CONVEX RELAXATION APPROACHES TO MAXIMUM LIKELIHOOD DOA ESTIMATION IN ULA’S AND UCA’S WITH UNKNOWN MUTUAL COUPLING

Kehu Yang†, Shu Cai‡ and Zhi-Quan Luo‡

† ISN Lab, Xidian University
‡ Dept of ECE, University of Minnesota
200 Union Street SE, Minneapolis, MN 55455
Email: †{yang001, caishu}@xidian.edu.cn, ‡luozq@ece.umn.edu

ABSTRACT

Direction of arrival (DOA) estimation using sensor array super-resolution techniques are known to suffer from array modeling errors including array element displacements, mutual coupling, and array gain/phase perturbations. In this paper, we consider maximum likelihood (ML) DOA estimation for multiple sources in the presence of unknown mutual coupling, and propose convex semidefinite relaxation approaches to this nonlinear and non-convex problem for uniform linear arrays (ULA’s) and uniform circular arrays (UCA’s), respectively. Simulation results show that the proposed method are effective to practical applications.

Index Terms— Maximum likelihood, DOA estimation, semidefinite relaxation, ULA, UCA

1. INTRODUCTION

Direction of arrival (DOA) estimation, or direction finding (DF), is one of the main jobs in radar, sonar, wireless communication systems, etc. With the exact knowledge of array manifold or precisely array modeling, the superresolution techniques, such as MUSIC and maximum likelihood (ML) algorithms [1, 2], can provide superior performance [3, 4] of superresolution and DOA estimation for closely spaced spatial sources over the conventional direction-finding techniques. However, array manifold is often not known exactly in practice due to unavoidable array modeling errors, such as array element displacements, mutual coupling, and array gain/phase perturbations. In this case, the MUSIC and ML algorithms [5–7] will be suffered from heavy degradation of superresolution for closely spaced sources and high deviation of the DOA estimates from the true values. The former reveals that those two methods could not be used in the cases where closely-spaced sources are present, while the latter implies that the DOA estimation results may be useless.

In this paper, we consider DOA estimation in ULA’s and UCA’s in the presence of unknown mutual coupling, where the other array modeling errors are well calibrated. According to the electromagnetic theory, the mutual coupling resistance (or called coefficients) depends on array geometry and the distance between array elements. Due to the symmetric property of geometry, the mutual coupling matrix (MCM) of a ULA (or a UCA) can be represented by banded and symmetric Toeplitz one (or symmetric circulant one) [8]. Notice that a symmetric Toeplitz matrix is formed by a row vector. This led us to formulate the joint maximum likelihood estimation of the DOA’s of multiple sources and the mutual coupling coefficient vector that form the MCM of ULA or UCA, and propose convex semidefinite relaxation approaches to this nonlinear and non-convex problem.

2. ARRAY SIGNAL MODEL

Consider an M-element array with $M(\leq N)$ sources illuminating on. In the presence of mutual coupling, the received array signal vector is denoted by

$$x(t) = CA_s(t) + n(t),$$

(1)

where $x(t) = [x_1(t), \ldots, x_N(t)]^T$, $s(t) = [s_1(t), \ldots, s_M(t)]^T$, $n(t) = [n_1(t), \ldots, n_N(t)]^T$. $C$ denotes the mutual coupling matrix, $A = A(\Theta) = [a(\theta_1), a(\theta_2), \ldots, a(\theta_M)]$ is the array manifold for one-dimensional array, and $a(\theta)$ is the steering vector. For a ULA with the element spacing of $d$, we have

$$a(\theta) = \left[1, e^{j \frac{2\pi}{N} d \sin(\theta_1)}, \ldots, e^{j \frac{2\pi}{N} (N-1)d \sin(\theta)} \right]^T$$

(2)

For DOA estimation, we assume that: the columns of $CA$ are linearly independent, the signals of sources and the noises are uncorrelated, and the array noises are spatially white and Gaussian distributed with the covariance matrix $\mathbb{E}(n(t)n^H(t)) = \sigma^2 I$.

