\mathbb{C}^{M\times K}$ is defined as

$$\begin{aligned} \mathbf{V}(\boldsymbol{z}) &= [\boldsymbol{z}^0 \ \boldsymbol{z}^1 \ \dots \ \boldsymbol{z}^{(M-1)}]^{\mathsf{T}}, \\ &= [\mathbf{v}(z_1) \ \dots \ \mathbf{v}(z_K)], \end{aligned} \tag{4}$$

where $\boldsymbol{z} = [z_1 \ \dots \ z_K]^{\mathsf{T}}$ are known as harmonics, and fully parameterize the Vandermonde matrix [47, p. 409]. A column of the Vandermonde matrix is defined as $\mathbf{v}(z) = [1 \ z^1 \ \dots \ z^{M-1}]^{\mathsf{T}}$.

The array manifold matrix of a ULA has Vandermonde structure. Sensor positions of a ULA are described as

$$\mathbf{r}_{\mathrm{ULA}} = \alpha[0 \ 1 \ \dots \ M-1]^{\mathsf{T}} + \beta, \tag{5}$$

where α is array spacing in wavelengths and β is an arbitrary offset. For ULAs the angles, $\boldsymbol{\theta}$, from (2) are related to the

harmonics, $\boldsymbol{z}$, of (4) by

$$z_k = e^{-j\pi(\alpha+\beta)\sin(\theta_k)}, \tag{6}$$

$$\theta_k = -\sin^{-1}\left(\frac{\angle z_k}{(\alpha+\beta)\pi}\right). \tag{7}$$

Conceptually, the values of $\mathbf{r}_{\text{ULA}}$ are scaled and shifted integers, and the classical choice is $(\alpha, \beta) = (1, 0)$, which corresponds to a ULA with half-wavelength sensor spacing.

The set of Toeplitz matrices, $\boldsymbol{\mathcal{T}}$, is defined as

$$\boldsymbol{\mathcal{T}} = \left\{ \mathbf{T} : \mathbf{T} = \begin{bmatrix} u_1 & u_2^* & \cdots & u_M^* \\ u_2 & u_1 & \cdots & u_{M-1}^* \\ \vdots & \vdots & \ddots & \vdots \\ u_M & u_{M-1} & \cdots & u_1 \end{bmatrix} \right\}. \tag{8}$$

The first column, $\mathbf{u}$, fully parameterizes a Toeplitz matrix $\mathbf{T} \in \boldsymbol{\mathcal{T}}$. Thus Toeplitz matrices are often expressed as $\mathbf{T}(\mathbf{u})$.

Any Toeplitz matrix, $\mathbf{T}(\mathbf{u})$, can be decomposed into Vandermonde components,

$$\mathbf{T}(\mathbf{u}) = \mathbf{V}(\boldsymbol{z})\mathbf{D}\mathbf{V}(\boldsymbol{z})^{\mathsf{H}}, \tag{9}$$

where $\mathbf{D} \in \mathbb{R}^{K\times K}$ is diagonal with positive entries, and the harmonics $\boldsymbol{z}$ have unit magnitude. The decomposition of (9) is unique when the Toeplitz matrix is not full rank [24].

The decomposition (9) will be referred to as Vandermonde decomposition.

C. Root-MUSIC and Vandermonde Decoposition

Root-MUSIC is a subspace based DOA algorithm [1], [2]. The algorithm consists of forming a polynomial from the ULA sample covariance matrix. The complex roots of the polynomial, $\boldsymbol{z}$, are used to estimate the DOAs through (7). Without noise, the sample covariance matrix is Toeplitz and the polynomial roots are the $\boldsymbol{z}$ harmonics of its Vandermonde decomposition (9). Root-MUSIC is valid only for measurements taken at a ULA.

The sample covariance matrix, $\mathbf{R}_{yy}$, is

$$\mathbf{R}_{yy} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}, \quad \mathbb{E}[\mathbf{R}_{yy}] = \left(\mathbf{A}_s\boldsymbol{\Sigma}_X\mathbf{A}_s^{\mathsf{H}} + \boldsymbol{\Sigma}_N\right),$$

$$\boldsymbol{\Sigma}_X = \frac{1}{L}\mathbb{E}\left[\mathbf{X}\mathbf{X}^{\mathsf{H}}\right], \quad \boldsymbol{\Sigma}_N = \frac{1}{L}\mathbb{E}\left[\mathbf{N}\mathbf{N}^{\mathsf{H}}\right] = \sigma_N^2\mathbf{I}, \tag{10}$$

where $\boldsymbol{\Sigma}_N$ is the expected noise covariance matrix, which is $\sigma_N^2\mathbf{I}$ for uncorrelated Gaussian noise with variance σ_N^2.

For noiseless measurements $\text{span}(\mathbf{R}_{yy}) = \text{span}(\mathbf{A}_s)$. In this case, the columnspace and left nullspace of $\mathbf{R}_{yy}$ are known as the *signal and noise subspaces* respectively. When noise is present an estimated basis to both spaces is found from the eigen-decomposition of $\mathbf{R}_{yy}$,

$$\mathbf{R}_{yy} = \mathbf{U}_S\boldsymbol{\Lambda}_S\mathbf{U}_S^{\mathsf{H}} + \mathbf{U}_N\boldsymbol{\Lambda}_N\mathbf{U}_N^{\mathsf{H}}, \tag{11}$$

where $\boldsymbol{\Lambda}_S \in \mathbb{R}^{K\times K}$ is diagonal with the K largest eigenvalues of $\mathbf{R}_{yy}$, $\mathbf{U}_S \in \mathbb{C}^{M\times K}$ contains the corresponding eigenvectors, and $\boldsymbol{\Lambda}_N \in \mathbb{R}^{(M-K)\times(M-K)}$ and $\mathbf{U}_N \in \mathbb{C}^{M\times(M-K)}$ contain the remaining noise eigenvalues and eigenvectors. $\mathbf{U}_S$ and $\mathbf{U}_N$ are bases to the signal and noise subspaces respectively, which can only be estimated when $\text{rank}(\mathbf{R}_{yy}) \geq K$. The sources must be incoherent because multiple coherent sources can occupy the same 1-D subspace, and $L \geq K$ because at least K snapshots are required to build a rank K approximation of $\mathbf{R}_{yy}$.

For noise free measurements, $\mathbf{U}_N$ is composed of eigenvectors whose corresponding eigenvalues are 0. In this case $\mathbf{U}_N$ exactly spans the left nullspace of $\mathbf{A}_s$ and

$$\mathbf{U}_N \perp \mathbf{A}_s \Rightarrow \mathbf{U}_N^{\mathsf{H}}\mathbf{A}_s = \mathbf{0}, \tag{12}$$

When noise is present, the eigenvectors composing $\mathbf{U}_N$ have non-zero corresponding eigenvalues and $\mathbf{U}_N$ approximates of the noise subspace.

For a ULA, all array manifold vectors $\mathbf{a}(\theta)$ have a corresponding Vandermonde vector, $\mathbf{v}(z)$. From (12), the *null spectrum* is formed [4, p.1159]

$$\begin{aligned} D(z) &= \|\mathbf{U}_N^{\mathsf{H}}\mathbf{a}(\theta)\|_2^2 = \mathbf{a}(\theta)^{\mathsf{H}}\mathbf{U}_N\mathbf{U}_N^{\mathsf{H}}\mathbf{a}(\theta), \\ &= \mathbf{v}\left(\frac{1}{z}\right)^{\mathsf{T}}\mathbf{U}_N\mathbf{U}_N^{\mathsf{H}}\mathbf{v}(z), \end{aligned} \tag{13}$$

for $|z| = 1$. Defining the $M \times M$ matrix

$$\mathbf{G} = \mathbf{U}_N\mathbf{U}_N^{\mathsf{H}}, \tag{14}$$

the null spectrum is expanded into a polynomial in z,

$$d_i = \sum_{m_1,m_2}^{M} \mathbf{G}_{m_1,m_2}, \quad (m_1 - m_2) = i,$$

$$D(z) = d_{-(M-1)}z^{-(M-1)} + \cdots + d_{M-1}z^{M-1}, \tag{15}$$

where d_i are the sums along the diagonals of $\mathbf{G}$. Some properties of $D(z)$, $z \in \mathbb{C}$, are as follows:

1) $d_i = d_{-i}^*$, because $\mathbf{G}$ is Hermitian.
2) From (12), K root pairs appear on the unit circle at phase angles $\boldsymbol{\theta}$ in the noise free case, [2]. Noise in the measurements causes deflection of the roots from their noise free positions proportional to the noise magnitude.
3) If one root exists at $\tilde{z}$, another root exists at $\frac{1}{\tilde{z}^*}$ because $D(z) = D(\frac{1}{z^*})$.

For root-MUSIC, the K roots inside the unit circle with largest magnitude, $|z|$, are taken as DOA estimates using the mapping of (7), [4]. These roots produce highly reliable DOA estimates. When noise is present, the accuracy of the DOA estimates is related to how well $\mathbf{U}_N$ approximates the true noise subspace.

Note that $D(z)$ was derived for z on the unit circle, but its roots located off the unit circle are used as DOA estimates. There is no physical reason behind this decision. Rather, it is a useful mathematical trick. The polynomial expansion of $D(z)$ is especially useful because its roots are calculated efficiently.

Fig. 1(a) and (b) depict an example null spectrum from a half-wavelength spaced ULA, which highlights points ii and iii. Even when the measurements are severely corrupted by noise, the null spectrum gives good estimates of the DOAs.

When $\mathbf{R}_{yy}$ is noise free, it is rank K with Toeplitz structure. Additionally, the K (double) roots on the unit circle of its null spectrum are the harmonics, $\boldsymbol{z}$, of the Vandermonde decomposition of $\mathbf{R}_{yy}$, see (9). Thus root-MUSIC doubles as a method to perform Vandermonde decomposition on a Toeplitz

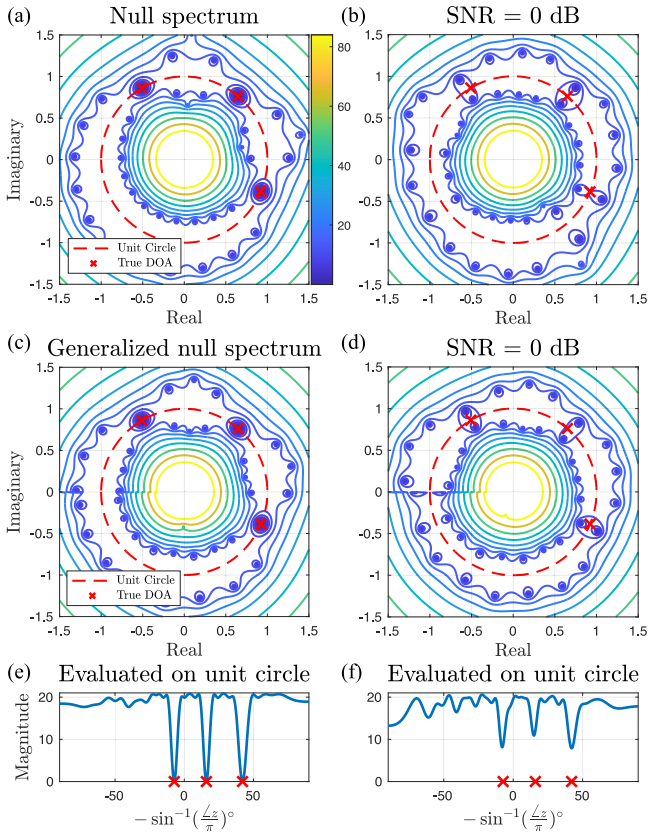

Fig. 1. Null spectrum (13) and irregular null spectrum (29) contours (dB) using $M = 21$, $K = 3$, $L = 10$. DOAs at $\boldsymbol{\theta} = [-7.2, 15.9, 42.1]^\circ$, marked by red x's, red dashed line marks complex unit circle. Measurements generated from (1). (a) Null spectrum of from noiseless ULA measurements. (b) Null spectrum from ULA measurements, SNR = 0 dB. (c) Irregular null spectrum for noiseless NUA measurements (d) Irregular null spectrum of for NUA measurements, SNR = 0 dB. (e) Evaluation of '(c)' on unit circle. (f) Evaluation of '(d)' on unit circle.

matrix. Because root-MUSIC also applies for sample covariance matrices corrupted by noise, it is a method for approximating the harmonics, $\boldsymbol{z}$, of a 'noisy' Toeplitz matrix.

