



HAL
open science

Optimal prior knowledge-based direction of arrival estimation

Petter Wirfält, Guillaume Bouleux, Magnus Jansson, Petre Stoica

► **To cite this version:**

Petter Wirfält, Guillaume Bouleux, Magnus Jansson, Petre Stoica. Optimal prior knowledge-based direction of arrival estimation. IET Signal Processing, 2012, 6, pp.731 - 742. 10.1049/iet-spr.2011.0453 . hal-01555277

HAL Id: hal-01555277

<https://hal.science/hal-01555277>

Submitted on 6 Jul 2017

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.

Optimal Prior Knowledge-Based Direction of Arrival Estimation

Petter Wirfält^{1†}, Guillaume Bouleux², Magnus Jansson¹ and Petre Stoica³

¹ ACCESS Linnaeus Center, Electrical Engineering, Signal Processing,
KTH Royal Institute of Technology, Osquidas v.10, SE-100 44 Stockholm, Sweden

[†] Corresponding author — Email: wirfalt@kth.se; Fax: +46-8-790 72 60

² University of Lyon and University of Saint Etienne, LASPI, 20 Avenue de Paris, 42334 Roanne Cedex, France

³Information Technology Department, Division of Systems and Control,
Uppsala University, PO Box 337, SE-751 05 Uppsala, Sweden

Abstract—In certain applications involving direction of arrival (DOA) estimation we may have a priori information on some of the DOAs. This information could refer to a target known to be present at a certain position, or to a reflection. In this paper we investigate a methodology for array processing that exploits the information on the known DOAs for estimating the unknown DOAs as accurately as possible. We present algorithms that can efficiently handle the case of both correlated and uncorrelated sources when the receiver is a uniform linear array. We find a major improvement in estimator accuracy in feasible scenarios, and we compare the estimator performance to the corresponding theoretical stochastic Cramér-Rao Bounds (CRBs) as well as to the performance of other methods capable of exploiting such prior knowledge. In addition, we apply the investigated estimators to real data from an ultra-sound array.

Keywords—Antenna arrays; array signal processing; direction of arrival estimation; Cramer-Rao bound

1 INTRODUCTION

In some direction of arrival (DOA) estimation applications there exists a priori information that is not exploited in the traditional DOA finding algorithms such as [1], [2], and [3]. For example, in a stationary RADAR scenario there can be a reflection that is always received. The DOA associated with this reflection is of no interest to the observer, since it is already known. However its presence might degrade the estimation of the unknown DOAs of interest.

A methodologically related problem is the frequency estimation, for example the diagnosis of rotating machines in industrial environments. The rotational behavior of such machines is well-known and hence some peaks of the related frequency spectrum are known, e.g. the power grid supply frequency [4]. Estimating this supply frequency provides no information about the equipment; rather, the known spectral peaks might obscure weaker interesting peaks.

The important feature of this prior information is that it is hidden in the received data; regular algorithms will simply treat the known directions or frequencies as part of the set

of unknown ones. It would then be advantageous if prior knowledge on some DOAs or frequencies could be used in order to increase the accuracy when estimating the remaining unknown parameters.

For angle estimation this has been done in various ways. The MUSIC method [1], [5], is a so-called subspace method in which the subspaces spanned by the desired signals and the noise, respectively, are separated. One approach, then, is to use the prior information to decrease the dimension of the space spanned by the signals such that this space only contains the unknown signals. This dimension decrease, or deflation, of the signal subspace can be viewed geometrically as an orthogonal projection onto a lower dimensional subspace. This type of approach is suggested in, e.g., [6], [7], and [8]: the space spanned by the array data is projected onto the orthogonal complement of the subspace containing the known DOAs, thus removing the subspace spanned by the known directions from the array manifold. Then, it is shown that the unknown directions are more easily found from the reduced-dimension data set.

A similar approach is to use oblique projections, see [9] for an introduction to this concept. In this case, instead of projecting orthogonally onto a certain subspace, a projection along the subspace that is to be removed is performed. This results in a different solution as compared to projecting orthogonally since the projection is along a different path. The benefit of this approach is that the entire contribution of the concerned subspace is cancelled — in the orthogonal projection case, only the portion perpendicular to the desired subspace is cancelled. Thus, if the two subspaces are not orthogonal, the oblique projection can be superior to the orthogonal one. However, the oblique projector may increase the contribution of noise terms; see [9] for more details on this subject.

Comparisons of these two approaches applied to MUSIC can be found in [10], where some theoretical bounds are studied as well. By studying the Cramér-Rao Bound (CRB) for the case when prior information is available, it can be seen that the previously mentioned projection based methods are not performing optimally in the sense that they are not achieving the theoretically best achievable accuracy. This should not be surprising — in [11] and [12] the asymptotic properties of

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement #228044

MUSIC were studied, and it was concluded that MUSIC is not efficient in the case when there is correlation between the source signals. On the other hand in the case of uncorrelated sources, MUSIC is efficient in the sense that it asymptotically achieves the CRB. However, in [13] it was shown that a tighter CRB exists when the sources are known to be uncorrelated; from the perspective of asymptotic accuracy, MUSIC is thus suboptimal in this scenario as well. The conclusion is that we cannot achieve asymptotically efficient estimates with MUSIC whether the sources are correlated or not, if we properly account for the model-imposed limits on accuracy.

Motivated by the above observation, the focus of this paper is on incorporating prior DOA knowledge into methods that are asymptotically statistically efficient. We are studying a scenario with a Uniform Linear Array (ULA) at the receiver, i.e. a sensor array consisting of identical omnidirectional sensors with uniform inter-sensor separation. We investigate two ULA DOA methods ([2], [3]), which are asymptotically efficient in the case of correlated and uncorrelated impinging signals, respectively. Both of these methods find their estimates through polynomial rooting, and this is where the approach [14] denoted PLEDGE (as in Prior knowLEDGE) advocated in this paper can be used: we convert the knowledge of some DOAs into knowledge of the corresponding roots of the polynomial to be factored. Doing so allows us to refine the estimates of the other unknown roots, and hence of the unknown DOAs. An important fact is that we incorporate the known information in a statistically optimal way.

As will be shown, both DOA estimation methods show significant accuracy improvements when modified to utilize known DOAs: the signal-to-noise-ratio (SNR) required to achieve the theoretical performance limit is reduced, and in addition this limit is improved as predicted by the respective CRB. When the sources are uncorrelated, the exploitation of prior DOA information as in [14] does not enhance the method in [2] as much as when the sources are correlated; this motivates the application of PLEDGE to the method in [3], which is usable only when the sources are uncorrelated. Being able to exploit the uncorrelatedness of the signals in conjunction with prior DOA knowledge allows greater accuracy increases. The realizable gain was illustrated in [15] and is further investigated in this paper. To accurately judge the performance of the two PLEDGE-based methods we compare them to other methods capable of exploiting prior DOA knowledge, [7] and [10].

The paper is structured as follows. In Section 2 we define the scenario and in Section 3 we review the two DOA finding algorithms [2] and [3]. In Section 4, we discuss different methods of incorporating prior DOA-knowledge: the approach introduced in [14] and advocated in this paper, and two other methods ([7] and [10]) for comparison. In Section 5 we derive the CRBs for the studied scenarios, and in Section 6 we perform numerical simulations to evaluate the potential of the proposed methodology. Finally, we apply the estimators to real data in Section 7.

The PLEDGE-concept was previously introduced in [14], and it was examined for uncorrelated source signals in [15]; the contribution of this paper is a more thorough examination of the potentially significant accuracy gains that result from

using prior information, comparing the algorithms to existing methods, and applying them to real data. We also derive the corresponding accuracy bounds in a more compact manner.

2 DATA MODEL

Consider d narrow-band, plane waves sensed by an m sensor ULA. We assume that d_k of the impinging waves originate from a priori known directions; hence, define the angles of arrival as $\bar{\boldsymbol{\theta}} \triangleq [\boldsymbol{\theta}^T \ \boldsymbol{\vartheta}^T]^T \triangleq [\theta_1 \ \dots \ \theta_{d_u} \ \vartheta_1 \ \dots \ \vartheta_{d_k}]^T$, where θ_i denotes the unknown direction i and ϑ_j denotes the known direction j ; henceforth, the subscripts u and k will denote unknown and known quantities, respectively. The angular measurements are referenced to 0° at array broadside, T is the transposition operation, and $d_u = d - d_k$.