3. ML DOA ESTIMATION IN ULA’S AND UCA’S WITH UNKNOWN MUTUAL COUPLING

3.1. ML DOA Estimation in ULA’s

In the presence of unknown mutual coupling, ML DOA estimation of $M$ sources can be performed by

$$\min_{\Theta, C, S} \sum_{k=1}^{N} \left\| x(t_k) - CA(\Theta)s(t_k) \right\|^2,$$

(3)

where $N$ denotes the number of snapshots, $\Theta = [\theta_1, \ldots, \theta_M]$, $S = [s(t_1), \ldots, s(t_N)]$, and $C$ is a banded and symmetric Toeplitz matrix for ULA’s [8]. Let $B(\Theta) = CA(\Theta)$, (3) can be equivalently written as the joint optimization problem:

$$\max_{\Theta, C} \operatorname{tr} \left\{ P_B(\Theta) \hat{R} \right\},$$

(4)
where “tr” stands for “trace” of a square matrix,
\[ P_{B(\Theta)} = B(\Theta)(B^H(\Theta)B(\Theta))^{-1}B^H(\Theta), \]
\[ \hat{R} = \frac{1}{N_t} \sum_{t_k=1}^{N_t} x(t_k)x^H(t_k). \]

The optimization problem (4) can be solved using the procedure of alternating coordinate optimization, that is, when \( \Theta \) is fixed, the estimate of \( \Theta \) can be obtained by solving (4), and vice versa. In [2], the alternating projection (AP) algorithm is proposed to estimate \( \Theta \). According to (4), for single source case and given \( \Theta \), the mutual coupling coefficient vector \( c \) that form the MCM of ULA’s (or UCA’s) can be directly estimated by
\[ \hat{c} = \arg \max_c \frac{c^HQ^H(\alpha(\Theta))\hat{R}Q(\alpha(\Theta))c}{c^HQ^H(\alpha(\Theta))Q(\alpha(\Theta))c}, \tag{5} \]
where \( Q(\alpha(\Theta))c \) is defined according to the MCM of ULA’s (or UCA’s) see Lemma 2 or Lemma 3 in [8]. The optimization problem (5) can be solved by finding the eigenvector of the objective function corresponding to the maximum eigenvalue.

