Sparsity-Aware Direction Finding for Strictly Non-Circular Sources Based on Rank Minimization

Jens Steinwandt [1], *Christian Steffens* [2], *Marius Pesavento* [2], *and Martin Haardt* [1]

[1] Communications Research Laboratory, TU Ilmenau, P.O. Box 100565, 98684 Ilmenau, Germany
[2] Communication Systems Group, TU Darmstadt, Merckstraße 25, 64283 Darmstadt, Germany
Emails: {jens.steinwandt, martin.haardt}@tu-ilmenau.de, {steffens, pesavento}@nt.tu-darmstadt.de

(invited paper)

***Abstract*—Exploiting the statistical properties of strictly non-circular (NC) signals in direction of arrival (DOA) estimation has long been an active area of research due to its associated performance improvements. Recently, this concept has been introduced to DOA estimation via sparse signal recovery (SSR), where similar benefits from processing NC signals are achieved. However, the standard approach to NC SSR requires solving a two-dimensional (2-D) SSR problem in the spatial and the phase rotation domain, which is not only associated with a high computational complexity itself but also with a 2-D off-grid problem. In this paper, we propose an entirely new NC SSR approach based on nuclear norm (rank) minimization after lifting the original bilinear optimization problem to a linear optimization problem in a higher-dimensional space. Thereby, the SSR-based 2-D estimation problem is reduced to a 1-D estimation problem only in the sampled spatial domain, which automatically provides gridless estimates of the rotation phases. In our second contribution, we present a simple closed-form grid offset estimator for a single NC source and a numerical joint grid offset estimation procedure for two closely-spaced NC sources assuming a uniform linear array (ULA). Simulations validate the effectiveness of the new approach.**

***Index Terms*—Compressed sensing, sparse signal recovery, non-circular sources, nuclear norm minimization.**

I. Introduction

Direction of arrival (DOA) estimation of impinging signals has a long-standing history [1] in the field of array signal processing as this problem arises in a wide range of applications including radar, channel sounding, and wireless communications. In some of these applications, the sources represent real-valued digital modulation signals, e.g., BPSK, ASK, PAM, Offset-QPSK-modulated signals that possess a strictly second-order (SO) non-circular (NC) structure [2]. As shown in previous work, exploiting this statistical property helps to improve the estimation accuracy and doubles the number of identifiable sources of the conventional DOA estimators [1]. This observation has led to the design of several DOA estimation algorithms [3]-[7] that take advantage of the signals' non-circularity.

Recently, a new perspective on the original DOA estimation problem has been provided by employing sparse signal recovery (SSR) [8], [9]. In this context, the array output is interpreted by means of a sparse spatial power spectrum, which represents the superposition of the received signal power from very few wavefronts in an overcomplete basis. Such a finite basis is obtained by sampling the continuous spatial domain with a predefined grid. Many sparsity-based DOA estimation algorithms [10]-[13] have been developed based on this concept. It has been observed that SSR methods exhibit super-resolution capabilities without facing the drawbacks of the conventional DOA methods [1], i.e., a performance degradation for high source correlation, unknown model order, a low sample size, etc. The recent work in [14] considers SSR for DOA estimation using partly-calibrated arrays. Therein, it was shown that the emerging bilinear optimization problem can be transformed to a linear optimization problem via an extended parametrization, which can be interpreted as lifting the optimization problem to a higher-dimensional space. Subsequently, the resulting linear problem is addressed by means of nuclear norm minimization.

This work was supported by the EXPRESS project within the DFG priority program CoSIP (DFG-SPP 1798).

Despite the various benefits of the SSR-based DOA estimation algorithms, they all face the well-known off-grid problem as the true DOAs usually lie off the predefined grid, which causes a performance degradation due to the model mismatch. Several solutions to the off-grid problem include an adaptive grid refinement [10], statistical fitting of the offset error [15], and a low-complexity analytical solution by directly estimating the grid offset [16].

The concept of exploiting the signals' NC structure has recently been applied to SSR-based DOA estimation in [17]. However, as NC signals usually have an unknown rotation phase, [17] requires solving a two-dimensional (2-D) SSR problem to estimate the support in the spatial domain as well as the rotation phase domain, which is computationally complex. In addition, this approach naturally leads to a 2-D off-grid problem that needs to be addressed.

In this paper, we propose a novel approach towards exploiting the NC signal structure in SSR-based DOA estimation. First, we show that the NC SSR problem in [17] can be equivalently formulated as a bilinear optimization problem, which admits the application of the extended parametrization from [14]. Thus, the bilinear NC SSR problem is addressed by solving a linear problem in a higher-dimensional space via nuclear norm minimization. As a result, the original SSR-based 2-D estimation problem is reduced to a 1-D estimation problem only in the spatial domain, while gridless estimates of the rotation phases are provided automatically. Therefore, the proposed method requires a significantly lower computational complexity. In our second contribution, we address the resulting 1-D NC off-grid problem for a uniform linear array (ULA) and present a simple closed-form grid offset estimator for a single NC source and a numerical joint grid offset estimation procedure for two closely-spaced NC sources inspired by [16] and [17]. Simulation results illustrate the effectiveness of the new approach.

II. System Model

Consider an arbitrary sensor array composed of M sensor elements receiving the narrowband signals from L stationary far-field sources. The N subsequent observations can be modeled as

$$\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{\mu})\,\boldsymbol{S} + \boldsymbol{W} \in \mathbb{C}^{M\times N}, \tag{1}$$

where $\boldsymbol{A}(\boldsymbol{\mu}) = [\boldsymbol{a}(\mu_1), \ldots, \boldsymbol{a}(\mu_L)] \in \mathbb{C}^{M\times L}$ is the array steering matrix, which is composed of the array steering vectors $\boldsymbol{a}(\mu_\ell)$, $\ell = 1, \ldots, L$, for the spatial frequencies contained in $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_L]^{\mathrm{T}}$. The matrix $\boldsymbol{S} \in \mathbb{C}^{L\times N}$ represents the matrix of source waveforms, and $\boldsymbol{W} \in \mathbb{C}^{M\times N}$ consists of additive white Gaussian noise samples with variance σ_W^2.

