Full Covariance Fitting DoA Estimation Using Partial Relaxation Framework

David Schenck*, Minh Trinh Hoang*, Xavier Mestre†,
Mats Viberg‡ and Marius Pesavento*
*Communication Systems Group, Technische Universität Darmstadt, Germany
†Centre Tecnològic de Telecomunicacions de Catalunya, Castelldefels, Spain
‡University Executive Unit, Blekinge Institute of Technology, Karlskrona, Sweden

***Abstract*—The so-called Partial Relaxation approach has recently been proposed to solve the Direction-of-Arrival estimation problem. In this paper, we extend the previous work by applying Covariance Fitting with a data model that includes the noise covariance. Instead of applying a single source approximation to multi-source estimation criteria, which is the case for MUSIC, the conventional beamformer, or the Capon beamformer, the Partial Relaxation approach accounts for the existence of multiple sources using a non-parametric modification of the signal model. In the Partial Relaxation framework, the structure of the desired direction is kept, whereas the sensor array manifold corresponding to the remaining signals is relaxed [1], [2]. This procedure allows to compute a closed-form solution for the relaxed signal part and to come up with a simple spectral search with a significantly reduced computational complexity. Unlike in the existing Partial Relaxed Covariance Fitting approach, in this paper we utilize more prior-knowledge on the structure of the covariance matrix by also considering the noise covariance. Simulation results show that, the proposed method outperforms the existing Partial Relaxed Covariance Fitting method, especially in difficult conditions with small sample size and low Signal-to-Noise Ratio. Its threshold performance is close to that of Deterministic Maximum Likelihood, but at significantly lower cost.**

I. Introduction

In the field of sensor array signal processing, Direction-of-Arrival (DoA) estimation has always been a major area of research due to its widespread applications in radar, sonar, seismic exploration, electronic surveillance and mobile communication [3]–[6]. Several high resolution algorithms, such as Multiple Signal Classification (MUSIC) [7], the minimum variance method of Capon [8], Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) [9] have been proposed [10], [11]. However, when two or multiple sources are closely spaced, the performance of conventional "low-cost" methods strongly degrades [12], [13]. This is due to the fact that conventional spectral search based approaches ignore the presence of interfering sources and therefore treat multi-source scenarios as single source scenarios. With increasing number of sources, the interference power increases and therefore, the performance of the conventional approaches degrades.

The Partial Relaxation (PR) framework was introduced to overcome the aforementioned disadvantages of the conventional spectral-based DoA methods [1], [2]. Instead of ignoring the presence of multiple sources, the PR approach considers both the signal impinging from the current direction of interest as well as the interfering ones. To reduce the computational demand, the manifold structure of the undesired signal components is relaxed, whereas the manifold structure of the desired signal component is kept unchanged. Closed-form solutions for the relaxed part of the array manifold are computed and substituted back into the initial optimization problem which considers multiple signals. The multi-dimensional optimization problem reduces to a one-dimensional problem which admits a simple spectral based grid search that can be applied to any array geometry.

Using the PR framework, we propose a new Partial Relaxed Full Covariance Fitting (PR-FCF) approach that utilizes more structure as compared to the existing Partial Relaxed Covariance Fitting (PR-CF) approach. The proposed PR-FCF method is of similar computational complexity as compared to conventional spectral-based techniques, and thus significantly less complex than optimal Maximum Likelihood (ML) based methods. In our simulation study, PR-FCF is found to outperform previous "low-cost" methods and its threshold Signal-to-Noise Ratio (SNR) remains close to that of Deterministic Maximum Likelihood (DML) even in low sample scenarios.

The paper is organized as follows. The signal model is introduced in Section II. The conventional covariance fitting DoA approach is introduced as non-linear Least Square (LS) problem in Section III followed by the PR-CF approach and the proposed PR-FCF approach in Section IV. Simulation results are presented in Section V. Finally, Section VI concludes this paper.