For multiple sources case and given \( \Theta \), it is evident to have the relationship:
\[ P_{\hat{A}(\Theta)} = I - P_{\hat{A}(\Theta)} = U(\Theta)U^H(\Theta), \tag{6} \]
where \( P_{\hat{A}(\Theta)} = A(\Theta)(A^H(\Theta)A(\Theta))^{-1}A^H(\Theta) \), and \( U(\Theta) \) is the eigenvector matrix with all the corresponding eigenvalues equal to 1’s. Define
\[ F(\Theta) = (C^H)^{-1}U(\Theta). \tag{7} \]
In the sense of (6), we can see that \( F(\Theta) \) is orthogonal to an arbitrary vector of \( B(\Theta) \). This means that the column space of \( F(\Theta) \) is the orthogonal complement of that of \( B(\Theta) \), and we have the relationship that \( P_{B(\Theta)} = I - P_{F(\Theta)} \) with \( P_{F(\Theta)} \) similarly defined as that of \( P_{B(\Theta)} \). Therefore, (4) can be equivalently written as
\[ \min_{c,T} \text{tr} \left\{ P_{F(\Theta)} \hat{R} \right\}. \tag{8} \]
Under given \( \Theta \), (8) becomes
\[ \min_{\Theta} \text{tr} \left\{ P_{F(\Theta)} \hat{R} \right\}. \tag{9} \]
Define
\[ C = (C^H)^{-1}, \quad U = U(\Theta). \tag{10} \]
Notice that \( C^HC = (C^H)^{-1}, \hat{R} = \hat{R}^H \hat{R}^{1/2}, \) and \( \hat{R}^{1/2} = (\hat{R}^{1/2})^H \), we have
\[ \text{tr} \left\{ P_{F(\Theta)} \hat{R} \right\} = \text{tr} \left\{ \hat{R}^{1/2}CU(U^HC^HCU)^{-1} (\hat{R}^{1/2}CU)^H \right\}. \]
The optimization problem (9) can thus be equivalently written as
\[ \min_{\Theta,T} \text{tr} \left\{ T \right\} \]
subject to
\[ \hat{R}^{1/2}CUZ^{-1}(\hat{R}^{1/2}CU)^H \preceq T, \quad Z = (CU)^HC^HCU, \quad T \succeq 0, \quad T = T^H, \quad C = C^T. \tag{11} \]
According to the matrix theory, the inversion of a banded and symmetric Toeplitz matrix is a symmetric one and the contrary is not always true. This means that \( C \) and \( C \) must be estimated simultaneously. Define
\[ W \triangleq C^HC, \tag{12} \]
we have
\[ Z = U^HW^{-1}U. \tag{13} \]
Substituting (12) and (13) into (11), we have the following equivalent form of (9):
\[ \min_{C,T} \text{tr} \{ T \} \]
subject to
\[ \hat{R}^{1/2}CUZ^{-1}(\hat{R}^{1/2}CU)^H \preceq T, \quad Z = (CU)^HC^HCU, \quad W = C^HC, \quad T \succeq 0, \quad T = T^H, \quad C = C^T. \tag{14} \]
Using the relaxations
\[ C^HC \preceq W \iff \begin{bmatrix} I & C \\ C^H & W \end{bmatrix} \succeq 0, \]
\[ U^HW^{-1}U \preceq Z \iff \begin{bmatrix} W & U \\ U^H & Z \end{bmatrix} \succeq 0, \]
The optimization problem (14) can be relaxed to the following SDP:
\[ \min_{C,T,Z,W} \text{tr} \{ T \} + \delta_1 \text{tr} \{ Z \} + \delta_2 \text{tr} \{ W \} \]
subject to
\[ \begin{bmatrix} Z & \hat{R}^{1/2}CU \\ \hat{R}^{1/2}CU & T \end{bmatrix} \succeq 0, \quad \begin{bmatrix} W & U \\ U^H & Z \end{bmatrix} \succeq 0, \quad \begin{bmatrix} I & C \\ C^H & W \end{bmatrix} \succeq 0, \quad T = T^H, \quad Z = Z^H, \quad W = W^H, \quad C = C^T, \quad Z \succeq 0, \quad T \succeq 0, \quad W \succeq 0, \quad Z \succeq 0, \quad T \succeq 0, \quad W \succeq 0, \quad \delta_1, \delta_2 \text{ are small positive numbers for penalization}. \tag{15} \]

According to (15), the ML DOA estimation with unknown mutual coupling can be performed using the following coordinate optimization procedure.

**Procedure 1:**

**Step 0:** Let \( l = 0, k = 1, \) and \( y = \alpha(\Theta) \). Given control accuracy \( \epsilon \geq 0 \), and the interval \([\theta_a, \theta_b] \) (called the field of view).

Choose \( N_p \) points of \( \theta \) uniformly in the interval, i.e., \( \Theta \in [\theta_a : \theta_b, \theta_b : \theta_a, \ldots, \theta_a : \theta_a + \frac{\theta_b - \theta_a}{N_p - 1} : \theta_a] \). For each point of \( \Theta \), solve (5) and find the maximum value of the objective function of (5). For all the \( N_p \) points of \( \Theta \), find the great maximum and call the \( \Theta \) corresponding to it as \( \Theta^{(0)} \). Let \( \Theta^{(0)} = \Theta^{(0)} \).