D. Gridless DOA for Uniform Linear Arrays

In gridless DOA estimation, the DOAs composing $\mathbf{Z}$, (1), are found by minimizing the atomic ℓ_0 norm of atoms defined by the array manifold matrix [23].

The noiseless signal $\mathbf{Z}$ is re-written as

$$\mathbf{Z} = \sum_{k=1}^{K} \mathbf{a}(\mathbf{r}, \theta_k)\mathbf{x}_k^{\mathsf{H}} = \sum_{k=1}^{K} c_k \mathbf{a}(\mathbf{r}, \theta_k)\mathbf{b}_k^{\mathsf{H}} \tag{16}$$

where $c_k = \|\mathbf{x}\|_2 > 0, \mathbf{b}_k = c_k^{-1}\mathbf{x}_k$, thus $\|\mathbf{b}_k\|_2 = 1$. We define the atomic set as

$$\mathcal{A} = \{\mathbf{a}(\mathbf{r}, \theta_k, \mathbf{b}_k) = \mathbf{a}(\mathbf{r}, \theta_k)\mathbf{b}_k^{\mathsf{H}}\}, \tag{17}$$

which is understood as the set of fixed norm rank 1 matrices constructed from the manifold $\mathbf{a}(\mathbf{r}, \theta)$ over all values of θ. Expressing $\mathbf{Z}$ as a linear combination of K atoms in $\mathcal{A}$ gives the definition of the atomic ℓ_0 norm,

$$\|\mathbf{Z}\|_{\mathcal{A},0} = \inf_{c_k, \theta_k, \mathbf{b}_k} \left\{ K : \mathbf{Z} = \sum_{k=1}^{K} c_k \mathbf{a}(\mathbf{r}, \theta_k)\mathbf{b}_k^{\mathsf{H}} \right\}. \tag{18}$$

Atomic Norm Minimization (ANM) is concerned with finding a solution to (18). Towards this goal, consider the matrix

$$\mathbf{S} = \sum_{k=1}^{K} c_k^2 \begin{bmatrix} \mathbf{a}(\mathbf{r}, \theta_k) \\ \mathbf{b}_k \end{bmatrix} \begin{bmatrix} \mathbf{a}(\mathbf{r}, \theta_k) \\ \mathbf{b}_k \end{bmatrix}^{\mathsf{H}} = \begin{bmatrix} \mathbf{T} & \mathbf{Z} \\ \mathbf{Z}^{\mathsf{H}} & \mathbf{Q} \end{bmatrix}, \tag{19}$$

where

$$\mathbf{T} = \mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta})\mathbf{D}\mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta})^{\mathsf{H}}, \tag{20}$$

$$\mathbf{Q} = \mathbf{X}^{\mathsf{H}}\mathbf{X}, \tag{21}$$

and $\mathbf{D} \in \mathbb{R}^{K \times K}$ is diagonal with elements c_k^2. By definition $\mathbf{S}$ is positive semi-definite (PSD).

For ULAs, the array manifold matrix is Vandermonde and $\mathbf{T} \in \mathcal{T}$. Theorem 11.4 of [23] says that the atomic ℓ_0 norm (18) in the ULA case is the optimal solution of the following rank constrained optimization problem,

$$\underset{\mathbf{T} \in \mathcal{T},\, \mathbf{Q}}{\text{minimize}} \quad \text{rank}\,(\mathbf{T}), \quad \text{subject to} \quad \mathbf{S} \succeq 0. \tag{22}$$

Once the optimal $\mathbf{T}$ is found, the DOAs are recovered through Vandermonde decomposition (or root-MUSIC) of $\mathbf{T}$.

There are two problems with (22). First, the rank constraint is non-convex, thus the optimization cannot be solved using convex methods. Second, if the measurements are corrupted by noise (i.e., $\mathbf{Z}$ replaced with $\mathbf{Y}$ in $\mathbf{S}$) then there will be no optimal low rank $\mathbf{T}$.

Both of the above problems are solved by substituting (22) by its convex relaxation. In practice, the non-convex optimization problem of (22) is substituted for its convex analog [23, Theorem 11.14],

$$\underset{\mathbf{T} \in \mathcal{T},\, \mathbf{Q}}{\text{minimize}} \quad \text{Tr}\,(\mathbf{T}) + \text{Tr}\,(\mathbf{Q}), \quad \text{subject to} \quad \tilde{\mathbf{S}} \succeq 0 \tag{23}$$

where

$$\tilde{\mathbf{S}} = \begin{bmatrix} \mathbf{T} & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \mathbf{Q} \end{bmatrix}. \tag{24}$$

By substituting matrix rank for matrix trace, the optimization of (23) is in the form of a semi-definite program (SDP), for which there are many available solvers [41]. The substitution also makes the optimization robust to noise. The most popular convex solver is the alternating directions method of multipliers (ADMM) algorithm [42], summarized in Appendix.

III. Extension to Non-Uniform Array Geometries

We introduce a generalization of the Vandermonde matrix, which we call the *irregular Vandermonde matrix*. Just as the Toeplitz matrix is constructed from Vandermonde components, we define an *irregular Toeplitz matrix* constructed from irregular Vandermonde components. It is then shown that irregular Toeplitz matrices can be decomposed back to their irregular Vandermonde components.

We deem this decomposition the *irregular Vandermonde decomposition* (IVD) because it is interpreted as the Vandermonde decomposition of an irregularly sampled signal (i.e., sensor positions act as sample locations of a spectrally sparse signal whose frequencies are related to the DOAs). Further, the array manifold matrix of a NUA has irregular Vandermonde structure. We propose an 'irregular' root-MUSIC algorithm, which is related to the IVD just as root-MUSIC is related to Vandermonde decomposition.

The IVD follows the procedure of Vandermonde decomposition for irregular Toeplitz matrices. Under this framework, the null spectrum no longer has polynomial structure and cannot be easily rooted. Regardless, the roots of interest are the $2\,K$ root pairs which lie on or near the unit circle. A method to recover the relevant information from these roots is presented.

The key difference between Vandermonde decomposition and IVD is the sensor position vector involved in the IVD. For any given sensor position vector there is a set of matrices akin to the Toeplitz set containing all matrices that can be decomposed exactly by the IVD. This set is used in Section IV to ANM to NUA measurements.

A. Irregular Vandermonde and Toeplitz Matrices

Define an irregular Vandermonde matrix $\mathbb{C}^{M\times K}$ as,

$$\begin{aligned}\mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z}) &= [\boldsymbol{z}^{\gamma_1}\ \dots\ \boldsymbol{z}^{\gamma_M}]^{\mathsf{T}},\\ &= [\mathbf{w}(\boldsymbol{\gamma},z_1)\ \dots\ \mathbf{w}(\boldsymbol{\gamma},z_K)],\end{aligned} \tag{25}$$

where γ_i is the ith element of $\boldsymbol{\gamma}\in\mathbb{R}^M$, $\boldsymbol{z}\in\mathbb{C}^K$, and $\mathbf{w}(\boldsymbol{\gamma},z) = [z^{\gamma_1}\ \dots\ z^{\gamma_M}]^{\mathsf{T}}$. This matrix has been previously studied for its efficient matrix operations [48].

For DOA estimation, $\mathbf{A}_s(\mathbf{r},\boldsymbol{\theta}) = \mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})$, using the mapping

$$\boldsymbol{\gamma}=\mathbf{r},\quad z_k = e^{-j\pi\sin\theta_k},\quad \theta_k = -\sin^{-1}(\frac{\angle z_k}{\pi}). \tag{26}$$

Because $\boldsymbol{\gamma}=\mathbf{r}$, we refer to both as the sensor positions. Note that (26) is a generalization of the mapping presented in (5–7).

Following (9), we define the set of irregular Toeplitz matrices as those constructed from irregular Vandermonde matrices,

$$\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}} = \{\mathbf{T}:\mathbf{T}=\mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})\mathbf{D}\mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})^{\mathsf{H}},|\boldsymbol{z}|=\mathbf{1}\}, \tag{27}$$

where $\mathbf{D}\in\mathbb{R}^{K\times K}$ is a diagonal matrix with elements c_k^2. $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ is the set of irregular Toeplitz matrices for known sensor positions $\boldsymbol{\gamma}$.

Each unique sensor position vector $\boldsymbol{\gamma}$ is associated with its own irregular Toeplitz set $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$. The Toeplitz set $\boldsymbol{\mathcal{T}}$ is a subset of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ reached by any $\boldsymbol{\gamma}$ corresponding to a ULA, i.e., $\boldsymbol{\gamma} = \alpha[0,\dots,M-1]^{\mathsf{T}}+\beta$, see (5).

The expected sample covariance matrix for NUA measurements with sensor positions $\mathbf{r}=\boldsymbol{\gamma}$ is a diagonally loaded member of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$,

$$\begin{aligned}\mathbb{E}[\mathbf{R}_{yy}] &= \mathbf{A}_s(\mathbf{r},\boldsymbol{\theta})\boldsymbol{\Sigma}_X\mathbf{A}_s(\mathbf{r},\boldsymbol{\theta})^{\mathsf{H}}+\boldsymbol{\Sigma}_N\\ &= \mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})\boldsymbol{\Sigma}_X\mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})^{\mathsf{H}}+\sigma_N^2\mathbf{I}.\end{aligned} \tag{28}$$

for $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_N$ in (10). When the measurements are noiseless $\boldsymbol{\Sigma}_N=\mathbf{0}$ and $\mathbf{R}_{yy}\in\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$.

