



**HAL**  
open science

# A PARTIALLY-RELAXED ROBUST DOA ESTIMATOR UNDER NON-GAUSSIAN LOW-RANK INTERFERENCE AND NOISE

Minh Trinh-Hoang, Mohammed Nabil El Korso, Marius Pesavento

► **To cite this version:**

Minh Trinh-Hoang, Mohammed Nabil El Korso, Marius Pesavento. A PARTIALLY-RELAXED ROBUST DOA ESTIMATOR UNDER NON-GAUSSIAN LOW-RANK INTERFERENCE AND NOISE. (ICASSP 2021) 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing Search form Search, Jun 2021, Toronto, Canada. hal-03144957

**HAL Id: hal-03144957**

**<https://hal.science/hal-03144957>**

Submitted on 18 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A PARTIALLY-RELAXED ROBUST DOA ESTIMATOR UNDER NON-GAUSSIAN LOW-RANK INTERFERENCE AND NOISE

Minh Trinh-Hoang<sup>\*</sup>    Mohammed Nabil El Korso<sup>‡</sup>    Marius Pesavento<sup>\*</sup>

<sup>\*</sup> Communication Systems Group, Technical University of Darmstadt, Germany

<sup>‡</sup> Laboratory for Energetic Mechanics and Electromagnetism, Paris-Nanterre University, France

## ABSTRACT

In practical applications, non-Gaussianity of the signal at the sensor array is detrimental to the performance of conventional Direction-of-Arrival (DOA) estimators developed under the Gaussian model. In this paper, we propose a novel robust DOA estimator from the data collected at the sensor array under the corruption of non-Gaussian interference and noise. Additionally, the Cramér-Rao bound for DOA parameters under the considered signal model is derived. Simulation results show that the proposed estimator exhibits near-optimal estimation performance under the assumed model while being robust to model mismatch and/or the presence of outliers.

**Index Terms**— DOA Estimation, EM Algorithm, Robust Estimation, Partial Relaxation

## 1. INTRODUCTION

Recently, a class of Direction-of-Arrival (DOA) estimators have been proposed under the Partial Relaxation (PR) framework [1, 2]. Similar to the MUSIC algorithm [3], PR estimators employ a one-dimensional spectral search to estimate the DOAs. Simulation results show that, under the Gaussian signal model, the threshold performance of PR estimators is comparable to that of conventional maximum likelihood estimators [4–6] while the computational complexity is in the same order of magnitude as the MUSIC algorithm.

As the derivation of PR estimators in [1] relies on the assumption of Gaussian source signals and spatially white Gaussian noise, the DOA estimation performance of the PR methods degrades in practical scenarios where the receive signal at the sensor array is not Gaussian-distributed. For example, in MIMO radar systems, the receive signal is corrupted by radar clutter, which is generally modeled as realizations of some stochastic processes with heavy-tailed distributions [7]. In the context of radio astronomical imaging, aside from thermal noise, low-rank terrestrial interferers [8] are imminent, which adversely impact the imaging quality, especially for subspace-based methods in general and estimators under the PR framework in particular [9, 10]. Therefore, deriving robust DOA estimators under consideration of the non-Gaussianity of both interference and noise [11–13] is of great interest.

In this paper, we propose a robust DOA estimator under the consideration of non-Gaussian observations which does not rely on the robust estimation of the covariance matrix as in [8, 14–17]. The receive signal at the sensor array is assumed to be corrupted by Student  $t$ -distributed interference and noise. The proposed algorithm, which is based on the Expectation Maximization (EM) algorithm [18] and the PR framework, jointly estimates the non-Gaussianity of the receive signal and the DOA parameters in each iteration. From the simulation results, the proposed algorithm exhibits favorable performance even in the case of model mismatch. The paper is organized as follows. The signal model is presented in Section 2. In Section 3, the Cramér-Rao bound (CRB) for the DOA parameters is derived in closed-form. Details on the derivation of the proposed DOA estimator is provided in Section 4. Numerical results in Section 5 illustrates the performance of the proposed algorithm under different scenarios. Lastly in Section 6, remarks and extensions for future research are discussed.