II. Signal Model

Let us consider an antenna array equipped with M sensors and N impinging narrowband signals that satisfy $M > N$. The source signal at time instant t is denoted by $\mathbf{s}(t) = [s_1(t), \ldots, s_N(t)]^T \in \mathbb{C}^{N\times 1}$. The DoAs of the signals are unknown and denoted by $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^T$. Furthermore, the full-rank steering matrix $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times N}$ is given by

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_N)],$$

where $\mathbf{a}(\theta_i) \in \mathbb{C}^M$ denotes the sensor array response for the i-th impinging signal. The number of sources N is assumed

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

to be known. The received baseband signal $\mathbf{x}(t) \in \mathbb{C}^M$ at the t-th time instant is given by

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\,\mathbf{s}(t) + \mathbf{n}(t), \qquad t = 1, \ldots, T, \tag{1}$$

where T denotes the number of snapshots and $\mathbf{n}(t) \in \mathbb{C}^M$ the sensor noise. Equation (1) can be equivalently expressed as

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\,\mathbf{S} + \mathbf{N}, \tag{2}$$

where $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(T)] \in \mathbb{C}^{M \times T}$ denotes the received baseband signal in matrix notation. The noise matrix is given by $\mathbf{N} = [\mathbf{n}(1), \ldots, \mathbf{n}(T)] \in \mathbb{C}^{M \times T}$ and the signal matrix is denoted by $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(T)] \in \mathbb{C}^{N \times T}$. Assuming that signals and noise variables are statistically independent zero-mean circularly Gaussian distributed, the covariance matrix of the received signal $\mathbf{R} \in \mathbb{C}^{M \times M}$ is given by

$$\mathbf{R} = \mathbb{E}\left\{\mathbf{x}(t)\,\mathbf{x}^H(t)\right\} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}_M, \tag{3}$$

where $\mathbf{R}_s = \mathbb{E}\left\{\mathbf{s}(t)\mathbf{s}^H(t)\right\}$ denotes the covariance of the transmitted signal and $\sigma_n^2\mathbf{I}_M$ is the noise covariance matrix. Since the true covariance matrix is unavailable in practice, the sample covariance $\hat{\mathbf{R}}$ is used instead

$$\hat{\mathbf{R}} = \frac{1}{T}\mathbf{X}\mathbf{X}^H. \tag{4}$$

The eigenvalue decomposition of (4) is given by $\hat{\mathbf{R}} = \hat{\mathbf{U}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{U}}^H$ where $\hat{\boldsymbol{\Lambda}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_M)$ contains the eigenvalues sorted in non-ascending order and $\hat{\mathbf{U}} \in \mathbb{C}^{M \times M}$ is the matrix containing the associated eigenvectors.

III. Conventional DoA Estimators

In the conventional DoA estimation framework the DoAs $\boldsymbol{\theta}$ are estimated by searching for the steering matrix $\mathbf{A}$ in the highly structured non-convex array manifold which is denoted by

$$\mathcal{A}_N = \left\{\mathbf{A} \mid \mathbf{A} = [\mathbf{a}(\vartheta_1), \ldots, \mathbf{a}(\vartheta_N)],\ \vartheta_1 < \ldots < \vartheta_N\right\}. \tag{5}$$

The multi-source DoA estimation problems are generally of the following form:

$$\left\{\tilde{\boldsymbol{\theta}}\right\} = \underset{\mathbf{A}(\boldsymbol{\theta}) \in \mathcal{A}_N}{\arg\min}\ f(\mathbf{A}(\boldsymbol{\theta})), \tag{6}$$

where $f(\mathbf{A}(\boldsymbol{\theta}))$ denotes the multi-source estimation cost function, following e.g. the DML [4], the Stochastic ML (SML) [14], or the Weighted Subspace Fitting (WSF) criteria [15]. However, the multi-source DoA estimation methods suffer from a very high computational complexity as they require a multi-dimensional grid search. In order to reduce the computational complexity, a common approach is to only consider a single source scenario. The multi-dimensional optimization problem in (6) is reduced to a multiple single source estimation problem

$$\left\{\tilde{\boldsymbol{\theta}}\right\} = {}^{N}\underset{\mathbf{a}(\theta) \in \mathcal{A}_1}{\arg\min}\ f(\mathbf{a}(\theta)),$$

where ${}^{N}\arg\min f(\cdot)$ denotes N arguments that correspond to the N deepest minima of the objective function $f(\mathbf{a}(\theta))$. It can be shown that the single source approximation is equivalent to the multidimensional search if and only if the steering vectors are orthogonal $\mathbf{A}^H\mathbf{A} = \mathbf{I}_N$. Defining the pseudo-spectrum $F(\theta) = \frac{1}{f(\mathbf{a}(\theta))}$ as the inverse of the single-source objective function, in the single source approximation the DoAs $\boldsymbol{\theta}$ are estimated by performing a grid search over the Field of View (FoV) and determining the N largest peaks of the pseudo-spectrum.

