



Grid-free compressive beamforming

Xenaki, Angeliki; Gerstoff, Peter

Published in:
Acoustical Society of America. Journal

Link to article, DOI:
[10.1121/1.4916269](https://doi.org/10.1121/1.4916269)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Xenaki, A., & Gerstoff, P. (2015). Grid-free compressive beamforming. *Acoustical Society of America. Journal*, 137(4), 1923-1935. <https://doi.org/10.1121/1.4916269>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Grid-free compressive beamforming

Angeliki Xenaki^{a)}

Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

Peter Gerstoft

Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093-0238

(Received 5 November 2014; revised 3 February 2015; accepted 2 March 2015)

The direction-of-arrival (DOA) estimation problem involves the localization of a few sources from a limited number of observations on an array of sensors, thus it can be formulated as a sparse signal reconstruction problem and solved efficiently with compressive sensing (CS) to achieve high-resolution imaging. On a discrete angular grid, the CS reconstruction degrades due to basis mismatch when the DOAs do not coincide with the angular directions on the grid. To overcome this limitation, a continuous formulation of the DOA problem is employed and an optimization procedure is introduced, which promotes sparsity on a continuous optimization variable. The DOA estimation problem with infinitely many unknowns, i.e., source locations and amplitudes, is solved over a few optimization variables with semidefinite programming. The grid-free CS reconstruction provides high-resolution imaging even with non-uniform arrays, single-snapshot data and under noisy conditions as demonstrated on experimental towed array data.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4916269>]

[ZHM]

Pages: 1923–1935

I. INTRODUCTION

Sound source localization with sensor arrays involves the estimation of the direction-of-arrival (DOA) of (usually a few) sources from a limited number of observations. Compressive sensing^{1,2} (CS) is a method for solving such underdetermined problems with a convex optimization procedure which promotes sparse solutions.

Solving the DOA estimation as a sparse signal reconstruction problem with CS, results in robust, high-resolution acoustic imaging,^{3–6} outperforming traditional methods⁷ for DOA estimation. Furthermore, in ocean acoustics, CS is shown to improve the performance of matched field processing,^{8,9} which is a generalized beamforming method for localizing sources in complex environments (e.g., shallow water), and of coherent passive fathometry in inferring the number and depth of sediment layer interfaces.¹⁰

One of the limitations of CS in DOA estimation is basis mismatch¹¹ which occurs when the sources do not coincide with the look directions due to inadequate discretization of the angular spectrum. Under basis mismatch, spectral leakage leads to inaccurate reconstruction, i.e., estimated DOAs deviating from the actual ones. Employing finer grids^{3,12} alleviates basis mismatch at the expense of increased computational complexity, especially in large two-dimensional or three-dimensional problems as encountered in seismic imaging, for example.^{13–15}

To overcome basis mismatch, we formulate the DOA estimation problem in a continuous angular spectrum and introduce a sparsity promoting measure for general signals, the atomic norm.¹⁶ The atomic norm minimization problem,

which has infinitely many unknowns, is solved efficiently over few optimization variables in the dual domain with semidefinite programming.¹⁷ Utilizing the dual optimal variables, we show that the DOAs are accurately reconstructed through polynomial rooting. It is demonstrated that grid-free CS gives robust, high-resolution reconstruction also with non-uniform arrays and noisy measurements, exhibiting great flexibility in practical applications.

Polynomial rooting is employed in several DOA estimation methods to improve the resolution. However, these methods involve the estimation of the cross-spectral matrix hence they require many snapshots and stationary incoherent sources and are suitable only for uniform linear arrays (ULA).¹⁸ Grid-free CS is demonstrated not to have these limitations.

Finally, we process acoustic data¹⁹ from measurements in the North-East (NE) Pacific with grid-free CS and demonstrate that the method provides high-resolution acoustic imaging even with single-snapshot data.

In this paper, vectors are represented by bold lowercase letters and matrices by bold uppercase letters. The symbols T , H denote the transpose and the Hermitian (i.e., conjugate transpose) operator, respectively, on vectors and matrices. The symbol $*$ denotes simple conjugation. The generalized inequality $\mathbf{X} \succeq 0$ denotes that the matrix \mathbf{X} is positive semidefinite. The ℓ_p -norm of a vector $\mathbf{x} \in \mathbb{C}^n$ is defined as $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$. By extension, the ℓ_0 -norm is defined as $\|\mathbf{x}\|_0 = \sum_{i=1}^n 1_{x_i \neq 0}$. The paper makes heavy use of convex optimization theory; for a summary see Appendix A.

II. DISCRETE DOA ESTIMATION

The DOA estimation problem involves the localization of usually a few sources from measurements on an array of

^{a)}Author to whom correspondence should be addressed. Electronic mail: anxe@dtu.dk

sensors. For simplicity, we assume that the sources are in the far-field of the array, such that the wavefield impinging on the array consists of a superposition of plane waves, that the processing is narrowband and the sound speed is known. Moreover, we consider the one-dimensional problem with a uniform linear array of sensors and the sources residing in the plane of the array.

The location of a source is characterized by the direction of arrival of the associated plane wave, $\theta \in [-90^\circ, 90^\circ]$, with respect to the array axis. The propagation delay from the i th potential source to each of the M array sensors is described by the steering (or replica) vector

$$\mathbf{a}(\theta_i) = e^{j2\pi(d/\lambda)[0, \dots, M-1]^T \sin \theta_i}, \quad (1)$$

where λ is the wavelength and d is the intersensor spacing.

Discretizing the half-space of interest, $\theta \in [-90^\circ, 90^\circ]$, into N angular directions the DOA estimation problem is expressed in a matrix-vector formulation

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (2)$$

where $\mathbf{y} \in \mathbb{C}^M$ is the vector of the wavefield measurements at the M sensors, $\mathbf{x} \in \mathbb{C}^N$ is the unknown vector of the complex source amplitudes at all N directions on the angular grid of interest, and \mathbf{A} is the sensing matrix which maps the signal to the observations

$$\mathbf{A}_{M \times N} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_N)]. \quad (3)$$

In the presence of additive noise $\mathbf{n} \in \mathbb{C}^M$, the measurement vector is described by

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}. \quad (4)$$

The noise is generated as independent and identically distributed complex Gaussian. The array signal-to-noise ratio (SNR) for a single-snapshot is used in the simulations, defined as $\text{SNR} = 20 \log_{10}(\|\mathbf{A}\mathbf{x}\|_2 / \|\mathbf{n}\|_2)$, which determines the noise ℓ_2 -norm, $\|\mathbf{n}\|_2 = \|\mathbf{A}\mathbf{x}\|_2 10^{-\text{SNR}/20}$.

A. Sparse signal reconstruction

Practically, we are interested in a fine resolution on the angular grid such that $M < N$ and problem (2) is underdetermined. A way to solve this ill-posed problem is to constrain the possible solutions with prior information.

Traditional methods solve the underdetermined problem (2) by seeking the solution with the minimum ℓ_2 -norm which fits the data as described by the minimization problem

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_2 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (5)$$

The minimization problem (5) is convex with analytic solution, $\hat{\mathbf{x}} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{y}$. However, it aims to minimize the energy of the signal rather than its sparsity, hence the resulting solution is non-sparse.

Conventional beamforming²⁰ (CBF) is the simplest source localization method and it is based on the ℓ_2 -norm method with the simplifying condition $\mathbf{A}\mathbf{A}^H = \mathbf{I}_M$. CBF

combines the sensor outputs coherently to enhance the signal at a specific look direction from the ubiquitous noise yielding the solution

$$\hat{\mathbf{x}}_{\text{CBF}} = \mathbf{A}^H \mathbf{y}. \quad (6)$$

CBF is robust to noise but suffers from low resolution and the presence of sidelobes.

A sparse solution \mathbf{x} is preferred by minimizing the ℓ_0 -norm leading to the minimization problem

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (7)$$

However, the minimization problem (7) is a non-convex combinatorial problem which becomes computationally intractable even for moderate dimensions. The breakthrough of compressive sensing^{1,2} (CS) came with the proof that for sufficiently sparse signals, $K \ll N$, $K < M$, and sensing matrices with sufficiently incoherent columns the minimization problem (7) is equivalent to the minimization problem

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (8)$$

where the ℓ_0 -norm is replaced with the ℓ_1 -norm. The problem (8) is the closest convex optimization problem to problem (7) and can be solved efficiently by convex optimization even for large dimensions.²¹

For noisy measurements Eq. (4), the constraint in Eq. (8) becomes $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$, where ϵ is the noise floor, i.e., $\|\mathbf{n}\|_2 \leq \epsilon$. Then, the solution is²²

$$\hat{\mathbf{x}}_{\text{CS}} = \underset{\mathbf{x} \in \mathbb{C}^N}{\text{argmin}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon, \quad (9)$$

which has the minimum ℓ_1 -norm while it fits the data up to the noise level.

Herein, we use the cvx toolbox for disciplined convex optimization which is available in the Matlab environment. It uses interior point solvers to obtain the global solution of a well-defined optimization problem.²³ Interior point methods solve an optimization problem with linear equality and inequality constraints by transforming it to a sequence of simpler linear equality constrained problems which are solved iteratively with the Newton's method (iterative gradient descent method) increasing the accuracy of approximation at each step.²⁴

B. Basis mismatch

CS offers improved resolution due to the sparsity constraint and it can be solved efficiently with convex optimization. However, CS performance in DOA estimation is limited by the coherence of the sensing matrix \mathbf{A} (see Ref. 5), described by the restricted isometry property,²⁵ and by basis mismatch^{11,12} due to inadequate discretization of the angular grid. Herein, we demonstrate a way to overcome the limitation of basis mismatch by solving the ℓ_1 -minimization problem on a grid-free, continuous spatial domain.