## 2. SIGNAL MODEL

Consider an array of  $M$  sensors receiving  $N$  narrowband signals emitted from sources with the corresponding unknown DOAs  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T$  and under corruption of  $R$  interferers. The baseband receive signal  $\boldsymbol{x}(t) = [x_1(t), \dots, x_M(t)]^T \in \mathbb{C}^M$  at the time instant  $t$  is modeled as

$$\boldsymbol{x}(t) = \mathbf{A}(\boldsymbol{\theta})\boldsymbol{s}(t) + \mathbf{B}\boldsymbol{j}(t) + \boldsymbol{n}(t) \text{ for } t = 1, \dots, T. \quad (1)$$

In (1), the steering matrix  $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_N)]$  is a full rank matrix whose  $n$ -th column  $\mathbf{a}(\theta_n)$  represents the sensor array response for the DOA  $\theta_n$ .  $\mathbf{B} \in \mathbb{C}^{M \times R}$  denotes the unstructured loading matrix of  $R$  interferers. In addition,  $\boldsymbol{s}(t) \in \mathbb{C}^N$ ,  $\boldsymbol{j}(t) \in \mathbb{C}^R$  and  $\boldsymbol{n}(t) \in \mathbb{C}^M$  denote the baseband source signal vector, the baseband interference vector, and the noise vector at the sensor array, respectively. While the source signal  $\boldsymbol{s}(t)$  is assumed to be deterministic but unknown, the interference vector  $\boldsymbol{j}(t)$  and the noise vector  $\boldsymbol{n}(t)$  admit the stochastic representations:

$$\boldsymbol{j}(t) \stackrel{d}{=} \frac{1}{\sqrt{\tau(t)}}\boldsymbol{d}(t) \quad \text{and} \quad \boldsymbol{n}(t) \stackrel{d}{=} \frac{1}{\sqrt{\tau(t)}}\boldsymbol{e}(t). \quad (2)$$

In (2), the speckles  $\boldsymbol{d}(t)$  and  $\boldsymbol{e}(t)$  are i.i.d. vectors drawn from  $\boldsymbol{d}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  and  $\boldsymbol{e}(t) \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ , respectively.

This work was supported by the DFG PRIDE Project PE 2080/2-1.

The noise power at the sensor array is denoted by  $\sigma^2$ . The textures  $\tau(t) \sim \text{Gamma}(\nu, \nu)$ <sup>1</sup> for  $t = 1, \dots, T$  are i.i.d. random scalars. The positive number of degrees of freedom (DOF) is denoted by  $\nu$ , which models the non-Gaussianity of the interferers and the noise. Equation (1) can be rewritten for multiple snapshots  $t = 1, \dots, T$  in a compact notation as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{B}\mathbf{J} + \mathbf{N}, \quad (3)$$

where  $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(T)] \in \mathbb{C}^{M \times T}$  is the baseband receive signal matrix. In a similar manner, we define the source signal matrix  $\mathbf{S} \in \mathbb{C}^{N \times T}$ , the interfering signal matrix  $\mathbf{J} \in \mathbb{C}^{R \times T}$  and the sensor noise matrix  $\mathbf{N} \in \mathbb{C}^{M \times T}$  as  $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(T)]$ ,  $\mathbf{J} = [\mathbf{j}(1), \dots, \mathbf{j}(T)]$  and  $\mathbf{N} = [\mathbf{n}(1), \dots, \mathbf{n}(T)]$ , respectively.

### 3. CRAMÉR-RAO BOUND

The derivation of the CRB for the DOA parameters  $\boldsymbol{\theta}$  under the signal model in (1) is outlined in this section. By marginalizing w.r.t. the texture  $\tau(t)$ , we can show that the observation vectors  $\mathbf{x}(t)$  for  $t = 1, \dots, T$  are i.i.d. drawn from a circularly complex Student's  $t$ -distribution  $\mathbf{x}(t) \sim \text{C}t_\nu(\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t), \boldsymbol{\Sigma})$  with the scale matrix  $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}^H + \sigma^2\mathbf{I}$ . As  $\boldsymbol{\Sigma}$  does not depend on the DOA parameters  $\boldsymbol{\theta}$ , the CRB for  $\boldsymbol{\theta}$  under the Student's  $t$ -Distribution case is proportional to that under the Gaussian case with the mean  $\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$  and the covariance matrix  $\boldsymbol{\Sigma}$  [19]. Therefore, combining [19, Eq. (22)] with [20, Eq. (31)], the CRB for the DOA parameters  $\boldsymbol{\theta}$  is given by

$$\mathbf{C}(\boldsymbol{\theta}) = \frac{\nu + M + 1}{2(\nu + M)T} \left[ \text{Re} \left\{ \left( \tilde{\mathbf{D}}^H \boldsymbol{\Pi}_A^\perp \tilde{\mathbf{D}} \right) \odot \mathbf{P}_s \right\} \right]^{-1}, \quad (4)$$

with  $\tilde{\mathbf{D}} = \boldsymbol{\Sigma}^{-1/2} \left[ \frac{\partial \mathbf{a}(\theta_1)}{\partial \theta}, \dots, \frac{\partial \mathbf{a}(\theta_N)}{\partial \theta} \right]$ ,  $\boldsymbol{\Pi}_A^\perp = \mathbf{I} - \mathbf{A}\mathbf{A}^\dagger$  and  $\mathbf{P}_s = \sum_{t=1}^T \mathbf{s}(t)\mathbf{s}(t)^H$ .