In the case of strictly SO non-circular sources, the complex signal amplitudes of each source lie on a rotated line through the origin in the complex plane. Thus, $\boldsymbol{S}$ can be written as

$$\boldsymbol{S} = \boldsymbol{\Phi}(\boldsymbol{\varphi})\boldsymbol{S}_0, \tag{2}$$

where $\boldsymbol{S}_0 = [\boldsymbol{s}_{0,1}, \ldots, \boldsymbol{s}_{0,L}]^{\mathrm{T}} \in \mathbb{R}^{L\times N}$ is a real-valued symbol matrix with rows $\boldsymbol{s}_{0,\ell}^{\mathrm{T}}$, and $\boldsymbol{\Phi}(\boldsymbol{\varphi}) = \mathrm{diag}([\mathrm{e}^{\mathrm{j}\varphi_1}, \ldots, \mathrm{e}^{\mathrm{j}\varphi_L}])$ is a diagonal matrix that contains the rotation phase shifts corresponding to the phases $\boldsymbol{\varphi} = [\varphi_1, \ldots, \varphi_L]^{\mathrm{T}}$ on its diagonal, which can be arbitrary for each received signal [7].

In order to take advantage of the signals' NC property, we apply the common preprocessing step to model (1) to obtain the $2M \times N$

augmented measurement matrix [7]

$$\boldsymbol{X}^{(\mathrm{nc})} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{\Pi}_M \boldsymbol{X}^* \end{bmatrix} = \boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi}) \boldsymbol{S}_0 + \boldsymbol{W}^{(\mathrm{nc})}, \tag{3}$$

where $\boldsymbol{\Pi}_M$ is the $M \times M$ exchange matrix with ones on its anti-diagonal and zeros elsewhere. The augmented steering matrix $\boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi}) \in \mathbb{C}^{2M \times L}$ can be decomposed as $\boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi}) = \bar{\boldsymbol{A}}^{(\mathrm{nc})}(\boldsymbol{\mu})\, \boldsymbol{\Phi}^{(\mathrm{nc})}(\boldsymbol{\varphi})$ with

$$\bar{\boldsymbol{A}}^{(\mathrm{nc})}(\boldsymbol{\mu}) = \begin{bmatrix} \boldsymbol{A}(\boldsymbol{\mu}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Pi}_M \boldsymbol{A}^*(\boldsymbol{\mu}) \end{bmatrix}, \quad \boldsymbol{\Phi}^{(\mathrm{nc})}(\boldsymbol{\varphi}) = \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\varphi}) \\ \boldsymbol{\Phi}^*(\boldsymbol{\varphi}) \end{bmatrix}.$$

The extended dimensions of $\boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi}) \in \mathbb{C}^{2M \times L}$ can be interpreted as a virtual doubling of the sensor elements, which improves the estimation performance and doubles the number of resolvable sources. Following [14], we reformulate model (3) as

$$\boldsymbol{X}^{(\mathrm{nc})} = \boldsymbol{B}(\boldsymbol{\mu})\, \boldsymbol{\Psi}(\boldsymbol{\varphi})\, \boldsymbol{S}_0 + \boldsymbol{W}^{(\mathrm{nc})} \tag{4}$$

where $\boldsymbol{B}(\boldsymbol{\mu}) = \bar{\boldsymbol{A}}^{(\mathrm{nc})}(\boldsymbol{\mu})\boldsymbol{J} \in \mathbb{C}^{2M \times 2L}$ is a column permuted version of the steering matrix $\bar{\boldsymbol{A}}^{(\mathrm{nc})}(\boldsymbol{\mu})$ with a permutation matrix $\boldsymbol{J}$ such that

$$\boldsymbol{B}(\boldsymbol{\mu}) = \begin{bmatrix} \boldsymbol{a}(\mu_1) & \boldsymbol{0} & \cdots & \boldsymbol{a}(\mu_L) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Pi}_M \boldsymbol{a}^*(\mu_1) & & \boldsymbol{0} & \boldsymbol{\Pi}_M \boldsymbol{a}^*(\mu_L) \end{bmatrix}. \tag{5}$$

Similarly, we define the row permuted phase shift matrix $\boldsymbol{\Psi} = \boldsymbol{J}^{\mathrm{T}} \boldsymbol{\Phi}^{(\mathrm{nc})} \in \mathbb{C}^{2L \times L}$ that possesses the block-diagonal structure

$$\boldsymbol{\Psi}(\boldsymbol{\varphi}) = \mathrm{blkdiag}(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_L), \tag{6}$$

where $\boldsymbol{\psi}_\ell = [\mathrm{e}^{\mathrm{j}\varphi_\ell},\ \mathrm{e}^{-\mathrm{j}\varphi_\ell}]^{\mathrm{T}} \in \mathbb{C}^{2\times 1}$ contains the rotation phase shift and its complex conjugate corresponding to the ℓ-th source as defined in (2). Furthermore, we define the parametrization

$$\boldsymbol{Q}(\boldsymbol{\varphi}) = \boldsymbol{\Psi}(\boldsymbol{\varphi}) \boldsymbol{S}_0 = [\boldsymbol{Q}_1^{\mathrm{T}}(\varphi_1), \ldots, \boldsymbol{Q}_L^{\mathrm{T}}(\varphi_L)]^{\mathrm{T}} \in \mathbb{C}^{2L \times N}, \tag{7}$$

which is partitioned into L rank-one matrices $\boldsymbol{Q}_\ell(\varphi_\ell) = \boldsymbol{\psi}_\ell\, \boldsymbol{s}_{0,\ell}^{\mathrm{T}} \in \mathbb{C}^{2\times N},\ \ell = 1, \ldots, L$. Additionally, due to the complex conjugate structure of the phase shift vectors $\boldsymbol{\psi}_\ell$, each sub-matrix $\boldsymbol{Q}_\ell(\varphi_\ell)$ exhibits a conjugate row structure, i.e., $\boldsymbol{Q}_\ell(\varphi_\ell) = \boldsymbol{\Pi}_2\, \boldsymbol{Q}_\ell^*(\varphi_\ell)$. Using the definition in (7), the signal models in (3) and (4) can be equivalently expressed as

$$\boldsymbol{X}^{(\mathrm{nc})} = \boldsymbol{B}(\boldsymbol{\mu})\, \boldsymbol{Q}(\boldsymbol{\varphi}) + \boldsymbol{W}^{(\mathrm{nc})}. \tag{8}$$

Note that the parametrization $\boldsymbol{Q}(\boldsymbol{\varphi})$ in (8) requires $2NL$ complex symbols and is therefore less compact than directly modeling $\boldsymbol{S}_0$ and $\boldsymbol{\varphi}$ in (3), which only contains $(N+1)L$ real-valued symbols. However, the advantage of the representation (8) is that, in contrast to (3), the bilinear nature is removed as the rotation phase matrix $\boldsymbol{\Phi}(\boldsymbol{\varphi})$ is embedded in the block signal matrix $\boldsymbol{Q}(\boldsymbol{\varphi})$. Moreover, $\boldsymbol{Q}(\boldsymbol{\varphi})$ is composed of rank-one sub-matrices with conjugate row structure. Both properties of the sub-matrices will be exploited by the proposed SSR algorithm for NC signals.