**Step 1:** Let \( \Theta = [\Theta^{(0)}] \). For each point of \( \Theta \in [\Theta^{(0)} : \Theta^{(0)} + \frac{\theta_b - \Theta^{(0)}}{N_p - 1} : \Theta^{(0)}] \), solve (15) by using SDP solver SeDuMi [9], obtain an estimate \( \hat{C} \) of \( C \), use this \( \hat{C} \) and \( \Theta \) as an initial point for further local search by applying a Newton type method (or any standard nonlinear optimization routine) to (4), and obtain the refined \( \hat{C}, \Theta \), and value of the objective function. For all the \( N_p \) points of \( \Theta \), find the great maximum and call the \( \Theta \) corresponding to it as \( \Theta^{(0)} \);

**Step 2:** Let \( k = \text{length}([\Theta^{(0)}]) \).

If \( k < M \) then go to Step 1

Else let \( l = 0 \) and go to Step 3;
Step 3: Let $\Theta(l) = [\theta_1^{(l)}, \ldots, \theta_M^{(l)}]$ and $\Theta = [\theta_1^{(0)}, \ldots, 
abla_{\theta_1}^{(0)}, \ldots, \hat{\theta}_0^{(0)}, \ldots, \theta_M^{(0)}]$. For each point of $\theta \in [\theta_0 : \frac{\theta_0 - \theta_1}{N_p} : \theta_1]$, solve (15) by using SeDuMi, obtain an estimate $\hat{C}$ of $C$, use this $\hat{C}$ and $\Theta$ as an initial point for further local search by applying a Newton type method to (4), and obtain the refined $\hat{C}$, $\hat{C}$, and value of the objective function. For all the $N_p$ points of $\theta$, find the great maximum and call the $\Theta$ corresponding to it as $\Theta(\hat{1})$.

Step F: For $i = 1, \ldots, M$, let $l = l + 1$ and repeat Step 3.

If $||\Theta(l+1) - \Theta(l-M+1)|| \leq \epsilon$ then stop.

Else let $l = l + 1$, go to step 3.

Notice that in Procedure 1, Step 0 ~ 2 are for the initialization, while Step 3 ~ F are for updating.

3.2. ML DOA Estimation in UCA's

The ML DOA estimation in UCA’s in the presence of unknown mutual coupling can be performed in the same way as that in ULA’s by (4) or its equivalent form (8):

$$\min_{\Theta, C} \text{tr} \left\{ P_F(\Theta) \hat{r} \right\},$$

where $F(\Theta)$ is defined by (7) and the corresponding $a(\theta)$ denotes the array steering vector of UCA’s. Under given $\Theta$, the optimization problem (16) becomes (9), which is equivalent to (11).

As mentioned in [8], the mutual coupling matrix $C$ of a UCA is circulant and symmetric. According to the matrix theory, the inversion of $C$ is also circulant and symmetric, and both are formed by the corresponding vectors with the same number of elements, respectively. This means that $C$ can be uniquely determined by its inversion $\hat{C}$, and there is no need to estimate $C$ and $\hat{C}$ simultaneously.

Using the relaxation

$$(CU)^HCU \preceq Z \iff \left[ \begin{array}{cc} I & CU \\ (CU)^H & Z \end{array} \right] \succeq 0,$$

the optimization problem (11) can be relaxed to the following SDP:

$$\min_{C, T, Z} \text{tr} \{ T \} + \delta \text{tr} \{ Z \}$$

subject to

$$\left[ \begin{array}{cc} Z & \hat{r}^{1/2}CU \\ \hat{r}^{1/2}CU & T \end{array} \right] \succeq 0,$$

$$\left[ \begin{array}{cc} I & CU \\ (CU)^H & Z \end{array} \right] \succeq 0, Z \succeq 0, T \succeq 0,$$

$$T = T^H, Z = Z^H, C = C^T,$$

where $\delta$ is a small positive number for penalization.

Similarly, the ML DOA estimation in UCA’s with unknown mutual coupling can be performed under the procedure similar to Procedure 1. By making minor modification to Procedure 1, we have the following procedure for ML DOA estimation in UCA’s.