Assume that N array response snapshots are recorded, with $t = 1, \dots, N$, according to the data model

$$\mathbf{y}(t) = \mathbf{A}(\bar{\boldsymbol{\theta}})\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{A}(\bar{\boldsymbol{\theta}}) = [\mathbf{a}(\bar{\theta}_1), \dots, \mathbf{a}(\bar{\theta}_d)] \in \mathbb{C}^{m \times d}$ is the array response matrix and $\mathbf{a}(\bar{\theta}_i) \in \mathbb{C}^{m \times 1}$ defined as

$$\mathbf{a}(\bar{\theta}_i) = [1 \ e^{j\phi_i} \ \dots \ e^{j(m-1)\phi_i}]^T \quad (2)$$

is the steering vector corresponding to the i -th angle, viz.

$$\phi_i = -2\pi\Delta \sin(\bar{\theta}_i), \quad (3)$$

with Δ being the ULA inter-sensor spacing in wavelengths. Further, in (1), $\mathbf{s}(t) \in \mathbb{C}^{d \times 1}$ and $\mathbf{n}(t) \in \mathbb{C}^{m \times 1}$ are the source signal and additive noise vectors at time t , respectively, and they are assumed to be mutually independent zero-mean circular Gaussian random sequences with second-order moments given by

$$\mathbb{E}[\mathbf{s}(t)\mathbf{s}^*(\tau)] = \mathbf{P}\delta(t, \tau) \quad (4)$$

$$\mathbb{E}[\mathbf{n}(t)\mathbf{n}^*(\tau)] = \sigma^2\mathbf{I}\delta(t, \tau). \quad (5)$$

Here, $*$ denotes conjugate transposition and $\delta(t, \tau) = 1$ for $t = \tau$, and 0 otherwise. Without loss of generality, we can additionally partition $\mathbf{A}(\bar{\boldsymbol{\theta}}) \triangleq [\mathbf{A}_u(\boldsymbol{\theta}) \ \mathbf{A}_k(\boldsymbol{\vartheta})]$ with the subscripts added for clarity. Further, let

$$\mathbf{R} = \mathbb{E}[\mathbf{y}(t)\mathbf{y}^*(t)] \quad (6)$$

be the data covariance matrix, and define

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{t=1}^N \mathbf{y}(t)\mathbf{y}^*(t) \quad (7)$$

as a sample version thereof. The superscript $\hat{\cdot}$ denotes a quantity estimated from data.

3 DOA ESTIMATION

This section describes the two algorithms to which we apply the PLEDGE framework [14]. None of these methods can exploit the prior information $\boldsymbol{\vartheta}$; hence these algorithms will estimate all the elements of the vector $\boldsymbol{\theta}$. Both algorithms are Maximum Likelihood (ML) based methods that minimize a subspace projection, and they effectively reduce (in the ULA

case) the DOA finding problem to polynomial rooting. These two methods should be used when the signal correlation is unknown [2] and, respectively, when the sources are known to be uncorrelated [3]. In these cases the methods can be shown to be asymptotically statistically efficient (see [2], [3] and references therein).

Let

$$\hat{\mathbf{R}} = \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{E}}_n^* \quad (8)$$

be the eigendecomposition of (7), where $\hat{\mathbf{\Lambda}}_s$ is a diagonal matrix comprising the d' largest eigenvalues of $\hat{\mathbf{R}}$, with d' being the rank of \mathbf{P} , and $\hat{\mathbf{E}}_s$ comprises the associated eigenvectors. Similarly, the $m-d'$ smallest eigenvalues are collected in the diagonal matrix $\hat{\mathbf{\Lambda}}_n$ and $\hat{\mathbf{E}}_n$ contains the associated eigenvectors.

3.1 General signal correlation

Here there is no assumed structure for \mathbf{P} . To review the method of [2], in [16] the authors define a polynomial, based on ϕ_1, \dots, ϕ_d that correspond to the true DOAs, by

$$b_0 z^d + b_1 z^{d-1} + \dots + b_d = b_0 \prod_{i=1}^d (z - e^{j\phi_i}). \quad (9)$$

All the roots of (9) are by definition on the unit-circle, and thus the coefficients can be chosen to be conjugate symmetric: $b_i = b_{d-i}^*$, $i = 0, 1, \dots, d$ [16]. While the conjugate symmetry does not guarantee unit-circle roots, when it comes to estimating the polynomial coefficients from data it is much simpler to enforce this constraint than actually constraining the roots to the unit circle.

Using the polynomial coefficients, define

$$\mathbf{B}^* = \begin{bmatrix} b_d & \dots & b_1 & b_0 & 0 & \dots & 0 \\ 0 & b_d & \dots & b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & b_d & \dots & b_1 & b_0 \end{bmatrix} \quad (10)$$

with $\mathbf{B}^* \in \mathbb{C}^{(m-d) \times m}$. The MODE (Method of Direction Estimation) estimate is found by minimizing the function [2]

$$V_{\text{MODE}}(\mathbf{b}) = \text{Tr}[\mathbf{B} (\hat{\mathbf{B}}^* \hat{\mathbf{B}})^{-1} \mathbf{B}^* \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^*], \quad (11)$$

where $\hat{\mathbf{\Lambda}} = \hat{\mathbf{\Lambda}}_s^{-1} (\hat{\mathbf{\Lambda}}_s - \hat{\sigma}^2 \mathbf{I})^2$, $\hat{\sigma}^2 = \frac{1}{m-d'} \text{Tr}[\hat{\mathbf{\Lambda}}_n]$ is a consistent estimate of the noise power, and $\hat{\mathbf{B}}$ is constructed from an estimate of $\mathbf{b} = [b_0 \ b_1 \ \dots \ b_d]^T$ — typically, the algorithm is initialized with $\hat{\mathbf{B}}^* \hat{\mathbf{B}} = \mathbf{I}$, giving a coarse $\hat{\mathbf{b}}$, and then it is iterated one more time. Due to the special structure of \mathbf{B} it can be seen that (11) is a quadratic function of the polynomial coefficients \mathbf{b} . Thus we can equivalently write (11) as

$$V_{\text{MODE}}(\mathbf{b}) = \|\mathbf{H}\mathbf{b}\|^2, \quad (12)$$

where \mathbf{H} is estimated from the data and a function of $\hat{\mathbf{b}}$. The exact structure of \mathbf{H} is somewhat involved and not in the primary scope of this article, but can be readily obtained from [16] along with the modifications given by [2].

Minimizing (12) with respect to \mathbf{b} subject to a non-triviality constraint (i.e. $\mathbf{b} \neq \mathbf{0}$) and the conjugate symmetry constraint previously mentioned (which is a simple least squares or eigen-decomposition problem) provides an estimate of the polynomial coefficients \mathbf{b} . Then, the angles of the roots of the corresponding polynomial (9) gives estimates of (3) from which the DOAs θ are estimated.

3.2 Uncorrelated sources

If we have the a priori information that the sources are uncorrelated we can devise a more accurate estimator. In what follows, we give a brief summary of a statistically efficient algorithm originally developed in [3]. This algorithm, which we will call “DOA UC” (where UC stands for uncorrelated), is based on minimizing the covariance-matching criterion

$$V_{\text{UC}}(\boldsymbol{\eta}) = (\hat{\mathbf{r}} - \mathbf{r}(\boldsymbol{\eta}))^* \mathbf{W} (\hat{\mathbf{r}} - \mathbf{r}(\boldsymbol{\eta})). \quad (13)$$

In (13) $\mathbf{r} = \text{vec}(\mathbf{R})$, with $\text{vec}(\cdot)$ stacking the columns of a matrix on top of each other, and $\hat{\mathbf{r}}$ is the sample estimate thereof (7). Also $\boldsymbol{\eta} = [\bar{\boldsymbol{\theta}}^T \ \mathbf{p}^T \ \sigma^2]^T$ parameterizes \mathbf{r} , with \mathbf{p} denoting the diagonal elements of (4), and \mathbf{W} is a suitably chosen weighting matrix. Following [3] and [17], $\mathbf{W} = \hat{\mathbf{R}}^{-T} \otimes \hat{\mathbf{R}}^{-1}$ minimizes the asymptotic variance of the (asymptotically unbiased) estimate given by minimizing (13) (here, \otimes denotes the Kronecker product).

In [3] it is shown how to successively solve (13) for the parameters \mathbf{p} and σ^2 , producing closed form estimates $\hat{\mathbf{p}}$ and $\hat{\sigma}^2$. Introducing these estimates in (13), and using the same polynomial parametrization \mathbf{b} of the DOAs as in (9), (10), it can be shown that (13) can be written as

$$V_{\text{UC}}(\mathbf{b}) = \|\tilde{\mathbf{H}}\mathbf{b}\|^2, \quad (14)$$

where $\tilde{\mathbf{H}}$ is given in [3]. Similarly to \mathbf{H} in (12) $\tilde{\mathbf{H}}$ is a function of the data and the DOA and (14) can obviously — using some initialization for the DOA-dependent terms in $\tilde{\mathbf{H}}$ — be minimized, in the same manner as (12).

4 USING PRIOR KNOWLEDGE

4.1 PLEDGE

As already mentioned the approach of incorporating prior DOA information into the estimation is referred to as PLEDGE (Prior knowLEDGE) [14]. This approach utilizes the fact that the minimization variables of the criterion functions in (12) and (14) are polynomial coefficients. Knowing some of the DOAs is equivalent to knowing some of the roots of the polynomial in (9); we can then partition this polynomial in two factors: one fully known factor corresponding to the known roots, and one factor that is unknown but has fewer unknown terms than the original polynomial. Note that no information is lost in this way: we are only constraining the polynomial to have some of its roots at certain locations, defined by the known DOAs. The estimates of the remaining unknown roots will then presumably be more accurate.