III. Review of Sparse Signal Recovery

The conventional approach of using SSR for DOA estimation relies on the formulation of a sparse representation of the signal model in (1). To this end, an overcomplete basis $\boldsymbol{A}(\tilde{\boldsymbol{\mu}}) \in \mathbb{C}^{M \times K}$ is constructed by discretizing the spatial frequency range $[0, 2\pi]$ at the K grid points $\tilde{\boldsymbol{\mu}} = [\tilde{\mu}_1, \ldots, \tilde{\mu}_K]$. Typically, $K = MP$, where $P > 1$ is the oversampling factor such that $K > M > L$. For simplicity, we consider uniform sampling with $\tilde{\mu}_k = (k-1)\Delta_\mu,\ k = 1, \ldots, K$, where $\Delta_\mu = 2\pi/K$ is the grid spacing, and assume that the true spatial frequencies lie exactly on the sampling grid, i.e., $\{\mu_\ell\}_{\ell=1}^L \in \{\tilde{\mu}_k\}_{k=1}^K$. Under this *on-grid* assumption, we obtain the $K \times N$ sparse signal matrix $\tilde{\boldsymbol{S}} = [\tilde{\boldsymbol{s}}_1, \ldots, \tilde{\boldsymbol{s}}_K]^{\mathrm{T}}$ with rows defined as

$$\tilde{\boldsymbol{s}}_k^{\mathrm{T}} = \begin{cases} \boldsymbol{s}_\ell^{\mathrm{T}} & \text{if } \tilde{\mu}_k = \mu_\ell \\ \boldsymbol{0} & \text{else,} \end{cases} \tag{9}$$

where $\boldsymbol{s}_\ell^{\mathrm{T}}$ denotes the ℓ-th row in the source symbol matrix $\boldsymbol{S}$ given in (2). Thus, $\tilde{\boldsymbol{S}}$ exhibits a row sparse structure, i.e., the rows $\tilde{\boldsymbol{s}}_k$, for $k = 1, \ldots, K$, are either jointly zero or jointly non-zero. The support, i.e., the indices of the non-zero rows, of $\tilde{\boldsymbol{S}}$ corresponds to the locations of the DOAs on the spatial grid. Then, the model in (1) admits the sparse representation

$$\boldsymbol{X} = \boldsymbol{A}(\tilde{\boldsymbol{\mu}})\, \tilde{\boldsymbol{S}} + \boldsymbol{W} \in \mathbb{C}^{M \times N}. \tag{10}$$

Given (10), the support can be estimated from the well known mixed-norm regularized optimization problem [10]

$$\hat{\tilde{\boldsymbol{S}}} = \arg\min_{\tilde{\boldsymbol{S}}} \sum_{k=1}^{K} \|\tilde{\boldsymbol{s}}_k\|_2 \quad \text{s.t. } \left\|\boldsymbol{X} - \boldsymbol{A}(\tilde{\boldsymbol{\mu}})\, \tilde{\boldsymbol{S}}\right\|_F^2 \leq \beta, \tag{11}$$

which can be computed by standard convex solvers. The regularization parameter β can be selected according to [10]. It is important to note that (11) does not account for the NC signal structure as it is based on model (1).

IV. Nuclear-Norm-Based Sparse NC Signal Recovery

In order to take advantage of the NC structure of the signals, the previous work in [17] relies on the sparse representation of the NC model (3). Based on the fact that the rotation phases in (2) can be arbitrary, the DOA estimation problem in [17] is considered as a two-dimensional estimation problem. This results in a 2-D SSR problem in the spatial domain and the rotation phase domain, which is addressed similarly to (11). However, the 2-D sampling of both the spatial and the rotation phase domain significantly increases the dimensions of the joint overcomplete basis, which leads to a prohibitive computational complexity.

A. Proposed NC Signal Recovery Algorithm

In this section, we present a new approach for exploiting the signals' NC structure in SSR-based DOA estimation. To overcome the bilinear structure of the signal model in (3) and to avoid a costly 2-D SSR approach, we propose to formulate the SSR-based estimation problem for the parameterized model in (8). As previously mentioned, the parameterization in (7) avoids the bilinear nature of the model but increases the number of estimation parameters, hence, in the context of optimization, it can be interpreted as lifting to a higher-dimensional space. The sparse representation of (8) is given by

$$\boldsymbol{X}^{(\mathrm{nc})} = \boldsymbol{B}(\tilde{\boldsymbol{\mu}})\, \tilde{\boldsymbol{Q}}(\boldsymbol{\varphi}) + \boldsymbol{W}^{(\mathrm{nc})}, \tag{12}$$

where the dictionary matrix $\boldsymbol{B}(\tilde{\boldsymbol{\mu}}) \in \mathbb{C}^{2M \times 2K}$ is computed according to (5) using the spatial grid in $\tilde{\boldsymbol{\mu}}$. Under the *on-grid* assumption $\{\mu_\ell\}_{\ell=1}^L \in \{\tilde{\mu}_k\}_{k=1}^K$, the block-sparse signal matrix $\tilde{\boldsymbol{Q}}(\boldsymbol{\varphi}) = [\tilde{\boldsymbol{Q}}_1^{\mathrm{T}}(\varphi_1), \ldots, \tilde{\boldsymbol{Q}}_K^{\mathrm{T}}(\varphi_K)]^{\mathrm{T}} \in \mathbb{C}^{2K \times N}$ contains the sub-matrices

$$\tilde{\boldsymbol{Q}}_k(\varphi_k) = \begin{cases} \boldsymbol{Q}_\ell(\varphi_\ell) & \text{if } \tilde{\mu}_k = \mu_\ell \\ \boldsymbol{0} & \text{else.} \end{cases} \tag{13}$$

It should be highlighted that the matrix $\tilde{\boldsymbol{Q}}(\boldsymbol{\varphi})$ exhibits two different levels of sparsity. Considering (13), it is block sparse, i.e., the elements of the blocks $\tilde{\boldsymbol{Q}}_k(\varphi_k)$, for $k = 1, \ldots, K$, are either jointly zero or jointly non-zero. Additionally, from (7), the blocks $\tilde{\boldsymbol{Q}}_k(\varphi_k)$ are rank sparse, i.e., they are either of rank one or rank zero. In what follows, we drop the argument $\boldsymbol{\varphi}$ in $\tilde{\boldsymbol{Q}}(\boldsymbol{\varphi})$ for notational convenience.

