JOINT DOA AND FUNDAMENTAL FREQUENCY ESTIMATION BASED ON RELAXED ITERATIVE ADAPTIVE APPROACH AND OPTIMAL FILTERING

Zhenhua Zhou[†] *M. G. Christensen*[‡] *J. R. Jensen*[‡] *H. C. So*[†]

[†] Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China
[‡] Audio Analysis Lab, AD:MT, Aalborg University, Aalborg, Denmark
E-mail: sjtu.zzh2010@gmail.com, mgc@create.aau.dk, jrj@create.aau.dk, hcso@ee.cityu.edu.hk

ABSTRACT

In this work, the problem of joint direction-of-arrival and fundamental frequency estimation for multi-channel harmonic sinusoidal signals is addressed. Different from the conventional optimal filtering method, we estimate the covariance matrix with the 2-D iterative adaptive approach, which is based on a single snapshot. In addition, to improve the estimation accuracy for the off-grid sources, a relaxation technique is utilized. Then, joint estimation is conducted on this covariance matrix estimate with the optimal filtering method. As a result, the relaxed iterative adaptive approach - optimal filtering method is devised. Statistical evaluation with synthetic signals shows the accurate performance of the proposed method compared with the Cramér-Rao lower bound.

Index Terms— joint DOA and fundamental frequency estimation, iterative adaptive approach, optimal filtering, multi-channel harmonic sinusoidal signal

1. INTRODUCTION

The problem of parameter estimation for speech sources, namely, the direction-of-arrival (DOA) and fundamental frequency (or pitch), has been of interest to the signal processing community for many years, finding its applications in source localization, speech signal enhancement, automatic transcription and classification of music [1]. Traditionally, estimation of these two parameters is conducted separately [2] - [4], [5] - [7]. Nevertheless, due to the issue of overlapping harmonics (spatial or temporal) [7], it is difficult to identify different sources with similar temporal or spatial frequencies accurately.

Recently, to facilitate the resolution of such sources, several kinds of joint estimation approaches for DOA and pitch have been developed. In [8], the nonlinear least squares (NLS) estimator is proposed for a single-pitch signal, which is the maximum likelihood estimator in the scenarios of white Gaussian noise and anechoic environment, while it bears poor performance for the multi-pitch signal due to the decoupling difficulty [9]. To overcome this problem, in [10] and [11], the optimal filtering and multi-channel harmonic MUSIC (MC-HMUSIC) estimators are presented, respectively. However, in both methods, the data matrix is used. Thus, the user-defined parameters such as the size of the data matrix affect the performance of these estimators.

To address these issues, we propose the relaxed iterative adaptive approach (RIAA) to estimate the covariance matrix, which is based on the 2-D iterative adaptive approach (IAA) and relaxation techniques. At first, we tackle joint estimation of DOA and pitch for the single-pitch signal. As shown in [12], no data partitioning is needed in the IAA-based method, i.e., the covariance matrix can be estimated iteratively from only a single snapshot. That is, there is no user-defined parameter when using this technique. On the other hand, the IAA covariance matrix model is defined on some frequency grid. To improve the estimation accuracy for the off-grid sources, the relaxation technique is utilized herein with the IAA covariance matrix estimate as an initial estimate, and the RIAA covariance matrix estimate is produced. The DOA and pitch estimates are computed with the harmonic optimal filtering method and the RIAA covariance matrix estimate. This algorithm is termed as the RIAA - optimal filtering (RIAA-OF) method. Furthermore, we extend the RIAA-OF method to the multi-pitch scenario. Simulation results show that the proposed scheme has good accuracy performance compared with the Cramér-Rao lower bound (CRLB) [13].

The rest of this paper is organized as follows. The proposed joint estimator of the fundamental frequency and DOA, namely, the RIAA-OF method, is detailed in Section 2, including the problem statement, covariance matrix estimation, parameter estimation, and extension of this method to the multi-pitch scenario. In Section 3, simulation results are shown to evaluate the performance of the proposed estimators by comparing with the CRLB. Finally, conclusions are drawn in Section 4.

2. ALGORITHM DEVELOPMENT

2.1. Spatial-Temporal Signal Model

To facilitate the derivation to follow, we first present the multi-channel model under consideration. Without multi-path propagation of sources, the multi-channel signal model is given as follows. The signal $x_i(n)$ received by the microphone element (or sensor[1]) i arranged in a uniform linear array (ULA) configuration, $i = 1, 2, \cdots, I$, is modeled as [15]:

$$x_i(n) = s_i(n) + q_i(n), \tag{1}$$

$$s_i(n) = \sum_{k=1}^{K} \sum_{l=1}^{L_k} A_{l,k} e^{j(\omega_k ln + \alpha_k l(i-1) + \phi_{l,k})}, \tag{2}$$

for $n = 1, 2, \cdots, N$, with ω_k, α_k, $\{A_{l,k}\}$ and $\{\phi_{l,k}\}$ denoting the unknown fundamental frequency, scaled DOA, amplitudes and initial phases of the k-th source, respectively. The number of sources K and the number of harmonics L_k of each source k, are assumed to be known, or found in some way such as [7]. The $q_i(n)$ is the noise (assumed white Gaussian if not mentioned) of the i-th microphone

This work was funded in part by the Villum Foundation.