Formulating this mathematically we re-write (9) as

$$b_0 \prod_{i=1}^d (z - e^{j\phi_i}) = P_k(z)P_u(z), \quad (15)$$

in which

$$P_u(z) = \tilde{b}_0 \prod_{i=1}^{d_u} (z - e^{j\phi_i}) = \tilde{b}_0 z^{d_u} + \dots + \tilde{b}_{d_u} \quad (16)$$

and

$$P_k(z) = c_0 \prod_{i=d_u+1}^d (z - e^{j\phi_i}) = c_0 z^{d_k} + \dots + c_{d_k}. \quad (17)$$

Here, $P_k(z)$ is the polynomial with d_k zeros corresponding to the known DOAs ϑ whereas $P_u(z)$ has $d_u = d - d_k$ zeros corresponding to the unknown DOAs θ . Note that the coefficients $\mathbf{c} = [c_0 \dots c_{d_k}]^T$ are given by ϑ via (3); thus the prior DOA knowledge defines \mathbf{c} . If (15) is expanded, it is obvious that the polynomial coefficients \mathbf{b} can be written as a convolution of $\tilde{\mathbf{b}} = [\tilde{b}_0 \tilde{b}_1 \dots \tilde{b}_{d_u}]^T$ and \mathbf{c} . Thus we can rewrite (15)-(17) in matrix form as

$$\mathbf{b} = \mathbf{C}\tilde{\mathbf{b}}, \quad (18)$$

if we introduce the (Toeplitz structured) convolution matrix

$$\mathbf{C}^T = \begin{bmatrix} c_0 & c_1 & \dots & c_{d_k} & 0 & \dots & 0 \\ 0 & c_0 & c_1 & \dots & c_{d_k} & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_0 & c_1 & \dots & c_{d_k} \end{bmatrix} \in \mathbb{C}^{(d_u+1) \times (d+1)}, \quad (19)$$

the elements of which are given by \mathbf{c} .

4.2 PLEDGE: General case

Substituting (18) in (12), we obtain the PLEDGE criterion function

$$V_P(\tilde{\mathbf{b}}) = \|\mathbf{H}\mathbf{C}\tilde{\mathbf{b}}\|^2 \quad (20)$$

(hereafter we let the subscript P stand for ‘‘PLEDGE’’).

The method of finding the unknown DOAs consists of the following steps:

- 1) Form the \mathbf{C} matrix from the known DOAs.
- 2) Perform an eigendecomposition of $\hat{\mathbf{R}}$ and, using $(\hat{\mathbf{B}}^* \hat{\mathbf{B}})^{-1} = \mathbf{I}$ as an initial estimate, form \mathbf{H} as in [16] and [2].
- 3) Find an estimate of $\tilde{\mathbf{b}}$ by minimizing (20) subject to $\|\mathbf{b}\| = 1$ and the conjugate symmetry constraint. Use this estimate in (18) to form $(\hat{\mathbf{B}}^* \hat{\mathbf{B}})^{-1}$, and update \mathbf{H} .
- 4) Minimize (20) once again subject to the same constraints to find $\tilde{\mathbf{b}}$, then find ϕ_i , $i = 1, \dots, d_u$, by rooting (16), and finally the sought DOAs through (3).

4.3 PLEDGE: Uncorrelated sources

Substituting (18) in (14), we obtain the following PLEDGE criterion function

$$V_{PUC} = \|\tilde{\mathbf{H}}\mathbf{C}\tilde{\mathbf{b}}\|^2 \quad (21)$$

(the subscript PUC denotes ‘‘PLEDGE uncorrelated’’). This criterion is minimized in the same way as for (20), that is:

- 1) Form \mathbf{C} in (19) from the known DOAs
- 2) Form an initial DOA-independent $\tilde{\mathbf{H}}$ based on the sample data according to [3].
- 3) Find an initial estimate of $\tilde{\mathbf{b}}$ by minimizing (21) subject to $\|\mathbf{b}\| = 1$ and the conjugate symmetry constraint; use this initial DOA estimate to find a more accurate $\tilde{\mathbf{H}}$.
- 4) Minimize (21), once again, subject to the same constraints to find $\tilde{\mathbf{b}}$, then obtain ϕ_i , $i = 1, \dots, d_u$, by rooting (16), and finally the sought DOAs through (3).

4.4 Other methods

For comparison purposes, we also consider two prior-based methods that are not using the PLEDGE concept. These are C-MUSIC (as in Constrained MUSIC, see [7]), which is a method based on an orthogonal projection. We also look at a method based on an oblique projection; the method was introduced in [10] as P-MUSIC (short for Prior MUSIC), which was modified in [14] to remove an observed bias. We denote this method MP-MUSIC (Modified P-MUSIC). These MUSIC-based algorithms find the respective DOAs by a rooting approach, along the lines of root-MUSIC [18].

4.4.1 C-MUSIC: This algorithm is described in full in [7]. The main idea is to use an orthogonal projection that removes the known part of the signal subspace; this is done by pre- and post-multiplying the sample covariance matrix with a projection matrix. Accordingly, we apply a regular MUSIC-algorithm to

$$\hat{\mathbf{R}}_c = \Pi_{\mathbf{A}_k}^\perp \hat{\mathbf{R}} \Pi_{\mathbf{A}_k} \quad (22)$$

where $\Pi_{\mathbf{X}}^\perp = \mathbf{I} - \Pi_{\mathbf{X}}$ denotes the orthogonal projection onto the null-space of \mathbf{X}^* , $\Pi_{\mathbf{X}} = \mathbf{X}\mathbf{X}^\dagger = \mathbf{X}(\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*$ is the orthogonal projection onto the range space of \mathbf{X} , and \mathbf{A}_k is composed of the steering vectors corresponding to the known DOAs. The contribution of the known sources to $\hat{\mathbf{R}}$ has thus been cancelled, whereas the eigenvectors of $\hat{\mathbf{R}}_c$ still contain information on the unknown sources; see [7] for further details.

4.4.2 MP-MUSIC: In [10] a criterion function is proposed based on an oblique projection; it is then shown that this criterion condenses to a minimization only involving an orthogonal projection. It can be inferred from [10] that the resulting algorithm suffers from a bias, rendering it worse than other methods, e.g., C-MUSIC. Thus a corrective measure was suggested in [14], and here we denote the resulting algorithm MP-MUSIC; we below briefly investigate the consequences of this correction, which takes the form of removing the noise from the sample covariance matrix. Thus let

$$\mathbf{R} - \sigma^2 \mathbf{I} = \mathbf{A}\mathbf{P}\mathbf{A}^* = [\mathbf{A}_u \quad \mathbf{A}_k] \begin{bmatrix} \mathbf{P}_{u^*} & \mathbf{P}_{uk} \\ \mathbf{P}_{uk}^* & \mathbf{P}_k \end{bmatrix} \begin{bmatrix} \mathbf{A}_u^* \\ \mathbf{A}_k^* \end{bmatrix} \quad (23)$$

in which \mathbf{A}_u contains the steering vectors corresponding to the unknown sources, \mathbf{P}_u and \mathbf{P}_k are the square signal covariance matrices related to the unknown and known emitters, respectively, and \mathbf{P}_{uk} is the cross-correlation matrix of the signals for the unknown and known emitters.

The obliquely projected criterion in [10] condensed to pre-multiplying the (now modified) sample covariance matrix with the same orthogonal projector as used in Section 4.4.1; hence, look at

$$\Pi_{\mathbf{A}_k}^\perp (\mathbf{R} - \sigma^2 \mathbf{I}) = \Pi_{\mathbf{A}_k}^\perp \mathbf{A}_u \mathbf{P}_u \mathbf{A}_u^* + \Pi_{\mathbf{A}_k}^\perp \mathbf{A}_u \mathbf{P}_{uk} \mathbf{A}_k^*. \quad (24)$$

If in addition the known and unknown signals are uncorrelated to one another, i.e. $\mathbf{P}_{uk} = \mathbf{0}$, and we use estimated quantities in (24) then:

$$\Pi_{\mathbf{A}_k}^\perp (\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I}) = \Pi_{\mathbf{A}_k}^\perp \mathbf{A}_u \mathbf{P}_u \mathbf{A}_u^* + \mathbf{E}_\epsilon \quad (25)$$

where \mathbf{E}_ϵ accounts for the error introduced by using estimated quantities. Using the right singular-vectors of (25) in a standard MUSIC algorithm then gives estimates of the unknown DOAs.

5 PLEDGE CRB

We now derive the CRB based on the prior-knowledge of some of the DOAs. We derive it both for the general case (which has been previously considered in [14]) and for the case when the sources are known to be uncorrelated.