As discussed in [14], this twofold sparsity structure can be exploited by means of nuclear norm minimization [18], which promotes solutions with low rank sub-matrices $\tilde{\boldsymbol{Q}}_k$. Hence, we formulate the convex minimization problem

$$\hat{\tilde{\boldsymbol{Q}}} = \arg\min_{\tilde{\boldsymbol{Q}}} \sum_{k=1}^{K} \|\tilde{\boldsymbol{Q}}_k\|_* \tag{14a}$$

$$\text{s.t. } \left\|\boldsymbol{X}^{(\mathrm{nc})} - \boldsymbol{B}(\tilde{\boldsymbol{\mu}})\, \tilde{\boldsymbol{Q}}\right\|_F^2 \leq \beta^{(\mathrm{nc})} \tag{14b}$$

$$\tilde{\boldsymbol{Q}}_k = \boldsymbol{\Pi}_2\, \tilde{\boldsymbol{Q}}_k^*, \text{ for } k = 1, \ldots, K. \tag{14c}$$

The nuclear norm in (14a) is defined by

$$\left\|\tilde{\boldsymbol{Q}}_k\right\|_* = \sum_{r=1}^{\min(2,N)} \sigma_r(\tilde{\boldsymbol{Q}}_k), \tag{15}$$

where $\sigma_r(\tilde{\boldsymbol{Q}}_k)$ denotes the r-th singular value of the sub-matrix $\tilde{\boldsymbol{Q}}_k$. Moreover, the constraint (14c) enforces the conjugate row structure in

each matrix block $\tilde{\boldsymbol{Q}}_k$. The regularization parameter $\beta^{(\mathrm{nc})}$ can again be chosen according to [10]. Note that (14) can be interpreted as a lifted version of the 2-D SSR problem in [17] due to the extended parametrization in (7).

Since the problem in (14) is convex, it can be solved, for instance, by semidefinite programming [18]. Given a solution $\hat{\tilde{\boldsymbol{Q}}}$ to problem (14), the support set is identified from the indices of the non-zero sub-matrices according to

$$\mathcal{S} \;=\; \{k \,\big|\, \hat{\tilde{\boldsymbol{Q}}}_k \neq \boldsymbol{0}\,\} \tag{16}$$

and the spatial frequency estimates are extracted as $\{\hat{\mu}_\ell\}_{\ell=1}^{\hat{L}} = \{\tilde{\mu}_k \,\big|\, k \in \mathcal{S}\}$. Given a sub-matrix estimate $\hat{\tilde{\boldsymbol{Q}}}_k$, $k \in \mathcal{S}$, the corresponding signal waveforms and their rotation phases are recovered through

$$\hat{\boldsymbol{s}}_{0,k} = \frac{\sigma_1\big(\hat{\tilde{\boldsymbol{Q}}}_k\big)}{\sqrt{2}} \cdot \hat{\boldsymbol{v}}_{k,1}, \qquad \hat{\varphi}_k = \arg\big([\hat{\boldsymbol{u}}_{k,1}]_1\big), \tag{17}$$

where $\hat{\boldsymbol{u}}_{k,1}$ and $\hat{\boldsymbol{v}}_{k,1}$ denote the principal left and right singular vectors of the sub-matrix $\hat{\tilde{\boldsymbol{Q}}}_k$, $[\boldsymbol{x}]_i$ denotes the i-th element of a vector $\boldsymbol{x}$ and $\arg(x)$ denotes the phase of a complex scalar x. Note that due to the conjugate row structure of the sub-matrices $\hat{\tilde{\boldsymbol{Q}}}_k = [\hat{\tilde{\boldsymbol{q}}}_k, \hat{\tilde{\boldsymbol{q}}}_k^*]^{\mathrm{T}} \in \mathbb{C}^{2\times N}$, there always exist real-valued right singular vectors $\hat{\boldsymbol{v}}_{k,1}$, since $\hat{\tilde{\boldsymbol{Q}}}_k^{\mathrm{H}}\hat{\tilde{\boldsymbol{Q}}}_k = \hat{\tilde{\boldsymbol{q}}}_k^*\hat{\tilde{\boldsymbol{q}}}_k^{\mathrm{T}} + \hat{\tilde{\boldsymbol{q}}}_k\hat{\tilde{\boldsymbol{q}}}_k^{\mathrm{H}} \in \mathbb{R}^{N\times N}$. Moreover, (17) enables the gridless estimation of the rotation phases φ_ℓ in contrast to [17]. A procedure for reducing the computational complexity of solving (14) and an appropriate choice of the regularization parameter $\beta^{(\mathrm{nc})}$ that provides proper estimation performance will be presented in the next section.

B. Signal Subspace Processing and Regularization Parameter

The computational complexity of solving problem (14) is mainly determined by the size of the sparse signal matrix $\tilde{\boldsymbol{Q}}$, i.e., the number of grid points K and the number of snapshots N. To reduce the effective number of snapshots, in this section, we present an estimation problem which operates on the signal subspace of the measurement matrix $\boldsymbol{X}^{(\mathrm{nc})}$ rather than on $\boldsymbol{X}^{(\mathrm{nc})}$ itself, similar to the ℓ_1-SVD method presented in [10].

Let the singular value decomposition (SVD) of the measurement matrix be given by

$$\boldsymbol{X}^{(\mathrm{nc})} = \boldsymbol{B}(\boldsymbol{\mu})\,\boldsymbol{\Psi}(\boldsymbol{\varphi})\,\boldsymbol{S}_0 + \boldsymbol{W}^{(\mathrm{nc})} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{H}}. \tag{18}$$