[1] Each sensor stands for a single channel.

element with variance σ^2. The objective is to estimate the nonlinear parameters ω_k and α_k accurately.

2.2. Single-Pitch Estimation

To begin with, we focus on the single-pitch estimation ($K = 1$), which consists of two steps: 1) covariance matrix estimation using the RIAA and 2) joint estimation of the pitch and DOA using the optimal filtering method. In the first step, the covariance matrix of the $(N \cdot I) \times 1$ single snapshot data, $\mathbf{x}_{N,I}$, $\mathbf{x}_{N,I}((i-1) \times N + n) = x_i(n)$, namely, $\mathbf{R} = E\{\mathbf{x}_{N,I} \cdot \mathbf{x}_{N,I}^H\}$, is approximated by the well-known IAA covariance matrix model [12], [16]:

$$\hat{\mathbf{R}} = \mathbf{Z}(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}})\hat{\mathbf{P}}\mathbf{Z}^H(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}}), \tag{3}$$

where the covariance matrix is defined at the $K_1 \times K_2$-point frequency grid $\overline{\boldsymbol{\varpi}} \times \overline{\boldsymbol{\alpha}}$, $\overline{\boldsymbol{\varpi}} = \left[2\pi \frac{1}{K_1}\ 2\pi \frac{2}{K_1}\ \cdots\ 2\pi\right]^T$, $\overline{\boldsymbol{\alpha}} = \left[2\pi \frac{1}{K_2}\ 2\pi \frac{2}{K_2}\ \cdots\ 2\pi\right]^T$, and the matrices $\mathbf{Z}(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}})$ and $\hat{\mathbf{P}}$ have the forms of [12], [16]:

$$\begin{aligned} \mathbf{Z}(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}}) &= [\mathbf{z}(\overline{\alpha}_1)\ \cdots\ \mathbf{z}(\overline{\alpha}_{K_2})] \\ &\quad \otimes [\mathbf{z}(\overline{\omega}_1)\ \cdots\ \mathbf{z}(\overline{\omega}_{K_1})], \qquad (4) \\ \mathbf{z}(\overline{\omega}_{k_1}) &= \left[e^{\overline{\omega}_{k_1}}, \cdots, e^{\overline{\omega}_{k_1} N}\right]^T, \qquad (5) \\ \mathbf{z}(\overline{\alpha}_{k_2}) &= \left[e^{\overline{\alpha}_{k_2}}, \cdots, e^{\overline{\alpha}_{k_2} I}\right]^T, \qquad (6) \end{aligned}$$

$\overline{\omega}_{k_1}$ $(1 \le k_1 \le K_1)$ and $\overline{\alpha}_{k_2}$ $(1 \le k_2 \le K_2)$ standing for the k_1-th and k_2-th elements of $\overline{\boldsymbol{\varpi}}$ and $\overline{\boldsymbol{\alpha}}$, respectively, $\otimes$ being the Kronecker product, and

$$\begin{aligned} \hat{\mathbf{P}} &= \mathrm{diag}(\hat{P}_{1,1}, \hat{P}_{2,1}, \cdots, \hat{P}_{K_1,K_2}) \\ &= \mathrm{diag}(|\hat{\beta}_{1,1}|^2, |\hat{\beta}_{2,1}|^2, \cdots, |\hat{\beta}_{K_1,K_2}|^2), \qquad (7) \end{aligned}$$

with $\hat{P}_{k_1,k_2}$ denoting the power estimate at each frequency point on the corresponding scanning grid, and equal to the magnitude square of the amplitude: $\hat{P}_{k_1,k_2} = |\hat{\beta}_{k_1,k_2}|^2$. Here the amplitude estimate is solved by minimizing the following weighted least squares cost function [17]:

$$\begin{aligned} \hat{\beta}_{k_1,k_2} &= \arg \min_{\tilde{\beta}_{k_1,k_2}} (\mathbf{x}_{N,I} - \mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\tilde{\beta}_{k_1,k_2})^H \\ &\quad \hat{\mathbf{Q}}^{-1}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})(\mathbf{x}_{N,I} - \mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\tilde{\beta}_{k_1,k_2}), \qquad (8) \end{aligned}$$

and

$$\begin{aligned} \mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2}) &= \mathbf{z}(\overline{\alpha}_{k_2}) \otimes \mathbf{z}(\overline{\omega}_{k_1}), \qquad (9) \\ \hat{\mathbf{Q}}(\overline{\omega}_{k_1}, \overline{\omega}_{k_2}) &= \hat{\mathbf{R}} - |\hat{\beta}_{k_1,k_2}|^2 \\ &\quad \mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2}). \qquad (10) \end{aligned}$$