Procedure 2:

Step 0: Let $l = 0, k = 1$, and $y = a(\theta)$. Given control accuracy $\epsilon > 0$, and the interval $[\theta_0, \theta_1]$ (called the field of view).

Choose $N_p$ points of $\theta$ uniformly in the interval, i.e., $\theta \in [\theta_0 : \frac{\theta_0 - \theta_1}{N_p} : \theta_1]$. For each point of $\theta$, solve (5) and find the maximum value of the objective function of (5). For all the $N_p$ points of $\theta$, find the great maximum and call the $\theta$ corresponding to it as $\theta^{(0)}_1$. Let $\Theta^{(0)} = [\theta^{(0)}_1]$.

Step 1: Let $\Theta = [\theta^{(0)}_1]$. For each point of $\theta \in [\theta_0 : \frac{\theta_0 - \theta_1}{N_p} : \theta_1]$, solve (17) by using SDP solver SeDuMi [9], obtain an estimate $\hat{C}$ and then $\hat{C}$ by inverting $\hat{C}$, use this $\hat{C}$ and $\Theta$ as an initial point for further local search by applying a Newton type method to (4), and obtain the refined $\hat{C}$, $\hat{C}$ and value of the objective function. For all the $N_p$ points of $\theta$, find the great maximum and call the $\Theta$ corresponding to it as $\Theta^{(0)}_i$.

Step 2: If $k = length(\Theta^{(0)})$.

Else let $k = 0$ and go to Step 3.

Step 3: Let $\Theta^{(1)} = [\theta_1^{(1)}, \ldots, \theta_M^{(1)}]$ and $\Theta = [\theta_1^{(1)}, \ldots, \theta^{(1)}_1, \theta^{(1)}_{i+1}, \ldots, \theta^{(1)}_M]$. For each point of $\theta \in [\theta_0 : \frac{\theta_0 - \theta_1}{N_p} : \theta_1]$, solve (17) by using SeDuMi, obtain an estimate $\hat{C}$ and then $\hat{C}$ by inverting $\hat{C}$, use this $\hat{C}$ and $\Theta$ as an initial point for further local search by applying a Newton type method to (4), and obtain the refined $\hat{C}$, $\hat{C}$ and value of the objective function. For all the $N_p$ points of $\theta$, find the great maximum and call the $\Theta$ corresponding to it as $\Theta^{(1)}_i$.

Step F: For $i = 1, \ldots, M$, let $l = l + 1$ and repeat Step 3.

If $||\Theta^{(1)}_i - \Theta^{(1)}_{i-M+1})|| \leq \epsilon$ then stop.

Else let $l = l + 1$, go to step 3.

4. NUMERICAL RESULTS

To demonstrate the performance of the proposed convex relaxation methods for ML DOA estimation in ULA’s and UCA’s with unknown mutual coupling, we perform two numerical experiments to compare the proposed methods with the alternating projection (AP) algorithm [2] without considering the mutual coupling. In the experiments, we assume that both the ULA and the UCA are of $N = 16$ isotropic elements with half wavelength spacing, there are two equal power narrow-band sources illuminating the arrays, respectively, and the array noise is spatially white and Gaussian distributed. The number of snapshots is set as 1000 for the two experiments.