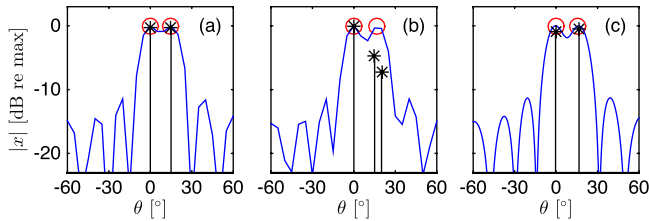


FIG. 1. (Color online) CS performance in DOA estimation in terms of the discretization of the angular space. A standard ULA is used with $M = 8$ sensors, $d/\lambda = 1/2$ and $\text{SNR} = 20$ dB. CBF and CS (*) reconstruction of two sources (a) at 0° and 15° on a grid $[-90^\circ : 5^\circ : 90^\circ]$, (b) at 0° and 17° on a grid $[-90^\circ : 5^\circ : 90^\circ]$, and (c) at 0° and 17° on a grid $[-90^\circ : 1^\circ : 90^\circ]$.

The fundamental assumption in CS is the sparsity of the underlying signal in the basis of representation, i.e., the sensing matrix \mathbf{A} . However, when the sources do not match with the selected angular grid, the signal might not appear sparse in the selected discrete Fourier transform basis.¹¹ Figure 1 shows the degradation of CS performance under basis mismatch due to inadequate discretization of the DOA domain in fast Fourier transform beamforming.

To increase the precision of the CS reconstruction, Malioutov *et al.*³ and Duarte and Baraniuk¹² propose an adaptive grid refinement. The adaptive grid refinement aims at improving the resolution of CS reconstruction without significant increase in the computational complexity by first detecting the regions where sources are present on a coarse grid and then refining the grid locally only at these regions. Grid refinement is an intuitive way of circumventing basis mismatch. However, the problem of basis mismatch is avoided only if the problem is solved in a continuous setting, particularly for moving sources.

III. CONTINUOUS DOA ESTIMATION

In the continuous approach, the K -sparse signal, x , is expressed as

$$x(t) = \sum_{i=1}^K x_i \delta(t - t_i), \quad (10)$$

where $x_i \in \mathbb{C}$ is the complex amplitude of the i th source, $t_i = \sin \theta_i$ is its support, i.e., the corresponding DOA, on the continuous sine spectrum $\mathbb{T} = [-1, 1]$ (with $T \subset \mathbb{T}$ the set of the DOAs of all K sources) and $\delta(t)$ is the Dirac delta function.

The sound pressure received at the m th sensor is expressed as a superposition of plane waves from all possible directions on the continuous sine spectrum \mathbb{T} ,

$$y_m = \int_{-1}^1 x(t) e^{j2\pi(d/\lambda)(m-1)t} dt = \sum_{i=1}^K x_i e^{j2\pi(d/\lambda)(m-1)t_i}, \quad (11)$$

and the measurement vector of the sensor array is

$$\mathbf{y}_{M \times 1} = \mathcal{F}_M x, \quad (12)$$

where \mathcal{F}_M is a linear operator (inverse Fourier transform) which maps the continuous signal x to the observations $\mathbf{y} \in \mathbb{C}^M$.

In the presence of additive noise, $\mathbf{n} \in \mathbb{C}^M$, the measurement vector is described by

$$\mathbf{y} = \mathcal{F}_M x + \mathbf{n}, \quad (13)$$

similarly to Eq. (4).

IV. GRID-FREE SPARSE RECONSTRUCTION

To solve the underdetermined problem (12) [or equivalently problem (13)] in favor of sparse solutions, we describe an optimization procedure which promotes sparsity on a continuous optimization variable.

A. Atomic norm

In the discrete formulation (4) of the DOA estimation problem, the prior information about the sparse distribution of sources is imposed through the ℓ_1 -norm of the vector \mathbf{x} to obtain sparse estimates (9). By extension, in the continuous formulation (13), we introduce the atomic norm,¹⁶ $\|\cdot\|_{\mathcal{A}}$, as a sparsity promoting measure for the continuous signal $x(t)$ in Eq. (10) defined as

$$\|x\|_{\mathcal{A}} = \sum_{i=1}^K |x_i|. \quad (14)$$

In other words, the atomic norm is a measure for continuous signals equivalent to the ℓ_1 -norm (which is defined only on vector spaces). Hence, the atomic norm is a convex function which promotes sparsity in a general framework. For a discrete grid the atomic norm corresponds to the ℓ_1 -norm.

To clarify the analogy between the ℓ_1 -norm and the atomic norm and justify the term *atomic*, consider that the vector $\mathbf{x} \in \mathbb{C}^N$ can be interpreted as a linear combination of N unit vectors. The unit vectors, in this case, are the smallest units, or *atoms*, in which the vector \mathbf{x} can be decomposed into. The ℓ_1 -norm is the sum of the absolute values of the weights of this linear combination of atoms.²⁴

Analogously, the continuous signal (10) can be interpreted as a linear combination of K delta functions $\delta(t - t_i)$, serving as atoms for the continuous signal $x(t)$ and the atomic norm is the sum of the absolute values of the weights of the linear combination of these atoms.¹⁶ Even though there are infinitely many atoms in the continuous case, only few of those, $K < M$, constitute the signal and the sum in Eq. (14) is finite.

B. Primal problem

Utilizing the convex measure of the atomic norm, the DOA estimation in the continuous angular space is solved with the sparsity promoting minimization problem,

$$\min_x \|x\|_{\mathcal{A}} \text{ subject to } \mathbf{y} = \mathcal{F}_M x. \quad (15)$$

Since the optimization variable x is a continuous parameter, the primal problem (15) is infinite dimensional and cannot be solved as such. It is possible to approximate the continuous variable x on a discrete grid and solve the ℓ_1 -norm optimization problem (8). This would increase the computational

complexity significantly when the discretization step is reduced to improve precision. An alternative to this, is to gradually refine the discretization step.³ However, we show that by solving the dual problem instead, there is no need to employ a discrete approximation of the continuous variable, x .

C. Dual problem

To formulate the dual problem to problem (15) (see Appendix A for details), we construct the Lagrangian by making the explicit equality constraints, $\mathbf{y} = \mathcal{F}_M x$, implicit in the objective function,

$$L(x, \mathbf{c}) = \|x\|_{\mathcal{A}} + \text{Re}[\mathbf{c}^H (\mathbf{y} - \mathcal{F}_M x)], \quad (16)$$

where $\mathbf{c} \in \mathbb{C}^M$ is the vector of dual variables.

The dual function $g(\mathbf{c})$ is the infimum, i.e., the greatest lower bound, of the Lagrangian, $L(x, \mathbf{c})$, over the primal optimization variable x ,

$$\begin{aligned} g(\mathbf{c}) &= \inf_x L(x, \mathbf{c}) \\ &= \text{Re}[\mathbf{c}^H \mathbf{y}] + \inf_x (\|x\|_{\mathcal{A}} - \text{Re}[\mathbf{c}^H \mathcal{F}_M x]). \end{aligned} \quad (17)$$

To evaluate the second term in Eq. (17) we note that for every x_i , $\text{Re}[(\mathbf{c}^H \mathcal{F}_M)_i x_i] = \text{Re}[(\mathcal{F}_M^H \mathbf{c})_i^H x_i] = |(\mathcal{F}_M^H \mathbf{c})_i| |x_i| \cos \phi_i$, where ϕ_i is the angle between x_i and $(\mathcal{F}_M^H \mathbf{c})_i$. Then,

$$\begin{aligned} |x_i| - \text{Re}[(\mathcal{F}_M^H \mathbf{c})_i^H x_i] &= |x_i| [1 - |(\mathcal{F}_M^H \mathbf{c})_i| \cos \phi_i] \\ &\geq |x_i| [1 - |(\mathcal{F}_M^H \mathbf{c})_i|]. \end{aligned} \quad (18)$$

The lower bound in Eq. (18) is non-negative if $|\mathcal{F}_M^H \mathbf{c}|$ is less than one, $\max_i |(\mathcal{F}_M^H \mathbf{c})_i| \leq 1$, and the infimum is zero. Otherwise, $|x_i| [1 - |(\mathcal{F}_M^H \mathbf{c})_i|] < 0$ and the infimum is attained at $-\infty$. Hence, the dual function is

$$g(\mathbf{c}) = \begin{cases} \text{Re}[\mathbf{c}^H \mathbf{y}], & \|\mathcal{F}_M^H \mathbf{c}\|_{\infty} \leq 1 \\ -\infty, & \text{otherwise.} \end{cases} \quad (19)$$

From Eq. (18), $|x_i| [1 - |(\mathcal{F}_M^H \mathbf{c})_i| \cos \phi_i] = 0$ at the infimum, which for every $x_i \neq 0$ yields $|(\mathcal{F}_M^H \mathbf{c})_i| \cos \phi_i = 1$, i.e., $|(\mathbf{c}^H \mathcal{F}_M)_i| = 1$ and $\phi_i = 0$, as both $|(\mathbf{c}^H \mathcal{F}_M)_i| \leq 1$ and $\cos \phi_i \leq 1$. Thus, for $x_i \neq 0$, $(\mathcal{F}_M^H \hat{\mathbf{c}})_i$ is a unit vector in the direction of x_i ,

$$\begin{aligned} (\mathcal{F}_M^H \hat{\mathbf{c}})_i &= x_i / |x_i|, \quad x_i \neq 0, \\ |\mathcal{F}_M^H \hat{\mathbf{c}}_i| &< 1, \quad x_i = 0. \end{aligned} \quad (20)$$

Maximizing the dual function (19) constitutes the dual problem

$$\max_{\mathbf{c} \in \mathbb{C}^M} \text{Re}[\mathbf{c}^H \mathbf{y}] \text{ subject to } \|\mathcal{F}_M^H \mathbf{c}\|_{\infty} \leq 1. \quad (21)$$

Since the primal problem (15) is convex with linear equality constraints (A11), strong duality holds assuring that the maximum of the dual problem (21) is equal to the minimum of the primal problem.

The dual problem (21) selects a vector $\mathbf{c} \in \mathbb{C}^M$ which is maximally aligned with the measurement vector $\mathbf{y} \in \mathbb{C}^M$ while

its beamformed amplitude $|\mathcal{F}_M^H \mathbf{c}|$ is bounded by unity across the whole angular spectrum. At the angular direction corresponding to the DOA of an existing source, the beamformed dual vector (20) is equal to the normalized source amplitude.