### 4. PROPOSAL OF THE EM-PR-DML ALGORITHM

In this section, we propose a novel maximum likelihood-based DOA estimator under the signal model in Section 2. The parameter vector  $\boldsymbol{\phi}$  which characterizes the density function of the observation vector  $\mathbf{x}(t)$  is given by

$$\boldsymbol{\phi} = [\boldsymbol{\theta}^T, \mathbf{s}(1)^T, \dots, \mathbf{s}(T)^T, \boldsymbol{\zeta}^T, \nu, \sigma^2]^T, \quad (5)$$

where  $\boldsymbol{\zeta}$  contains the real and imaginary part of non-redundant entries of  $\boldsymbol{\Sigma}$ . The log-likelihood  $\mathcal{L} = \log p(\mathbf{X}|\boldsymbol{\phi})$  is given by

$$\mathcal{L} = -T \log \det \boldsymbol{\Sigma} - (M + \nu) \sum_{t=1}^T \log \left( 1 + \frac{1}{\nu} \mathbf{r}(t)^H \boldsymbol{\Sigma}^{-1} \mathbf{r}(t) \right) \quad (6)$$

with the residuals  $\mathbf{r}(t) = \mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$ . As the direct maximization of the log-likelihood function in (6) w.r.t. the parameter vector  $\boldsymbol{\phi}$  in (5) appears to be challenging, we adopt the EM algorithm that estimates the DOA parameters in an iterative manner: first, in the E-step (expectation step) of the  $k$ -th iteration, the lower-bound function  $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)})$  of the log-likelihood  $\log p(\mathbf{X}|\boldsymbol{\phi})$  at the current iteration point  $\boldsymbol{\phi}^{(k)}$  is constructed as

<sup>1</sup>Gamma( $\kappa_1, \kappa_2$ ) denotes the Gamma distribution with the shape parameter  $\kappa_1$  and the inverse-scale parameter  $\kappa_2$

$$Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) = \mathbb{E}_{\mathbf{Y}|\mathbf{X}, \boldsymbol{\phi}^{(k)}} \{ \log p(\mathbf{Y}|\boldsymbol{\phi}) \}. \quad (7)$$

In (7),  $\mathbf{Y}$  denotes the so-called (unobserved) complete data matrix. Then, in the M-step (maximization step) in the  $k$ -th iteration, the parameter vector  $\boldsymbol{\phi}$  is updated by maximizing the lower-bound function as follows:

$$\boldsymbol{\phi}^{(k+1)} = \arg \max \boldsymbol{\phi} Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}). \quad (8)$$

The two steps in (7) and (8) are alternatively performed until convergence. Details on the E-steps and M-steps are provided in the following subsections.

#### 4.1. Expectation Step

By choosing the complete data as  $\mathbf{Y} = [\mathbf{y}(1) \dots, \mathbf{y}(T)]$  with  $\mathbf{y}(t) = [\mathbf{x}(t)^T, \mathbf{j}(t)^T, \tau(t)]^T$ , the lower-bound function  $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)})$  in (7) is rewritten as:

$$Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) = \sum_{t=1}^T Q_t(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) \quad (9)$$

$$\begin{aligned} \text{with} \\ Q_t(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) &= \mathbb{E}_{\mathbf{j}(t), \tau(t)|\mathbf{x}(t), \boldsymbol{\phi}^{(k)}} \{ \log p(\tau(t)|\boldsymbol{\phi}) \} \\ &+ \mathbb{E}_{\mathbf{j}(t), \tau(t)|\mathbf{x}(t), \boldsymbol{\phi}^{(k)}} \{ \log p(\mathbf{j}(t)|\tau(t), \boldsymbol{\phi}) \} \\ &+ \mathbb{E}_{\mathbf{j}(t), \tau(t)|\mathbf{x}(t), \boldsymbol{\phi}^{(k)}} \{ \log p(\mathbf{x}(t)|\mathbf{j}(t), \tau(t), \boldsymbol{\phi}) \}. \end{aligned} \quad (10)$$

For notational simplicity, we drop the time instant  $t$  and the DOA parameters  $\boldsymbol{\theta}$  in the remainder of this subsection. From (1) and (2), the expression in (10) is computed as

$$\begin{aligned} Q_t(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) &= -(M + R) \log \pi + \nu \log \nu - \log \Gamma(\nu) \\ &- M \log \sigma^2 + (M + R + \nu - 1) \overline{\log \tau} - (\|\mathbf{r}\|^2 / \sigma^2 + \nu) \bar{\tau} \\ &+ \frac{2}{\sigma^2} \text{Re} \{ \mathbf{r}^H \mathbf{B} \bar{\tau} \bar{\mathbf{j}} \} - \frac{1}{\sigma^2} \text{tr} \{ \mathbf{B}^H \mathbf{B} \bar{\tau} \bar{\mathbf{j}} \bar{\mathbf{j}}^H \} - \text{tr} \{ \bar{\tau} \bar{\mathbf{j}} \bar{\mathbf{j}}^H \}, \end{aligned} \quad (11)$$