Note that $\boldsymbol{X}^{(\mathrm{nc})^{\mathrm{H}}}\boldsymbol{X}^{(\mathrm{nc})} = \boldsymbol{X}^{\mathrm{H}}\boldsymbol{X} + \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}^* \in \mathbb{R}^{N\times N}$ such that there exists a real-valued unitary basis $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ for the row space of $\boldsymbol{X}^{(\mathrm{nc})}$. Then, we can perform the preprocessing step

$$\begin{aligned}\boldsymbol{X}_{\mathrm{sv}}^{(\mathrm{nc})} &= \boldsymbol{X}^{(\mathrm{nc})}\boldsymbol{V}\boldsymbol{K}\\ &= \boldsymbol{B}(\boldsymbol{\mu})\,\boldsymbol{\Psi}(\boldsymbol{\varphi})\,\boldsymbol{S}_0\,\boldsymbol{V}\boldsymbol{K} + \boldsymbol{W}^{(\mathrm{nc})}\,\boldsymbol{V}\boldsymbol{K}\\ &= \boldsymbol{B}(\boldsymbol{\mu})\,\boldsymbol{\Psi}(\boldsymbol{\varphi})\,\boldsymbol{S}_{\mathrm{sv}} + \boldsymbol{W}_{\mathrm{sv}}^{(\mathrm{nc})},\end{aligned} \tag{19}$$

where the selection matrix $\boldsymbol{K} = [\boldsymbol{I}_L,\ \boldsymbol{0}_{L\times(N-L)}]^{\mathrm{T}}$ extracts the L dominant right singular vectors in $\boldsymbol{V}$. As both $\boldsymbol{S}_0$ and $\boldsymbol{V}$ are real-valued, so is the matrix product $\boldsymbol{S}_{\mathrm{sv}} = \boldsymbol{S}_0\,\boldsymbol{V}\boldsymbol{K}$. Thus, the matrix $\tilde{\boldsymbol{Q}}_{\mathrm{sv}} = \boldsymbol{\Psi}(\boldsymbol{\varphi})\boldsymbol{S}_{\mathrm{sv}} \in \mathbb{C}^{2K\times L}$ has the same conjugate row structure as $\tilde{\boldsymbol{Q}} = \boldsymbol{\Psi}(\boldsymbol{\varphi})\boldsymbol{S}_0 \in \mathbb{C}^{2K\times N}$, but has a significantly reduced number of columns from N to L. With the aforementioned definitions, we formulate the new optimization problem

$$\begin{aligned}\hat{\tilde{\boldsymbol{Q}}}_{\mathrm{sv}} = \arg\min_{\tilde{\boldsymbol{Q}}_{\mathrm{sv}}} \sum_{k=1}^{K} \|\tilde{\boldsymbol{Q}}_{\mathrm{sv},k}\|_* \\ \text{s.t. } \big\|\boldsymbol{X}_{\mathrm{sv}}^{(\mathrm{nc})} - \boldsymbol{B}(\tilde{\boldsymbol{\mu}})\,\tilde{\boldsymbol{Q}}_{\mathrm{sv}}\big\|_F^2 \leq \beta_{\mathrm{sv}}^{(\mathrm{nc})} \\ \tilde{\boldsymbol{Q}}_{\mathrm{sv},k} = \boldsymbol{\Pi}_2\,\tilde{\boldsymbol{Q}}_{\mathrm{sv},k}^*,\ \text{for}\ \ k = 1,\ldots,K.\end{aligned} \tag{20}$$

Considering the selection of the regularization parameter $\beta_{\mathrm{sv}}^{(\mathrm{nc})}$, we follow the idea in [10]. We choose $\beta_{\mathrm{sv}}^{(\mathrm{nc})}$ according to the noise statistics such that it provides an upper bound on the noise power with high probability γ, i.e.,

$$P\big(\|\boldsymbol{W}_{\mathrm{sv}}^{(\mathrm{nc})}\|_F^2 \leq \beta_{\mathrm{sv}}^{(\mathrm{nc})}\big) = \gamma. \tag{21}$$

In the case that $\boldsymbol{W}$ contained in $\boldsymbol{W}^{(\mathrm{nc})} = [\boldsymbol{W}^{\mathrm{T}}, (\boldsymbol{\Pi}_M\boldsymbol{W})^{\mathrm{H}}]^{\mathrm{T}}$ has independent identically distributed (i.i.d.) Gaussian entries and for a moderate to a high SNR, $\|\boldsymbol{W}_{\mathrm{sv}}^{(\mathrm{nc})}\|_F^2 = \|\boldsymbol{W}^{(\mathrm{nc})}\,\boldsymbol{V}\boldsymbol{K}\|_F^2$ follows approximately a χ^2-distribution with ML degrees of freedom upon its normalization by the noise variance $2\sigma_W^2$. The reason that this holds only approximately is that the SVD in (18) depends on the particular realization of the noise, and hence, the matrix $\boldsymbol{V}$ is a function of $\boldsymbol{W}^{(\mathrm{nc})}$. However, when the noise is small, the term $\boldsymbol{B}(\boldsymbol{\mu})\,\tilde{\boldsymbol{Q}}$ dominates the SVD, the effect of $\boldsymbol{W}^{(\mathrm{nc})}$ becomes small, and $\boldsymbol{W}_{\mathrm{sv}}^{(\mathrm{nc})}$ has a χ^2-distribution, such that, according to (21), we can compute $\beta_{\mathrm{sv}}^{(\mathrm{nc})}$ by an inverse χ^2-distribution for some predefined probability γ.

V. Off-Grid Estimation

In practice, the *on-grid* assumption from the previous sections is not fulfilled and we generally have $\{\mu_\ell\}_{\ell=1}^{L} \notin \{\tilde{\mu}_k\}_{k=1}^{K}$, which leads to the well-known *off-grid* problem. In [16], an efficient offset estimation procedure is proposed based on the observation that SSR algorithms concentrate the most signal power at the two adjacent grid points around an off-grid source. Thus, using the relative height of these dominant peaks, each off-grid source can be well approximated. In [17], this concept was extended to the NC case to address the emerging 2-D NC off-grid problem in the spatial and the rotation phase domain. However, as a result of the proposed NC parametrization procedure in the previous section, we are left with a less complex spatial 1-D NC off-grid problem, such that [17] is no longer applicable. Therefore, here, we present two analytical 1-D NC grid-offset estimation schemes for a single NC source and two closely-spaced NC sources that can be applied after estimating the support via SSR. For the derivation of the proposed grid-offset estimators, we assume a ULA with centered phase reference, where the steering vectors are given by $\boldsymbol{a}(\mu) = [\mathrm{e}^{-\mathrm{j}\frac{M-1}{2}\mu}, \ldots, \mathrm{e}^{\mathrm{j}\frac{M-1}{2}\mu}]^{\mathrm{T}}$, such that the augmented NC steering vector is given as [7] $\boldsymbol{a}^{(\mathrm{nc})}(\mu,\varphi) = [\mathrm{e}^{\mathrm{j}\varphi}, \mathrm{e}^{-\mathrm{j}\varphi}]^{\mathrm{T}} \otimes \boldsymbol{a}(\mu)$.