D. Dual problem using semidefinite programming

The dual problem (21) is a semi-infinite programming problem with a finite number of optimization variables, $\mathbf{c} \in \mathbb{C}^M$, and infinitely many inequality constraints, which is still intractable.

Define the dual polynomial

$$H(z) = \mathcal{F}_M^H \mathbf{c} = \sum_{m=0}^{M-1} c_m z^m = \sum_{m=0}^{M-1} c_m e^{-j[2\pi(d/\lambda)t]m}. \quad (22)$$

Note that $\mathcal{F}_M^H \mathbf{c}$ is a trigonometric polynomial (B1), of the variable $z(t) = e^{-j2\pi(d/\lambda)t}$, $t \in \mathbb{T}$, with the dual variables $\mathbf{c} = [c_0, \dots, c_{M-1}]^T$ as coefficients and degree $M-1$.

The inequality constraint in Eq. (21) implies that the dual polynomial has amplitude uniformly bounded for all $t \in \mathbb{T}$; see Eq. (B7). Making use of the approximation in Eq. (B6) for bounded trigonometric polynomials, the constraint in Eq. (21) can be replaced with finite dimensional linear matrix inequalities. Thus, the dual problem is solved with semidefinite programming,^{23,24} i.e., a convex optimization problem where the inequality constraints are linear matrix inequalities with semidefinite matrices

$$\begin{aligned} \max_{\mathbf{c}, \mathbf{Q}} \text{Re}(\mathbf{c}^H \mathbf{y}) \text{ subject to } & \begin{bmatrix} \mathbf{Q}_{M \times M} & \mathbf{c}_{M \times 1} \\ \mathbf{c}_{1 \times M}^H & 1 \end{bmatrix} \succeq 0, \\ \sum_{i=1}^{M-j} \mathbf{Q}_{i,i+j} &= \begin{cases} 1, & j = 0 \\ 0, & j = 1, \dots, M-1. \end{cases} \end{aligned} \quad (23)$$

The number of optimization variables of the dual problem (23) is $(M+1)^2/2$ equal to half the number of elements of the Hermitian matrix in the inequality constraint. Thus, a problem with infinitely many unknown parameters, Eq. (15), is solved over a few optimization variables.

E. Support detection through the dual polynomial

Strong duality assures that by solving the dual problem (21), or equivalently Eq. (23), we obtain the minimum of the primal problem (15). However, the dual problem provides an optimal dual vector, $\hat{\mathbf{c}}$, but not the primal solution, \hat{x} . Since the corresponding dual polynomial, $H(z) = \mathcal{F}_M^H \hat{\mathbf{c}}$, has the properties in Eq. (20), the support \hat{T} of the primal solution \hat{x} can be estimated by locating the angular directions t_i where the amplitude of the dual polynomial is one (i.e., the angular directions at the maxima of the beamformed dual vector),

$$|H(z)| \leq 1, \quad \forall t \in \mathbb{T} \rightarrow \begin{cases} |H[z(t_i)]| = 1, & t_i \in \hat{T} \\ |H[z(t)]| < 1, & t \in \mathbb{T} \setminus \hat{T}. \end{cases} \quad (24)$$

Following Appendix B4, this is done by locating the roots of the non-negative polynomial which lie on the unit circle $|z| = 1$ (see also Sec. VIII B),

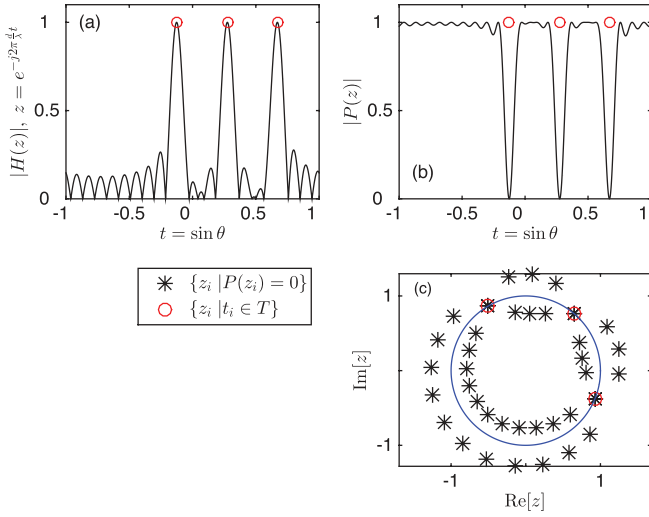


FIG. 2. (Color online) Support detection through the dual polynomial. A ULA is used with $M = 21$ sensors and $d/\lambda = 1/2$ to localize three sources with support set $T = [-0.126, 0.275, 0.67]$. (a) The dual polynomial $|H(z)|$. (b) The non-negative polynomial $P(z)$. (c) The support T is estimated by the angle of the roots, z_i , of $P(z)$ for which $|z_i| = 1$.

$$P(z) = 1 - R(z) = 1 - \sum_{m=-(M-1)}^{M-1} r_m z^m, \quad (25)$$

where $R(z) = H(z)H(z)^H = |H(z)|^2$ with coefficients $r_m = \sum_{l=0}^{M-1-m} \hat{c}_l \hat{c}_{l+m}^*$, $m \geq 0$ and $r_{-m} = r_m^*$, i.e., the autocorrelation of $\hat{\mathbf{c}}$.

Note that the polynomial of degree $2(M-1)$,

$$P_+(z) = z^{M-1}P(z) = (1 - r_0)z^{M-1} - \sum_{m=-(M-1), m \neq 0}^{M-1} r_m z^{(m+M-1)}, \quad (26)$$

which has only positive powers of the variable z , has the same roots as $P(z)$, besides the trivial root $z = 0$. Thus, the support \hat{T} of $\hat{\mathbf{x}}$, i.e., the DOAs of the sources, is recovered by locating the roots of $P_+(z)$ on the unit circle (see Fig. 2),

$$\hat{T} = \left\{ t_i = \frac{\lambda}{2\pi d} \arg z_i \mid P_+(z_i) = 0, |z_i| = 1 \right\}. \quad (27)$$

F. Reconstruction of the primal solution \mathbf{x}

Once the support is recovered by locating the roots of the polynomial in Eq. (26) that lie on the unit circle (27), the source amplitudes [the complex weights in Eq. (10)] are recovered from

$$\hat{\mathbf{x}}_{\text{CS,dual}} = \mathbf{A}_T^+ \mathbf{y}, \quad (28)$$

where $^+$ denotes the pseudoinverse of \mathbf{A}_T with columns $\mathbf{a}(t_i) = e^{j2\pi(d/\lambda)[0, \dots, M-1]^T t_i}$ for $t_i \in \hat{T}$.

Figure 3 shows the DOA estimation with grid-free CS following the procedure described in this section (see Appendix C for a MATLAB implementation). The dual polynomial attains unit amplitude, $|H(z)| = 1$, at the support of the

solution, i.e., the DOAs of the existing sources; see Fig. 3(a). Figure 3(b) compares the grid-free CS (28) and the CBF reconstruction in DOA estimation. The grid-free CS offers very accurate localization, while CBF is characterized by low resolution. Moreover, CBF fails to detect the weak source at 15.962° since it is totally masked by the sidelobes.

V. MAXIMUM RESOLVABLE DOAS

The maximum number of resolvable DOAs with grid-free CS is determined by the maximum number of roots of $P_+(z)$ in Eq. (26) which can be on the unit circle, $|z| = 1$. Since the coefficients of the polynomial $P_+(z)$ are conjugate symmetric around the term z^{M-1} , the roots appear in pairs at the same angular direction t_l , one inside the unit circle, $z_{\text{in}} = r_l e^{-j2\pi(d/\lambda)t_l}$ at radius $r_l < 1$, and the other outside of the unit circle $z_{\text{out}} = (1/r_l)e^{-j2\pi(d/\lambda)t_l} = 1/(z_{\text{in}})^H$. This implies that the roots on the unit circle have double multiplicity. The polynomial $P_+(z)$ has in total $2(M-1)$ roots, as determined by its degree. Hence, there are at most $M-1$ (double) roots on the unit circle.

The necessary condition for the dual polynomial (24) to satisfy the condition $|H(z)| < 1$ for some $t \in \mathbb{T}$, thus avoid the non-informative case of a constant dual polynomial, is that the number of sources should not exceed^{17,26}

$$K_{\text{max}} = \left\lfloor \frac{M-1}{2} \right\rfloor, \quad (29)$$

where $\lfloor \cdot \rfloor$ is the largest integer not greater than the argument. In other words, at least half of the (paired) $M-1$ roots should lie off the unit circle alternating with the roots on the unit circle leading to the bound (29).

For positive source amplitudes, $x_i \in \mathbb{R}_+$, the condition (29) is sufficient and no separation condition is required for the resolvable sources.²⁶ However, for complex amplitudes, $x_i \in \mathbb{C}$, the sources are resolved uniquely only if the corresponding DOAs are separated by at least,^{17,27}

$$\min_{t_i, t_j \in T} |t_i - t_j| = \frac{\lambda}{Md}, \quad (30)$$

where $|t_i - t_j|$ is a wrap-around distance meaning that we identify the points $-1, 1$, in $\mathbb{T} = [-1, 1]$.

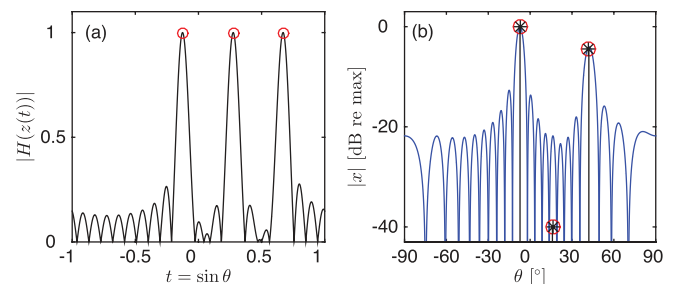


FIG. 3. (Color online) Grid-free sparse reconstruction. A standard ULA is used with $M = 21$ sensors and $d/\lambda = 1/2$ to localize three sources (O) at $\theta = [-7.2385^\circ, 15.962^\circ, 42.0671^\circ]$ with amplitudes $|\mathbf{x}| = [1, 0.01, 0.6]$. (a) The dual polynomial. (b) Reconstruction with grid-free CS (*) and CBF.