The first experiment is for ML DOA estimation in the ULA by using Procedure 1, where the array elements are monopole and vertically polarized with 10n long and 20n spacing corresponding to the central frequency $\Sigma MHz$. Accordingly, the 3dB beamwidth of the ULA, defined as the reciprocal of its aperture in wavelength, is about 7.6°. We assume that the two emitters are located at $-2^\circ$ and $1^\circ$, whose difference is smaller than one half the beamwidth. The theoretical MCM $C$ of the ULA is obtained according to the electromagnetic theory, which is a symmetric Toeplitz matrix and formed by the corresponding 16 x 1 multipole coupling coefficient vector $c = [49.3603 + 52.3976i - 14.238 - 15.4087i - 8.3972 + 8.3972i - 6.4842 + 8.471i - 5.4764 + 2.7916i - 4.7418 - 1.3907i + 4.1124 + 0.3703i - 3.5175 + 0.8345i + 2.9612 + 0.9521i - 2.4035 + 1.3548i + 1.8712 - 1.4356i - 1.3677 + 1.8021i - 0.8871 + 1.8718i - 0.4501 + 1.8561i - 0.9506 + 1.7814i - 0.2943 + 1.6363i]$. Notice that the $C$ formed by this $c$ is not banded. Here we consider two cases for the ML DOA estimation in this experiment. In Case #1, the simulation data are generated according to (1), where $C$ is formed as a symmetric and Toeplitz matrix by the abovementioned $c$, while in Case #2, the simulation data are generated according to (1), where $C$ is formed as a banded and symmetric Toeplitz matrix by a truncated $c$ that is formed by the first 8 elements of the abovementioned $c$. In this experiment, 100 Monte-Carlo runs are performed for each case.

For the experiment of Case #2, we directly use Procedure 1 to estimate the DOA’s of the two sources and the unknown banded and
symmetric Toeplitz matrix $C$, the RMSE of the DOA estimates versus the signal to noise ratio (SNR) is plotted in Fig.1, where the black sold line named by “SDRT1” is corresponding to the source with $\theta_1 = -2^\circ$ and the dashed line named by “SDRT2” is for the source with $\theta_2 = 1^\circ$. For the experiment of Case #1, we still use Procedure 1 to estimate the DOA’s of the two sources by assuming that the unknown $C$ is banded and symmetric Toeplitz and formed by 8 unknown mutual coupling coefficients except that the local search in all the steps are with respect to the unknown symmetric Toeplitz $C$. The corresponding RMSE curves red marked and named by “SDR1” and “SDR2” are plotted in Fig.1. For comparison, we also plotted in Fig.1 the RMSE curves by AP algorithm and named by ‘AP1” and “AP2”. It is seen from Fig.1 that the proposed convex relaxation methods provide superior performance as compared to the AP algorithm [2], whereas the performance of the AP algorithm [8] without considering the mutual coupling becomes worse no matter what the input SNR is.

The second experiment is for ML DOA estimation in the UCA by using Procedure 2, where the array elements are monopole and vertically polarized with 10$m$ long and 20$m$ spacing corresponding to the central frequency 8$MHz$. The 3dB beamwidth is therefore about 22.5$. The theoretic MCM $C$ of the UCA is obtained according to the electromagnetic theory, which is a symmetric circulant matrix [8] and formed by the corresponding $9 \times 1$ vector $c = [50.5411 + 52.4967i, -14.2338 - 15.4987i, 7.3927 + 9.4967i, -3.5097 - 7.7558i, -0.9866 + 6.6993i, 5.3217 - 2.2457i, -2.0591 - 4.7851i, 10.8265 + 37.1243i, -4.7731 + 0.6508i]$ in the manner as that in the Appendix I of [8]. We also consider two cases for the second experiment. In Case #1, we assume the DOA’s of the two sources are that $\theta_1 = -4^\circ$ and $\theta_2 = 2^\circ$, while in Case #2, the DOA’s of the two sources are assumed as that $\theta_1 = -4^\circ$ and $\theta_2 = 4^\circ$. For both the cases, the simulation data are generated according to (1). In this experiment, 300 Monte-Carlo runs are performed for each case.

Fig.2 shows the RMSE of the DOA estimates versus the input SNR for both cases are plotted in Fig.2. For comparison, we also plot in Fig.2 the RMSE of the DOA estimates by AP algorithm [2]. In Fig.2, the “SDR” stands for our proposed relaxation method, while the “AP” stands for AP algorithm. It is seen from the figure that our proposed relaxation method considering the mutual coupling significantly outperforms the AP algorithm that does not consider the effect of mutual coupling.

5. REFERENCES