where  $\overline{(\cdot)} = \mathbb{E}_{\mathbf{j}, \tau|\mathbf{x}, \boldsymbol{\phi}^{(k)}} \{ \cdot \}$  denotes the expectation of the argument conditioned to the observed data  $\mathbf{x}$  and the previously estimated parameters  $\boldsymbol{\phi}^{(k)}$ . Using Bayes theorem and after some algebraic manipulation, the conditional expectation quantities in (11) are computed as follows

$$\bar{\tau} = \frac{\alpha^{(k)}}{\beta^{(k)}}, \quad \overline{\log \tau} = \psi(\alpha^{(k)}) - \log \beta^{(k)}, \quad (12)$$

$$\bar{\tau} \bar{\mathbf{j}} = \bar{\tau} \bar{\mathbf{j}}, \quad \overline{\tau \mathbf{j} \bar{\mathbf{j}}^H} = \bar{\tau} \bar{\mathbf{j}} \bar{\mathbf{j}}^H + (\sigma^{(k)})^2 \mathbf{F}^{(k)}, \quad (13)$$

with  $\mathbf{r}^{(k)} = \mathbf{x} - \mathbf{A}^{(k)}\mathbf{s}^{(k)}$ ,  $\boldsymbol{\Sigma}^{(k)} = \mathbf{B}^{(k)}(\mathbf{B}^{(k)})^H + (\sigma^{(k)})^2 \mathbf{I}$ ,  $\mathbf{F}^{(k)} = ((\mathbf{B}^{(k)})^H \mathbf{B}^{(k)} + (\sigma^{(k)})^2 \mathbf{I})^{-1}$ ,  $\alpha^{(k)} = M + \nu^{(k)}$ ,  $\beta^{(k)} = \nu^{(k)} + (\mathbf{r}^{(k)})^H (\boldsymbol{\Sigma}^{(k)})^{-1} \mathbf{r}^{(k)}$ ,  $\bar{\mathbf{j}} = \mathbf{F}^{(k)}(\mathbf{B}^{(k)})^H \mathbf{r}^{(k)}$  and  $\psi(\cdot)$  denotes the digamma function.

#### 4.2. Maximization Step

In the M-step, the parameters  $\boldsymbol{\theta}$ ,  $\mathbf{S}$ ,  $\mathbf{B}$ ,  $\nu$  and  $\sigma^2$  are updated by maximizing the expression in (11). Neglecting constant terms, the function  $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)})$  in (9) is rewritten as

$$Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(k)}) \propto -MT \log \sigma^2 - \frac{f(\mathbf{A}, \mathbf{S}, \mathbf{B}|\boldsymbol{\phi}^{(k)})}{\sigma^2} - g(\nu|\boldsymbol{\phi}^{(k)}) \quad (14)$$

where two auxiliary functions  $g = g(\nu|\boldsymbol{\phi}^{(k)})$  and  $f = f(\mathbf{A}, \mathbf{S}, \mathbf{B}|\boldsymbol{\phi}^{(k)})$  are defined by:

$$g = \nu \sum_{t=1}^T (\bar{\tau}(t) - \log \tau(t)) + T \log \Gamma(\nu) - \nu T \log \nu, \quad (15)$$

$$\begin{aligned} f &= \|(\mathbf{X} - \mathbf{A}\mathbf{S})\mathbf{D}\|_F^2 + \text{tr} \{ \mathbf{B}^H \mathbf{B} \mathbf{V} \} \\ &- 2\text{Re} \left\{ \text{tr} \left\{ \mathbf{B} \mathbf{U} (\mathbf{X} - \mathbf{A}\mathbf{S})^H \right\} \right\}, \end{aligned} \quad (16)$$

respectively. Regarding the definition of the function  $f$  in (16), the auxiliary matrices

$$\mathbf{D} = \text{diag}(\sqrt{\tau(1)}, \dots, \sqrt{\tau(T)}), \quad (17)$$

$$\mathbf{U} = \begin{bmatrix} \tau(1)\mathbf{j}(1), \dots, \tau(T)\mathbf{j}(T) \end{bmatrix}, \quad (18)$$

$$\mathbf{V} = \sum_{t=1}^T \tau(t)\mathbf{j}(t)\mathbf{j}(t)^H, \quad (19)$$

are constant. We remark from (14) that the number of DOF  $\nu$  are decoupled from the remaining parameters. In the following, we provide the update of  $\nu$  from the minimization of  $g = g(\nu|\phi^{(k)})$  in (15).