B. Irregular Root-MUSIC and Vandermonde Decomposition

Consider the sample covariance matrix, $\mathbf{T}\in\mathbb{C}^{M\times M}$, from a NUA constructed as in (28). Following (13), the irregular null spectrum of $\mathbf{T}$ is

$$\begin{aligned}\tilde{D}(z) &= \|\mathbf{U}_N^{\mathsf{H}}\mathbf{a}(\mathbf{r},\theta)\|_2^2 = \mathbf{a}(\mathbf{r},\theta)^{\mathsf{H}}\mathbf{U}_N\mathbf{U}_N^{\mathsf{H}}\mathbf{a}(\mathbf{r},\theta)\\ &= \mathbf{w}\left(\boldsymbol{\gamma},\frac{1}{z}\right)^{\mathsf{T}}\mathbf{G}\mathbf{w}(\boldsymbol{\gamma},z),\end{aligned} \tag{29}$$

for $\mathbf{r}=\boldsymbol{\gamma}$ and $\mathbf{G}$ in (14). Because (29) is constructed from irregular Vandermonde matrices, we refer to it as the *irregular null spectrum*. The expansion of $\tilde{D}(z)$ is

$$\tilde{D}(z) = \sum_{m=1}^{M}\sum_{n=1}^{M} g_{m,n}z^{\gamma_m-\gamma_n}, \tag{30}$$

where $g_{m,n}$ is element (m,n) of $\mathbf{G}$. We treat the domain of $\tilde{D}(z)$ as the set of all complex numbers, despite deriving $\tilde{D}(z)$ for z on the unit circle.

The behavior of the irregular null spectrum $\tilde{D}(z)$ in Fig. 1(c) and (d) is similar to that of the null spectrum from a half-wavelength spaced ULA in Fig. 1(a) and (b). The irregular null spectrum has a discontinuity at $\angle z=\pi$, because for nearly all array geometries $\tilde{D}(z)|_{\angle z=\pi_+}\neq\tilde{D}(z)|_{\angle z=\pi_-}$.

To understand this, consider $\angle z=\pi_\pm$ corresponds to DOAs $\theta=\pm\frac{\pi}{2}$, (26). The discontinuity is due to the inequality of the NUA array manifold at the extreme values of θ. The half-wavelength spaced ULA is one of several exceptional array geometries which produce the same array pattern across the array for signals arriving at $\pm90^\circ$,

$$e^{-j\pi(r+\beta)\sin(\frac{\pi}{2})} = e^{-j\pi(r+\beta)\sin(-\frac{\pi}{2})},\quad r\in\mathbb{Z}. \tag{31}$$

Properties i-iii from Section II-C hold for the irregular null spectrum, (the equivalent property i for the irregular null spectrum is $g_{m,n}=g_{n,m}^*$). Unlike (15), the irregular null spectrum no longer expands to a polynomial, thus the roots are not easily solvable. Instead, the expansion of (30) is a non-linear equation with no (known) closed form solution for its roots. Special NUAs such as the minimum redundancy and co-prime arrays still expand to rootable polynomials (30).

We exploit the following facts about $\tilde{D}(z)$:

1) The roots of interest are those that appear on or near the unit circle, as seen by (12).
2) Only the phase angle of the roots of interest are used to generate DOA estimates.

Thus the local minimums of $\tilde{D}(z)$ evaluated on the unit circle give DOA estimates with similar accuracy as those given by the actual roots. $\tilde{D}(z)$ evaluated on the unit circle is the inverse of the MUSIC spectrum [49].

Because each root of $\tilde{D}(z)$ has a root pair on its radial line (i.e., $\angle\tilde{z}=\angle\frac{1}{\tilde{z}^*}$), a saddle point exists on each of these radial lines. These saddle points manifest as local minima of $\tilde{D}(z)$ evaluated over $|z|=1$, and provide good estimates of the DOAs. The K smallest local minima of $\tilde{D}(z)$ over $|z|=1$ estimate the DOAs using (26). An example is in Fig. 1.

When $\mathbf{T}\in\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$, K roots appear on the unit circle because $\mathbf{G}\perp\mathbf{W}(\boldsymbol{\gamma},\boldsymbol{z})$. Evaluating $\tilde{D}(z)$ on the unit circle yields all

root locations, $\boldsymbol{z}$, which can be used to perfectly reconstruct $\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})$ from (27). The signal powers, $\mathbf{c} = [c_1^2 \ \dots \ c_K^2]^{\mathsf{T}}$ are found from,

$$\mathbf{c} = \operatorname{diag}\left(\mathbf{W}^{\dagger}\mathbf{T}(\mathbf{W}^{\dagger})^{\mathsf{H}}\right),$$

$$\mathbf{W}^{\dagger} = (\mathbf{W}^{\mathsf{H}}\mathbf{W})^{-1}\mathbf{W}^{\mathsf{H}}, \tag{32}$$

where $\mathbf{D} = \operatorname{diag}(\mathbf{c})$, and $\mathbf{W} = \mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})$. By this procedure an irregular Toeplitz matrix is decomposed to irregular Vandermonde components. Pseudocode for the IVD and irregular root-MUSIC are given in Section III-D

C. Irregular Toeplitz Structure

The unique structure shared by all members of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ is discussed here. Consider a noiseless NUA sample covariance matrix, this time constructed as an irregular Toeplitz matrix,

$$\mathbf{R}_{zz} = \mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta})\boldsymbol{\Sigma}_X\mathbf{A}_s(\mathbf{r}, \boldsymbol{\theta})^{\mathsf{H}} = \mathbf{W}(\boldsymbol{\gamma}, z)\mathbf{D}\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})^{\mathsf{H}}, \tag{33}$$

where $\boldsymbol{\Sigma}_X = \mathbf{D} = \operatorname{diag}(\mathbf{c})$ hold the uncorrelated source powers, and $(\boldsymbol{\gamma}, z)$ is related to $(\mathbf{r}, \theta)$ by (26). Each element of $\mathbf{R}_{zz}$ is written as,

$$(\mathbf{R}_{zz})_{m,n} = \sum_{k=1}^{K} c_k^2 z_k^{(\gamma_m - \gamma_n)}. \tag{34}$$

For a ULA, $\boldsymbol{\gamma} = \alpha[0, 1, \dots, (M-1)]^{\mathsf{T}} + \beta$, as (5), and $\gamma_m - \gamma_n$ is constant across each diagonal to produce a Toeplitz structure. In the NUA case element (m, n) of $\mathbf{R}_{zz}$ represents a complex function sampled at locations given by elements (m, n) of the Euclidean distance matrix of $\boldsymbol{\gamma}$ [43].

The irregular Toeplitz set for any $\boldsymbol{\gamma}$ is convex. Consider two irregular Toeplitz matrices, $\mathbf{T}_1, \mathbf{T}_2 \in \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ composed according to (33) with harmonics $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$,

$$\mathbf{T}_1 = \mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z}_1)\mathbf{D}_1\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z}_1)^{\mathsf{H}}, \tag{35}$$

$$\mathbf{T}_2 = \mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z}_2)\mathbf{D}_2\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z}_2)^{\mathsf{H}}. \tag{36}$$

Then their convex combination is

$$\lambda\mathbf{T}_1 + (1-\lambda)\mathbf{T}_2 = \mathbf{W}\left(\boldsymbol{\gamma}, \begin{bmatrix}\boldsymbol{z}_1 \\ \boldsymbol{z}_2\end{bmatrix}\right)\begin{bmatrix}\lambda\mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\mathbf{D}_2\end{bmatrix}\mathbf{W}\left(\boldsymbol{\gamma}, \begin{bmatrix}\boldsymbol{z}_1 \\ \boldsymbol{z}_2\end{bmatrix}\right)^{\mathsf{H}}, \tag{37}$$

for $0 \le \lambda \le 1$. The right hand side of (37) is a member of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ composed of irregular Vandermonde matrices with harmonics $\begin{bmatrix}\boldsymbol{z}_1 \\ \boldsymbol{z}_2\end{bmatrix}$.

D. IVD and Irregular Root-MUSIC Pseudocode

The IVD is a computationally efficient extension of root-MUSIC to NUA measurements. Pseudocode is in Algorithm 1.

The IVD takes as input the sensor positions, $\boldsymbol{\gamma}$, the dimension of the noise subspace, K, and the matrix to be decomposed, $\mathbf{T}$.

Algorithm 1: Irregular Vandermonde Decomposition (IVD).

$(\boldsymbol{z}, \mathbf{c}) = \texttt{IVD}(\mathbf{T}, \boldsymbol{\gamma}, K)$

Require: $\mathbf{T} \in \mathbb{C}^{M\times M}$, $\boldsymbol{\gamma} \in \mathbb{R}^M$, $K \in \mathbb{Z}$
$[\mathbf{U}, \boldsymbol{\Sigma}, \mathbf{V}] = \texttt{svd}(\mathbf{T})$
$\mathbf{U}_N = \mathbf{U}(:, K+1:M)$
$\boldsymbol{z} = \texttt{find}(\ \texttt{argmin}(\ \tilde{D}(z)\ ), |z| = 1\ )$, see (30) and (43)
$\mathbf{c} = \operatorname{diag}(\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})^{\dagger}\mathbf{T}(\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})^{\dagger})^{\mathsf{H}})$, see (32)

Algorithm 2: Irregular Root-MUSIC.

$\hat{\boldsymbol{\theta}} = \texttt{IrregularRootMusic}(\mathbf{Y}, \boldsymbol{\gamma}, K)$

Require: $\mathbf{Y} \in \mathbb{C}^{M\times L}$, $\boldsymbol{\gamma} \in \mathbb{R}^M$, $K \in \mathbb{Z}$, $L \ge K$
$\mathbf{T} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$
$(\boldsymbol{z}, \mathbf{c}) = \texttt{IVD}(\mathbf{T}, \boldsymbol{\gamma}, K)$, see Alogrithm 1
$\hat{\boldsymbol{\theta}} = -\texttt{asin}(\texttt{angle}(\boldsymbol{z})/\pi)$

The outputs are the harmonics, $\boldsymbol{z}$, and the diagonal matrix, $\mathbf{D}$. If $\mathbf{T}$ is a sample covariance matrix, then the harmonics are related to the DOAs by (26).

In Algorithm 1, $\texttt{svd}()$ is the singular value decomposition of a matrix and $\texttt{find}(\texttt{argmin}(\tilde{D}(z)), |z| = 1)$ outputs the values of z producing the K smallest local minima of $\tilde{D}(z)$ on the unit circle. For simulations in Section V, a gridded search over $10\,M$ evenly spaced points on the unit circle finds intervals containing minima, then the estimated minima locations are iteratively refined using golden section search [50].

The IVD allows for the extension of root-MUSIC to NUAs. We deem this *irregular root-MUSIC*, which is given in Algorithm 2.

The procedure of irregular root-MUSIC is to form the sample covariance matrix, $\mathbf{R}_{yy} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ and perform the IVD to produce DOA estimates. The number of measurement snapshots must be at least as large as the number of DOAs present ($L \ge K$) to accurately estimate the noise subspace and the resulting DOA estimates will be incorrect. The sensor positions, $\boldsymbol{\gamma}$, input to Algorithm 2 are in half-wavelength units.

IV. Gridless DOA for Non-Uniform Arrays

The irregular Toeplitz set enables ANM (see Section II-D) to be generalized to the NUA case. In this section we modify the optimization for gridless DOA to extend its use to measurements from NUAs, then propose an alternating projections (AP) based algorithm for solving said optimization. The AP solution is based on similar AP based algorithms that were recently examined in the context of CCS [44]–[46].