Let us first consider the case of a single source with spatial frequency μ_1. We assume that the support $\mathcal{S}$ defined in (16) consists of the indices $\mathcal{S} = \{k_1, k_1+1\}$ such that $\tilde{\mu}_{\mathcal{S}} = [\tilde{\mu}_{k_1}\ \tilde{\mu}_{k_1+1}]^{\mathrm{T}}$ contains the two neighboring grid points closest to μ_1, from the left and right, respectively. Let ϵ_1 denote the grid offset with $0 \leq \epsilon_1 \leq 1$ such that the off-grid model $\mu_1 = \tilde{\mu}_{k_1} + \epsilon_1\Delta_\mu$ holds, where Δ_μ is the grid spacing. In the single source scenario, the 1-D estimator for the non-NC case from [16] is applicable. This follows from the fact that a single NC source provides no performance benefits over a non-NC source [7], i.e., its rotation phase φ_1 is irrelevant. Thus, in the absence of noise, the true steering vector $\boldsymbol{a}^{(\mathrm{nc})}(\mu_1,\varphi_1) = \boldsymbol{a}^{(\mathrm{nc})}(\tilde{\mu}_{k_1} + \epsilon_1\Delta_\mu, \varphi_1)$ can be approximated by the linear model

$$\boldsymbol{a}^{(\mathrm{nc})}(\mu_1,\varphi_1) \approx \Big[\boldsymbol{a}^{(\mathrm{nc})}(\tilde{\mu}_{k_1}, \hat{\varphi}_{k_1}), \boldsymbol{a}^{(\mathrm{nc})}(\tilde{\mu}_{k_1+1}, \hat{\varphi}_{k_1+1})\Big]\boldsymbol{\alpha}(\epsilon_1),$$

where the unknown coefficients in $\boldsymbol{\alpha}(\epsilon_1) = [\alpha_{k_1}(\epsilon_1), \alpha_{k_1+1}(\epsilon_1)]^{\mathrm{T}}$ provide a good representation of the true steering vector $\boldsymbol{a}^{(\mathrm{nc})}(\mu_1,\varphi_1)$, and $\hat{\varphi}_{k_1}$, $\hat{\varphi}_{k_1+1}$ are the two corresponding continuous rotation phase estimates reconstructed from (17). After computing the coefficients in $\boldsymbol{\alpha}(\epsilon)$ as presented in [16], the simple closed-form estimator derived in [16]

$$\hat{\epsilon}_1 = \frac{\alpha_{k_1+1}(\epsilon_1)}{\alpha_{k_1}(\epsilon_1) + \alpha_{k_1+1}(\epsilon_1)} \tag{22}$$

can also be applied in the NC case presented here.

In the case of two closely-spaced NC off-grid sources, we propose a numerical joint grid-offset estimation procedure, which was inspired by [16], [17]. Let $\boldsymbol{\mu} = [\mu_1, \mu_2]^{\mathrm{T}}$ and $\boldsymbol{\varphi} = [\varphi_1, \varphi_2]^{\mathrm{T}}$ contain the true spatial frequencies and the rotation phases of the two impinging source signals. Furthermore, let the support set $\mathcal{S} = \{k_1, k_1+1, k_2, k_2+1\}$ found from (16) contain the grid indices such that $\tilde{\mu}_{k_\ell}$ and $\tilde{\mu}_{k_\ell+1}$ are the nearest grid points left and right of the source direction μ_ℓ, for $\ell = 1, 2$. We summarize the grid points as $\tilde{\boldsymbol{\mu}}_{\mathcal{S}} = [\tilde{\mu}_{k_1}, \tilde{\mu}_{k_1+1}, \tilde{\mu}_{k_2}, \tilde{\mu}_{k_2+1}]^{\mathrm{T}}$ and the corresponding gridless rotation phase estimates obtained from (17) as $\hat{\boldsymbol{\varphi}}_{\mathcal{S}} = [\hat{\varphi}_{k_1}, \hat{\varphi}_{k_1+1}, \hat{\varphi}_{k_2}, \hat{\varphi}_{k_2+1}]^{\mathrm{T}}$. Moreover, we define the grid offsets $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2]^{\mathrm{T}}$ in $\boldsymbol{\mu} = [\tilde{\mu}_{k_1}, \tilde{\mu}_{k_2}]^{\mathrm{T}} + \boldsymbol{\epsilon}\Delta_\mu$ and the free parameters

$\boldsymbol{\delta} = [\delta_1, \delta_2]^{\mathrm{T}}$ in $\boldsymbol{\varphi} = [\bar{\varphi}_{k_1}, \bar{\varphi}_{k_2}]^{\mathrm{T}} + \boldsymbol{\delta}$, where $\bar{\varphi}_{k_\ell} = (\hat{\varphi}_{k_\ell} + \hat{\varphi}_{k_\ell+1})/2$ is the average of the two rotation phase estimates corresponding to $\tilde{\mu}_{k_\ell}$ $\tilde{\mu}_{k_\ell+1}$ for the ℓ-th source.

Then, following the arguments for the noiseless case $\boldsymbol{X}_0^{(\mathrm{nc})} = \boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi})\boldsymbol{S}_0$ in [17], we define the matrix

$$\boldsymbol{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta}) = \boldsymbol{A}^{(\mathrm{nc})^+}(\tilde{\boldsymbol{\mu}}_{\mathcal{S}}, \hat{\boldsymbol{\varphi}}_{\mathcal{S}})\boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi})\boldsymbol{S}_0 \in \mathbb{R}^{4\times N}, \tag{23}$$

which is the correlation of the reconstructed steering matrix based on the spatial frequencies in $\tilde{\boldsymbol{\mu}}_{\mathcal{S}}$ and the phase estimates $\hat{\boldsymbol{\varphi}}_{\mathcal{S}}$ with the true noise-free measurement matrix. Furthermore, we make use of the fact that we can rewrite (23) as $\boldsymbol{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta}) = \boldsymbol{D}_0^{-1}\boldsymbol{D}(\boldsymbol{\epsilon}, \boldsymbol{\delta})\boldsymbol{S}_0$, where we define