The minimum separation condition (30) is a consequence of the coherence of the sensing process which is related to the beampattern; see Sec. IVD in Ref. 5. To guarantee a well-posed sparse signal reconstruction, it is required that the columns of the inverse Fourier operator \mathcal{F}_M , the steering vectors (1), are sufficiently uncorrelated. The continuous formulation (12) implies that adjacent steering vectors are in arbitrarily close directions, hence fully coherent. However, the requirement (30) inhibits closely spaced (i.e., highly correlated) steering vectors, hence prevents the sparse reconstruction problem from being too ill-posed due to coherence.

Figure 4 shows the reconstruction for the maximum number of sources possible. For positive source amplitudes, $x_i \in \mathbb{R}_+$, the bound (29) suffices to ensure a unique solution. Grid-free CS achieves super-resolution even for DOAs in general position; see Figs. 4(a)–4(b). Inserting an additional source at 71.81° , thus exceeding the maximum number of resolvable sources (29), results in a non-informative dual polynomial, $|H(z)| \approx 1$, for all $t \in \mathbb{T}$, Fig. 4(c), and inaccurate reconstruction where only seven out of the 11 sources are resolved, Fig. 4(d). For complex source amplitudes, $x_i \in \mathbb{C}$, an additional constraint (30) on the minimum

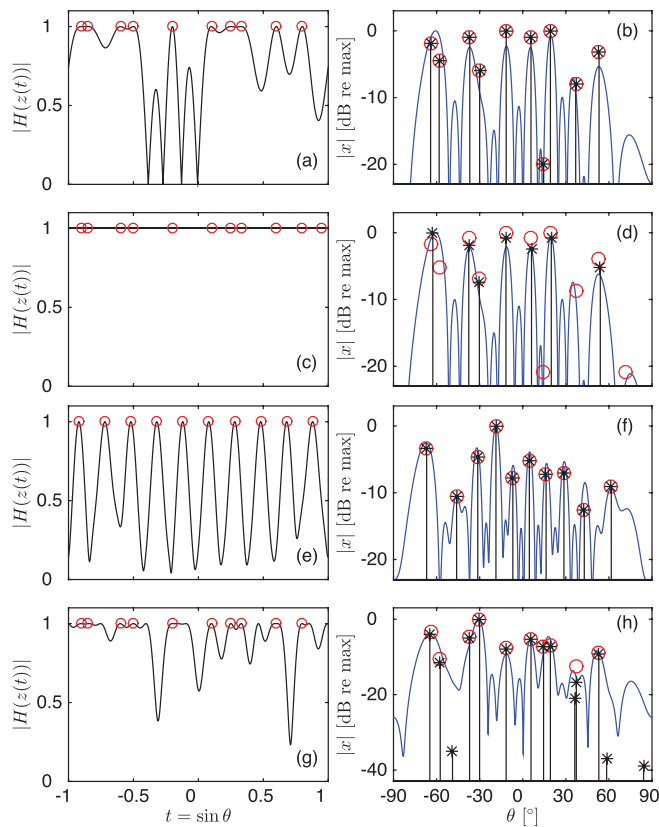


FIG. 4. (Color online) Grid-free sparse reconstruction. A ULA is used with $M = 21$ sensors and $d/\lambda = 1/2$ to localize the possible maximum number of sources (\circ) , $\lfloor (M-1)/2 \rfloor = 10$. (a) The dual polynomial and (b) reconstruction with grid-free CS $(*)$ and CBF for sources with positive amplitudes, $x_{10, \mathbb{R}} = [0.8, 0.6, 0.9, 0.5, 1, 0.9, 0.1, 1, 0.4, 0.7]$. (c) The dual polynomial and (d) reconstruction for 11 sources with positive amplitudes, $x_{11, \mathbb{R}} = [x_{10, \mathbb{R}}, 0.1]$. (e) The dual polynomial and (f) reconstruction for sources with complex amplitudes, $x_{10, \mathbb{C}} = x_{10, \mathbb{R}} + i[-1.6, 0.5, -1.3, -2.6, 0.4, -1.2, -1.2, -0.6, -0.5, 0.6]$, separated by the condition (30). (g) The dual polynomial and (h) reconstruction for sources with complex amplitudes, $x_{10, \mathbb{C}}$, but locations violating the condition (30).

separation of DOAs is required along with the bound on the number of sources (29) to ensure a unique solution, Figs. 4(e)–4(f). Violating the minimum separation condition, the CS DOA estimation becomes extremely ill-posed due to the coherence of the underlying steering vectors resulting in inaccurate reconstruction characterized by the presence of spurious sources, Figs. 4(g)–4(h).

VI. NON-UNIFORM ARRAYS

The method is also applicable to non-uniform arrays, constructed by randomly choosing sensors from a standard ULA configuration, by adding an additional constraint in the optimization problem (23).²⁷ The additional constraint ensures that coefficients of the dual polynomial corresponding to inactive sensors on the ULA, $c_{m_{\text{null}}}$, are annihilated.

The dual problem in a semidefinite programming formulation (23) is augmented with an additional constraint and takes the form

$$\max_{\mathbf{c}, \mathbf{Q}} \text{Re}(\mathbf{c}^H \mathbf{y}) \text{ subject to } \begin{bmatrix} \mathbf{Q}_{M \times M} & \mathbf{c}_{M \times 1} \\ \mathbf{c}_{1 \times M}^H & 1 \end{bmatrix} \succeq 0, \quad (31)$$

$$\sum_{i=1}^{M-j} \mathbf{Q}_{i, i+j} = \begin{cases} 1, & j=0 \\ 0, & j=1, \dots, M-1, c_{m_{\text{null}}} = 0. \end{cases}$$

Figure 5 shows the DOA estimation with grid-free CS and compares it with the CBF reconstruction in the case of a random array. Even though CBF performance degrades significantly due to the increased sidelobe levels introduced by the random array and the strong source towards endfire, CS still offers exact reconstruction.

VII. GRID-FREE RECONSTRUCTION WITH NOISE

The problem of grid-free DOA estimation with CS extends to noisy measurements making the framework useful for practical applications. Assuming that the measurements [Eq. (13)] are contaminated with additive noise $\mathbf{n} \in \mathbb{C}^M$, such that $\|\mathbf{n}\|_2 \leq \epsilon$, the atomic norm minimization problem (15) is reformulated as²⁸

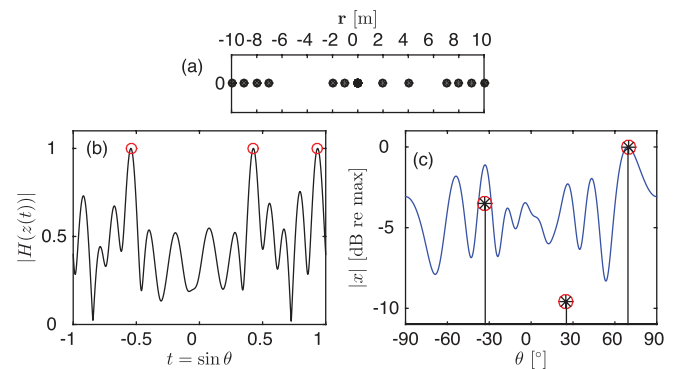


FIG. 5. (Color online) Grid-free sparse reconstruction. (a) A random array constructed by randomly selecting $M = 13$ sensors out of a standard ULA with 21 sensors and $d/\lambda = 1/2$. The sources (\circ) are at $\theta = [-32.8881^\circ, 25.2773^\circ, 69.3903^\circ]$ with amplitudes $|x| = [0.67, 0.33, 1]$. (b) The dual polynomial. (c) Reconstruction with grid-free CS $(*)$ and CBF.

$$\min_x \|x\|_{\mathcal{A}} \text{ subject to } \|\mathbf{y} - \mathcal{F}_M x\|_2 \leq \epsilon. \quad (32)$$

To solve the infinite dimensional primal problem (32) we formulate the equivalent dual problem (see Appendix D)

$$\max_{\mathbf{c}} \text{Re}(\mathbf{c}^H \mathbf{y}) - \epsilon \|\mathbf{c}\|_2 \text{ subject to } \|\mathcal{F}_M^H \mathbf{c}\|_{\infty} \leq 1, \quad (33)$$

and we replace the infinite-dimensional constraints with finite matrix inequalities,

$$\begin{aligned} \max_{\mathbf{c}, \mathbf{Q}} \text{Re}(\mathbf{c}^H \mathbf{y}) - \epsilon \|\mathbf{c}\|_2 \text{ subject to } & \begin{bmatrix} \mathbf{Q}_{M \times M} & \mathbf{c}_{M \times 1} \\ \mathbf{c}_{1 \times M}^H & 1 \end{bmatrix} \succeq 0, \\ \sum_{i=1}^{M-j} \mathbf{Q}_{i,i+j} = & \begin{cases} 1, & j = 0 \\ 0, & j = 1, \dots, M-1. \end{cases} \end{aligned} \quad (34)$$

Problem (34) is a convex optimization problem which can be solved efficiently with semidefinite programming²³ to obtain an estimate for the coefficients, $\mathbf{c} \in \mathbb{C}^M$, of the dual polynomial. The support of the solution, i.e., the DOAs of the existing sources is found by locating the points where the dual polynomial has unit amplitude following the methodology in Sec. IV E. Once the support is recovered the source amplitudes are estimated by solving a discrete overdetermined problem, Eq. (28).