#### Updating the number of degrees of freedom $\nu$ :

A necessary condition for the updated number of DOF  $\nu^{(k+1)}$  is that the derivative  $g'(\nu|\phi^{(k)}) = h(\nu)$  in (15) vanishes at  $\nu = \nu^{(k+1)}$ . The function  $h(\nu)$  for  $\nu > 0$  is given by

$$h(\nu) = \sum_{t=1}^T (\overline{\tau(t)} - \log \tau(t)) + T\psi(\nu) - T \log \nu - T \quad (20)$$

and its first derivative of  $h'(\nu)$  is given by

$$h'(\nu) = T\psi_1(\nu) - T/\nu, \quad (21)$$

where  $\psi_1(\cdot)$  denotes the trigamma function. By noting that  $\tau - \log \tau \geq 1$  for all  $\tau > 0$ , we obtain

$$\sum_{t=1}^T (\overline{\tau(t)} - \log \tau(t)) > T. \quad (22)$$

From (22) and [21, Sec. 2],  $h(\nu)$  is monotonically increasing from negative to positive. Therefore, the root of  $h(\nu)$ , which corresponds to the minimum of  $g(\nu|\phi^{(k)})$ , is unique [22]. In order to obtain the minimum numerically, we adopt the rational approximation method [23]: At a given iteration point  $\nu_q$ , the function  $h(\nu)$  is approximated by a rational function of type  $u(\nu) = a - \frac{b}{\nu}$  with two parameters  $a$  and  $b$  such that  $h(\nu_q) = u(\nu_q)$  and  $h'(\nu_q) = u'(\nu_q)$ . (23)

Then, the next iteration point  $\nu_{q+1}$  is computed as the root of the rational approximant, i.e.,  $u(\nu_{q+1}) = 0$ . The expression to compute the next iteration point  $\nu_{q+1}$  from the current iteration point  $\nu_q$  is given in Step 4 of Algorithm 1. Using arguments similar to the proof in [23, Sec. 3], the convergence rate of Algorithm 1 is quadratic, which is identical to the Newton's method. In practice, the number of iterations of Algorithm 1 is approximately one third of that using the Newton's method for a given accuracy. This phenomenon can be explained as follows: Newton's method, which is based on linear approximation of the derivative, requires several restarts if the constraint  $\nu_{q+1} > 0$  is violated. In contrast, if the current point  $\nu_q$  is positive, it can be proven that the pole  $\nu = 0$  of the rational approximant  $u(\nu)$  prevents the next iteration point  $\nu_{q+1}$  from being negative. As a result, restart is not required in Algorithm 1 for any starting point  $\nu_0 > 0$ .

#### Updating the remaining parameters $\{\theta, \mathbf{S}, \mathbf{B}, \sigma^2\}$

Taking the derivative of the cost function in (14) w.r.t.  $\sigma^2$  while keeping  $\mathbf{A} = \mathbf{A}(\theta)$ ,  $\mathbf{B}$  and  $\mathbf{S}$  fixed, the optimal value for  $\sigma^2$  is given by

$$\left(\sigma^{(k+1)}\right)^2 = \frac{f(\mathbf{A}, \mathbf{S}, \mathbf{B}|\phi^{(k)})}{MT}, \quad (24)$$

---

#### Algorithm 1 Updating the number of DOF $\nu$

---

- 1: **Initialization:** Inner iteration index  $q = 0$ , starting point  $\nu_q = \nu^{(k)} > 0$  from the  $k$ -th EM iteration.
  - 2: **while** not converged **do**
  - 3:   Compute  $h(\nu_q)$  in (20) and  $h'(\nu_q)$  in (21).
  - 4:   Compute  $\nu_{q+1} = \frac{\nu_q^2 h'(\nu_q)}{h(\nu_q) + \nu_q h'(\nu_q)}$ .
  - 5:    $q \leftarrow q + 1$ .
  - 6: **end while**
  - 7: **return** Number of DOF  $\nu^{(k+1)} = \nu_q$ .
- 

where  $f(\mathbf{A}, \mathbf{S}, \mathbf{B}|\phi^{(k)})$  is defined in (16). Substituting (24) back to (14), the maximization step in the EM algorithm is equivalent to the minimization of the function  $f = f(\mathbf{A}, \mathbf{S}, \mathbf{B}|\phi^{(k)})$ . As the function  $f$  is quadratic w.r.t.  $\mathbf{B}$ , the minimizer  $\mathbf{B}^{(k+1)}$  is given by

$$\mathbf{B}^{(k+1)} = (\mathbf{X} - \mathbf{A}\mathbf{S})\mathbf{U}^H\mathbf{V}^{-1}. \quad (25)$$