A. Extension to Non-Uniform Arrays

For NUAs the optimization of (22) is modified such that the optimal matrix belongs to $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ for $\boldsymbol{\gamma} = \mathbf{r}$,

$$\underset{\mathbf{T}\in\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}},\,\mathbf{Q}}{\text{minimize}} \quad \operatorname{rank}(\mathbf{T}), \quad \text{subject to } \tilde{\mathbf{S}} \succeq 0. \tag{38}$$

For $\tilde{\mathbf{S}}$ defined in (24).

The rank constrained set is defined as the set of matrices with $\text{rank} \leq K$. Taking the intersection between the rank constrained set and $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$, we get the rank constrained irregular Toeplitz set,

$$\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K} = \{\mathbf{T} : \mathbf{T} \in \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}},\ \text{rank}(\mathbf{T}) \leq K\}. \tag{39}$$

The rank constrained set is non-convex thus $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ is also non-convex.

Using $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$, the NUA gridless DOA problem simplifies to

$$\underset{\mathbf{T}\in\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K},\,\mathbf{Q}}{\text{minimize}}\ \ \|\tilde{\mathbf{S}}\|_F \quad \text{subject to } \tilde{\mathbf{S}} \succeq 0. \tag{40}$$

Unfortunately (40) cannot be solved using traditional convex solvers, and is difficult to cast into a convex formulation. Instead we present a solution to (40) based on the alternating projections (AP) algorithm. Once the optimal $\mathbf{T}$ is known, the IVD (Algorithm 1) is applied to recover the DOAs and source powers.

B. Projections for NUA Gridless DOA

Before formulating the optimization of (40) using AP, we define some useful projections.

1) Projection Onto Toeplitz Set: For ULAs, the projection onto $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ is a projection onto $\boldsymbol{\mathcal{T}}$. The projection of a matrix, $\mathbf{T} \in \mathbb{C}^{M\times M}$, to the Toeplitz set is [51],

$$P_{\boldsymbol{\mathcal{T}}}(\mathbf{T}) = \boldsymbol{\mathcal{T}}(\mathbf{u})$$

$$u_i = \frac{1}{2(M-i)} \sum_{j=1}^{(M-i)} \mathbf{T}_{j,j+i-1} + \mathbf{T}^{*}_{j+i-1,j}. \tag{41}$$

In words, the Toeplitz projection, $P_{\boldsymbol{\mathcal{T}}}(\mathbf{T})$, is achieved by replacing the elements along each diagonal with the mean of that diagonal.

2) Projection Onto Irregular Toeplitz Set: Recall the definition of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ from (39). Consider a matrix $\mathbf{T} \notin \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$. The goal is to project $\mathbf{T}$ onto $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$,

$$P_{\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}}(\mathbf{T}) = \mathbf{W}(\boldsymbol{\gamma}, \tilde{\boldsymbol{z}})\mathbf{D}\mathbf{W}(\boldsymbol{\gamma}, \tilde{\boldsymbol{z}})^{\mathsf{H}}, \tag{42}$$

for some set of harmonics $\tilde{\boldsymbol{z}} \in \mathbb{C}^{K}$. A projection onto $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ is achieved by constructing an irregular Toeplitz matrix using approximate $\tilde{\boldsymbol{z}}$ retrieved from the local minima of the irregular null spectrum on the unit circle,

$$\tilde{\boldsymbol{z}} = \arg\min_{|z|=1}^{k} \tilde{D}(z), \quad k = 1, \ldots, K, \tag{43}$$

where $\arg\min_z^k$ denotes the argument, z, which produces the kth smallest local minima.

The null spectrum roots for a given $\mathbf{T} \notin \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ are generally not on the unit circle. Instead, roots $\boldsymbol{z}$ are approximated from the irregular null spectrum (30). From $\tilde{\boldsymbol{z}}$, the corresponding signal powers are estimated (32).

Pseudocode for projection onto $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ is given in Algorithm 3. The algorithm involves computing the IVD for given $\mathbf{T}$ and reconstructing an approximation $\tilde{\mathbf{T}} \in \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$.

Algorithm 3: Projection Onto Irregular Toeplitz Set.

$\tilde{\mathbf{T}} = \mathtt{P}_{\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}}(\mathbf{T}, \boldsymbol{\gamma}, K)$

Require: $\mathbf{T} \in \mathbb{C}^{M\times M}$, $\boldsymbol{\gamma} \in \mathbb{R}^{M}$, $K \in \mathbb{Z}$, $M \geq K \geq 1$
$(\boldsymbol{z}, \mathbf{c}) = \mathtt{IVD}(\mathbf{T}, \boldsymbol{\gamma}, K)$, see Algorithm 1
$\tilde{\mathbf{T}} = \mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})\text{diag}(\mathbf{c})\mathbf{W}(\boldsymbol{\gamma}, \boldsymbol{z})^{\mathsf{H}}$, see (42)

3) Projection Onto Positive Semi-Definite Cone: The set of PSD matrices is

$$\mathcal{S}_{\succeq 0} = \{\mathbf{M} \in \mathbb{C}^{N\times N} : \lambda_i \in \mathbb{R},\ \lambda_i \geq 0, \quad \forall\, i\}, \tag{44}$$

where λ_i is the ith eigenvalue of $\mathbf{M}$, with eigenvector $\mathbf{e}_i$. The projection onto $\mathcal{S}_{\succeq 0}$ is [52]

$$P_{\mathcal{S}_{\succeq 0}}(\mathbf{M}) = \sum_{i=1}^{N} \max(0, \lambda_i)\mathbf{e}_i\mathbf{e}_i^{\mathsf{H}}. \tag{45}$$

The set $\mathcal{S}_{\succeq 0}$ forms a cone and is convex [42].

C. Alternating Projections for NUA Gridless DOA

The AP algorithm is an optimization scheme with notable success when applied to the structured low rank matrix completion problem [44], [45], [53]–[55]. In particular, it has been demonstrated that AP can be used to solve gridless DOA problems for co-prime arrays [56]. The concept of AP is that a solution at the intersection of two or more sets can be found by iteratively projecting an estimate between the sets. The algorithm is guaranteed to converge when all sets are convex, but convergence is not guaranteed when one or more sets is non-convex (i.e., the rank constrained set) [57]. AP can be used to solve the NUA formulation of ANM in (40).

Define the set

$$\mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}, \mathbf{Y})} = \left\{\mathbf{M} : \mathbf{M} = \begin{bmatrix} \mathbf{T} & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \mathbf{Q} \end{bmatrix}, \mathbf{T} \in \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}\right\}, \tag{46}$$

where $\boldsymbol{\gamma}$, $\mathbf{Y} \in \mathbb{C}^{M\times L}$, and K are known and $\mathbf{Q} \in \mathbb{C}^{L\times L}$ is a free matrix. Here $\mathbf{Y}$ are the array measurements. The projection onto $\mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}, \mathbf{Y})}$ is performed as

$$\mathbf{M} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{Q} \end{bmatrix},$$

$$P_{\mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}, \mathbf{Y})}}(\mathbf{M}) = \begin{bmatrix} P_{\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}}(\mathbf{B}) & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \mathbf{Q} \end{bmatrix}. \tag{47}$$

In words, the projection is achieved by replacing the top left submatrix of $\mathbf{M}$ with its projection to $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}$ (Algorithm 3), and replacing the corner submatrices with $\mathbf{Y}$ and $\mathbf{Y}^{\mathsf{H}}$.

The AP algorithm for solving (40) is achieved by projecting an initial estimate between $\mathcal{S}_{\succeq 0}$ and $\mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}, \mathbf{Y})}$,

$$\begin{aligned} \mathbf{H}^{(i+1)} &= P_{\mathcal{S}_{\succeq 0}}(\mathbf{L}^{(i)}), \\ \mathbf{L}^{(i+1)} &= P_{\mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^{K}, \mathbf{Y})}}(\mathbf{H}^{(i+1)}), \end{aligned} \tag{48}$$

Algorithm 4: AP Gridless.

$\hat{\boldsymbol{\theta}} = \texttt{AP\_Gridless}(\mathbf{Y}, \boldsymbol{\gamma}, K)$

Require: $\boldsymbol{\gamma} \in \mathbb{R}^M$, $K \in \mathbb{Z}$, $0 < K < M$
  $[M, L] = \texttt{size}(\mathbf{Y})$
  $\mathbf{L}^{(0)} = [\mathbf{0}, \mathbf{Y}; \mathbf{Y}^{\mathsf{H}}, \frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}]$
  for $i = 1 : \texttt{max\_iterations}$ **do**
    $\mathbf{H}^{(i)} = P_{\mathcal{S}_{\succeq 0}}(\mathbf{L}^{(i-1)})$
    $\mathbf{T} = \mathbf{H}^{(i)}(1 : M, 1 : M)$
    $\mathbf{L}^{(i)} = [\mathtt{P}_{\mathcal{T}_{\boldsymbol{\gamma}}}(\mathbf{T}, \boldsymbol{\gamma}, K), \mathbf{Y}; \mathbf{Y}^{\mathsf{H}}, \frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}]$, see Algorithm 3
    if $\|\mathbf{L}^{(i)} - \mathbf{L}^{(i-1)}\|_F \leq \texttt{1e} - \texttt{7}$ **then**
      break
    end if
  end for
  $\mathbf{T} = \mathbf{L}(1 : M, 1 : M)$
  $[\boldsymbol{z}, \sim] = \texttt{IVD}(\mathbf{T}, \boldsymbol{\gamma}, K)$, see Algorithm 1
  $\hat{\boldsymbol{\theta}} = -\texttt{asin}(\texttt{angle}(\boldsymbol{z})/\pi)$

Algorithm 5: AP Gridless for ULA.

$\hat{\boldsymbol{\theta}} = \texttt{AP\_ULA}(\mathbf{Y}, K)$

Require: $\boldsymbol{\gamma} \in \mathbb{R}^M$, $K \in \mathbb{Z}$, $0 < K < M$
  $[M, L] = \texttt{size}(\mathbf{Y})$
  $\mathbf{L}^{(0)} = [\mathbf{0}, \mathbf{Y}; \mathbf{Y}^{\mathsf{H}}, \frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}]$
  for $i = 1 : \texttt{max\_iterations}$ **do**
    $\mathbf{H}^{(i)} = P_{\mathcal{S}_{\succeq 0}}(\mathbf{L}^{(i-1)})$
    $\mathbf{T} = \mathbf{H}^{(i)}(1 : M, 1 : M)$
    $\mathbf{L}^{(i)} = [P_{\mathcal{T}}(\mathbf{T}), \mathbf{Y}; \mathbf{Y}^{\mathsf{H}}, \frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}]$, see (41)
    if $\|\mathbf{L}^{(i)} - \mathbf{L}^{(i-1)}\|_F \leq \texttt{1e} - \texttt{7}$ **then**
      break
    end if
  end for
  $\mathbf{T} = \mathbf{L}(1 : M, 1 : M)$
  $[\boldsymbol{z}, \sim] = \texttt{IVD}(\mathbf{T}, \boldsymbol{\gamma}, K)$, see Algorithm 1
  $\hat{\boldsymbol{\theta}} = -\texttt{asin}(\texttt{angle}(\boldsymbol{z})/\pi)$

for arbitrary initial estimate $\mathbf{L}^{(0)} \in \mathbb{C}^{(M+L)\times(M+L)}$ containing $\mathbf{Y}$. Upon convergence, the IVD of the $M \times M$ upper right submatrix of $\tilde{\mathbf{S}}$ retrieves the DOAs. Pseudocode for the proposed AP based gridless DOA algorithm is given in Algorithm 4. We deem this algorithm, "*AP gridless*".