Figure 6 shows the DOA estimation for three sources with grid-free CS when the array measurements are contaminated with additive noise [Eq. (13)] such that SNR = 20 dB. Grid-free CS improves significantly the resolution in the reconstruction compared to CBF, even though some weak spurious sources appear as artifacts due to the noise in the measurements.

VIII. DOA ESTIMATION WITH POLYNOMIAL ROOTING

Polynomial rooting can increase performance and achieve super-resolution in several DOA estimation methods, such as the minimum variance distortionless response (MVDR) beamformer, the multiple signal classification (MUSIC) method and the minimum-norm method. All these methods involve the estimation or the eigendecomposition of the cross-spectral matrix both in their spectral and root version.

The cross-spectral matrix estimated from L snapshots (i.e., observations of \mathbf{y} at a particular frequency) is defined as

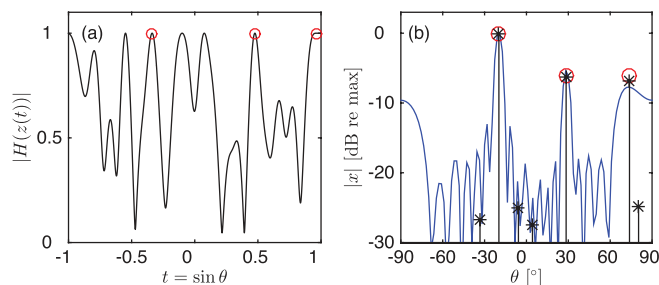


FIG. 6. (Color online) Grid-free sparse reconstruction. A ULA is used with $M = 21$ sensors and $d/\lambda = 1/2$ to localize three sources (\odot) at $\theta = [-19.6942^\circ, 28.3594^\circ, 73.9457^\circ]$ with amplitudes $|x| = [0.6, 0.3, 0.3]$. (a) The dual polynomial. (b) Reconstruction with grid-free CS ($*$) and CBF. The SNR is 20dB.

$$\hat{\mathbf{C}}_y = \frac{1}{L} \sum_{l=1}^L \mathbf{y}_l \mathbf{y}_l^H. \quad (35)$$

The eigendecomposition of the cross-spectral matrix separates the signal and the noise subspaces

$$\hat{\mathbf{C}}_y = \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{U}}_n^H, \quad (36)$$

where $\hat{\mathbf{U}}_s$ comprises the signal eigenvectors, which correspond to the largest eigenvalues $\hat{\mathbf{\Lambda}}_s$, and $\hat{\mathbf{U}}_n$ comprises the noise eigenvectors. The signal eigenvectors are in the same subspace as the steering vectors (1), while the noise eigenvectors are orthogonal to the subspace of the steering vectors, thus $\mathbf{a}(\theta)^H \hat{\mathbf{U}}_n = \mathbf{0}$.

A. Spectral version of DOA estimation methods

MVDR (Ref. 29) aims to minimize the output power of the beamformer under the constraint that the signal from the look direction remains undistorted. The MVDR beamformer power spectrum is

$$P_{\text{MVDR}}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \hat{\mathbf{C}}_y^{-1} \mathbf{a}(\theta)}. \quad (37)$$

MUSIC (Ref. 30) uses the orthogonality between the signal and the noise subspace to locate the maxima in the spectrum

$$P_{\text{MUSIC}}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \mathbf{a}(\theta)}. \quad (38)$$

The minimum-norm is also an eigendecomposition based method but, unlike MUSIC which utilizes all noise eigenvectors, it uses a single vector, $\mathbf{v} = [v_0, \dots, v_{M-1}]^T$, which resides in the noise subspace [compare with the dual vector $\hat{\mathbf{c}}$ Eq. (20) which resides in the signal subspace] such that

$$\mathbf{a}(\theta_i)^H \mathbf{v} = 0, \quad i = 1, \dots, K, \quad (39)$$

where K is the number of sources.

All the noise subspace eigenvectors, i.e., the columns of $\hat{\mathbf{U}}_n$ have the property in Eq. (39). However, if the vector \mathbf{v} is chosen as a linear combination of the noise subspace eigenvectors the algorithm tends to be more robust.^{18,31,32}

The minimum-norm method selects a vector, \mathbf{v} , in the noise subspace with minimum ℓ_2 -norm and unit first element, $v_0 = 1$. The vector \mathbf{v} can be constructed from the noise eigenvectors as³¹

$$\mathbf{v} = \hat{\mathbf{U}}_n \mathbf{d}^H / \|\mathbf{d}\|_2^2, \quad (40)$$

where the vector \mathbf{d} is the first row of $\hat{\mathbf{U}}_n$. Equivalently, the vector \mathbf{v} can be constructed from the signal eigenvectors as

$$\mathbf{v} = \hat{\mathbf{U}}_s \frac{\mathbf{b}^H}{1 - \|\mathbf{b}\|_2^2}, \quad (41)$$

where the vector \mathbf{b} is the first row of $\hat{\mathbf{U}}_s$.

The minimum-norm spectrum is

$$P_{\min\text{-norm}}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \mathbf{v} \mathbf{v}^H \mathbf{a}(\theta)}. \quad (42)$$

B. Root version of DOA estimation methods

The root version of the DOA estimation methods is based on the fact that for ULAs the null spectrum has the form of the trigonometric polynomial in Eq. (B2) with $\omega = 2\pi(d/\lambda) \sin \theta$ (since $\sin \theta \in [-1, 1]$, then for a standard ULA $\omega \in [-\pi, \pi]$). Thus, evaluating the spectrum is equivalent to evaluating the roots of the polynomial on the unit circle.³³

More analytically, let $N(\theta) = \mathbf{a}(\theta)^H \mathbf{\Psi} \mathbf{a}(\theta)$ be the null spectrum, such that the spectrum is $S(\theta) = N(\theta)^{-1}$. For MVDR, $\mathbf{\Psi} = \hat{\mathbf{C}}_y^{-1}$ (Ref. 18, p. 1147), for MUSIC, $\mathbf{\Psi} = \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$ (Ref. 18, p. 1159) and for the minimum-norm method, $\mathbf{\Psi} = \mathbf{v} \mathbf{v}^H$ (Ref. 18, p. 1163). Then

$$\begin{aligned} N(\theta) &= \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} e^{-j2\pi m(d/\lambda) \sin \theta} \Psi_{mn} e^{-j2\pi n(d/\lambda) \sin \theta} \\ &= \sum_{l=-(M-1)}^{M-1} \psi_l e^{-j2\pi l(d/\lambda) \sin \theta}, \\ N(z) &= \sum_{l=-(M-1)}^{M-1} \psi_l z^{-l}, \end{aligned} \quad (43)$$

where $\psi_l = \sum_{m=-n=l} \Psi_{mn}$ is the sum of the elements of the Hermitian matrix $\mathbf{\Psi}$ along the l th diagonal and $z = e^{j2\pi(d/\lambda) \sin \theta}$.

The set of DOAs, \hat{T} , is estimated from the roots of the polynomial $N(z)$, or equivalently the polynomial $N_+(z) = z^{M-1} N(z)$, which lie on the unit circle, $z_i = e^{j \arg(z_i)}$ as

$$\hat{T} = \left\{ \sin \theta_i = \frac{\lambda}{2\pi d} \arg z_i \mid N_+(z_i) = 0, |z_i| = 1 \right\}. \quad (44)$$

After the support is recovered, the amplitudes can be estimated through an overdetermined problem as in Eq. (28).

Even though the root forms of DOA estimation methods have, often, more robust performance than the corresponding spectral forms,³⁴ they require a regular array geometry to form a trigonometric polynomial and detect its roots behavior. To achieve a robust estimate of the cross-spectral matrix many snapshots are required, $L > M$, i.e., stationary sources. Furthermore, eigendecomposition based methods fail to discern coherent arrivals. Forward/backward smoothing techniques^{35,36} can be employed to mitigate this problem and make eigendecomposition based methods suitable for identification of coherent sources as well, but they still require a regular array geometry and an increased number of sensors.

IX. EXPERIMENTAL RESULTS

The high-resolution capabilities of sparse signal reconstruction methods, i.e., CS for DOA estimation, and the robustness of grid-free sparse reconstruction even under

noisy conditions and with random array configurations are demonstrated on ocean acoustic measurements. The interest is on single-snapshot reconstruction for source tracking and the results are compared with CBF.

The data set is from the long range acoustic communications (LRAC) experiment¹⁹ recorded from 10:00–10:30 UTC on 16 September 2010 in the NE Pacific and is the same as in Ref. 5 to allow comparison of the results. The data are from a horizontal uniform linear array towed at 3.5 knots at 200 m depth. The array has $M = 64$ sensors, with intersensor spacing $d = 3$ m. The data were acquired with a sampling frequency of 2000 Hz and the record is divided in 4 s non-overlapping snapshots. Each snapshot is Fourier transformed with 2^{13} samples.