By solving the above minimizing problem with the matrix inversion lemma, it is derived that

$$\hat{\beta}_{k_1,k_2} = \frac{\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\hat{\mathbf{R}}^{-1}\mathbf{x}_{N,I}}{\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\hat{\mathbf{R}}^{-1}\mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})}. \tag{11}$$

Note that (11) is related to the unknown covariance matrix estimate $\hat{\mathbf{R}}$. As a result, the covariance matrix estimation is performed in an iterative way initialized by the periodogram estimate. In most applications, 15 iterations are enough [16]. The steps of the above IAA procedure are listed in Table 1.

$\hat{\beta}_{k_1,k_2} = \left(\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\mathbf{x}_{N,I}\right)/\left(N \cdot I\right)$, $k_1 = 1, \cdots, K_1, k_2 = 1, \cdots, K_2$.
repeat $\hat{\mathbf{R}} = \mathbf{Z}(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}})\hat{\mathbf{P}}\mathbf{Z}^H(\overline{\boldsymbol{\varpi}}, \overline{\boldsymbol{\alpha}})$. for $k_1 = 1, \cdots, K_1, k_2 = 1, \cdots, K_2$ • $\hat{\beta}_{k_1,k_2} = \frac{\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\hat{\mathbf{R}}^{-1}\mathbf{x}_{N,I}}{\mathbf{z}^H(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})\hat{\mathbf{R}}^{-1}\mathbf{z}(\overline{\omega}_{k_1}, \overline{\alpha}_{k_2})}$, • $\hat{P}_{k_1,k_2} = \|\hat{\beta}_{k_1,k_2}\|^2$. end until convergence

Table 1. IAA-based estimation of covariance matrix

As seen from (3), the IAA-based covariance matrix estimation is related to the definition of the scanning grid $\overline{\boldsymbol{\varpi}} \times \overline{\boldsymbol{\alpha}}$. When the sources are off grid, the performance will be degraded. To overcome this problem, the relaxation technique [18] is utilized to refine the estimate of the covariance matrix. The basic idea of relaxation is to recover the signal part corresponding to some sinusoidal component, and then to refine its location in a maximum likelihood (ML) way under the assumption that the residual is white Gaussian. Now that the location corresponding to the sinusoidal component is unknown *a priori*, such refinement is conducted in an iterative way. The relaxation approach is outlined in Table 2, and empirically, 5 iterations are sufficient for its convergence.

$L = \sum_{k=1}^{K} L_k$: Number of the sinusoidal components from the K sources; $(\overline{\omega}'_l, \overline{\alpha}'_l)$ $(l = 1, \cdots, L)$: Locations of the sinusoidal peaks obtained from IAA; $\hat{\beta}_l$ $(l = 1, \cdots, L)$: Amplitude estimates corresponding to the sinusoidal peaks.
repeat for $l = 1, \cdots, L$ • $\mathbf{x}_l = \mathbf{x}_{N,I} - \sum_{i=1, i \neq l}^{L} \mathbf{z}(\overline{\omega}'_i, \overline{\alpha}'_i)\hat{\beta}_i$, • $(\overline{\omega}'_l, \overline{\alpha}'_l) = \arg \max_{\tilde{\overline{\omega}}, \tilde{\overline{\alpha}}} \|\mathbf{x}_l^H \cdot \mathbf{z}(\tilde{\overline{\omega}}, \tilde{\overline{\alpha}})\|^2$, • $\hat{\beta}_l = \frac{1}{N \cdot I}\mathbf{z}^H(\overline{\omega}'_l, \overline{\alpha}'_l) \cdot \mathbf{x}_l$. end until convergence