Conventional Covariance Fitting (CF)

The CF method aims to reduce noise power while preserving the N target signals. The CF approach is based on the LS problem [16]

$$\left\{\tilde{\mathbf{A}}, \tilde{\mathbf{R}}_s\right\} = \underset{\mathbf{A} \in \mathcal{A}_N,\ \mathbf{R}_s \succeq \mathbf{0}}{\arg\min} \left\|\hat{\mathbf{R}} - \mathbf{A}\mathbf{R}_s\mathbf{A}^H\right\|_F^2. \tag{7}$$

Considering the single source approach and using $\mathbf{a} = \mathbf{a}(\theta)$, the optimization problem reduces to

$$\left\{\tilde{\boldsymbol{\theta}}_{\mathrm{CF}}\right\} = {}^{N}\underset{\mathbf{a} \in \mathcal{A}_1}{\arg\min} \min_{\sigma_s^2 \geq 0} \left\|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right\|_F^2, \tag{8}$$

where σ_s^2 denotes the signal power of the non-relaxed desired signal part. The source power estimate that minimizes the inner optimization problem in (8) is given by the conventional beamformer [17]:

$$\tilde{\sigma}_s^2 = \underset{\sigma_s^2 \geq 0}{\arg\min} \left\|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right\|_F^2 = \frac{\mathbf{a}^H\hat{\mathbf{R}}\mathbf{a}}{(\mathbf{a}^H\mathbf{a})^2}. \tag{9}$$

Including a positive semidefiniteness constraint in the inner optimization problem of (8) to admit a simple closed-form solution results in the restricted problem formulation

$$\begin{aligned} \tilde{\sigma}_s^2 = \underset{\sigma_s^2 \geq 0}{\arg\min} \left\|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right\|_F^2 \\ \text{subject to } \hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H \succeq \mathbf{0}. \end{aligned} \tag{10}$$

From the positive semi-definitness constraint in (10), it follows that all eigenvalues of $\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H$ are larger than or equal to zero. Therefore, the minimizer $\tilde{\sigma}_s^2$ is given in closed-form by the Capon spectrum [17], [18]:

$$\tilde{\sigma}_s^2 = \frac{1}{\mathbf{a}^H\hat{\mathbf{R}}^{-1}\mathbf{a}}. \tag{11}$$

IV. Partial Relaxation (PR) Framework

In the framework of PR [2], we do not only consider the signals from the desired directions but also from the interfering directions. However, the structure of the interfering signals is relaxed and the computational complexity is greatly reduced. The steering matrix $\mathbf{A}$ no longer describes the highly structured array manifold $\mathcal{A}_N$ as parameterized in (5). Instead, we assume that $\mathbf{A}$ describes the partially relaxed array manifold

$$\bar{\mathcal{A}}_N = \left\{\mathbf{A} \mid \mathbf{A} = [\mathbf{a}(\vartheta), \mathbf{B}],\ \mathbf{a}(\vartheta) \in \mathcal{A}_1,\ \mathbf{B} \in \mathbb{C}^{M \times (N-1)}\right\},$$

which still retains some geometric structure of the sensor array [1]. In the minimization procedure in (6), the objective function $f([\mathbf{a}(\theta), \mathbf{B}])$ is minimized with respect to $\mathbf{B}$. A

closed-form solution for $\mathbf{B}$ is obtained and substituted back into the objective function. Afterwards a grid search on $\mathbf{a}(\theta)$ is applied to locate the N smallest minima [2], which can be summarized as

$$\left\{\hat{\boldsymbol{\theta}}\right\} = {}^{N}\underset{\mathbf{a}(\theta)\in\mathcal{A}_1}{\arg\min} \min_{\mathbf{B}} f\left([\mathbf{a}(\theta), \mathbf{B}]\right).$$

A. Partial Relaxed Covariance Fitting (PR-CF)

Let us decompose $\mathbf{A} = [\mathbf{a}(\theta), \mathbf{B}]$ and $\mathbf{S} = \left[\mathbf{s}, \mathbf{J}^T\right]^T$, where $\mathbf{s} \in \mathbb{C}^T$ denotes the signals impinging from the desired source and $\mathbf{J} \in \mathbb{C}^{(N-1)\times T}$ denotes the signals of the remaining sources. Applying the decomposition to the system model in (2) results in

$$\mathbf{X} = \mathbf{a}\mathbf{s}^T + \mathbf{B}\mathbf{J} + \mathbf{N},$$

where the relaxed received signal satisfies $\operatorname{rank}(\mathbf{B}\mathbf{J}) \leq N-1$. The sample covariance matrix $\hat{\mathbf{R}}$ is assumed to be positive definite for $T \geq M$, and the signals from other sources are assumed to be uncorrelated with the signals from the direction of interest $\mathbf{a}$. By applying the decoupling approach to the optimization problem in (7) we obtain the PR-CF method [1]