The data are post-processed with CBF and CS on a discrete DOA grid $[-90^\circ : 1^\circ : 90^\circ]$ as well as grid-free CS at frequency $f = 125$ Hz ($d/\lambda = 1/4$). To facilitate the comparison of the results, the grid-free CS reconstruction is also presented on the grid $[-90^\circ : 1^\circ : 90^\circ]$ by rounding the estimated DOAs to the closest integer angle and using the maximum power within each bin. The results are depicted in Fig. 7 both with all $M = 64$ sensors active, Figs. 7(a)–7(d) and by retaining only $M = 16$ sensors active in a non-uniform

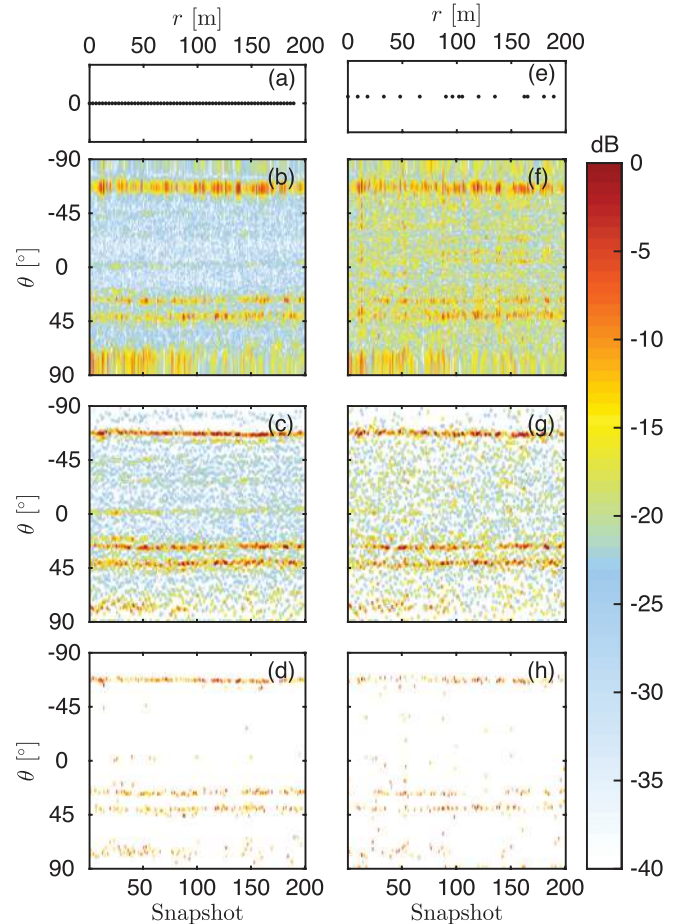


FIG. 7. (Color online) Data from LRAC. (a) Uniform array with $M = 64$ sensors and the corresponding (b) CBF, (c) CS on a discrete grid, $[-90^\circ : 1^\circ : 90^\circ]$, and (d) grid-free CS reconstruction. (e) Non-uniform array with $M = 16$ sensors and the corresponding (f) CBF, (g) CS on a discrete grid and (h) grid-free CS reconstruction.

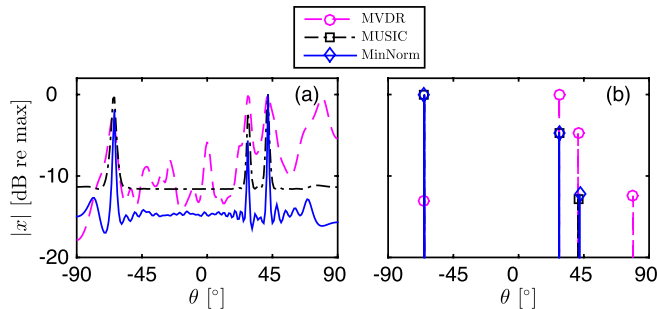


FIG. 8. (Color online) Data from LRAC, combining the 200 snapshots to estimate the cross-spectral matrix and processing with MVDR, MUSIC, and the minimum-norm method. (a) Spectral version and (b) root version. The ULA with $M = 64$ sensors and $d/\lambda = 1/4$ is used.

configuration, Figs. 7(e)–7(h). Both array configurations, Figs. 7(a) and 7(e), have the same aperture thus the same resolution.

The CBF map (6) in Fig. 7(b) indicates the presence of three stationary sources at around 45° , 30° , and -65° . The two arrivals at 45° and 30° are attributed to distant transiting ships, even though a record of ships in the area was not kept. The broad arrival at -65° is from the tow ship R/V Melville. The CBF map suffers from low resolution and artifacts due to sidelobes and noise. The CS reconstruction (9) [$\epsilon = 3.5$, Fig. 7(c)] results in improved resolution in the localization of the three sources by promoting sparsity and significant reduction of artifacts in the map. The grid-free CS solution (28), Fig. 7(d), provides high resolution and further artifact reduction due to polynomial rooting.

Retaining only 1/4 of the sensors on the array in a non-uniform configuration degrades the resolution of CBF due to increased sidelobe levels, Fig. 7(f). However, both CS on a discrete DOA grid, Fig. 7(g), and grid-free CS, Fig. 7(h), provide high-resolution DOA estimation without a significant reconstruction degradation.

The single-snapshot processing, Fig. 7, indicates that the sources are adequately stationary. Therefore, the 200 snapshots can be combined to estimate the cross-spectral matrix (35) and employ cross-spectral methods for DOA estimation. Figure 8(a) compares the power spectra of MVDR (37), MUSIC (38), and the minimum-norm method (42) and Fig. 8(b) the corresponding root versions.

The root versions of cross-spectral methods, especially the root MUSIC and the root minimum-norm method, provide improved resolution compared to the corresponding spectral forms. However, the root cross-spectral methods require both many snapshots (i.e., stationary sources) for a robust estimate of the cross-spectral matrix and uniform arrays. Grid-free CS does not have these limitations.

X. CONCLUSION

DOA estimation with sensor arrays is a sparse signal reconstruction problem which can be solved with CS. Discretization of the problem involves a compromise between the quality of reconstruction and the computational complexity, especially for high-dimensional problems. Grid-free CS assures that the sparsity promoting optimization

problem in CS can be solved in the dual domain with semi-definite programming even when the unknowns are infinitely many. Grid-free CS achieves high-resolution DOA estimation through the polynomial rooting method.

In contrast to established DOA estimation methods, CS provides high-resolution acoustic imaging even with non-uniform array configurations and robust performance under noisy measurements and single-snapshot data. Finally, the grid-free CS has the same performance both with coherent and incoherent, stationary or moving sources while other DOA estimation methods based on polynomial rooting fail to discern coherent arrivals and have degraded resolution for moving sources as they require many snapshots.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, under Grant No. N00014-11-1-0320.

APPENDIX A: CONVEX OPTIMIZATION PROBLEMS

This section summarizes the basic notions and formulations encountered in convex optimization problems, as presented analytically in Ref. 24.

1. Primal problem

A generic optimization problem has the form

$$\begin{aligned} \min_{\mathbf{x}} f_0(\mathbf{x}) \\ \text{subject to } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ h_j(\mathbf{x}) = 0, \quad j = 1, \dots, q, \end{aligned} \quad (\text{A1})$$

where $\mathbf{x} \in \mathbb{C}^N$ is the optimization variable, the function $f_0 : \mathbb{C}^N \rightarrow \mathbb{R}$ is the objective (or cost) function, the functions $f_i : \mathbb{C}^N \rightarrow \mathbb{R}$ are the inequality constraint functions and the functions $h_j : \mathbb{C}^N \rightarrow \mathbb{C}$ are the equality constraint functions. The optimization problem (A1) is convex when f_0, \dots, f_m are convex functions and h_1, \dots, h_q are affine (linear) functions.

The set of points for which the objective and all constraint functions in Eq. (A1) are defined is called the domain of the optimization problem

$$\mathcal{D} = \bigcap_{i=0}^m \text{dom } f_i \cap \bigcap_{j=1}^q \text{dom } h_j. \quad (\text{A2})$$

A point $\tilde{\mathbf{x}} \in \mathcal{D}$ is called feasible if it satisfies the constraints in Eq. (A1).

The optimal value p^* of the optimization problem (A1), achieved at the optimal variable \mathbf{x}^* , is

$$\begin{aligned} p^* &= \inf \{f_0(\mathbf{x}) \mid f_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0\} \\ &= \{f_0(\mathbf{x}^*) \mid f_i(\mathbf{x}^*) \leq 0, h_j(\mathbf{x}^*) = 0\}, \end{aligned} \quad (\text{A3})$$

for all $i = 1, \dots, m$ and $j = 1, \dots, q$.

2. The Lagrangian

The Lagrangian, L , of an optimization problem is obtained by augmenting the objective function with a weighted sum of the constraint functions. The Lagrangian of the generic optimization problem (A1) is

$$L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{v}) = f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}) + \operatorname{Re} \left[\sum_{j=1}^q \nu_j h_j(\mathbf{x}) \right], \quad (\text{A4})$$

where λ_i is the Lagrange multiplier associated with the i th inequality constraint, $f_i(\mathbf{x}) \leq 0$, and ν_j is the Lagrange multiplier associated with the j th equality constraint, $h_j(\mathbf{x}) = 0$. The vectors $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{C}^q$ are the dual variables of the problem (A1).

3. The dual function

The dual function of the problem (A1) is the minimum value of the Lagrangian (A5) over $\mathbf{x} \in \mathcal{D}$ for $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{C}^q$,

$$g(\boldsymbol{\lambda}, \mathbf{v}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{v}). \quad (\text{A5})$$

Since the dual function is the pointwise infimum of a family of affine functions of $(\boldsymbol{\lambda}, \mathbf{v})$, it is concave, even when problem (A1) is not convex.

The dual function (A5) yields lower bounds on the optimal value p^* Eq. (A3) for any $\boldsymbol{\lambda} \succeq 0$ (where \succeq represents componentwise inequality) and any \mathbf{v} ,

$$g(\boldsymbol{\lambda}, \mathbf{v}) \leq p^*, \quad (\text{A6})$$

since $g(\boldsymbol{\lambda}, \mathbf{v}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{v}) \leq L(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \mathbf{v}) \leq f_0(\tilde{\mathbf{x}})$ for every feasible point $\tilde{\mathbf{x}}$.

4. Dual problem

The dual function (A5) gives a lower bound on the optimal value p^* of the optimization problem (A1), which depends on the dual variables $(\boldsymbol{\lambda}, \mathbf{v})$ with $\boldsymbol{\lambda} \succeq 0$; see Eq. (A6). The best lower bound, i.e., the lower bound with the greatest value, is obtained through the optimization problem

$$\max_{\boldsymbol{\lambda}, \mathbf{v}} g(\boldsymbol{\lambda}, \mathbf{v}) \text{ subject to } \boldsymbol{\lambda} \succeq 0, \quad (\text{A7})$$

which is the dual problem to the optimization problem (A1).

The dual problem (A7) is a convex optimization problem, since the objective function to be maximized is concave and the constraints are convex, irrespectively whether the primal problem (A1) is convex or not.