The derivation starts out by following [19]. Let

$$\boldsymbol{\alpha} = [\boldsymbol{\theta}^\top \quad \boldsymbol{\varrho}^\top \quad \sigma^2]^\top \quad (26)$$

be the vector of unknown parameters in the model. Here, $\boldsymbol{\theta}$ is again the vector of unknown DOAs, $\boldsymbol{\varrho}$ is the vector made from $\{\mathbf{P}_{ii}\}$ and $\{\text{Re}(\mathbf{P}_{ij}), \text{Im}(\mathbf{P}_{ij}); i > j\}$, and σ^2 is the noise variance. Under the assumptions in Section 2, the (k, l) th element of the Fisher Information Matrix for the parameter vector $\boldsymbol{\alpha}$ is given by

$$\text{FIM}_{k,l} = N \text{Tr} \left(\frac{\partial \mathbf{R}}{\partial \alpha_k} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \alpha_l} \mathbf{R}^{-1} \right). \quad (27)$$

This can be written in matrix form as

$$\frac{1}{N} \text{FIM} = \left(\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}^\top} \right)^* (\mathbf{R}^{-\text{T}} \otimes \mathbf{R}^{-1}) \left(\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}^\top} \right), \quad (28)$$

where

$$\mathbf{r} = \text{vec}(\mathbf{R}) = (\mathbf{A}^c \otimes \mathbf{A}) \text{vec}(\mathbf{P}) + \sigma^2 \text{vec}(\mathbf{I}), \quad (29)$$

and where the superscript c denotes complex conjugation. Introduce \mathbf{G} , \mathbf{V} , and \mathbf{u} via

$$\mathbf{W}^{1/2} \begin{bmatrix} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\theta}^\top} & \left| & \frac{\partial \mathbf{r}}{\partial \boldsymbol{\varrho}^\top} & \left| & \frac{\partial \mathbf{r}}{\partial \sigma^2} \right. \\ \hline \mathbf{G} & \left| & \mathbf{V} & \left| & \mathbf{u} \right. \end{bmatrix}, \quad (30)$$

where $\mathbf{W}^{1/2} = \mathbf{R}^{-\text{T}/2} \otimes \mathbf{R}^{-1/2}$, and also let

$$[\mathbf{V} \mid \mathbf{u}] \triangleq \Delta. \quad (31)$$

Then we can write (28) as

$$\frac{1}{N} \text{FIM} = \begin{bmatrix} \mathbf{G}^* \\ \Delta^* \end{bmatrix} [\mathbf{G} \quad \Delta]. \quad (32)$$

Since we are interested in a bound on the angle estimates, we want an expression for the top-left block of the inverse of the partitioned matrix in (32). Using a standard result on partitioned matrix inversion gives

$$\begin{aligned} \frac{1}{N} \text{CRB}_{\mathbf{P}}^{-1}(\boldsymbol{\theta}) &= \mathbf{G}^* \mathbf{G} - \mathbf{G}^* \Delta (\Delta^* \Delta)^{-1} \Delta^* \mathbf{G} \\ &= \mathbf{G}^* \Pi_{\Delta}^\perp \mathbf{G}. \end{aligned} \quad (33)$$

5.1 Correlated sources

Since only some of the DOAs are considered unknown, the \mathbf{G} in the present derivation differs from the one in [19]. This difference propagates to the final result and managing it is in essence the contribution of this section.

From (33), following [19], it can be shown that

$$\Pi_{\Delta}^\perp = \Pi_{\mathbf{V}}^\perp - \frac{\Pi_{\mathbf{V}}^\perp \mathbf{u} \mathbf{u}^* \Pi_{\mathbf{V}}^\perp}{\mathbf{u}^* \Pi_{\mathbf{V}}^\perp \mathbf{u}}. \quad (34)$$

By evaluating the derivatives in (30), and through some further algebra, one finds

$$\mathbf{u}^* \Pi_{\mathbf{V}}^\perp \mathbf{g}_k = \mathbf{0}, \quad (35)$$

where \mathbf{g}_k is the k th column of \mathbf{G} . This result allows us to rewrite the individual elements of (33) as

$$\frac{1}{N} [\text{CRB}_{\mathbf{P}}^{-1}(\boldsymbol{\theta})]_{k,l} = \mathbf{g}_k^* \Pi_{\mathbf{V}}^\perp \mathbf{g}_l. \quad (36)$$

By further calculations we arrive at

$$\begin{aligned} \frac{1}{N} [\text{CRB}_{\mathbf{P}}^{-1}(\boldsymbol{\theta})]_{k,l} \\ = \frac{2}{\sigma^2} \text{Re} \left((\mathbf{D}_u^* \Pi_{\mathbf{A}}^\perp \mathbf{D}_u)_{k,l} (\bar{\mathbf{P}}_u^* \mathbf{A}^* \mathbf{R}^{-1} \mathbf{A} \bar{\mathbf{P}}_u)_{l,k} \right), \end{aligned} \quad (37)$$

where

$$\mathbf{D}_u = \begin{bmatrix} \frac{\partial \mathbf{a}(\theta_1)}{\partial \theta_1} & \dots & \frac{\partial \mathbf{a}(\theta_{d_u})}{\partial \theta_{d_u}} \end{bmatrix}, \quad (38)$$

and $\bar{\mathbf{P}}_u$ is the $d \times d_u$ sub-matrix of (4) associated with the unknown sources.

Finally, we can write (37) in matrix form as

$$\begin{aligned} \text{CRB}_{\mathbf{P}}(\boldsymbol{\theta}) \\ = \frac{\sigma^2}{2N} \left[\text{Re} \left((\mathbf{D}_u^* \Pi_{\mathbf{A}}^\perp \mathbf{D}_u) \odot (\bar{\mathbf{P}}_u^* \mathbf{A}^* \mathbf{R}^{-1} \mathbf{A} \bar{\mathbf{P}}_u)^\text{T} \right) \right]^{-1} \end{aligned} \quad (39)$$

where \odot is the Schur (element-wise) product.

5.2 Uncorrelated sources

Assuming the sources to be uncorrelated leads to a diagonal \mathbf{P} , which reduces the number of unknown parameters. We then have that $\boldsymbol{\varrho}$ in (26) is equal to the diagonal entries \mathbf{p} of (4), and the vector of unknown parameters becomes

$$\boldsymbol{\alpha}_{\text{UC}} = [\boldsymbol{\theta}^\top \quad \mathbf{p}^\top \quad \sigma^2]^\top. \quad (40)$$

In this case the derivatives of (30) corresponding to α_{UC} are given by:

$$\begin{aligned}\mathbf{G}_{UC} &= \mathbf{W}^{1/2} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\theta}^T} \\ &= \mathbf{W}^{1/2} [(\mathbf{D}_u^c \circ \mathbf{A}_u) + (\mathbf{A}_u^c \circ \mathbf{D}_u)] \mathbf{P}_u \\ &\triangleq \mathbf{W}^{1/2} \mathbf{D}_P,\end{aligned}\quad (41)$$

$$\begin{aligned}\boldsymbol{\Delta}_{UC} &= \mathbf{W}^{1/2} \left[\frac{\partial \mathbf{r}}{\partial \mathbf{p}^T} \quad \middle| \quad \frac{\partial \mathbf{r}}{\partial \sigma^2} \right] \\ &= \mathbf{W}^{1/2} [\mathbf{A}^c \circ \mathbf{A} \quad | \quad \text{vec}(\mathbf{I})] \\ &\triangleq \mathbf{W}^{1/2} \mathbf{D}_n,\end{aligned}\quad (42)$$

where \circ is the Khatri-Rao product (column-wise Kronecker product), and \mathbf{P}_u is the square matrix with the unknown source powers on the diagonal. Using these results in (33) gives

$$\text{CRB}_{PUC}(\boldsymbol{\theta}) = \frac{1}{N} \left[\mathbf{D}_P^* \mathbf{W}^{1/2} \boldsymbol{\Pi}_{\mathbf{W}^{1/2} \mathbf{D}_n}^\perp \mathbf{W}^{1/2} \mathbf{D}_P \right]^{-1}. \quad (43)$$

5.3 Comparison to no-prior bounds

Due to the model structures being nested, we immediately have, for sources of arbitrary correlation, that

$$\text{CRB}_{SML} \geq \text{CRB}_P, \quad (44)$$

where the notation \geq implies that the difference $\text{CRB}_{SML} - \text{CRB}_P$ is positive semi-definite. Here, CRB_{SML} denotes the bound when there is no prior information on the DOAs [19].

Similarly, if the sources are known to be uncorrelated, the relation

$$\text{CRB}_{SML} \geq \text{CRB}_{UC} \geq \text{CRB}_{PUC} \quad (45)$$

is true, where CRB_{UC} denotes the bound without prior DOA knowledge, but when the sources are known to be uncorrelated [13]. These relations follow from the general theory behind the CRB by noting the following. The unknown parameters for the tighter bounds represent a strict subset of the parameter vector for the wider bound; each successive reduction of the parameter set must produce a bound which is smaller than or equal to a bound obtained for the entire parameter vector. A more quantitative comparison of the different CRBs is quite involved and beyond the scope of this paper.

6 SIMULATED DATA EXAMPLES

Reproducible research: to reproduce the results of this simulation section, please see the first author's homepage: <http://www.ee.kth.se/~wirfalt/>. We perform Monte Carlo simulations to evaluate the performance of the studied estimators and compare the results to the standard MODE algorithm [2] and the DOA UC [3]. As previously stated, the latter methods are asymptotically optimal in the case of no prior DOA knowledge. For reference, we also include the prior-based methods mentioned in Section 4.4, from [10] and [7].

In each realization, N simulated snapshots of the data model (1) are used to form the sample covariance matrix $\hat{\mathbf{R}}$ in (7). For

each snapshot, $\mathbf{s}(t)$ is a realization of a zero-mean temporally-white complex-Gaussian noise process with spatial covariance matrix \mathbf{P} . Likewise, $\mathbf{n}(t)$ is a realization of a similarly defined process but with spatial covariance matrix $\sigma^2 \mathbf{I}$.