$$\begin{aligned}\left\{\hat{\boldsymbol{\theta}}_{\text{PR-CF}}\right\} = {}^{N}\underset{\mathbf{a}(\theta)\in\mathcal{A}_1}{\arg\min} \min_{\sigma_s^2\geq 0, \mathbf{D}} \left\|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{D}\mathbf{D}^H\right\|_F^2 \\ \text{subject to: } \hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{D}\mathbf{D}^H \succeq \mathbf{0} \\ \operatorname{rank}(\mathbf{D}) \leq N-1,\end{aligned} \tag{12}$$

where $\mathbf{D}\mathbf{D}^H = \mathbf{B}\mathbb{E}\left\{\mathbf{J}\mathbf{J}^H\right\}\mathbf{B}^H$. Keeping $\left\{\sigma_s^2, \mathbf{a}\right\}$ fixed and minimizing (12) with respect to $\mathbf{D}$ results in the low-rank approximation problem [19]

$$\begin{aligned}\min_{\mathbf{D}} \left\|\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \mathbf{D}\mathbf{D}^H\right\|_F^2 = \sum_{k=N}^{M} \lambda_k^2\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right) \\ \text{subject to: } \operatorname{rank}(\mathbf{D}) \leq N-1,\end{aligned} \tag{13}$$

where $\lambda_k(\cdot)$ denotes the k-th largest eigenvalue of the matrix in the argument. Computing the eigenvalue decomposition of

$$\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H = \tilde{\mathbf{U}}_s\tilde{\boldsymbol{\Lambda}}_s\tilde{\mathbf{U}}_s^H + \tilde{\mathbf{U}}_n\tilde{\boldsymbol{\Lambda}}_n\tilde{\mathbf{U}}_n^H, \tag{14}$$

where $\tilde{\boldsymbol{\Lambda}}_s$ contains the $N-1$ largest eigenvalues, a minimizer $\tilde{\mathbf{D}}$ of (13) obviously satisfies

$$\tilde{\mathbf{D}}\tilde{\mathbf{D}}^H = \tilde{\mathbf{U}}_s\tilde{\boldsymbol{\Lambda}}_s\tilde{\mathbf{U}}_s^H. \tag{15}$$

Since (12) corresponds to a classical low-rank approximation problem [19], the $(M-N)$ smallest eigenvalues of $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H$ and $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H-\tilde{\mathbf{D}}\tilde{\mathbf{D}}^H$ are identical. Additionally, $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H$ is positive semi-definite if and only if $\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \tilde{\mathbf{D}}\tilde{\mathbf{D}}^H$ is positive semi-definite, therefore the initial optimization problem in (12) is equivalent to [2]:

$$\begin{aligned}\min_{\sigma_s^2\geq 0} \sum_{k=N}^{M} \lambda_k^2\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right) \\ \text{subject to: } \hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H \succeq \mathbf{0}.\end{aligned}$$

The optimal σ_s^2 is computed by the Capon spectrum in (11) [1] and the pseudo-spectrum is therefore given by

$$F_{\text{PR-CF}}(\theta) = \frac{1}{\sum_{k=N}^{M}\lambda_k^2\left(\hat{\mathbf{R}} - \frac{1}{\mathbf{a}(\theta)^H\hat{\mathbf{R}}^{-1}\mathbf{a}(\theta)}\mathbf{a}(\theta)\mathbf{a}(\theta)^H\right)}.$$

B. Partial Relaxed Full Covariance Fitting (PR-FCF)

In the PR-FCF approach we propose to relax the structure of the covariance matrix in (3) to decrease the computational complexity. However, compared to the partial relaxation approach in (12) we will introduce a model that utilizes more structure by also considering the noise covariance matrix

$$\tilde{\mathbf{R}} = \sigma_s^2\mathbf{a}\mathbf{a}^H + \mathbf{D}\mathbf{D}^H + \sigma_n^2\mathbf{I}_M.$$

The optimization problem is given by

$$\begin{aligned}\left\{\hat{\boldsymbol{\theta}}_{\text{PR-FCF}}\right\} = {}^{N}\underset{\mathbf{a}(\theta)\in\mathcal{A}_1}{\arg\min} \min_{\sigma_s^2\geq 0, \sigma_n^2\geq 0, \mathbf{D}} \left\|\hat{\mathbf{R}} - \tilde{\mathbf{R}}\right\|_F^2 \\ \operatorname{rank}(\mathbf{D}) \leq N-1.\end{aligned} \tag{16}$$