5. Weak duality

The optimal value d^* of the dual problem (A7), achieved at the dual optimal variables $(\boldsymbol{\lambda}^*, \mathbf{v}^*)$ is

$$d^* = \sup \{g(\boldsymbol{\lambda}, \mathbf{v}) \mid \boldsymbol{\lambda} \succeq 0\} = \{g(\boldsymbol{\lambda}^*, \mathbf{v}^*) \mid \boldsymbol{\lambda}^* \succeq 0\}. \quad (\text{A8})$$

The dual maximum d^* is the best lower bound on the minimum of the primal problem (A3), which can be obtained from the Lagrange dual function. The inequality

$$d^* \leq p^*, \quad (\text{A9})$$

holds even if the primal problem (A1) is non-convex and is called weak duality.

The non-negative difference $p^* - d^*$ is called the duality gap for the optimization problem (A1), since it gives the gap between the minimum of the primal problem and the maximum of the dual problem.

6. Slater's condition and strong duality

When the duality gap, $p^* - d^*$, is zero, strong duality holds characterized by the equality

$$d^* = p^*. \quad (\text{A10})$$

Strong duality holds when the optimization problem (A1) is convex and there exists a strictly feasible point, i.e., the inequality constraints hold with strict inequalities. The constraint qualification which implies strong duality for convex problems is called Slater's condition,

$$\begin{aligned} f_i(\mathbf{x}) &< 0, \quad i = 1, \dots, m, \\ \mathbf{A}_{q \times N} \mathbf{x} &= \mathbf{y}. \end{aligned} \quad (\text{A11})$$

When the primal problem is convex and Slater's condition holds there exist a dual feasible $(\boldsymbol{\lambda}^*, \mathbf{v}^*)$ such that $g(\boldsymbol{\lambda}^*, \mathbf{v}^*) = d^* = p^*$, i.e., the optimal value of the primal problem can be obtained by solving the dual problem.

The Slater's condition holds also with a weaker constraint qualification, when some of the inequality constraint functions, f_1, \dots, f_k , are affine (instead of convex)

$$\begin{aligned} f_i(\mathbf{x}) &\leq 0, \quad i = 1, \dots, k, \\ f_i(\mathbf{x}) &< 0, \quad i = k + 1, \dots, m, \\ \mathbf{A}_{q \times N} \mathbf{x} &= \mathbf{y}. \end{aligned} \quad (\text{A12})$$

The weaker constraint qualifications (A12) imply that strong duality reduces to feasibility when both the inequality and the equality constraints are linear.

7. Schur complement

Let \mathbf{X} be a square Hermitian matrix partitioned as

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix}, \quad (\text{A13})$$

where \mathbf{A} is also square Hermitian. If $\det \mathbf{A} \neq 0$ then the matrix

$$\mathbf{S} = \mathbf{C} - \mathbf{B}^H \mathbf{A}^{-1} \mathbf{B} \quad (\text{A14})$$

is called the Schur complement of \mathbf{A} in \mathbf{X} .

A useful property related to the Schur complement is that if $\mathbf{A} \succ 0$ then $\mathbf{X} \succeq 0$ if and only if $\mathbf{S} \succeq 0$.

APPENDIX B: BOUNDED TRIGONOMETRIC POLYNOMIALS

This section presents useful results for bounded trigonometric polynomials and their roots as presented in Ref. 37.

1. Trigonometric polynomials

Let $\mathbf{a}(\omega) = [1, e^{j\omega}, \dots, e^{j\omega(L-1)}]^T$ be a $L \times 1$ basis vector for trigonometric polynomials of degree $L-1$ with $\omega \in [-\pi, \pi]$. A (causal) trigonometric polynomial can be written in terms of the basis vector as

$$H(\omega) = \sum_{l=0}^{L-1} h_l e^{-j\omega l} = \mathbf{a}(\omega)^H \mathbf{h}, \quad (\text{B1})$$

where $\mathbf{h} = [h_0, \dots, h_{L-1}]^T \in \mathbb{C}^L$ is the vector of the polynomial coefficients.

2. Non-negative trigonometric polynomials

Let $R(\omega) = |H(\omega)|^2 = H(\omega)H(\omega)^H$. From Eq. (B1), the non-negative trigonometric polynomial $R(\omega)$ has the form

$$R(\omega) = \sum_{k=-(L-1)}^{L-1} r_k e^{-j\omega k}, \quad (\text{B2})$$

where $r_k = \sum_{l=0}^{L-1-k} h_l h_{l+k}^*$ for $k \geq 0$ and $r_{-k} = r_k^*$, i.e., the coefficients are conjugate symmetric thus $R(\omega)$ is Hermitian. Equivalently, the coefficients r_k can be calculated as the sum of the k th diagonal elements of the autocorrelation matrix $\mathbf{Q}_{L \times L} = \mathbf{h}\mathbf{h}^H$ as

$$r_k = \sum_{i=1}^{L-k} \mathbf{Q}_{i, i+k}. \quad (\text{B3})$$

3. Bounded trigonometric polynomials

Let two polynomials $H(\omega)$ and $B(\omega)$ fulfill the inequality

$$|H(\omega)| \leq |B(\omega)|, \quad \forall \omega \in [-\pi, \pi], \quad (\text{B4})$$

which implies $|H(\omega)|^2 \leq |B(\omega)|^2$, $\forall \omega \in [-\pi, \pi]$. Defining $R_H(\omega) = |H(\omega)|^2$ and $R_B(\omega) = |B(\omega)|^2$ as in Eq. (B2), yields $R_H(\omega) \leq R_B(\omega)$. From Lemma 4.23 in Ref. 37, $R_H(\omega) \leq R_B(\omega)$ implies $\mathbf{Q}_H \preceq \mathbf{Q}_B$, where $\mathbf{Q}_H = \mathbf{h}\mathbf{h}^H$ and $\mathbf{Q}_B = \mathbf{b}\mathbf{b}^H$ are the autocorrelation matrices of the coefficient vectors $\mathbf{h} = [h_0, \dots, h_{L-1}]^T$ and $\mathbf{b} = [b_0, \dots, b_{L-1}]^T$ of the polynomials $H(\omega)$ and $B(\omega)$, respectively. Through a Schur complement (see Appendix A 7), $\mathbf{Q}_B - \mathbf{h}\mathbf{h}^H \succeq 0$ is equivalent to semidefinite matrix

$$\begin{bmatrix} \mathbf{Q}_B & \mathbf{h}_{L \times 1} \\ \mathbf{h}_{1 \times L}^H & 1 \end{bmatrix} \succeq 0. \quad (\text{B5})$$

Let the polynomial $H(\omega)$ have amplitude uniformly bounded for all $\omega \in [-\pi, \pi]$ such that, $|H(\omega)| \leq \gamma$, where $\gamma \in \mathbb{R}_+$ is a given positive real number. As a special case of the results for bounded trigonometric polynomials in Eqs.

(B4), (B5), with $|B(\omega)| = \gamma$, Theorem 4.24 and corollary 4.25 in Ref. 37 states that the inequality $|H(\omega)| \leq \gamma$ can be approximated by two linear matrix inequalities

$$\begin{aligned} & \begin{bmatrix} \mathbf{Q}_{L \times L} & \mathbf{h}_{L \times 1} \\ \mathbf{h}_{1 \times L}^H & 1 \end{bmatrix} \succeq 0, \\ & \sum_{i=1}^{L-j} \mathbf{Q}_{i, i+j} = \begin{cases} \gamma^2, & j = 0 \\ 0, & j = 1, \dots, L-1. \end{cases} \end{aligned} \quad (\text{B6})$$

The latter constraint follows from the autocorrelation matrix of the constant polynomial $R_B(\omega) = \gamma^2$.

The results for bounded trigonometric polynomials can be used in relation to the ℓ_∞ -norm, since setting an upper bound for the maximum amplitude of a polynomial implies that the polynomial has amplitude uniformly bounded for all $\omega \in [-\pi, \pi]$,

$$\begin{aligned} \|H\|_\infty &= \max_{\omega \in [-\pi, \pi]} |H(\omega)| \leq \gamma, \\ |H(\omega)| &\leq \gamma, \quad \forall \omega \in [-\pi, \pi]. \end{aligned} \quad (\text{B7})$$

4. Roots of real non-negative trigonometric polynomials

For a bounded trigonometric polynomial $|H(\omega)| \leq 1$, we can construct a polynomial

TABLE I. MATLAB code for Sec. IV.

Given $\mathbf{y} \in \mathbb{C}^M$, d, λ
Solve dual problem with CVX (Ref. 23), Eq. (23)

```

1: cvx_solver sdpt3
2: cvx_begin sdp
3:   variable S(M+1, M+1) hermitian
4:   S >= 0;
5:   S(M+1, M+1) == 1;
6:   trace(S) == 2;
7:   for j = 1 : M-1
8:     sum(diag(S, j)) == S(M+1-j, M+1);
9:   end
10:  maximize(real(S(1 : M, M+1)' * y))
11: cvx_end
12: c = S(1 : M, M+1);
```

Find the roots of P_+ , Eq. (26)

```

13: r = conv(c, flipud(conj(c)));
14: r(M) = 1 - r(M);
15: roots_P = roots(r);
```

Isolate roots on the unit circle, Eq. (27)

```

16: roots_uc = roots_P (abs(1-abs(roots_P)) < 1e-2);
17: [aux, ind] = sort(real(roots_uc));
18: roots_uc = roots_uc(ind);
19: t = angle(roots_uc(1:2:end))/(2 * pi * d / lambda);
```

Amplitude estimation, Eq. (28)

```

20: A_T = exp(1i * 2 * pi * d / lambda * [0 : (M-1)]' * t');
21: x_CS_dual = A_T \ y;
```

$$P(\omega) = 1 - |H(\omega)|^2 = 1 - R(\omega), \quad (\text{B8})$$

which is by definition real-valued and non-negative, thus it cannot have single roots on the unit circle. The degree of the polynomial $P(\omega)$ is $2(L-1)$. Therefore, the polynomial $P(\omega)$ has at most $L-1$ distinct roots on the unit circle. At a root, ω_0 , we have $P(\omega_0) = 0$ and subsequently $|H(\omega_0)| = 1$.