AP gridless requires prior knowledge of source number K. There are many works on source number estimation [58], [59]. Alternatively, AP gridless can be applied for a range of K and the sparsest solution which adequately reconstructs the measurements taken as optimal.

The algorithm is called "*AP ULA*" for ULA measurements because the irregular Toeplitz projection, $P_{\mathcal{T}_{\boldsymbol{\gamma}}^K}$ (Algorithm 3), is substituted for a Toeplitz projection, $P_{\mathcal{T}}$ (41). AP ULA is computationally more efficient, but can only be used for ULA. AP ULA is given in Algorithm 5.

D. Initialization and Convergence

Convergence of the AP Gridless algorithm falls into the branch of research known as low rank matrix approximation, which has been active for several decades [57], [60]. The AP algorithm is guaranteed to converge to the global optimum at a linear rate when applied between two closed convex sets [61], however global convergence is not guaranteed when one or more sets is non-convex.

The AP Gridless algorithm applies AP between a convex set, $\mathcal{S}_{\succeq 0}$, and a non-convex set, $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^K$, which is composed of the intersection between a convex set and the quasi-convex rank constrained set. Thus convergence depends on the behavior of AP applied between the rank constrained set and one or more convex sets.

Progress towards a convergence proof was achieved in [62, Theorem 3.2] and is summarized as it applies to (48) as follows:

Theorem: Let $\mathbf{L}^{(i)}, \mathbf{H}^{(i)}$ be generated according to (48), and $\mathbf{L}^{(0)} \in \mathcal{S}_{(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}^K, \mathbf{Y})}$ then

1) Either $\|\mathbf{L}^{(i)} - \mathbf{H}^{(i)}\|_F^2 \to \infty$ as $i \to \infty$ or $\|\mathbf{L}^{(i)} - \mathbf{H}^{(i)}\|_F^2$ converges to 0.
2) If $\mathbf{L}^{(0)}$ is sufficiently close to the global minimizer of (40), $\mathbf{L}^{(0)}$ converges to the global minimum.

Theorem 1 does not guarantee convergence to the global minimizer unless the initialization point is sufficiently nearby the optimal solution. Additional work on alternating projections for low rank matrix approximation problems can be found in [53], [63]–[65]. Due to the relatively weak theoretical guarantees currently available for the proposed algorithm, we rely on extensive simulations in Section V to determine the quality of the proposed algorithm.

All simulations in Section V used the initialization,

$$\mathbf{L}^{(0)} = \begin{bmatrix} \mathbf{0} & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{H}} & \frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y} \end{bmatrix}, \tag{49}$$

which was always observed to converge to a critical point rather than diverging. The ideal initialization of upper left submatrix spans the columnspace of the true solution. Such an initialization converges to the correct solution in one iteration, but finding such a matrix requires prior knowledge of the solution. We choose to initialize the upper left submatrix to $\mathbf{0}$ to avoid biasing the initialization. The ideal initialization of the lower right submatrix is $\mathbf{X}^{\mathsf{H}}\mathbf{X}$ [see (21)], which is well approximated by $\frac{1}{L}\mathbf{Y}^{\mathsf{H}}\mathbf{Y}$ and can be tuned based on prior knowledge of the received signals. Because the AP Gridless algorithm is not guaranteed to converge to a global minimum, the accuracy of AP for gridless DOA is based on numerical simulation performed in Section V.

E. Computational Complexity

The AP gridless (48) (Algorithm 4), and AP ULA (41) (Algorithm 5) are both iterative algorithms with similar computational complexity as ADMM algorithm. The per iteration computational complexity of each algorithm is limited by the eigen-decomposition of an $M + L$ dimensional matrix during

projection onto the positive semi-definite cone (45). In practice, this Eigen-decomposition has complexity $\mathcal{O}((M+Y)^3)$.

AP Gridless utilizes an additional Eigen-decomposition of an M dimensional matrix for projection to the irregular Toeplitz set. The computational complexity of the three algorithms are:

1) ADMM: $\mathcal{O}(N(M+Y)^3)$
2) AP ULA: $\mathcal{O}(N(M+Y)^3)$
3) AP Gridless: $\mathcal{O}(N[(M+Y)^3+M^3])$

where N is the number of iterations. Note that the proposed AP ULA algorithm has identical per iteration computational complexity to ADMM. AP Gridless achieves the solution to a non-convex optimization for only a minor penalty in complexity.

V. Simulation

AP gridless (48) (Algorithm 4), and AP ULA (41) (Algorithm 5) were applied to solve the gridless DOA problem (40) for simulated measurements modeled by (1). For each simulation and trial, K randomly generated DOAs were chosen such that the minimum angular separation was $\frac{1}{M}$ for DOAs randomly drawn between $[-90^\circ, 90^\circ)$. Signals contained in the K rows of $\mathbf{X}$ were generated as complex values with uniformly distributed phase. The signal from each DOA had random amplitude σ_s^x, $x \in \mathcal{U}(0,1)$ to model sources with different powers. Noise, $\mathbf{N}$, was drawn from a complex Gaussian distribution with mean zero and variance $\mathbf{I}$, then scaled to fit the desired signal to noise ratio (SNR) determined by

$$\text{SNR} = 10\log_{10}\left(\frac{\|\mathbf{Z}\|_F^2}{\|\mathbf{N}\|_F^2}\right). \tag{50}$$

Measurements were simulated for an M element ULA and NUA. A different NUA geometry realization was used in each simulation and trial. NUA geometries were generated by adding random offsets from a uniform distribution between $[-.5, .5)$ (in half-wavelength units) to each sensor position. No restriction was placed on spacing between NUA elements. DOAs and received signals were identical between ULA and NUA simulations, thus all differences between ULA and NUA simulations are attributable to sensor positions. The Cramér-Rao bound (CRB) was calculated as the average CRB over all trials of a simulation, using the conditional/deterministic model given by 2.13 of [66].

For ULA, AP gridless (Algorithm 4) and AP ULA (Algorithm 5) were both applied. It was assumed K was known for each simulation. AP gridless and AP ULA were considered converged when $\|\mathbf{L}^{(i)} - \mathbf{L}^{(i-1)}\|_F \leq 10^{-7}$, or terminated after $10\,K$ iterations.

Solution accuracy was gauged by the root mean square error (RMSE) between true, θ_k, and recovered, $\hat{\theta}_k$, DOAs

$$\text{RMSE} = \sqrt{\mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}(\theta_k - \hat{\theta}_k)^2\right]}. \tag{51}$$

A maximum threshold of 10° was used to provide an even error penalty for incorrect DOA estimates.

Sparse Bayesian learning (SBL) [67]–[71], and least absolute shrinkage and selector operator (LASSO) [7], [9], [72], [73] methods were given a dictionary of the array manifold with 1° separation between entries (180 total). The LASSO tuning parameter was selected such that the solution was K sparse (which requires knowledge of $\|\mathbf{N}\|_F$ [74]). The ADMM algorithm was applied to ULA measurements using parameters $\tau = .01$, and $\rho = 1$ (see App. A).

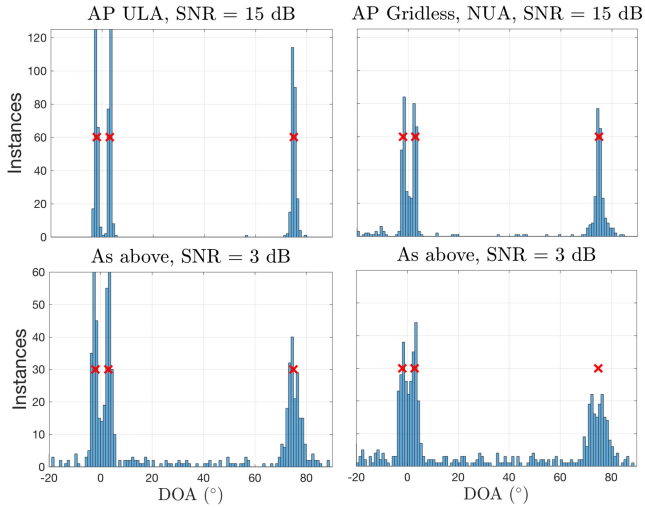

Fig. 2. Histogram of recovered DOAs for AP ULA (Algorithm 5) and AP gridless (Algorithm 4), $M = 20$, $L = 1$, $\sigma_s = 5$, $\boldsymbol{\theta} = [-2, 3, 75]^\circ$ (red x). 250 trials per histogram. Left- AP ULA. Right- AP Gridless (NUA).

Enhanced Matrix Completion (EMaC), a gridless method related to ANM [23] is included for comparison against AP ULA in the single snapshot ULA case. Fourier Domain root-MUSIC (FD root-MUSIC), which extends root-MUSIC to NUAs, is included for comparison against irregular root-MUSIC in the multiple snapshot case. FD root-MUSIC was implemented with interpolation factor M_2 equal to four times the array aperture [31].

Results from AP ULA and AP gridless are detailed in Fig. 2. The AP ULA algorithm never misclassifies a DOA when the measurement SNR is high (Fig. 2, top left), and decreased SNR results in few misclassifications (Fig. 2, bottom left). In contrast, there are some NUA geometries which cause AP gridless to misclassify a DOA, even for high SNR (Fig. 2, top right). The rate of misclassification for the low SNR NUA case is slightly higher than that of the ULA case (Fig. 2, bottom right).

Gridless DOA excels over gridded methods in the high SNR, low snapshot case. This is due to quantization error from grid mismatch limiting the maximum accuracy of gridded techniques. Fig. 3 compares gridless and gridded algorithms for high SNR, single snapshot measurements when all algorithms are run to convergence. The best achievable accuracy for a gridded technique using 1° DOA separation is RMSE $= .25^\circ$. In the ULA case (Fig. 3, top), all gridless techniques achieve accuracy proportional to SNR. ADMM is not as precise due to limitations in parameter tuning. For NUA (Fig. 3, bottom) AP gridless maintains excellent performance.

The performance of each algorithm in lower SNR scenarios is detailed in Fig. 4. For ULA, AP gridless and AP ULA achieve similar performance to ADMM, particularly in the single snapshot case (Fig. 4, top left), indicating that convex relaxation is

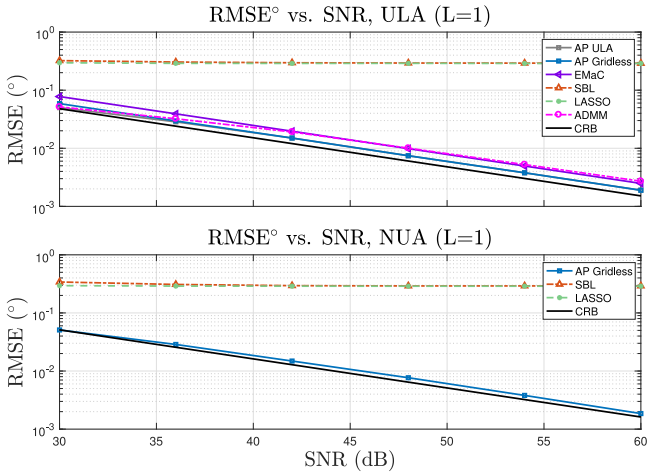

Fig. 3. RMSE vs. SNR on high SNR single snapshot measurements using AP Gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM (App. A$\tau = 10^{-5}$), EMaC, SBL, and LASSO. $M = 20, K = 3, L = 1, \sigma_s = 5$. Each point represents average over 250 trials. All algorithms run to convergence.