Substitute  $\mathbf{B} = \mathbf{B}^{(k+1)}$  in (25) back to (16), the concentrated cost function  $f$  w.r.t.  $\mathbf{A}$  and  $\mathbf{S}$  reads

$$f = \|(\mathbf{X} - \mathbf{A}\mathbf{S})\mathbf{D}\|_F^2 - \left\|(\mathbf{X} - \mathbf{A}\mathbf{S})\mathbf{U}^H\mathbf{V}^{-1/2}\right\|_F^2. \quad (26)$$

From (13) and (16), it is not difficult to show that

$$\mathbf{W} = \mathbf{D}\mathbf{D}^H - \mathbf{U}^H\mathbf{V}^{-1}\mathbf{U} \succeq \mathbf{D}\mathbf{\Pi}_{\mathbf{D}^{-1}\mathbf{U}^H}^{\perp}\mathbf{D}^H \succeq \mathbf{0}. \quad (27)$$

From (27) and direct algebraic manipulation, the expression in (26) is reformulated as follows:

$$f = \left\|(\mathbf{X} - \mathbf{A}\mathbf{S})\mathbf{W}^{1/2}\right\|_F^2. \quad (28)$$

The minimizer  $\mathbf{S}^{(k+1)}$  of (28) in dependence of  $\mathbf{A}$  is given by

$$\mathbf{S}^{(k+1)} = \mathbf{A}^\dagger \mathbf{X}. \quad (29)$$

Substitute  $\mathbf{S} = \mathbf{S}^{(k+1)}$  back to (28), the DOA vector  $\theta^{(k+1)}$  is obtained from the following optimization problem

$$\theta^{(k+1)} = \arg \min_{\theta} \text{tr} \left\{ \mathbf{\Pi}_{\mathbf{A}(\theta)}^{\perp} \mathbf{X}\mathbf{W}\mathbf{X}^H \right\}. \quad (30)$$

We remark that the DOA estimation step in (30) resembles the conventional DML estimator under the spatially uniform Gaussian noise [4]. Nevertheless, a key difference is the introduction of the weighting matrix  $\mathbf{W}$  in (27), which accounts for the expected texture realization in each iteration. As the DOA estimation problem in (30) requires an  $N$ -dimensional search, we propose the use of the PR-DML estimator in [1], which is a relaxed variant of the conventional DML estimator. Specifically, the DOA parameters are estimated from the location of the  $N$ -deepest local minima of the PR-DML one-dimensional null-spectrum:

$$F_{\text{PR-DML}}(\theta) = \sum_{m=N}^M \lambda_m(\mathbf{\Pi}_{\mathbf{A}(\theta)}^{\perp} \mathbf{X}\mathbf{W}\mathbf{X}^H). \quad (31)$$

In (31),  $\lambda_m(\cdot)$  denotes the  $m$ -th largest eigenvalue of the matrix argument. We remark that the null-spectrum in (31) can be efficiently implemented as in [1, Sec. V.] without the full eigenvalue decomposition. Furthermore, the estimated DOAs from (31) results in a decrease of the  $Q$  function in (9) in practice, which implies the convergence of the EM Algorithm. To summarize, the proposed EM algorithm for the DOA estimation under non-Gaussian, spatially-colored interference-plus-noise is presented in Algorithm 2.

---

**Algorithm 2** EM-PR-DML Algorithm
 

---

- 1: **Initialization:** Outer iteration index  $k = 0$ , initial quantities  $\boldsymbol{\theta}^{(0)}$ ,  $\mathbf{S}^{(0)}$ ,  $\mathbf{B}^{(0)}$ ,  $(\sigma^{(0)})^2$  and  $\nu^{(0)}$
  - 2: **while** not converged **do**
  - 3:    $k \leftarrow k + 1$ .
  - 4:   **E-step:** Compute necessary expectation quantities in (12)-(13) and update the matrices  $\mathbf{D}$ ,  $\mathbf{U}$ ,  $\mathbf{V}$  and  $\mathbf{W}$  in (17)-(19) and (27), respectively.
  - 5:   **M-step:**
  - 6:     Update  $\nu^{(k+1)}$  as presented in **Algorithm 1**.
  - 7:     Update  $\boldsymbol{\theta}^{(k+1)}$  from the local minima of the null-spectrum in (31) and  $\mathbf{A}^{(k+1)} = \mathbf{A}(\boldsymbol{\theta}^{(k+1)})$ .
  - 8:     Update  $\mathbf{S}^{(k+1)} = (\mathbf{A}^{(k+1)})^\dagger \mathbf{Y}$ .
  - 9:     Update  $\mathbf{B}^{(k+1)} = (\mathbf{Y} - \mathbf{A}^{(k+1)} \mathbf{S}^{(k+1)}) \mathbf{U}^H \mathbf{V}^{-1}$ .
  - 10:     Update  $(\sigma^{(k+1)})^2 = \frac{\text{tr} \left\{ \mathbf{\Pi}_{\mathbf{A}^{(k+1)}}^\perp \mathbf{X} \mathbf{W} \mathbf{X}^H \right\}}{MT}$ .
  - 11: **end while**
  - 12: **return** The estimated DOAs  $\boldsymbol{\theta}^{(k)}$
- 