$$\begin{aligned}
\boldsymbol{D}_0 &= \boldsymbol{A}^{(\mathrm{nc})^{\mathrm{H}}}(\tilde{\boldsymbol{\mu}}_{\mathcal{S}}, \hat{\boldsymbol{\varphi}}_{\mathcal{S}})\boldsymbol{A}^{(\mathrm{nc})}(\tilde{\boldsymbol{\mu}}_{\mathcal{S}}, \hat{\boldsymbol{\varphi}}_{\mathcal{S}}) \\
&= \begin{bmatrix} \boldsymbol{D}(0) & \cos(\Delta\hat{\varphi}) \cdot \boldsymbol{D}(d) \\ \cos(\Delta\hat{\varphi}) \cdot \boldsymbol{D}(d)^{\mathrm{T}} & \boldsymbol{D}(0) \end{bmatrix} \in \mathbb{R}^{4\times 4}, \\
\boldsymbol{D}(\boldsymbol{\epsilon}, \boldsymbol{\delta}) &= \boldsymbol{A}^{(\mathrm{nc})^{\mathrm{H}}}(\tilde{\boldsymbol{\mu}}_{\mathcal{S}}, \hat{\boldsymbol{\varphi}}_{\mathcal{S}})\boldsymbol{A}^{(\mathrm{nc})}(\boldsymbol{\mu}, \boldsymbol{\varphi}) \in \mathbb{R}^{4\times 2} \\
&= \begin{bmatrix} \cos(\delta_1) \cdot \boldsymbol{d}(\epsilon_1) & \cos(\delta_2 + \Delta\hat{\varphi}) \cdot \boldsymbol{d}(\epsilon_2 + d) \\ \cos(\delta_1 - \Delta\hat{\varphi}) \cdot \boldsymbol{d}(\epsilon_1 - 1) & \cos(\delta_2)\boldsymbol{d}(\epsilon_2) \end{bmatrix}
\end{aligned}$$

with $\boldsymbol{D}(x) = [\boldsymbol{d}(x), \boldsymbol{d}(x+1)] \in \mathbb{R}^{2\times 2}$, $\boldsymbol{d}(x) = [D(x\Delta_\mu), D((x-1)\Delta_\mu)]^{\mathrm{T}} \in \mathbb{R}^{2\times 1}$. The function $D(y)$ is defined as [17]

$$D(y) = \begin{cases} M & \text{if } y = 0 \\ \sin(yM/2)/\sin(y/2) & \text{else.} \end{cases} \tag{24}$$

Moreover, $d = k_2 - k_1$ denotes the grid distance of the estimated support and $\Delta\hat{\varphi} = |\bar{\varphi}_{k_2} - \bar{\varphi}_{k_1}|$ is the rotation phase difference. In the noisy case, we obtain the estimate $\hat{\boldsymbol{G}} = \boldsymbol{A}^{(\mathrm{nc})^+}(\tilde{\boldsymbol{\mu}}_{\mathcal{S}}, \hat{\boldsymbol{\varphi}}_{\mathcal{S}})\boldsymbol{X}^{(\mathrm{nc})}$, and define $\bar{\boldsymbol{G}} = \boldsymbol{D}_0\hat{\boldsymbol{G}}$. As the columns of $\boldsymbol{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ are a linear combination of the columns of $\boldsymbol{D}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$, we estimate $\boldsymbol{\epsilon}$ and $\boldsymbol{\delta}$ by minimizing the projection of $\bar{\boldsymbol{G}}$ onto the complement of the column space of $\boldsymbol{D}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$, i.e., we minimize the cost function

$$J(\boldsymbol{\epsilon}, \boldsymbol{\delta}) = \| \left(\boldsymbol{I}_4 - \boldsymbol{D}(\boldsymbol{\epsilon}, \boldsymbol{\delta})\boldsymbol{D}^{+}(\boldsymbol{\epsilon}, \boldsymbol{\delta})\right) \bar{\boldsymbol{G}} \|_F^2, \tag{25}$$

which is smooth and convex for $0 \leq \epsilon_\ell, \delta_\ell \leq 1$. Thus, any local optimization algorithm, e.g., the gradient descent method, can be used to minimize (25). It should be highlighted that if more than two NC sources are present, the presented grid-offset estimation schemes can still be applied after clustering the sources into groups of single and two NC sources.

VI. Simulation Results

This section provides simulation results to demonstrate the performance of the proposed NC SSR method that exploits the NC structure via nuclear norm minimization. Specifically, we compare the SSR solution of (20) combined with the proposed offset estimator (25) termed "NC NUC Joint" to its corresponding non-NC version (11) in combination with the offset estimator [16] termed "L1-SVD Joint". Moreover, we include the more complex NC SSR algorithm "NC L1-SVD Joint" from [17], Standard ESPRIT "SE" [1], NC Standard ESPRIT "NC SE" [5], the deterministic Cramér-Rao bound (CRB) "Det CRB" [19], and the deterministic NC CRB "Det NC CRB" [20]. The regularization parameters for the SSR methods are chosen according to (21) and [10] for $\gamma = 0.99$. The mean square error (MSE) is only computed based on the spatial frequency estimates. The simulation setup consists of a ULA with $M = 8$ isotropic sensors half-wavelength spacing apart, where the phase reference is at the array centroid. We assume $L = 2$ sources that transmit symbols drawn from a real-valued Gaussian distribution, while the sensor noise is circularly symmetric white complex Gaussian. 1000 Monte Carlo trials have been used to generate the plots.

In Fig. 1, the MSE versus the SNR is depicted for $P = 8$, $N = 10$, and the $L = 2$ uncorrelated sources are located at $\mu_1 = 15.3\Delta_\mu$ and $\mu_2 = 17.7\Delta_\mu$. As NC L1-SVD Joint [17] requires a rotation phase grid, we choose a uniform grid spacing defined by $\Delta_\varphi = \pi/MP_\varphi$. Then, the rotation phases are given by $\varphi_1 = 5.1\Delta_\varphi$ and $\varphi_2 = 29.1\Delta_\varphi$ with $P_\varphi = 6$. We emphasize again that the proposed algorithm NC NUC Joint does not require a rotation phase grid. Note that $\Delta\varphi = \varphi_2 - \varphi_1 = 24\Delta_\varphi = \pi/2$, which provides the maximum NC gain. It is apparent that NC NUC Joint and NC L1-SVD Joint both achieve the Det NC CRB and clearly outperform the L1-SVD Joint algorithm that does not exploit the NC structure.