Keeping $\left\{\sigma_s^2, \sigma_n^2, \mathbf{a}\right\}$ fixed and minimizing (16) with respect to $\mathbf{D}$ results in a low-rank approximation problem. Similar to the optimization problem in (13), the optimal $\mathbf{D}$ that minimizes (16) is given by the best rank-$(N-1)$ approximation of $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H-\sigma_n^2\mathbf{I}_M$ as described in (14) and (15). Since (16) corresponds to a classical low-rank approximation problem [19] with optimizer $\tilde{\mathbf{D}}$, the $(M-N)$ smallest eigenvalues of $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H-\sigma_n^2\mathbf{I}_M$ and $\hat{\mathbf{R}}-\sigma_s^2\mathbf{a}\mathbf{a}^H-\tilde{\mathbf{D}}\tilde{\mathbf{D}}^H-\sigma_n^2\,\mathbf{I}_M$ are identical. Therefore, the initial optimization problem in (16) can also be expressed as

$$\min_{\substack{\sigma_s^2\geq 0,\\ \sigma_n^2\geq 0}} f\left(\sigma_s^2, \sigma_n^2\right) = \min_{\substack{\sigma_s^2\geq 0,\\ \sigma_n^2\geq 0}} \sum_{k=N}^{M}\lambda_k^2\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H - \sigma_n^2\mathbf{I}_M\right). \tag{17}$$

Computing the first derivative of $f\left(\sigma_s^2, \sigma_n^2\right)$ with respect to σ_n^2 and equating it to zero [20] [21], we obtain the optimal value for σ_n^2

$$\hat{\sigma}_n^2 = \frac{\sum_{k=N}^{M}\lambda_k\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right)}{M-N+1}. \tag{18}$$

Substituting $\hat{\sigma}_n^2$ back into the objective function in (17) yields

$$\min_{\sigma_s^2\geq 0} f\left(\theta, \sigma_s^2\right) = \min_{\sigma_s^2\geq 0} \sum_{k=N}^{M}\overline{\lambda}_k^2(\theta) - \frac{\left(\sum_{k=N}^{M}\overline{\lambda}_k(\theta)\right)^2}{M-N+1}, \tag{19}$$

where we have introduced the short form notation $\overline{\lambda}_k(\theta) = \lambda_k\left(\hat{\mathbf{R}} - \sigma_s^2\mathbf{a}\mathbf{a}^H\right)$ and omit the positivity constraint on σ_n^2. The PR-FCF method can be interpreted as covariance fitting approach with diagonally-loaded sample covariance matrix. Depending on the choice of σ_s^2, the diagonal loading factor σ_n^2 in (18) can either be positive or negative. Noting that (19) is continuously differentiable, a local minimizer for σ_s^2 is determined by applying a local numerical minimization. In Algorithm IV.1 a Newton-Raphson procedure is used to estimate σ_s^2 for every grid-point within the FoV. The procedure utilizes the first and second order derivative of the cost function in (19) w.r.t. σ_s^2. Before the first iteration we initialize σ_s^2 using the conventional beamformer in (9). The pseudo-spectrum is given by a bi-dimensional function

$$F_{\text{PR-FCF}}\left(\theta, \hat{\sigma}_s^2\right) = \frac{1}{\sum_{k=N}^{M}\overline{\lambda}_k^2(\theta) - \frac{\left(\sum_{k=N}^{M}\overline{\lambda}_k(\theta)\right)^2}{M-N+1}}.$$

Computational Aspects: Using the eigenvalue decomposition of the sample covariance matrix in (4) and the unitary property of the eigenvector matrix $\hat{\mathbf{U}}$, we can equivalently express $\overline{\lambda}_k(\theta)$ as

$$\overline{\lambda}_k(\theta) = \lambda_k\left(\hat{\mathbf{\Lambda}} - \sigma_s^2 \hat{\mathbf{U}}^H \mathbf{a}\mathbf{a}^H \hat{\mathbf{U}}\right).$$

Initially proposed to allow parallel computation of eigenvalues and eigenvectors of symmetric tridiagonal matrices by Bunch, Nielsen and Sorensen in [22], the interlacing theorem for a rank-one modified Hermitian matrix states the following [20]:

Theorem 1: Let $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_K) \in \mathbb{R}^{K\times K}$ denote a diagonal matrix with distinct elements that are sorted in descending order. Furthermore assume the non-zero scalar ρ to be positive and $\mathbf{z} \in \mathbb{C}^{K\times 1}$ to contain only non-zero entries. If the eigenvalues $\left\{\overline{d}_1, \ldots, \overline{d}_K\right\}$ of matrix $\mathbf{D} - \rho\mathbf{z}\mathbf{z}^H$ are sorted in descending order as well, then:

- $\left\{\overline{d}_1, \ldots, \overline{d}_K\right\}$ correspond to the K roots of the secular function

$$\begin{aligned} p(x) =& 1 - \rho\mathbf{z}^H\left(\mathbf{D} - x\mathbf{I}_K\right)^{-1}\mathbf{z} \\ =& 1 - \rho\sum_{k=1}^{K}\frac{|z_k|^2}{d_k - x}. \end{aligned} \tag{20}$$

- $\left\{\overline{d}_1, \ldots, \overline{d}_K\right\}$ satisfy the interlacing property

$$d_1 > \overline{d}_1 > d_2 > \overline{d}_2 > \ldots > d_K > \overline{d}_K. \tag{21}$$

- The eigenvector $\overline{\mathbf{u}}_k$ that corresponds to the k-th eigenvalue $\overline{d}_k$ is a multiple of $\left(\mathbf{D} - \overline{d}_k\mathbf{I}_K\right)^{-1}\mathbf{z}$.

For an efficient computation of the eigenvalues $\overline{\lambda}_k(\theta)$, the secular function in (20) and the interlacing property in (21) of Theorem 1 are used [1], where $\mathbf{z} = \hat{\mathbf{U}}^H\mathbf{a}$, $\rho = \sigma_s^2$ and $\mathbf{D} = \hat{\mathbf{\Lambda}}$. In order to compute the minimizer σ_s^2 of (19) numerically, the first and second order derivatives of the objective function in (19) w.r.t. σ_s^2 have to be computed. Therefore, the first and the second derivatives of the modified eigenvalues w.r.t. σ_s^2 are required [20], [21]. The first derivative of $\overline{\lambda}_k(\theta)$ w.r.t σ_s^2 is given by

$$\begin{aligned} \overline{\lambda}_k'(\theta) =& -\frac{1}{\sigma_s^4\mathbf{z}^H\left(\mathbf{D} - \overline{\lambda}_k(\theta)\mathbf{I}_m\right)^{-2}\mathbf{z}} \\ =& -\frac{1}{\sigma_s^4\sum_{j=1}^{M}\frac{|z_j|^2}{\left(d_j - \overline{\lambda}_k(\theta)\right)^2}}, \end{aligned}$$

where we have applied Theorem 1. Computing the second order derivative of $\overline{\lambda}_k(\theta)$ w.r.t. σ_s^2 is straightforward and corresponds to

$$\begin{aligned} \overline{\lambda}_k''(\theta) =& \frac{2}{\sigma_s^8\left(\sum_{j=1}^{K}\frac{|z_j|^2}{\left(d_j - \overline{\lambda}_k(\theta)\right)^2}\right)^2} \\ &\times\left(\sigma_s^2\sum_{j=1}^{K}\frac{|z_j|^2}{\left(d_j - \overline{\lambda}_k(\theta)\right)^2} - \frac{\sum_{j=1}^{K}\frac{|z_j|^2}{\left(d_j - \overline{\lambda}_k(\theta)\right)^3}}{\sum_{j=1}^{K}\frac{|z_j|^2}{\left(d_j - \overline{\lambda}_k(\theta)\right)^2}}\right). \end{aligned}$$

Algorithm IV.1 Newton-Raphson Procedure to Estimate σ_s^2

1: **Initialization:** Iteration index $\tau = 0$, initial value $x^{(0)} = \sum_{j=1}^{M} d_j |z_j|^2$, tolerance ϵ
2: **repeat**
3: Compute $f'\left(\theta, x^{(\tau)}\right) = \left.\frac{\partial f\left(\theta,\sigma_s^2\right)}{\partial\sigma_s^2}\right|_{\sigma_s^2 = x^{(\tau)}}$
4: Compute $f''\left(\theta, x^{(\tau)}\right) = \left.\frac{\partial^2 f\left(\theta,\sigma_s^2\right)}{\partial^2\sigma_s^2}\right|_{\sigma_s^2 = x^{(\tau)}}$
5: Update $x^{(\tau+1)} = x^{(\tau)} - \frac{f'\left(\theta, x^{(\tau)}\right)}{f''\left(\theta, x^{(\tau)}\right)}$
6: $\tau \leftarrow \tau + 1$
7: **until** $\left|x^{(\tau+1)} - x^{(\tau)}\right| < \epsilon$ **or** $\left|f'\left(\theta, x^{(\tau)}\right)\right| < \epsilon$
8: **if** $x^{(\tau+1)} < 0$ **or** $f\left(\theta, x^{(\tau+1)}\right) > f(\theta, 0)$ **then**
9: $\hat{\sigma}_s^2 = 0$
10: **else**
11: $\hat{\sigma}_s^2 = x^{(\tau+1)}$
12: **end if**
13: **return** $\hat{\sigma}_s^2$