APPENDIX C: IMPLEMENTATION IN MATLAB

The algorithm in Table I for the implementation of the method described in Sec. IV is an adaptation of the code by Fernandez-Granda in Ref. 17.

APPENDIX D: DUAL PROBLEM WITH NOISE

In the case that the measurements Eq. (13) are contaminated with additive noise $\mathbf{n} \in \mathbb{C}^M$ such that $\|\mathbf{n}\|_2 \leq \epsilon$, the primal problem of atomic norm minimization (15) is reformulated to problem (32) or equivalently

$$\min_x \|x\|_{\mathcal{A}} \text{ subject to } \begin{cases} \mathbf{y} = \mathcal{F}_M x + \mathbf{n}, \\ \|\mathbf{n}\|_2 \leq \epsilon. \end{cases} \quad (\text{D1})$$

The Lagrangian for Eq. (D1) is formulated by augmenting the objective function with a weighted sum of the constraints

$$L(x, \mathbf{c}, \xi) = \|x\|_{\mathcal{A}} + \text{Re}[\mathbf{c}^H (\mathbf{y} - \mathcal{F}_M x - \mathbf{n})] + \xi (\mathbf{n}^H \mathbf{n} - \epsilon^2), \quad (\text{D2})$$

where $\mathbf{c} \in \mathbb{C}^M$ are the dual variables related to the equality constraints, $\mathbf{y} - \mathcal{F}_M x - \mathbf{n} = 0$, and $\xi \in \mathbb{R}^+$ is a Lagrange multiplier related to the inequality constraint, $\|\mathbf{n}\|_2 - \epsilon \leq 0$.

The dual function $g(\mathbf{c}, \xi)$ is the infimum of the Lagrangian, $L(x, \mathbf{c}, \xi)$, over the optimization variable x ,

$$\begin{aligned} g(\mathbf{c}, \xi) &= \inf_x L(x, \mathbf{c}, \xi) \\ &= \text{Re}[\mathbf{c}^H \mathbf{y} - \mathbf{c}^H \mathbf{n}] + \xi (\mathbf{n}^H \mathbf{n} - \epsilon^2) \\ &\quad + \inf_x (\|x\|_{\mathcal{A}} - \text{Re}[\mathbf{c}^H \mathcal{F}_M x]). \end{aligned} \quad (\text{D3})$$

Minimizing over the unknown noise $\mathbf{n} \in \mathbb{C}^M$

$$\frac{\partial g(\mathbf{c}, \xi)}{\partial \mathbf{n}} = -\mathbf{c} + 2\xi \mathbf{n} = 0, \quad (\text{D4})$$

yields the optimal noise vector, $\mathbf{n}_0 = \mathbf{c}/(2\xi)$. The dual function evaluated at \mathbf{n}_0 is

$$\begin{aligned} g(\mathbf{c}, \xi)|_{\mathbf{n}_0} &= \text{Re}[\mathbf{c}^H \mathbf{y}] - \frac{\mathbf{c}^H \mathbf{c}}{2\xi} + \xi \left(\frac{\mathbf{c}^H \mathbf{c}}{4\xi^2} - \epsilon^2 \right) \\ &\quad + \inf_x (\|x\|_{\mathcal{A}} - \text{Re}[\mathbf{c}^H \mathcal{F}_M x]). \end{aligned} \quad (\text{D5})$$

Further, maximizing over the dual variable ξ ,

$$\frac{\partial g(\mathbf{c}, \xi)|_{\mathbf{n}_0}}{\partial \xi} = \frac{\mathbf{c}^H \mathbf{c}}{4\xi^2} - \epsilon^2 = 0, \quad (\text{D6})$$

we obtain the optimal value for the dual variable $\xi_0 = \|\mathbf{c}\|_2 / (2\epsilon)$.

Finally, the dual function evaluated at the optimal values \mathbf{n}_0 and ξ_0 becomes

$$g(\mathbf{c})|_{\mathbf{n}_0, \xi_0} = \text{Re}[\mathbf{c}^H \mathbf{y}] - \epsilon \|\mathbf{c}\|_2 + \inf_x (\|x\|_{\mathcal{A}} - \text{Re}[\mathbf{c}^H \mathcal{F}_M x]), \quad (\text{D7})$$

and the dual problem is formulated by maximizing the dual function, $g(\mathbf{c})|_{\mathbf{n}_0, \xi_0}$, over the dual variables $\mathbf{c} \in \mathbb{C}^M$ similarly to the process detailed in Sec. IV C,

$$\begin{aligned} \max_{\mathbf{c}} g(\mathbf{c})|_{\mathbf{n}_0, \xi_0} &\equiv \max_{\mathbf{c}} \text{Re}[\mathbf{c}^H \mathbf{y}] - \epsilon \|\mathbf{c}\|_2 \\ \text{subject to } &\|\mathcal{F}_M^H \mathbf{c}\|_{\infty} \leq 1. \end{aligned} \quad (\text{D8})$$

¹M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing Pages* (Springer, New York, 2010), pp. 1–359.

²S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing* (Springer, New York, 2013), pp. 1–589.

³D. Malioutov, M. Çetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Trans. Signal Process.* **53**(8), 3010–3022 (2005).

⁴G. F. Edelmann and C. F. Gaumont, “Beamforming using compressive sensing,” *J. Acoust. Soc. Am.* **130**(4), 232–237 (2011).

⁵A. Xenaki, P. Gerstoft, and K. Mosegaard, “Compressive beamforming,” *J. Acoust. Soc. Am.* **136**(1), 260–271 (2014).

⁶C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, and M. Viberg, “Sequential Bayesian sparse signal reconstruction using array data,” *IEEE Trans. Signal Process.* **61**(24), 6344–6354 (2013).

⁷H. Krim and M. Viberg, “Two decades of array signal processing research: The parametric approach,” *IEEE Signal Proc. Mag.* **13**(4), 67–94 (1996).

⁸W. Mantzel, J. Romberg, and K. Sabra, “Compressive matched-field processing,” *J. Acoust. Soc. Am.* **132**(1), 90–102 (2012).

⁹P. A. Forero and P. A. Baxley, “Shallow-water sparsity-cognizant source-location mapping,” *J. Acoust. Soc. Am.* **135**(6), 3483–3501 (2014).

¹⁰C. Yardim, P. Gerstoft, W. S. Hodgkiss, and J. Traer, “Compressive geoaoustic inversion using ambient noise,” *J. Acoust. Soc. Am.* **135**(3), 1245–1255 (2014).

¹¹Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, “Sensitivity to basis mismatch in compressed sensing,” *IEEE Trans. Signal Process.* **59**(5), 2182–2195 (2011).

¹²M. F. Duarte and R. G. Baraniuk, “Spectral compressive sensing,” *Appl. Comput. Harmon. Anal.* **35**(1), 111–129 (2013).

¹³H. Yao, P. Gerstoft, P. M. Shearer, and C. Mecklenbräuker, “Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes,” *Geophys. Res. Lett.* **38**(20), 1–5, doi:10.1029/2011GL049223 (2011).

¹⁴H. Yao, P. M. Shearer, and P. Gerstoft, “Compressive sensing of frequency-dependent seismic radiation from subduction zone megathrust ruptures,” *Proc. Natl. Acad. Sci. U.S.A.* **110**(12), 4512–4517 (2013).

¹⁵W. Fan, P. M. Shearer, and P. Gerstoft, “Kinematic earthquake rupture inversion in the frequency domain,” *Geophys. J. Int.* **199**(2), 1138–1160 (2014).

¹⁶V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, “The convex geometry of linear inverse problems,” *Found. Comput. Math.* **12**(6), 805–849 (2012).

¹⁷E. J. Candes and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Comm. Pure Appl. Math.* **67**(6), 906–956 (2014).

¹⁸H. L. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)* (Wiley-Interscience, New York, 2002), Chap. 1–10.

¹⁹H. C. Song, S. Cho, T. Kang, W. S. Hodgkiss, and J. R. Preston, “Long-range acoustic communication in deep water using a towed array,” *J. Acoust. Soc. Am.* **129**(3), 71–75 (2011).

²⁰D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Prentice Hall, Englewood Cliffs, NJ, 1993), pp. 1–512.

- ²¹R. G. Baraniuk, "Compressive sensing," *IEEE Signal Proc. Mag.* **24**(4), 118–121 (2007).
- ²²J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inf. Theory* **52**(3), 1030–1051 (2006).
- ²³M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta. <http://cvxr.com/cvx>, September 2013.
- ²⁴S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004), pp. 1–684.
- ²⁵E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *C. R. Math. Acad. Sci.* **346**(9), 589–592 (2008).
- ²⁶J. J. Fuchs, "Sparsity and uniqueness for some specific under-determined linear systems," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'05* (IEEE, New York, 2005), Vol. 5, pp. 729–732.
- ²⁷G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory* **59**(11), 7465–7490 (2013).
- ²⁸E. J. Candes and C. Fernandez-Granda, "Super-resolution from noisy data," *J. Fourier Anal. Appl.* **19**(6), 1229–1254 (2013).
- ²⁹J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE* **57**(8), 1408–1418 (1969).
- ³⁰R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.* **34**(3), 276–280 (1986).
- ³¹R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.* **19**(1), 134–139 (1983).
- ³²R. Kumaresan, "On the zeros of the linear prediction-error filter for deterministic signals," *IEEE Trans. Acoust., Speech, Signal Process.* **31**(1), 217–220 (1983).
- ³³A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'83* (IEEE, New York, 1983), Vol. 8, pp. 336–339.
- ³⁴B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.* **37**(12), 1939–1949 (1989).
- ³⁵S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Process.* **37**(1), 8–15 (1989).
- ³⁶B. D. Rao and K. V. S. Hari, "Effect of spatial smoothing on the performance of MUSIC and the minimum-norm method," *IEE Proc. Radar and Signal Process.* **137**(6), 449–458 (1990).
- ³⁷B. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications* (Springer, Dordrecht, Netherlands, 2007), Chap. 4.3.