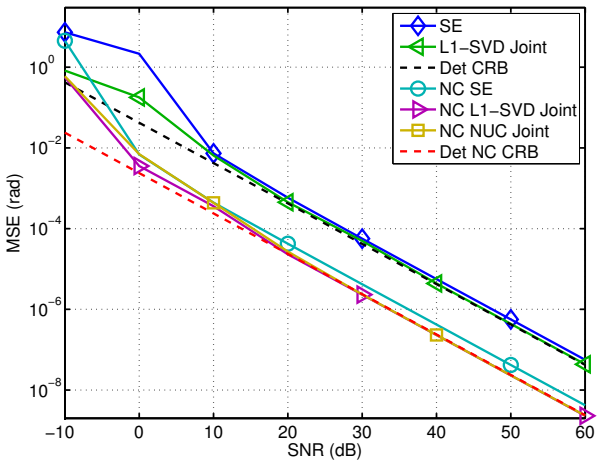

Fig. 1. MSE versus SNR for $M = 8$, $P = 8$, $N = 10$, $L = 2$ sources at $\mu_1 = 15.3\Delta_\mu$, $\mu_2 = 17.7\Delta_\mu$ with $\varphi_1 = 5.1\Delta_\varphi$, $\varphi_2 = 29.1\Delta_\varphi$ ($\Delta\varphi = \pi/2$) for $P_\varphi = 6$.

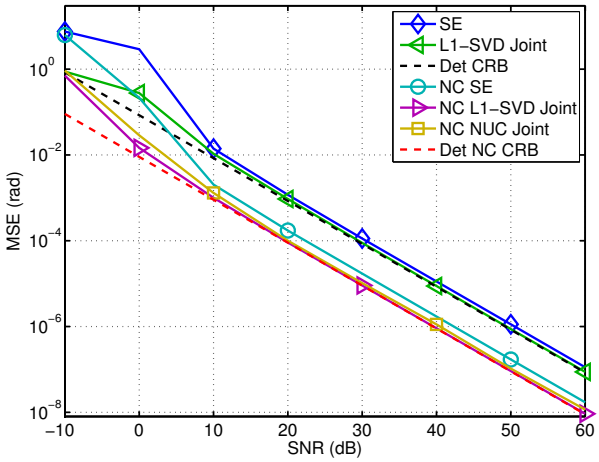

Fig. 2. MSE versus SNR for $M = 8$, $P = 8$, $N = 5$, $L = 2$ sources at $\mu_1 = 14.7\Delta_\mu$, $\mu_2 = 17.1\Delta_\mu$ with $\varphi_1 = 10.3\Delta_\varphi$, $\varphi_2 = 22.3\Delta_\varphi$ ($\Delta\varphi = \pi/4$) for $P_\varphi = 6$.

Fig. 2 shows the MSE versus the SNR for a different scenario, where $N = 5$ and $L = 2$ sources are positioned at $\mu_1 = 14.7\Delta_\mu$ and $\mu_2 = 17.1\Delta_\mu$ with $\varphi_1 = 10.3\Delta_\varphi$ and $\varphi_2 = 22.3\Delta_\varphi$ such that $\Delta\varphi = 12\Delta_\varphi = \pi/4$. The remaining parameters are kept the same. Again, the same behavior of the algorithms can be observed. NC NUC Joint and NC L1-SVD Joint perform identical and provide a clear gain over the L1-SVD Joint method.

VII. Conclusion

In this paper, we have proposed a novel approach towards exploiting the structure of NC signals in SSR-based DOA estimation. Existing NC SSR algorithms require solving a 2-D SSR problem followed by a 2-D grid offset estimation in the spatial domain and the rotation phase domain. The proposed method, however, is based on lifting the bilinear NC SSR optimization problem to a linear optimization problem in a higher-dimensional space, which is solved via nuclear norm minimization. Thus, the 2-D NC SSR problem is reduced to a 1-D NC SSR problem in the spatial domain including an automatic gridless rotation phase estimation with a significantly lower computational complexity. In our second contribution, assuming a ULA, we have presented a simple closed-form 1-D NC grid-offset estimator for a single NC source and a numerical joint grid-offset estimation procedure for two closely-spaced NC sources. The efficacy of the new NC approach has been illustrated through simulations.

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67-94, July 1996.

[2] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge, U.K.: Cambridge Univ. Press, 2010.

[3] H. Abeida and J. P. Delmas, "MUSIC-like estimation of direction of arrival for noncircular sources," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2678-2690, July 2006.

[4] P. Chargé, Y. Wang, and J. Saillard, "A non-circular sources direction finding method using polynomial rooting," *Signal Processing*, vol. 81, no. 8, pp. 1765-1770, Aug. 2001.

[5] A. Zoubir, P. Chargé, and Y. Wang, "Non circular sources localization with ESPRIT," in *Proc. European Conference on Wireless Technology (ECWT)*, Munich, Germany, Oct. 2003.

[6] M. Haardt and F. Roemer, "Enhancements of Unitary ESPRIT for non-circular sources," in *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Montreal, Canada, May 2004.

[7] J. Steinwandt, F. Roemer, M. Haardt, and G. Del Galdo, "R-dimensional ESPRIT-type algorithms for strictly second-order non-circular sources and their performance analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4824-4838, Sept. 2014.

[8] D. Donoho, "Superresolution via sparsity constraints," *SIAM Journal on Mathematical Analysis*, vol. 23, pp. 1309-1331, Sep. 1992.

[9] E. Candès, "Compressive sampling," in *Proc. of the International Congress of Mathematicians: Madrid*, August 22-30, 2006: invited lectures, 2006, pp. 1433-1452.

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

[11] M. M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed L2,0 norm approximation," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4646-4655, Sep. 2010.

[12] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 629-638, Feb. 2011.

[13] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624-638, Mar. 2006.

[14] C. Steffens, P. Parvazi, and M. Pesavento, "Direction finding and array calibration based on sparse reconstruction in partly calibrated arrays," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jun. 2014.

[15] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38-43, Jan. 2013.

[16] M. Ibrahim, F. Roemer, R. Alieiev, G. Del Galdo, and R. S. Thomä, "On the estimation of grid offsets in CS-based direction of arrival estimation," in *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[17] J. Steinwandt, F. Roemer, and M. Haardt, "Sparsity-based direction-of-arrival estimation for strictly non-circular sources," in *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, March 2016.

[18] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proc. of the American Control Conference*, vol. 6, pp. 4734-4739, 2001.

[19] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 5, pp. 720-741, May 1989.

[20] J. Steinwandt, F. Roemer, M. Haardt, and G. Del Galdo, "Deterministic Cramér-Rao bound for strictly non-circular sources and analytical analysis of the achievable gains," *IEEE Transactions on Signal Processing*, Apr. 2016, accepted.