In the simulated scenarios, we vary the parameters of the model (1); we vary the number of sources d and their locations $\boldsymbol{\theta}$. We also choose \mathbf{P} as to study different properties of the estimators, e.g. SNR and correlation dependence. We restrict ourselves to studying DOAs that are closely spaced since this is the more difficult estimation scenario. In the case of widely spaced DOAs, traditional methods already produce accurate estimates; indeed, as evidenced in e.g. Fig. 3(a), when the DOA separation grows large the prior information does not significantly enhance the accuracy of the estimators.

The estimated DOAs $\hat{\boldsymbol{\theta}}$ are compared to the true values $\boldsymbol{\theta}$, and for each Monte Carlo simulation the squared error is recorded. The performance measure in our simulations is the root-mean-square error, which for the unknown angle θ_i is computed as

$$\text{RMSE}_i = \sqrt{\frac{1}{L} \sum_{k=1}^L (\hat{\theta}_{i,k} - \theta_i)^2}, \quad i = 1, \dots, d_u, \quad (46)$$

where $\hat{\theta}_{i,k}$ denotes the k th estimate of the i th angle, and $L = 1000$ is the number of Monte Carlo realizations.

Since some of the algorithms are not designed to exploit any prior information, they will inherently attempt to estimate all d DOAs $\boldsymbol{\theta}$, including the known ones. Thus, when calculating the RMSEs of $\hat{\boldsymbol{\theta}}$ for those algorithms, the d_k estimates closest to $\boldsymbol{\vartheta}$ are associated with these known angles, and the remaining d_u angles are compared to the unknown angles $\boldsymbol{\theta}$.

Unless mentioned otherwise, we let all source signals have the same power (i.e., the diagonal elements of \mathbf{P} are the same); this is to simplify the scenarios such that the amount of parameters that are changed are kept limited. In most of the simulations below, the source SNR is given by the abscissa.

6.1 Uncorrelated sources

In the first simulation we investigate a scenario with closely spaced, uncorrelated sources located at $\bar{\boldsymbol{\theta}} = [10^\circ \ 15^\circ \ 12^\circ \ 25^\circ]^T \triangleq [\theta_1 \ \theta_2 \ \vartheta_1 \ \vartheta_2]^T$; $\boldsymbol{\vartheta} = [12^\circ \ 25^\circ]^T$ is assumed to be known a priori. The number of sensors is $m = 10$, and the number of snapshots $N = 1000$.

We compare the accuracy of all the estimators presented in the article: "MODE" from Section 3.1 and [2]; "MODE PLEDGE" as in Section 4.2 and [14]; "C-MUSIC" from Section 4.4.1 and [7]; "DOA UC" from Section 3.2 and [3]; "PLEDGE UC" as described in Section 4.3 (see also [15]); and "MP-MUSIC" as outlined in Section 4.4.2, [10], and [14]. In addition to the above methods, we also show the bounds corresponding to different amounts of prior knowledge: "CRB_{SML}" [19] corresponding to no prior knowledge; "CRB_P" which assumes some source DOAs to be known, see Section 5.1 and [14]; "CRB_{UC}" corresponding to the knowledge of sources being uncorrelated from [13]; and finally "CRB_{PUC}" which assumes knowledge of all sources

being uncorrelated and some DOAs known, given by (43) in Section 5.2.

For uncorrelated sources, we have

$$\mathbf{P} = \text{diag}(p_1, \dots, p_d), \quad (47)$$

where $\text{diag}(\cdot)$ denotes the diagonal matrix with the arguments on the diagonal. For simplicity, $p_i = p_j$, $i, j = 1, \dots, d$, corresponding to all sources being equipowered; we also define $\text{SNR}_i = \frac{p_i}{\sigma^2}$.

The benefits of using all the available prior information are clearly illustrated in Fig. 1(a) — for a given accuracy, the required SNR is reduced by roughly 15 dB when comparing PLEDGE UC to DOA UC, and by 20 dB when comparing to MODE.

Note from the same figure that MP-MUSIC gives poor estimates; in particular, the RMSEs of its estimates are not decreasing with increasing SNR. The method exhibits such a saturation behavior in all scenarios. This may be due to the fact that the method relies heavily on the assumption that $\mathbf{P}_{u_k} = 0$ (see (24)-(25)). This may well hold in theory but not in finite samples and in such cases the right singular vectors of (25) are perturbed by contributions from \mathbf{A}_k (cf. the last term in (24)). In other scenarios (not shown), the threshold at which the RMSE saturates is quite low, and the method can then be quite accurate. However, due to the behavior shown in Fig. 1(a) we will not consider this method any further.

MODE PLEDGE, which does not exploit the prior information about the correlation state, exhibits a slightly lower resolution threshold than MODE. C-MUSIC is somewhat less accurate than MODE PLEDGE.

The method DOA UC, which uses the prior information that the sources are uncorrelated, behaves very similarly to MODE, except it reaches the uncorrelated, lower, CRB_{UC} ; we thus see the benefits of utilizing the correlation state information in the estimator.

At high SNRs, the benefit of using the prior knowledge diminishes as is visible in the figure; when the SNR grows very large (not shown) all bounds converge to the same variance, and hence the estimators are equally accurate. Note however that, e.g., CRB_{PUC} always has a negative slope — a higher unknown-source SNR is always beneficial.

In Fig. 1(b) we repeat the simulation, but this time fix the SNR (for all sources) at 25 dB and vary the number of snapshots, N . PLEDGE UC is markedly superior to the other methods and can, as compared to MODE, decrease the number of samples required for a specific accuracy by roughly a factor of 50. Note that the simulated scenario is difficult for MODE: it requires about $N = 5000$ samples in order to be efficient. In addition, we see that C-MUSIC reaches the performance bound for slightly lower values of N than MODE PLEDGE.

6.2 Unknown source correlation

In the second case we explore a situation with coherence between the sources; accordingly, the spatial covariance matrix

of the emitted signals is

$$\mathbf{P} = \begin{bmatrix} p_1 & p_{12} & p_{13} \\ p_{12}^c & p_2 & p_{23} \\ p_{13}^c & p_{23}^c & p_3 \end{bmatrix} \quad (48)$$

where $p_{ij} = \rho_{ij}\sqrt{p_i p_j}$, with ρ_{ij} being the correlation coefficient between the sources i and j ($|\rho_{ij}| \leq 1$). The sources are now located at $\bar{\boldsymbol{\theta}} = [10^\circ \ 15^\circ \ 12^\circ]^\text{T}$ with $\vartheta = 12^\circ$ assumed to be known a priori. In (48), the entries are ordered as in $\bar{\boldsymbol{\theta}}$; hence, for example, p_{12} corresponds to the correlation between the sources at 10 and 15 degrees. The correlation coefficient ρ_{13} in Fig. 2(a) and Fig. 2(b) is chosen to maximize the difficulty of this specific simulation scenario in the sense that the CRB_{SML} is large for every angle. The correlation is assumed unknown; we thus omit the estimators explicitly assuming the sources to be uncorrelated from the analysis of this scenario, along with the corresponding bounds. For simplicity we again assume the sources to be equipowered, giving $p_1 = p_2 = p_3$. In the interest of brevity, in Fig. 2 we only present the results for θ_1 ; the results for θ_2 are similar.

It can be seen from Fig. 2 that MODE PLEDGE outperforms the other methods, and that MODE PLEDGE is significantly more accurate than MODE in both cases. This can be contrasted to Fig. 1(a), where the gain was not so pronounced. Hence it can be inferred that the major benefit of MODE PLEDGE is the removal of the correlation between known and unknown sources, rather than that the estimator has one less parameter to estimate.

It appears as if the correlation between the unknown sources and the known source has been cancelled through the exploitation of prior knowledge in C-MUSIC as well — in Fig. 2(a) C-MUSIC can accurately estimate the unknown source(s) even though $|\rho_{13}| = 1$. On the other hand, C-MUSIC completely fails in Fig. 2(b) as in that case there exists source correlation even after nullifying the known source — this is in agreement with the general behavior of the MUSIC-algorithm, which is well-known to fail when the sources are coherent. The fact that the correlation has indeed been cancelled can be realized by comparing (24) and (22); the parts of \mathbf{P} related to the known sources, including potential correlation with the unknown sources, vanish due to the orthogonal projectors operating on both sides of $\hat{\mathbf{R}}$.

As expected (and desired) MODE PLEDGE can handle both cases optimally. C-MUSIC can be seen in Fig. 2(a) to suffer from a higher threshold SNR than MODE PLEDGE; however, it passes quickly through the transition region to an asymptotic accuracy equivalent to that of MODE PLEDGE for larger SNR values. Additionally, the large decrease in CRB_{P} as compared to CRB_{SML} can be traced to the correlation cancellation; the estimates using prior knowledge are not perturbed by the coherence between the sources at 10 and 12 degrees.

Note also that the theoretical performance bound is attained by MODE PLEDGE already at low SNRs — however, the CRB at the lowest SNRs is too large to make the estimates meaningful (presumably due to the closely spaced sources).