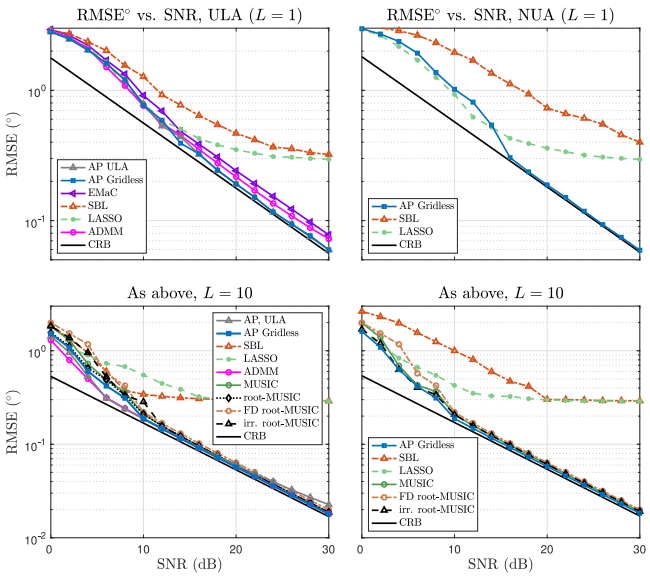

Fig. 4. RMSE vs. SNR for AP gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM (App. A), EMaC, SBL, LASSO, root-MUSIC, FD root-MUSIC, irregular root-MUSIC (Algorithm 2), and associated CRB. $M = 20$, $K = 3$, $\sigma_s = 5$. Each point represents 250 trials. Top left- ULA measurements, $L = 1$. Top right- NUA measurements, $L = 1$. Bottom left- ULA measurements, $L = 10$. Bottom right- NUA measurements, $L = 10$.

unnecessary to solve rank minimization problems. AP gridless and AP ULA perform slightly worse than ADMM for low SNR multi-snapshot measurements (Fig. 4, bottom left), suggesting that ADMM is more robust to noise. AP gridless is superior to gridded techniques for NUA measurements (Fig. 4, top and bottom right), except in low SNR scenarios where LASSO can attain similar performance. Irregular root MUSIC (Algorithm 2) attains near identical performance to AP gridless in the multiple snapshot NUA case (Fig. 4, bottom right), for less computational cost.

Fig. 5 gives the probability of resolution of two nearby DOAs located at $\pm 1.5°$. A trial was considered to have resolved the

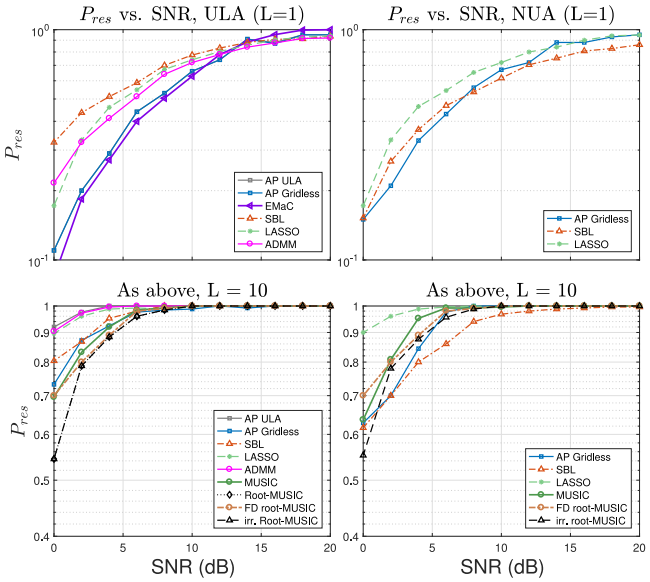

Fig. 5. Probability of Resolution vs. SNR for AP gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM (App. A), EMaC, SBL, LASSO, root-MUSIC, FD root-MUSIC, irregular root-MUSIC (Algorithm 2), and CRB. DOAs positioned at $\pm 1.5°$, $M = 20$, $\sigma_s = 5$. Each point represents 250 trials. Successful resolution defined trials where RMSE $\leq \sqrt{3}$. Top left- ULA measurements, $L = 1$. Top right- NUA measurements, $L = 1$. Bottom left- ULA measurements, $L = 10$. Bottom right- NUA measurements, $L = 10$.

pair of DOAs if RMSE $\leq \sqrt{3}$. In the ULA case, AP ULA (Algorithm 5) exhibited the best overall probability of resolution, while AP Gridless (Algorithm 4) was the least likely to resolve the DOAs. This indicates that the projection onto the irregular Toeplitz set (Section IV-B2) leads to less accurate results compared to the projection onto the Toeplitz set (Section IV-B1). For ULAs, irregular root-MUSIC (algorithm 2) had near identical performance to root-MUSIC. In the NUA case, AP Gridless performs well in high SNR simulations, but becomes the least likely algorithm to resolve nearby DOAs in low SNR simulations.

Performance of the algorithms versus array elements, M, is given in Fig. 6. AP Gridless and ADMM are superior in the ULA case (Fig. 6, top left), as gridded methods are limited in accuracy by grid resolution. For NUAs, AP gridless and irregular root-MUSIC are superior for the same reason. Root-MUSIC and irregular root-MUSIC attain equal performance to AP gridless and ADMM in the multi-snapshot case (Fig. 6, bottom left and right) for lower computational complexity. This is also shown by Fig. 7, which compares each method vs. number of snapshots, L. Once the number of snapshots is greater than the number of sources ($L \geq K$), there is no benefit to choosing a gridless technique because root-MUSIC and irregular root-MUSIC attain the same performance as gridless methods for reduced computational cost.

To get an indication of the high resolution capability of each algorithm, Fig. 8 compares AP gridless and AP ULA to conventional beamforming (CBF) and root-MUSIC for noiseless measurements with 2 DOAs at $\pm\theta°$. For the single snapshot case (Fig. 8, top), AP gridless has difficulty resolving nearby DOAs

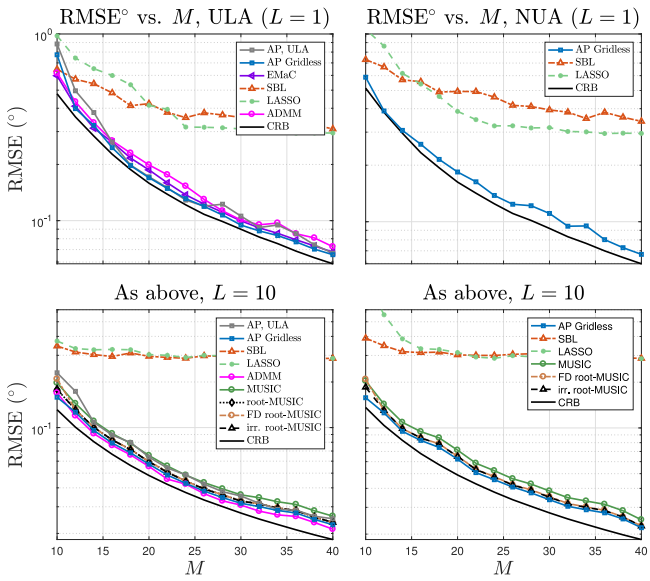

Fig. 6. RMSE vs. M for AP gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM A, EMaC, SBL, LASSO, root-MUSIC, FD root-MUSIC, irregular root-MUSIC (Algorithm 2), and CRB. SNR $= 20$ dB, $K = 3$, $\sigma_s = 5$. Each point represents 250 trials. Top left- ULA measurements, $L = 1$. Top right- NUA measurements, $L = 1$. Bottom left- ULA measurements, $L = 10$. Bottom right- NUA measurements, $L = 10$.

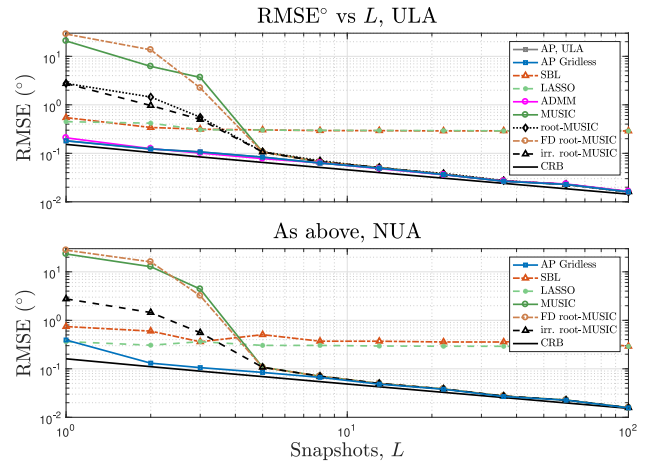

Fig. 7. RMSE vs. L for AP gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM (App. A), SBL, LASSO, root-MUSIC, FD root-MUSIC, irregular root-MUSIC (Algorithm 2), and CRB. $M = 20$, $K = 3$, $\sigma_s = 5$, SNR $=$ 20 dB. Each point represents 250 trials. Top- ULA measurements. Bottom- NUA measurements.

while ADMM and AP ULA have outstanding performance. The performance difference is due to ADMM and AP ULA utilizing a projection to $\mathcal{T}$ in place of projection to $\mathcal{T}_\gamma^K$.

The projection to $\mathcal{T}$ is an orthogonal projection which outputs the nearest Toeplitz matrix in Frobenius norm to its input (orthogonal projection), which is generally a full rank Toeplitz matrix. In contrast, $P_{\mathcal{T}_\gamma^K}$ is not an orthogonal projection, and relies on the IVD (Algorithm 3), which involves evaluation of a null spectrum based on a noise subspace estimate. As a result, situations where the noise subspace cannot be accurately

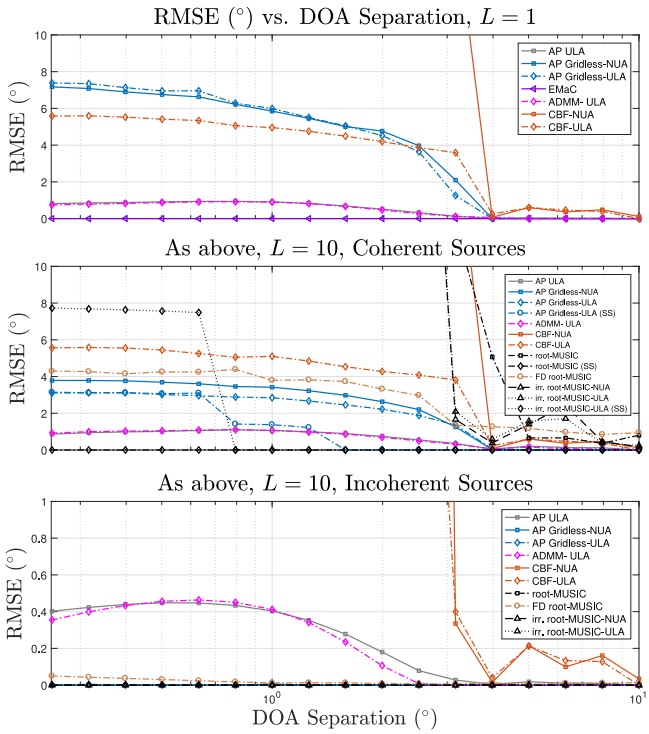

Fig. 8. RMSE vs. DOA separation of conventional (Bartlett) beamformer (CBF), AP gridless (Algorithm 4), AP ULA (Algorithm 5), ADMM (App. A), root-MUSIC, FD root-MUSIC, and irregular root-MUSIC (Algorithm 2) on noiseless simulated ULA and NUA measurements using $K = 2$ equal amplitude sources at $\pm\theta^\circ$, $M = 20$. Spatially smoothed estimates denoted by (SS). No maximum MSE penalty, each point represents 250 trials. Top- $L = 1$. Middle- $L = 10$, coherent sources. Bottom- $L = 10$, incoherent sources.

estimated (i.e., the single snapshot case or coherent sources case) can lead to poor results if the sources are not well separated.