V. Simulation Results

In this section, simulation results regarding the Root-Mean-Squared-Error (RMSE) performance of different DoA estimators are presented and compared with the stochastic Cramer-Rao Bound (CRB). All simulations are carried out for $N_R = 1000$ Monte-Carlo runs. The RMSE

$$\text{RMSE} = \sqrt{\frac{1}{N_R N}\sum_{r=1}^{N_R}\sum_{n=1}^{N}\left(\hat{\theta}_n^{(r)} - \theta_n\right)^2}, \tag{22}$$

is used as performance indicator, where both the estimated DoAs $\hat{\boldsymbol{\theta}}^{(r)} = [\hat{\theta}_1^{(r)}, \ldots, \hat{\theta}_N^{(r)}]^T$ and the true DoAs $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^T$ are sorted in ascending order. After computing the RMSE, 1% of the estimates with the largest errors are removed to avoid outliers in the RMSE caused by misdetection in all methods. We consider a Uniform Linear Array (ULA) with $M = 10$ antennas. The antennas have an antenna spacing equal to half of the wavelength. We consider two uncorrelated sources at $\boldsymbol{\theta} = [45°, 50°]^T$. The transmitted signals are zero-mean and statistically independent with unit power. The SNR is given by $\text{SNR} = 1/\sigma_n^2$.

In Figure 1 we investigate the case, where the number of snapshots $T = 8$ is smaller than the number of sensors $M = 10$. For $T < M$, the sample covariance matrix in (4) is singular and not invertible which makes PR-CF non-applicable. Therefore, the diagonal loading technique with loading factor $\gamma = 10^{-4}$ on the sample covariance matrix is applied [23] [24]. It can be observed that the proposed PR-FCF provides better SNR threshold performance than the existing PR-CF method. PR-FCF exhibits better performance than PR-CF also for high SNRs.

Figure 2 depicts a scenario for $T = 15$ snapshots. Since $T > M$, no diagonal loading is necessary. The proposed PR-FCF provides slightly better threshold performance than PR-CF and yields better RMSE performance for high SNRs.