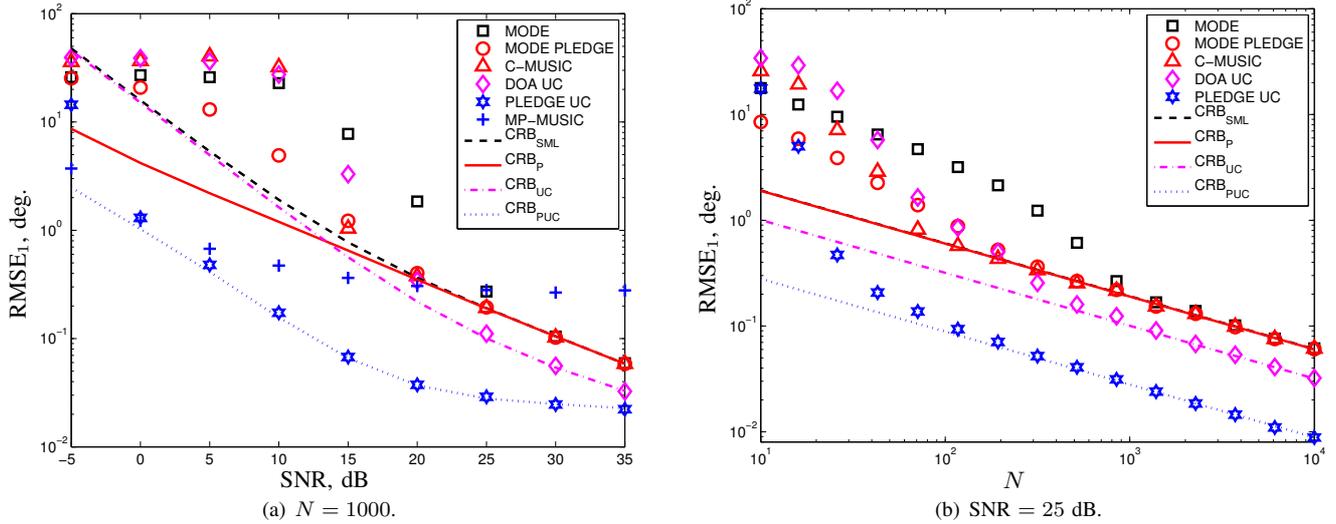


Fig. 1. Uncorrelated sources: The plots show the RMSE (of $\hat{\theta}_1$) for the methods presented in the article, along with the CRBs for the different scenarios. $\theta = [10^\circ \ 15^\circ]^T$, $\vartheta = [12^\circ \ 25^\circ]^T$. Equipowered sources, $m = 10$, and averages of 1000 independent Monte Carlo realizations.

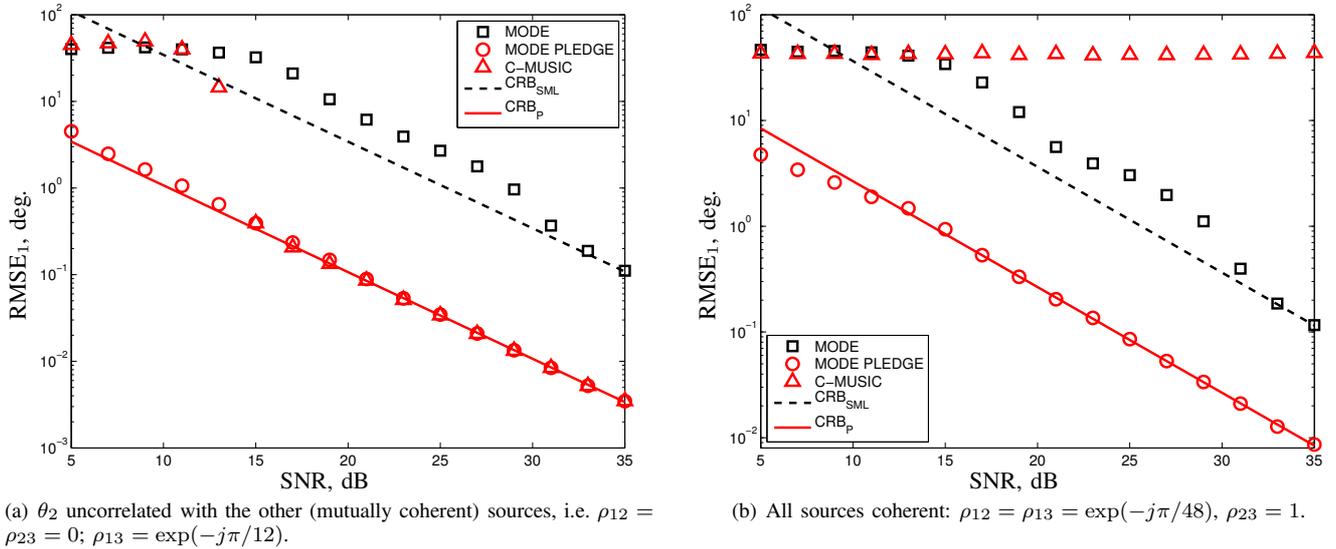


Fig. 2. Unknown source correlation: sources located at $\bar{\theta} = [10^\circ \ 15^\circ \ 12^\circ]^T \triangleq [\theta_1 \ \theta_2 \ \vartheta]^T$. RMSE of $\hat{\theta}_1$. Number of snapshots $N = 1000$, $m = 6$, equipowered sources, 1000 Monte Carlo realizations. Non-zero correlation coefficients are chosen to maximize the difficulty of the estimation scenario.

6.3 Error in prior knowledge

Here we study the case when there is an error in the prior DOA knowledge. To circumvent some of the issues of DOA association in this case we consider only two sources located at $\theta = 10^\circ$ and $\vartheta = 12^\circ$, or $\vartheta = 11^\circ$ depending on the correlation state — when the sources are coherent (with $\rho = \exp(-j\pi/8)$), we increase the separation to simplify the scenario. In Fig. 3 we vary the assumed location of ϑ and study the impact on performance. As before, $m = 6$, SNR= 25 dB, and $N = 1000$.

Even though there are errors in the prior knowledge, the

PLEDGE methods can still give enhanced estimates as long as the error is small when compared to the standard RMSE; this is seen in both Fig. 3(a) and Fig. 3(b). Thus, since the studied methods are asymptotically efficient, it is possible to judge whether it is advantageous to utilize a prior DOA-knowledge, with a given accuracy, based on the CRB of the source to be estimated. A prior information with a smaller variance than the estimate's CRB gives a strict statistical performance increase (in the sense of a smaller variance) in the estimated DOA. As can be seen from e.g. Fig. 3(a), in some cases errors larger than this bound can also translate into an increased estimator

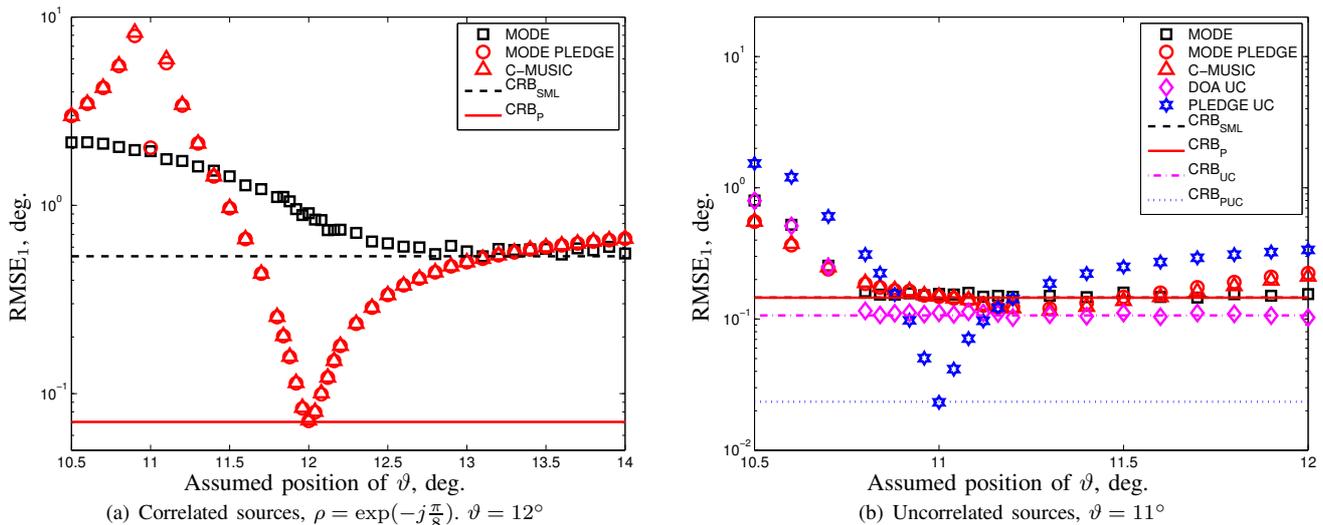


Fig. 3. Error in prior knowledge: RMSE of the estimates of $\theta = 10^\circ$ when varying the assumed known location as shown. Equipowered sources with $\text{SNR} = 25$ dB, $m = 6$, $N = 1000$ and 1000 Monte Carlo realizations.

accuracy.