## 5. SIMULATION RESULTS

In this section, simulation results regarding the estimation performance of different DOA estimators are presented and compared with the CRB in (4). The number of Monte-Carlo trials is  $N_R = 5000$ . The Root-Mean-Squared-Error (RMSE) is computed as  $\left( \sum_{r=1}^{N_R} \sum_{n=1}^N (\hat{\theta}_n^{(r)} - \theta_n)^2 / (N_R N) \right)^{1/2}$ , where the estimated DOAs in the  $r$ -th Monte-Carlo run  $\hat{\boldsymbol{\theta}}^{(r)} = [\hat{\theta}_1^{(r)}, \dots, \hat{\theta}_N^{(r)}]^T$  and the true DOAs  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T$  are sorted in ascending order. In our simulations we assume a ULA of  $M = 10$  antennas with the spacing equal to half of the wavelength. The number of snapshots is set to  $T = 100$ . The matrix  $\mathbf{B}$  associated with the interfering signals is chosen such that  $\|\mathbf{B}\|_F^2 = MR$ . The Signal-to-Noise-Ratio (SNR) is calculated as  $\text{SNR} = 1/\sigma^2$ .

To illustrate the robustness of the proposed EM-PR-DML algorithm, in the first setup, we assume two uncorrelated source signals with unit power located at  $\boldsymbol{\theta} = [10^\circ, 35^\circ]^T$ . The SNR is varied between  $-10$  and  $10$  dB. The receive signal at the sensor array is additionally corrupted by  $R = 2$  interfering signals with  $\nu = 2$ . However, we assume an incorrect number of interferers  $R' = 4$  in all considered algorithms. In Fig. 1, we observe that the AML estimator [10] cannot detect two source signals due to the corruption of spatially-correlated low-rank interference. The estimation performances of the conventional MUSIC [3] and PR-DML [1] slightly degrades due to the non-Gaussianity of the interference and noise. The RG-MUSIC [16] and our proposed EM-PR-DML algorithm exhibit similar performances. Nevertheless, only our proposed estimator achieves the CRB in the high SNR regime.

In Fig. 2, we assume no interference corrupting the receive signal. The sources are closely-spaced at  $\boldsymbol{\theta} = [10^\circ, 15^\circ]^T$ . However, we assume the presence of a Gaussian noise

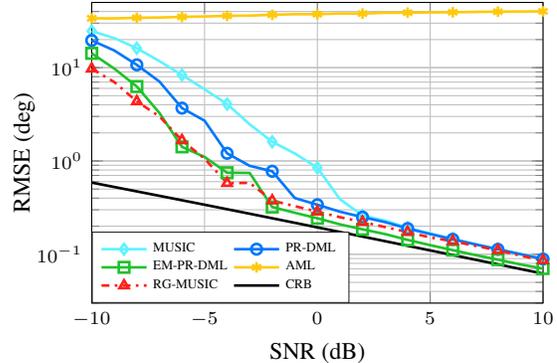


Fig. 1: RMSE vs SNR, overestimate the interference rank

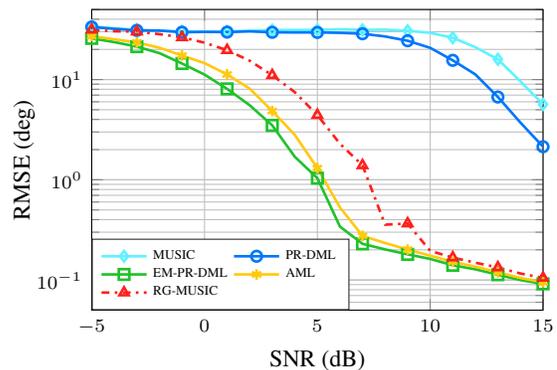


Fig. 2: RMSE vs SNR, model mismatch

$\mathbf{n}(t) \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$  in 90% of the snapshots. The remaining 10% of the snapshots contains outliers (specifically, outliers are generated as compound Gaussian noise with  $1/\tau(t) = 100$ ). We assume the number of interferers in our proposed EM-PR-DML estimator to be  $R' = 2$  to account for the model mismatch [24] of the noise when compared with that under the signal model in Section 2. We observe that the non-robust MUSIC and PR-DML estimator cannot resolve the DOAs of two sources in all Monte-Carlo trials. The proposed EM-PR-DML estimator outperforms RG-MUSIC and AML methods in both threshold and asymptotic regimes.