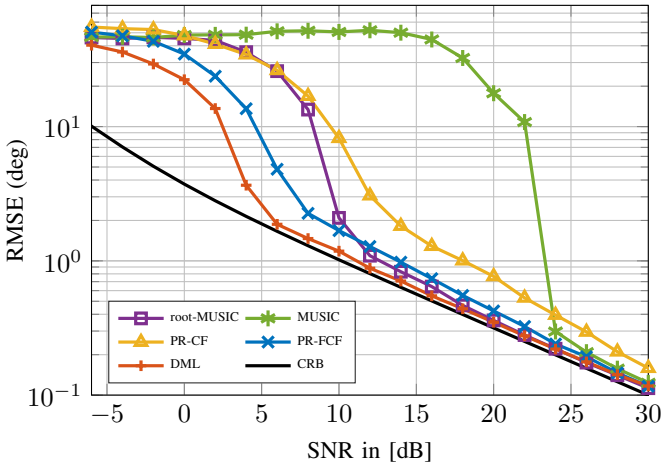

Fig. 1. Uncorrelated sources, number of snapshots $T = 8$

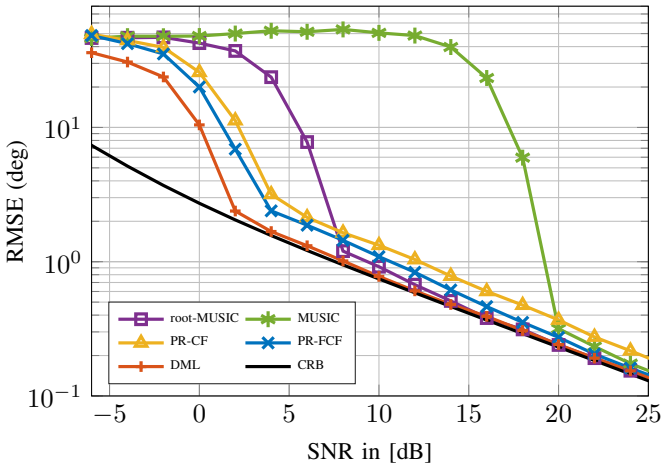

Fig. 2. Uncorrelated sources, number of snapshots $T = 15$

VI. Conclusion

In this paper, a new CF DoA estimator that utilizes the PR framework is proposed. In comparison to the existing PR-CF method, more prior-knowledge on the signal structure is used. Since the PR-FCF does not require a matrix inversion, it is also applicable for the rank deficient case where $T < M$. Therefore, PR-FCF does not require the computation of an optimal loading factor. However, PR-FCF has slightly higher computational complexity compared to PR-CF, as a Newton-Raphson procedure has to be carried out for each grid point. Nevertheless, it provides greatly reduced computational demand as compared to optimal DoA estimators such as DML that require a multi-dimensional search. Simulation results show that PR-FCF exhibits better performance than PR-CF, especially in difficult scenarios with small number of snapshots and low SNR. In the future work we intend to address outlier detection for the proposed methods.

Acknowledgment

This work was supported in part by the Catalan Government under grant 2017-SGR-01479 and Professor Dr.-Ing. Erich Müller-Stiftung.

References

[1] M. Trinh-Hoang, M. Viberg, and M. Pesavento, "Partial Relaxation Approach: An Eigenvalue-Based DOA Estimator Framework," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6190–6203, Dec 2018.

[2] ——, "An Improved DoA Estimator Based on Partial Relaxation Approach," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018, pp. 3246–3250.

[3] A. Rembovsky, A. Ashikhmin, V. Kozmin, and S. Smolskiy, *Radio Monitoring: Problems, Methods and Equipment*, ser. Lecture Notes in Electrical Engineering. Springer US, 2009.

[4] H. V. Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. Wiley, 2004.

[5] 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, Jul 1996.

[6] "DOA Estimation Methods and Algorithms," ser. Academic Press Library in Signal Processing, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, vol. 3.

[7] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar 1986.

[8] J. Capon, "High-Resolution Frequency-Wavenumber Spectrum Analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug 1969.

[9] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul 1989.

[10] L. C. Godara, "Application of Antenna Arrays to Mobile Communications. II. Beam-forming and Direction-of-Arrival Considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, Aug 1997.

[11] M. Haardt, M. Pesavento, F. Roemer, and M. N. E. Korso, "Subspace Methods and Exploitation of Special Array Structures," in *Academic Press Library in Signal Processing*, ser. Academic Press Library in Signal Processing, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, ch. 15, pp. 651 – 717.

[12] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood and Cramer-Rao Bound: Further Results and Comparisons," in *International Conference on Acoustics, Speech, and Signal Processing,*, May 1989, pp. 2605–2608 vol.4.

[13] C. Vaidyanathan and K. M. Buckley, "Performance Analysis of the MVDR Spatial Spectrum Estimator," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1427–1437, Jun 1995.

[14] P. Stoica and A. Nehorai, "On the Concentrated Stochastic Likelihood Function in Array Signal Processing," *Circuits, Systems and Signal Processing*, vol. 14, no. 5, pp. 669–674, Sep 1995.

[15] M. Viberg and B. Ottersten, "Sensor Array Processing Based on Subspace Fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1110–1121, May 1991.

[16] B. Ottersten, P. Stoica, and R. Roy, "Covariance Matching Estimation Techniques for Array Signal Processing Applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185 – 210, 1998.

[17] P. Stoica and R. Moses, "Spectral Analysis of Signals," 2005.

[18] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading," *Trans. Sig. Proc.*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[19] T. Hayden and J. Wells, "Approximation by Matrices Positive Semidefinite on a Subspace," *Linear Algebra and its Applications*, vol. 109, no. Supplement C, pp. 115 – 130, 1988.

[20] G. H. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins University Press, 1996.

[21] J. R. Magnus, "On Differentiating Eigenvalues and Eigenvectors," *Econometric Theory*, vol. 1, no. 2, p. 179–191, 1985.

[22] J. R. Bunch, C. P. Nielsen, and D. C. Sorensen, "Rank-One Modification of the Symmetric Eigenproblem," *Numerische Mathematik*, vol. 31, no. 1, pp. 31–48, 1978.

[23] S. A. Vorobyov, "Adaptive and Robust Beamforming," in *Academic Press Library in Signal Processing*, ser. Academic Press Library in Signal Processing, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, ch. 12, pp. 503 – 552.

[24] P. Stoica, J. Li, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading," in *IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP*, vol. 5, April 2003, pp. V–337.