Note that in Fig. 3(a) MODE and in Fig. 3(b) DOA UC suffer when the known DOA is believed to be closer to the unknown one than it actually is — this is due to the angle association as explained in the fifth paragraph of Section 6.

6.4 Varying known-source properties

In this scenario we vary the separation between the known source and the unknown one, and also the known source power.

In Fig. 4(a) we vary the known DOA and, as can be seen, PLEDGE UC can offer dramatic accuracy increases. This method can correctly estimate and resolve the two sources in the studied case for very small separations.

In Fig. 4(b), we vary the power of the known source. When the known source is significantly weaker than the unknown one, the methods not exploiting prior angular information have problems: they easily find the unknown source, but may associate it with the known one. The prior-based methods however do not suffer from this dramatic loss of accuracy. Once again, PLEDGE UC is significantly outperforming the other methods, and attains the accuracy bound for when there is only one source present (which we denoted CRB_{1S} in Fig. 4(b)). It can thus be seen that PLEDGE UC correctly accounts for the influence the known source has on the estimation problem; if the known source is too weak to affect the estimation of the unknown source, PLEDGE UC has in effect reduced the number of sources in the received data to only contain the unknown one. This is obviously a very desirable property for a prior-using method.

6.5 A note on the computational complexity of the algorithms

The algorithms investigated all showed execution times on the order of milliseconds on standard PCs. See Table I where we compare the mean time each algorithm needed to estimate

the unknown parameters in the scenario shown in Fig. 1(a). The methods are implemented in MATLAB to conceptually agree with their source publications; this has the effect that the implementations are not optimized for computational efficiency. Note that the algorithms scale differently with the scenario-specific parameters (i.e., m , d , etc). To illustrate this effect we also include simulation times for when $m = 20$ in Table I.

TABLE I. MEAN EXECUTION TIMES, MILLISECONDS.

	$m = 10$	$m = 20$
MODE:	0.8	2.2
MODE PLEDGE:	1.1	2.4
DOA UC:	2.1	8.5
PLEDGE UC:	2.2	8.6
C-MUSIC:	1.0	8.1
MP-MUSIC:	0.5	4.3

In Table I it can be seen that the PLEDGE methodology introduces a small computational burden; using PLEDGE in the MODE algorithm increased the execution time by about 0.3 ms or 40%, whereas for $m = 20$ the increase was about the same, 0.2 ms, which in this case translates to about 10%. The extra computational time required when using PLEDGE in DOA UC was 0.1 ms in both cases which translates to an increase in computational burden of 5% and 1%, respectively.

7 EXPERIMENTAL DATA APPLICATION

In this section we use experimental data from the University of Wyoming Source Tracking Array Testbed (UW STAT) [20] (“dataset 4”) in order to test the estimators on actual data. The data are collected using a six sensor ultrasonic ULA, operating in a narrow-band setup, with two uncorrelated sources one of which is stationary and that we here consider

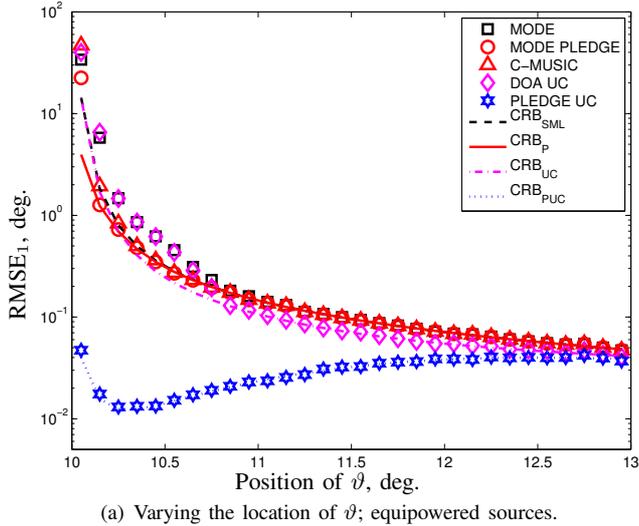
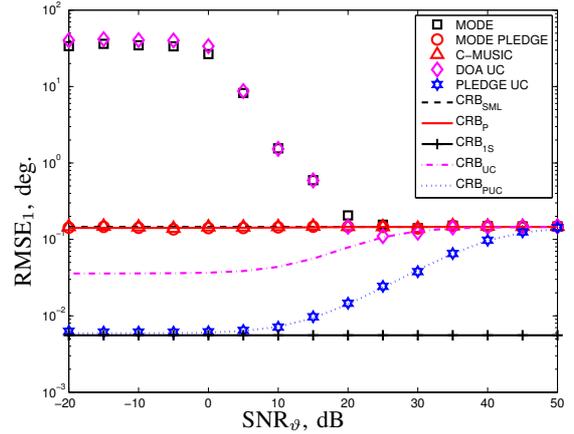
(a) Varying the location of ϑ ; equipowered sources.(b) Varying the strength of source at $\vartheta = 11^\circ$.

Fig. 4. Varying known-source properties: RMSE of the estimates of $\theta = 10^\circ$ when varying the properties of the known source. Uncorrelated sources, $\text{SNR}_\theta = 25$ dB, $m = 6$, $N = 1000$ and 1000 Monte Carlo realizations.

known. The sources are each emitting 200 Hz bandlimited noise, and the inter-sensor spacing is 2.1 signal wavelengths. Such a spacing provides an unambiguous array response in the region of interest (approximately $\pm 15^\circ$), while enhancing resolution in this region compared to a six-sensor ULA with half-plane field of view. With a sample rate of 800 Hz, but a signal bandwidth of only 200 Hz, we in effect do not have independent samples. This means that the effective number of snapshots, as compared to a theoretical analysis, is reduced by a factor of 2.

The data-dependent input to all the studied algorithms is the covariance matrix $\hat{\mathbf{R}}$ in (7). A problem in this case is that each data vector from the sensors represents a certain DOA, since one target is non-stationary. Thus when forming $\hat{\mathbf{R}}$, data vectors representing different DOAs are combined into $\hat{\mathbf{R}}$. From this $\hat{\mathbf{R}}$, the algorithms will of course find a single DOA for each source. Hence, in order to achieve meaningful results the source movement, or the change in the true DOA, has to be small over the interval of snapshots used to form $\hat{\mathbf{R}}$. In order to take this fact into account a recursive update of $\hat{\mathbf{R}}$ is used. Accordingly, $\hat{\mathbf{R}}$ is updated as

$$\hat{\mathbf{R}}(n) = \lambda \hat{\mathbf{R}}(n-1) + (1-\lambda) \mathbf{y}(n) \mathbf{y}^*(n), \quad \hat{\mathbf{R}}(0) = \mathbf{0}. \quad (49)$$

Thus, at time instant n , the DOA vector $\hat{\boldsymbol{\theta}}(n)$ is estimated based on $\hat{\mathbf{R}}(n)$. In (49) the forgetting parameter λ is chosen based upon how many past samples $\hat{\mathbf{R}}(n)$ is desired to effectively contain, according to the approximative relation: $\lambda = 1 - \frac{2}{N}$. We use $\lambda = 0.96$, approximately corresponding to a rectangular window size of $N = 50$.

The true source trajectories as a function of time can be seen in Fig. 5(a), and the absolute error $|\hat{\theta}(n) - \theta(n)|$ of the estimates in Fig. 5(b) and Fig. 5(c), with Fig. 5(b) corresponding to the algorithms that do not exploit information on the source correlation, and Fig. 5(c) showing the ones

utilizing this information. We plot the absolute error of the estimated angle at each time instant. We do not consider MP-MUSIC, as it was concluded in Section 6.1 that it is not robust.

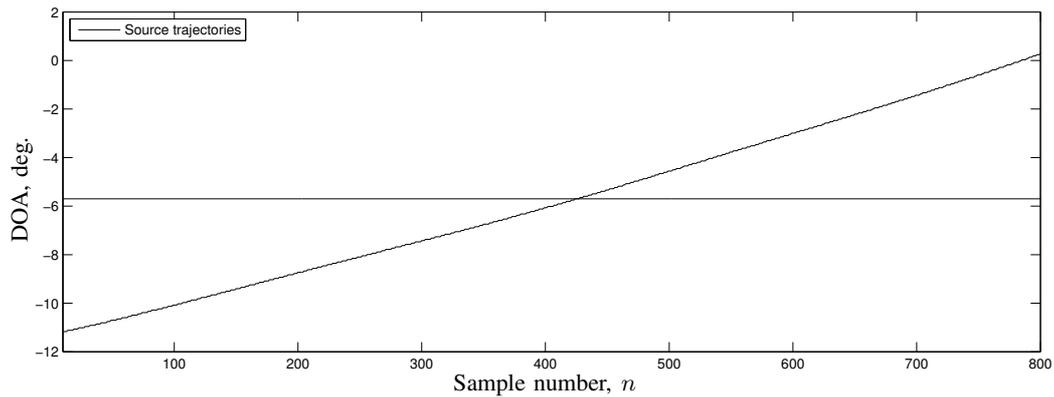
In the beginning of the data ($n < 300$) all algorithms show a relatively small error. As we proceed in time the moving source is approaching the stationary one. The source separation decreases and we can see that the algorithms suffer in accuracy; at first, this manifests as an increasing bias, which shows up as a gradual increase in absolute error, but eventually the detection breaks down and no useful information is provided by the estimators.