It is well established that spatial smoothing (SS) [75], [76] can be used to help resolve coherent sources, however is applicable only to ULA measurements. The application of the SS technique to AP Gridless, irregular root-MUSIC, and root-MUSIC algorithms (Fig. 8, middle) had the effect of increasing each algorithm's ability to resolve nearby DOAs in the coherent sources case.

When the sources are made incoherent (Fig. 8, bottom), there is no longer any issue with resolving the noise subspace, therefore AP gridless, root-MUSIC, and irregular root-MUSIC recover exact DOAs. In contrast, ADMM and AP ULA are not as precise because $P_{\mathcal{T}}$ does not project to a specifically rank K solution, resulting in 'noisy' DOA estimates.

The computational bottleneck of AP gridless is the projection to the irregular Toeplitz set, $P_{\mathcal{T}_\gamma^K}$, where a spectral search over $\tilde{D}(z)|_{|z|=1}$ is executed. In contrast, the AP ULA and ADMM algorithms use the relatively efficient projection to the Toeplitz set, $P_{\mathcal{T}}$ (41), and are limited by the eigen-decomposition of the $\tilde{\mathbf{S}}$ matrix in its projection to the positive semi-definite cone, $P_{\mathcal{S}\succeq 0}$. Both algorithms are relatively fast, having runtime under one second on a regular CPU for problems with $(M + L) < 100$.

VI. Conclusion

Gridless direction of arrival (DOA) estimation for arbitrary non-uniform array (NUA) geometries was considered. Towards this goal, the irregular Vandermonde decomposition (IVD) was introduced, which is a generalization of the Vandermonde decomposition for irregularly sampled signals, such as those sampled from a NUA. From the perspective of the IVD, any covariance matrix can be seen as an irregular Toeplitz matrix, and can be decomposed back into its irregular Vandermonde components. The decomposition can be used to extend gridless DOA to NUAs, as well as extending the larger continuous compressed sensing (CCS) problem to irregularly sampled signals. The decomposition also leads to a new algorithm for extending root-MUSIC to NUAs, which was previously achieved by algorithms such as 'Fourier Domain root-MUSIC'.

An intuitive non-convex method of solving gridless DOA for NUAs based on the alternative projections (AP) algorithm was proposed. The proposed algorithm was named AP gridless. Simulation on uniform linear array (ULA) measurements found AP gridless attains similar performance to its convex counterpart, ADMM. For NUAs, ADMM cannot be generalized, and only AP gridless applies. AP gridless was found to be robust to noise, high resolution, and has superior performance compared to grid based techniques for high SNR.

Appendix
ADMM for Gridless DOA

The alternating directions method of multipliers (ADMM) formulation of gridless DOA is summarized as it applies to multiple snapshot ULA measurements. A complete review of ADMM is given in [77]. Other sources which provide additional details on ADMM specifically for gridless DOA are [23], [40], [78]–[80].

Start with the rank minimization problem of (22). Because the the rank constraint is non-convex the problem must be cast to its convex relaxation before ADMM is applicable,

$$\begin{aligned} \min_{\mathbf{Q},\tilde{\mathbf{S}},\hat{\mathbf{Y}},\mathbf{u}} \quad & \frac{1}{2}\|\hat{\mathbf{Y}}-\mathbf{Y}\|_2^2+\frac{\tau}{2}\left(\mathrm{Tr}(\mathbf{Q})+\mathrm{Tr}(\mathbf{T}(\mathbf{u}))\right) \\ \text{subject to} \quad & \tilde{\mathbf{S}}=\begin{bmatrix}\mathbf{T}(\mathbf{u}) & \hat{\mathbf{Y}}\\ \hat{\mathbf{Y}}^{\mathsf{H}} & \mathbf{Q}\end{bmatrix}, \quad \tilde{\mathbf{S}}\succeq 0, \end{aligned} \tag{A-1}$$

where $\mathbf{Y}\in\mathbb{C}^{M\times L}$ is the measurement matrix, $\mathbf{T}(\mathbf{u})\in\mathbb{C}^{M\times M}$ is the Toeplitz matrix whose first column is $\mathbf{u}$, $\mathbf{Q}\in\mathbb{C}^{L\times L}$, and τ is a user defined parameter.

Next, the augmented Lagrangian is formed by collecting the constraints into the optimization function [78], [79],

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{Q},\tilde{\mathbf{S}},\hat{\mathbf{Y}},\mathbf{\Lambda},\mathbf{u})=&\frac{1}{2}\|\hat{\mathbf{Y}}-\mathbf{Y}\|_2^2+\frac{\tau}{2}\left(\mathrm{Tr}(\mathbf{Q})+\mathrm{Tr}(\mathbf{T}(\mathbf{u}))\right)\\ &+\left\langle\mathbf{\Lambda},\tilde{\mathbf{S}}-\begin{bmatrix}\mathbf{T}(\mathbf{u}) & \hat{\mathbf{Y}}\\ \hat{\mathbf{Y}}^{\mathsf{H}} & \mathbf{Q}\end{bmatrix}\right\rangle+\frac{\rho}{2}\left\|\tilde{\mathbf{S}}-\begin{bmatrix}\mathbf{T}(\mathbf{u}) & \hat{\mathbf{Y}}\\ \hat{\mathbf{Y}}^{\mathsf{H}} & \mathbf{Q}\end{bmatrix}\right\|_F^2, \end{aligned} \tag{A-2}$$

where ρ is a user defined parameter and $\langle .\,,\,.\rangle$ is the matrix inner product, and $\mathbf{\Lambda}$ is the Lagrange multiplier variable. ADMM is performed by iteratively optimizing (A-2) over each variable independent of the the other variables. The update steps at iteration i are [78], [79],

$$(\mathbf{Q}^{(i+1)},\mathbf{u}^{(i+1)},\hat{\mathbf{Y}}^{(i+1)})\leftarrow\arg\min_{\mathbf{Q},\hat{\mathbf{Y}},\mathbf{u}}\mathcal{L}_\rho(\mathbf{Q},\tilde{\mathbf{S}}^{(i)},\hat{\mathbf{Y}},\mathbf{\Lambda}^{(i)},\mathbf{u}), \tag{A-3}$$

$$\tilde{\mathbf{S}}^{(i+1)}\leftarrow\arg\min_{\tilde{\mathbf{S}}\succeq 0}\mathcal{L}_\rho(\mathbf{Q}^{(i+1)},\tilde{\mathbf{S}},\hat{\mathbf{Y}}^{(i+1)},\mathbf{\Lambda}^{(i)},\mathbf{u}^{(i+1)}), \tag{A-4}$$

$$\mathbf{\Lambda}^{(i+1)}\leftarrow\mathbf{\Lambda}^{(i)}+\rho\left(\tilde{\mathbf{S}}^{(i+1)}-\begin{bmatrix}\mathbf{T}(\mathbf{u}^{(i+1)}) & \hat{\mathbf{Y}}^{(i+1)}\\ \hat{\mathbf{Y}}^{(i+1)\mathsf{H}} & \mathbf{Q}^{(i+1)}\end{bmatrix}\right). \tag{A-5}$$

The update steps for $\mathbf{Q}$, $\mathbf{u}$, and $\hat{\mathbf{Y}}$ are computable in closed form. First define the partitions

$$\tilde{\mathbf{S}}^{(i)}=\begin{bmatrix}\tilde{\mathbf{S}}_{\mathbf{T}} & \tilde{\mathbf{S}}_{\mathbf{Y}}\\ \tilde{\mathbf{S}}_{\mathbf{Y}^{\mathsf{H}}} & \tilde{\mathbf{S}}_{\mathbf{Q}}\end{bmatrix} \quad\text{and}\quad \mathbf{\Lambda}^{(i)}=\begin{bmatrix}\mathbf{\Lambda}_{\mathbf{T}} & \mathbf{\Lambda}_{\mathbf{Y}}\\ \mathbf{\Lambda}_{\mathbf{Y}^{\mathsf{H}}} & \mathbf{\Lambda}_{\mathbf{Q}}\end{bmatrix}, \tag{A-6}$$

where the dimensions of each partition are the dimensions of $\mathbf{T}$, $\mathbf{Y}$, and $\mathbf{Q}$. The update steps are [78], [79],

$$\mathbf{Q}^{(i+1)}=\frac{1}{2}\tilde{\mathbf{S}}_{\mathbf{Q}}^{(i)}+\frac{1}{2}(\tilde{\mathbf{S}}_{\mathbf{Q}}^{(i)})^{\mathsf{H}}+\frac{1}{\rho}\left(\mathbf{\Lambda}_{\mathbf{Q}}^{(i)}-\frac{\tau}{2}\mathbf{I}\right), \tag{A-7}$$

$$\mathbf{u}^{(i+1)}=\mathcal{T}^{-1}\left(\tilde{\mathbf{S}}_{\mathbf{T}}^{(i)}+\frac{1}{\rho}\mathbf{\Lambda}_{\boldsymbol{T}}^{(i)}\right)-\frac{\tau}{2\rho}\mathbf{e}_1, \tag{A-8}$$

$$\hat{\mathbf{Y}}^{(i+1)}=\frac{1}{2\rho+1}\left(\mathbf{Y}+\rho\tilde{\mathbf{S}}_{\mathbf{Y}}^{(i)}+\rho\tilde{\mathbf{S}}_{\mathbf{Y}^{\mathsf{H}}}^{(i)}+2\mathbf{\Lambda}_{\mathbf{Y}}^{(i)}\right), \tag{A-9}$$

where $\mathcal{T}^{-1}$ is a function giving the first column of the nearest Toeplitz matrix to its input (41), and $\mathbf{e}_1$ is the canonical basis vector in M dimensional space, $[1\,0\,\ldots\,0]^{\mathsf{T}}$.