## 6. CONCLUSION AND OUTLOOK

In this paper, we propose a novel robust DOA estimator based on the EM algorithm and the PR framework. The proposed estimator jointly estimates the DOAs and the non-Gaussianity of the receive signal in each EM iteration. In order to maintain a low computational complexity, the PR framework is applied to substitute for the conventional deterministic ML. Simulation results show that the proposed DOA estimator possesses near-optimal performances under the assumed model without sacrificing robustness against model mismatch.

For future work, the theoretical explanation of the robustness of the EM-PR-DML estimator requires further comprehensive investigation. In addition, extensions to consider more realistic scenarios, e.g., decentralized DOA estimation among local subarrays or missing data, are of great interest.

## 7. 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, 2018.
- [2] M. Trinh-Hoang, M. Viberg, and M. Pesavento, "Cramér-Rao Bound for DOA Estimators Under the Partial Relaxation Framework: Derivation and Comparison," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3194–3208, 2020.
- [3] R.O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [4] J. Böhme, "Estimation of Source Parameters by Maximum Likelihood and Nonlinear Regression," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1984, vol. 9, pp. 271–274.
- [5] M. Viberg and B. Ottersten, "Sensor Array Processing Based on Subspace Fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1110–1121, 1991.
- [6] P. Stoica and A. Nehorai, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [7] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex Elliptically Symmetric Distributions: Survey, New Results and Applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, 2012.
- [8] R. B. Abdallah, A. Breloy, M.N. El Korso, and D. Lautru, "Bayesian Signal Subspace Estimation with Compound Gaussian Sources," *Signal Processing*, vol. 167, pp. 107310, 2020.
- [9] A. Leshem and A. J. van der Veen, "Radio-Astronomical Imaging in the Presence of Strong Radio Interference," *IEEE Transactions on Information Theory*, vol. 46, no. 5, pp. 1730–1747, 2000.
- [10] O. Besson, Y. Abramovich, and B. Johnson, "Direction-of-Arrival Estimation in a Mixture of K-distributed and Gaussian Noise," *Signal Processing*, vol. 128, pp. 512 – 520, 2016.
- [11] M. Rangaswamy, F. C. Lin, and K. R. Gerlach, "Robust Adaptive Signal Processing Methods for Heterogeneous Radar Clutter Scenarios," in *Proceedings of the 2003 IEEE Radar Conference*, 2003, pp. 265–272.
- [12] G. Ginolhac, P. Forster, F. Pascal, and J. Ovarlez, "Performance of Two Low-Rank STAP Filters in a Heterogeneous Noise," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 57–61, 2013.
- [13] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer, and P. Larzabal, "Robust Distributed Calibration of Radio Interferometers with Direction Dependent Distortions," *Signal Processing*, vol. 153, pp. 348 – 354, 2018.
- [14] P. J. Huber and E. M. Ronchetti, *Robust Statistics*, Wiley Series in Probability and Statistics. Wiley, 1981.
- [15] D. B. Williams and D. H. Johnson, "Robust Estimation of Structured Covariance Matrices," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2891–2906, 1993.
- [16] R. Couillet, "Robust Spiked Random Matrices and a Robust G-MUSIC Estimator," *Journal of Multivariate Analysis*, vol. 140, pp. 139 – 161, 2015.
- [17] A. Breloy, G. Ginolhac, F. Pascal, and P. Forster, "Robust Covariance Matrix Estimation in Heterogeneous Low Rank Context," *IEEE Transactions on Signal Processing*, vol. 64, no. 22, pp. 5794–5806, 2016.
- [18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [19] O. Besson and Y. I. Abramovich, "On the Fisher Information Matrix for Multivariate Elliptically Contoured Distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.
- [20] M. Pesavento and A. B. Gershman, "Maximum-Likelihood Direction-of-Arrival Estimation in the Presence of Unknown Nonuniform Noise," *IEEE Transactions on Signal Processing*, vol. 49, no. 7, pp. 1310–1324, 2001.
- [21] H. Alzer, "On Some Inequalities for the Gamma and Psi Functions," *Mathematics of Computation*, vol. 66, no. 217, pp. 373–389, 1997.
- [22] C. Liu and D. B. Rubin, "ML Estimation of the  $t$ -Distribution using EM and its extensions, ECM and ECME," *Statistica Sinica*, vol. 5, no. 1, pp. 19–39, 1995.
- [23] 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.
- [24] A. Mennad, S. Fortunati, M. N. El Korso, A. Younsi, A. M. Zoubir, and A. Renaux, "Slepian-Bangs-Type Formulas and the Related Misspecified Cramér-Rao Bounds for Complex Elliptically Symmetric distributions," *Signal Processing*, vol. 142, pp. 320 – 329, 2018.