In Fig. 5(b), it can be seen that usage of the prior information through the MODE PLEDGE algorithm has a noticeable effect but is not as beneficial as one might have hoped. Comparing to Fig. 1, this result is not so surprising: the region of substantial improvement is small for MODE PLEDGE when the sources are uncorrelated.

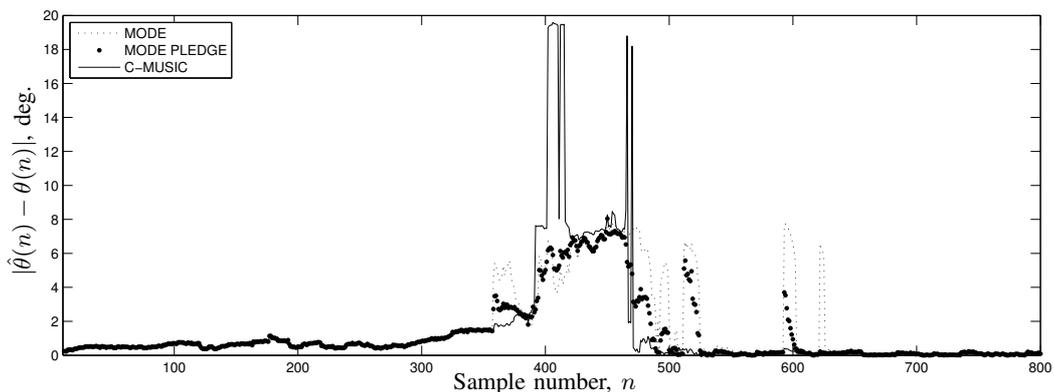
PLEDGE UC, as seen in Fig. 5(c), improves the estimation accuracy in some parts of the data sample — note that while its detection breaks down at approximately the same angular separation as for the other algorithms, it resumes tracking sooner giving a larger region of reasonable DOA estimation.

C-MUSIC is about as accurate as MODE PLEDGE; at some time instances it is better, while worse at others.

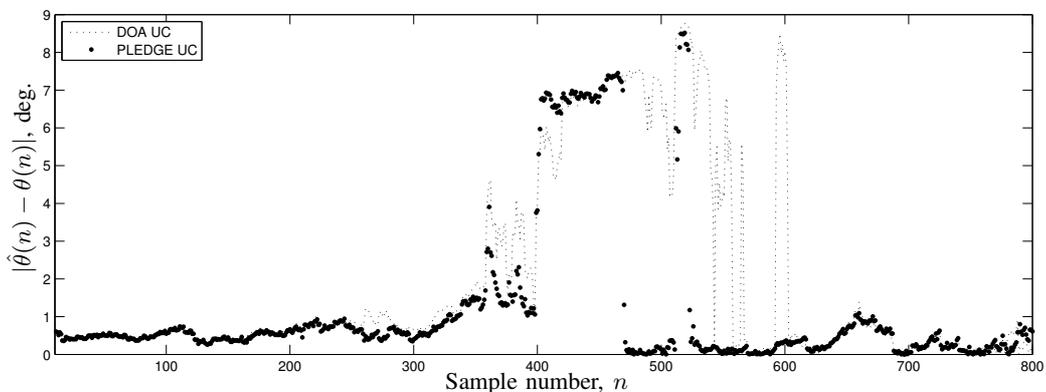
When the source DOAs are rather close to each other, we see from the figures that the prior-based methods fail similarly to the conventional methods. The PLEDGE methods remove, or filters out, the source at the known direction; however as the unknown source comes too close to the known one, the data from the unknown source is filtered out as well. Thus we cannot circumvent the fact that too closely spaced sources are indistinguishable. In other words the PLEDGE method causes the unknown source SNR to decrease when it comes too close to the known source. This can be contrasted to the situation of the conventional methods, in which the received



(a) True DOAs of the two sources, as a function of sample number. Dataset 4 from [20].



(b) Methods not exploiting correlation information.



(c) Methods exploiting the fact that the sources are uncorrelated.

Fig. 5. Real-data scenario: (a) true positions, and (b), (c) absolute error of DOA estimates ($\lambda = 0.96$). Known source fixed at $\vartheta = -5.7^\circ$, as evident from (a).

data transitions into an apparent scenario with a single source of double power in the direction of the merged sources.

defined according to

$$\text{ERMSE} = \sqrt{\frac{1}{M} \sum_{k=1}^M (\hat{\theta}_k - \theta)^2} \quad (50)$$

By using the Empirical Root Mean Square Error, ERMSE,

with M being the number of measured DOAs, we can aggre-

gate the results showed in Fig. 5. We thus find that MODE, MODE PLEDGE, and C-MUSIC gives ERMSE values of 2.7° , 2.2° , and 3.4° , respectively. The methods exploiting that the signals are uncorrelated, DOA UC and PLEDGE UC, gives ERMSE values of 3.3° and 2.3° , respectively. We can thus quantify the improvement in estimation accuracy given by the PLEDGE framework; however, in this particular case, the methods exploiting the uncorrelated nature of the source signals (i.e., DOA UC and PLEDGE UC) failed to improve the estimation, as compared to the methods that disregard this information. This is likely due to the real signals not satisfying the data model (1) — the data were acquired from moving sources and averaged according to (49), which is not what the methods were designed for. It is plausible that the UC methods, which are exploiting more of the structure in the received data, are more sensitive to such model imperfections.

8 CONCLUSIONS

We have closely examined a prior-knowledge-based framework that uses known positions to facilitate the estimation of unknown ones. We have shown how to exploit such prior DOA-knowledge in a manner that gives asymptotically optimal estimates whether the impinging signals are correlated or not. Through numerical simulations we have shown that in most scenarios a significant accuracy increase can be expected; this benefit is especially pronounced in the region where the estimator transitions to its asymptotic accuracy.

We also applied the methodology to a real-world scenario with one known DOA and one unknown time-varying DOA. Even though the methods are designed to analyze data received from a stationary target, we saw improvements in the angular estimates.

REFERENCES

- [1] R. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. of RADC Spectrum Estimation Workshop*, Griffiss AFB, New York, USA, Oct. 1979, pp. 243–258.
- [2] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.
- [3] M. Jansson and B. Ottersten, "Structured covariance matrix estimation: A parametric approach," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Istanbul, Turkey, Jun. 2000, pp. 3172–75.
- [4] P. Wirfält, G. Bouleux, M. Jansson, and P. Stoica, "Subspace-based frequency estimation utilizing prior information," in *IEEE Stat. Signal Process. Workshop*, Nice, France, Jun. 2011, pp. 533–536.
- [5] G. Bienvenu and L. Kopp, "Adaptivity to background noise spatial coherence for high resolution passive methods," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, vol. 5, Denver, Colorado, USA, Apr. 1980, pp. 307–310.
- [6] R. DeGroat, E. Dowling, and D. Linebarger, "The constrained MUSIC problem," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1445–1449, Mar. 1993.
- [7] D. Linebarger, R. DeGroat, E. Dowling, P. Stoica, and G. Fudge, "Incorporating a priori information into MUSIC-algorithms and analysis," *Signal Process.*, vol. 46, no. 1, pp. 85–104, 1995.
- [8] P. Stoica, P. Händel, and A. Nehorai, "Improved sequential MUSIC," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 4, pp. 1230–1239, Oct. 1995.
- [9] R. Behrens, L. Scharf, C. Inc, and C. Broomfield, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1413–1424, Jun. 1994.
- [10] R. Boyer and G. Bouleux, "Oblique projections for direction-of-arrival estimation with prior knowledge," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1374–1387, Apr. 2008.
- [11] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.
- [12] —, "MUSIC, maximum likelihood, and Cramér-Rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.
- [13] M. Jansson, B. Göransson, and B. Ottersten, "A subspace method for direction of arrival estimation of uncorrelated emitter signals," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 945–956, Apr. 1999.
- [14] G. Bouleux, P. Stoica, and R. Boyer, "An optimal prior knowledge-based DOA estimation method," in *17th European Signal Process. Conf.*, Glasgow, UK, Aug. 2009, pp. 869–873.
- [15] P. Wirfält, M. Jansson, G. Bouleux, and P. Stoica, "Prior-exploiting direction-of-arrival algorithm for partially uncorrelated source signals," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Vancouver, Canada, May 2013, pp. submitted, pre-print available at.
- [16] P. Stoica and K. Sharman, "Novel eigenanalysis method for direction estimation," *IEE Proc. Part F Radar and Signal Process.*, vol. 137, no. 1, pp. 19–26, Feb. 1990.
- [17] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1314–23, May 1999.
- [18] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, vol. 8, Boston, Massachusetts, USA, Apr. 1983, pp. 336–339.
- [19] P. Stoica, E. Larsson, and A. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.
- [20] J. Pierre, E. Scott, and M. Hays, "A sensor array testbed for source tracking algorithms," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Munich, Germany, Apr. 1997, pp. 3769–3772.