Finally, the $\tilde{\mathbf{S}}$ update is a projection onto the positive semidefinite cone (45),

$$\tilde{\mathbf{S}}^{(i+1)}=P_{\mathcal{S}\succeq 0}\left(\tilde{\mathbf{S}}^{(i)}-\begin{bmatrix}\mathbf{T}(\mathbf{u}^{(i+1)}) & \hat{\mathbf{Y}}^{(i+1)}\\ \hat{\mathbf{Y}}^{(i+1)\mathsf{H}} & \mathbf{Q}^{(i+1)}\end{bmatrix}+\frac{1}{\rho}\mathbf{\Lambda}^{(i)}\right). \tag{A-10}$$

ADMM is known to converge slowly in practice [77].

References

[1] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1983, pp. 336–339.

[2] B. D. Rao and K. V. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[3] M. Pesavento, A. B. Gershman, and M. Haardt, "Unitary root-MUSIC with a real-valued eigendecomposition: A theoretical and experimental performance study," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1306–1314, May 2000.

[4] H. L. V. Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2004.

[5] W. J. Zeng, X. L. Li, and X. D. Zhang, "Direction-of-arrival estimation based on the joint diagonalization structure of multiple fourth-order cumulant matrices," *IEEE Signal Process. Lett.*, vol. 16, no. 3, pp. 164–167, Mar. 2009.

[6] W. J. Zeng and X. L. Li, "High-resolution multiple wideband and non-stationary source localization with unknown number of sources," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3125–3136, Jun. 2010.

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

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

[9] P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker, "Multiple and single snapshot compressive beamforming," *J. Acoust. Soc. Amer.*, vol. 138, no. 4, pp. 2003–2014, 2015.

[10] 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, May 2011.

[11] 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, Nov. 2013.

[12] 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.

[13] P. Pal and P. P. Vaidyanathan, "A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 737–741, Jun. 2014.

[14] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, Jun. 2015.

[15] X. Wu, W. P. Zhu, and J. Yan, "A toeplitz covariance matrix reconstruction approach for direction-of-arrival estimation," *IEEE Trans. Veh.*, vol. 66, no. 9, pp. 8223–8237, Sep. 2017.

[16] C. Zhou, Y. Gu, X. Fan, Z. Shi, G. Mao, and Y. D. Zhang, "Direction-of-arrival estimation for coprime array via virtual array interpolation," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 5956–5971, Nov. 2018.

[17] Y. Park, P. Gerstoft, and W. Seong, "Grid-free compressive mode extraction," *J. Acoust. Soc. Amer.*, vol. 145, no. 3, pp. 1427–1442, 2019.

[18] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM J. Matrix Anal.*, vol. 34, no. 3, pp. 946–977, 2013.

[19] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.

[20] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5145–5157, Oct. 2016.

[21] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 995–1006, Feb. 2016.

[22] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6576–6601, Oct. 2014.

[23] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing: Array, Radar and Communications Engineering*, vol. 7, Academic Press, 2018.

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

[25] C. Zhou, Y. Gu, Z. Shi, and Y. D. Zhang, "Off-grid direction-of-arrival estimation using coprime array interpolation," *IEEE Signal Process. Lett.*, vol. 25, no. 11, pp. 1710–1714, Nov. 2018.

[26] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb. 2011.

[27] M. Wang, Z. Zhang, and A. Nehorai, "Gridless DOA estimation using sparse linear arrays based on Wasserstein distance," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 838–842, Jun. 2019.

[28] S. Semper, F. Roemer, T. Hotz, and G. D. Galdo, "Grid-free direction-of-arrival estimation with compressed sensing and arbitrary antenna arrays," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2018, pp. 3251–3255.

[29] A. G. Raj and J. H. McClellan, "Super-resolution DOA estimation for arbitrary array geometries using a single noisy snapshot," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2019, pp. 4145–4149.

[30] Y. Wang, Y. Zhang, Z. Tian, G. Leus, and G. Zhang, "Super-resolution channel estimation for arbitrary arrays in hybrid millimeter-wave massive mimo systems," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 947–960, Sep. 2019.

[31] M. Rubsamen and A. B. Gershman, "Direction-of-arrival estimation for nonuniform sensor arrays: From manifold separation to fourier domain MUSIC methods," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 588–599, Feb. 2009.

[32] B. Friedlander and A. J. Weiss, "Direction finding using spatial smoothing with interpolated arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 2, pp. 574–587, Apr. 1992.

[33] B. Friedlander, "The root-MUSIC algorithm for direction finding with interpolated arrays," *Signal Process.*, vol. 30, no. 1, pp. 15–29, 1993.

[34] M. A. Doran, E. Doron, and A. J. Weiss, "Coherent wide-band processing for arbitrary array geometry," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 414–417, Jan. 1993.

[35] A. J. Weiss, B. Friedlander, and P. Stoica, "Direction-of-arrival estimation using MODE with interpolated arrays," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 296–300, Jan. 1995.

[36] A. B. Gershman and J. F. Bohme, "A note on most favorable array geometries for DOA estimation and array interpolation," *IEEE Signal Process. Lett.*, vol. 4, no. 8, pp. 232–235, Aug. 1997.

[37] P. Hyberg, M. Jansson, and B. Ottersten, "Array interpolation and bias reduction," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2711–2720, Oct. 2004.

[38] F. Belloni, A. Richter, and V. Koivunen, "DOA estimation via manifold separation for arbitrary array structures," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4800–4810, Oct. 2007.

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

[40] S. Semper and F. Römer, "ADMM for ND line spectral estimation using grid-free compressive sensing from multiple measurements with applications to DOA estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 4130–4134.

[41] M. Grant, and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," Sep. 2013. [Online]. Available: http://cvxr.com/cvx

[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[43] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*. Palo Alto, CA, USA: Meboo Publishing USA, 2010.

[44] L. Condat and A. Hirabayashi, "Cadzow denoising upgraded: A new projection method for the recovery of dirac pulses from noisy linear measurements," *Sampling Theory Signal Image Process.*, vol. 14, no. 1, pp. 17–47, 2015.

[45] M. Cho, J. F. Cai, S. Liu, Y. C. Eldar, and W. Xu, "Fast alternating projected gradient descent algorithms for recovering spectrally sparse signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 4638–4642.

[46] S. Liu, "Projected Wirtinger gradient descent for spectral compressed sensing," Ph.D. dissertation, Univ. Iowa, 2017.

[47] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ, USA: Prentice hall, 2000.

[48] J. Demmel and P. Koev, "The accurate and efficient solution of a totally positive generalized vandermonde linear system," *SIAM J. Matrix Anal.*, vol. 27, no. 1, pp. 142–152, 2005.

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

[50] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ, USA: Prentice Hall, 1977.

[51] M. G. Eberle and M. C. Maciel, "Finding the closest toeplitz matrix," *Comput. Appl. Math.*, vol. 22, no. 1, pp. 1–18, 2003.

[52] S. Boyd and J. Dattorro, "Alternating projections," *EE 392, Stanford Univ.*, 2003.

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

[54] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, Apr. 2015.

[55] M. Wagner, P. Gerstoft, and Y. Park, "Gridless DOA estimation via alternating projections," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 4215–4219.

[56] Y. Park and P. Gerstoft, "Alternating projections gridless covariance-based estimation for DOA," 2021, *arXiv:2102.06372*.

[57] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, "Structured low rank approximation," *Linear Algebra Appl.*, vol. 366, pp. 157–172, 2003.

[58] S. Aouada, A. M. Zoubir, and C. M. See, "A comparative study on source number detection," in *Proc. 7th Int. Symp. Signal Process. Appl.*, 2003, pp. 173–176.

[59] K. Han and A. Nehorai, "Improved source number detection and direction estimation with nested arrays and ULAs using jackknifing," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6118–6128, Dec. 2013.
[60] Y. Chi, Y. M. Lu, and Y. Chen, "Nonconvex optimization meets low-rank matrix factorization: An overview," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5239–5269, Oct. 2019.
[61] H. H. Bauschke and J. M. Borwein, "On the convergence of von-Neumann's alternating projection algorithm for two sets," *Set Valued Anal.*, vol. 1, no. 2, pp. 185–212, 1993.
[62] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, 2010.
[63] A. S. Lewis, D. R. Luke, and J. Malick, "Local linear convergence for alternating and averaged nonconvex projections," *Found. Comput. Math.*, vol. 9, no. 4, pp. 485–513, 2009.
[64] X. Jiang, Z. Zhong, X. Liu, and H. C. So, "Robust matrix completion via alternating projection," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 579–583, May 2017.
[65] X. Jiang, J. Chen, X. Liu, and A. M. Zoubir, "Phase-only robust minimum dispersion beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 4340–4344.
[66] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.
[67] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.
[68] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.
[69] P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru, "Multi-snapshot sparse Bayesian learning for DOA," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1469–1473, Oct. 2016.
[70] K. L. Gemba, S. Nannuru, P. Gerstoft, and W. S. Hodgkiss, "Multi-frequency sparse Bayesian learning for robust matched field processing," *J. Acoust. Soc. Amer.*, vol. 141, no. 5, pp. 3411–3420, 2017.
[71] K. L. Gemba, S. Nannuru, and P. Gerstoft, "Robust ocean acoustic localization with sparse Bayesian learning," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 1, pp. 49–60, Mar. 2019.
[72] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Statist. Soc. B.*, pp. 267–288, 1996.
[73] G. F. Edelmann and C. F. Gaumond, "Beamforming using compressive sensing," *J. Acoust. Soc. Amer.*, vol. 130, no. 4, pp. 232–237, 2011.
[74] 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.
[75] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Signal Process.*, vol. 33, no. 4, pp. 806–811, Aug. 1985.
[76] A. Das, W. S. Hodgkiss, and P. Gerstoft, "Coherent multipath direction-of-arrival resolution using compressed sensing," *IEEE J. Ocean. Eng.*, vol. 42, no. 2, pp. 494–505, Apr. 2017.
[77] S. Boyd *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[78] 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, Dec. 2013.
[79] 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, Mar. 2016.
[80] Y. Yang, Z. Chu, and G. Ping, "Two-dimensional multiple-snapshot grid-free compressive beamforming using alternating direction method of multipliers," *Shock Vib.*, vol. 2020, pp. 524–540, 2020.

Mark Wagner (Member, IEEE) received the Ph.D. degree from the University of California San Diego, San Diego, CA, USA, in 2020. Since graduation, he has been a Systems Engineer with Northrop Grumman Aerospace Systems. His research interests includes electromagnetic propagation through abnormal atmospheric conditions, wireless communication system design, direction of arrival estimation, and nonconvex optimization algorithms.

Yongsung Park (Member, IEEE) received the Ph.D. degree from Seoul National University (SNU), Seoul, South Korea, in 2019. In 2019, he was also a Researcher with the Research Institute of Marine Systems Engineering, SNU. Since 2019, he has been with the University of California San Diego, San Diego, CA, USA, where he is currently a Postdoctoral Researcher. His research interests include underwater acoustics, ocean engineering, array signal processing, sparse signal recovery, statistical signal processing, bayesian inference, and machine learning.

Peter Gerstoft (Senior Member, IEEE) received the Ph.D. degree from the Technical University of Denmark, Lyngby, Demark, in 1986. Since 1997, he has been with the University of California San Diego, San Diego, CA, USA. His current research interests include signal processing and machine learning applied to acoustic, seismic, and electromagnetic signals.