Table 2. Relaxation of IAA

Now the covariance matrix estimation is achieved with the RIAA method, and we proceed to conduct the joint estimation of DOA and pitch with the optimal filtering method and the covariance matrix estimate. The application of the optimal filtering method to such joint estimation is introduced in [10], and is based on an optimal harmonic linearly constrained minimum variance (LCMV) filter. Consider the single snapshot $\mathbf{x}_{N,I}$, and introduce the FIR filter impulse response vector $\mathbf{h} = [h(0)\ h(1)\ \cdots\ h(N \cdot I - 1)]^T$, from which the output is given by:

$$y = \mathbf{h}^H \mathbf{x}_{N,I}. \tag{12}$$

The output power of the filter is defined as:

$$E\{|y|^2\} = \mathbf{h}^H \mathbf{R} \mathbf{h}, \tag{13}$$

where $\mathbf{R} = E\{\mathbf{x}_{N,I} \cdot \mathbf{x}_{N,I}^H\}$ is the covariance matrix of $\mathbf{x}_{N,I}$. The optimal filter response is found by using the LCMV principle, that is, we design the filter to have unit gain at the harmonic frequencies while having maximum interference suppression:

$$\hat{\mathbf{h}} = \min_{\tilde{\mathbf{h}}} \tilde{\mathbf{h}}^H \mathbf{R} \tilde{\mathbf{h}}, \qquad \text{s.t. } \tilde{\mathbf{h}}^H \mathbf{a}(l\omega_1, l\alpha_1) = 1, \quad \text{for } l = 1, \cdots, L_1, \tag{14}$$

where $\mathbf{a}(\omega_1, \alpha_1) = \left[1\ e^{j\alpha_1}\ \cdots\ e^{j(I-1)\alpha_1}\right]^T \otimes \left[1\ e^{j\omega_1}\ \cdots\ e^{j(N-1)\omega_1}\right]^T$. The well-known solution to this optimization problem is:

$$\hat{\mathbf{h}} = \mathbf{R}^{-1}\mathbf{A}(\omega_1, \alpha_1)\left(\mathbf{A}^H(\omega_1, \alpha_1)\mathbf{R}^{-1}\mathbf{A}(\omega_1, \alpha_1)\right)^{-1}\mathbf{1}_{L_1},$$

with $\mathbf{A}(\omega_1, \alpha_1) = [\mathbf{a}(\omega_1, \alpha_1)\ \cdots\ \mathbf{a}(L_1\omega_1, L_1\alpha_1)]$, $\mathbf{1}_{L_1}$ being an $L_1 \times 1$ vector with all the elements equal to 1. Then, we obtain the estimate of pitch and DOA jointly by maximizing the output power as:

$$(\hat{\omega}_1, \hat{\alpha}_1) = \arg\max_{\tilde{\omega}_1, \tilde{\alpha}_1} \mathbf{1}_{L_1}^H \left[\mathbf{A}^H(\tilde{\omega}_1, \tilde{\alpha}_1)\mathbf{R}^{-1}\mathbf{A}(\tilde{\omega}_1, \tilde{\alpha}_1)\right]^{-1}\mathbf{1}_{L_1}.$$

In most cases, it is sufficient to obtain coarse estimates of both the fundamental frequency and DOA from the fine grids of their admissible ranges. However, for some applications it is necessary to refine estimates, which is achieved by using a gradient method, that is, to select the coarse estimates as initial values, and calculate the estimates iteratively as:

$$\begin{bmatrix} \hat{\omega}_1^{(i+1)} \\ \hat{\alpha}_1^{(i+1)} \end{bmatrix} = \begin{bmatrix} \hat{\omega}_1^{(i)} \\ \hat{\alpha}_1^{(i)} \end{bmatrix} + \delta \cdot \nabla J(\tilde{\omega}_1, \tilde{\alpha}_1)\big|_{\tilde{\omega}_1 = \hat{\omega}_1^{(i)}, \tilde{\alpha}_1 = \hat{\alpha}_1^{(i)}},$$

where i and $i+1$ are the iteration indexes, $\delta > 0$ is a small constant which is found using a line search algorithm [19], and $\nabla J(\tilde{\omega}_1, \tilde{\alpha}_1) = \left[\frac{\partial J(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\omega}_1}\ \frac{\partial J(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\alpha}_1}\right]^T$ is the gradient of $J(\tilde{\omega}_1, \tilde{\alpha}_1)$. According to the rules of matrix derivative [20], $\nabla J(\tilde{\omega}_1, \tilde{\alpha}_1)$ is expressed as:

$$\begin{aligned} & \nabla J(\tilde{\omega}_1, \tilde{\alpha}_1) \\ = & \begin{bmatrix} \frac{\partial J(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\omega}_1} \\ \frac{\partial J(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\alpha}_1} \end{bmatrix} \\ = & -2\mathfrak{Re}\left\{\begin{bmatrix} \mathbf{1}_{L_1}^H \mathbf{\Lambda}_1 \mathbf{A}^H(\tilde{\omega}_1, \tilde{\alpha}_1)\mathbf{R}^{-1}\mathbf{B}_{1,1}\mathbf{\Lambda}_1\mathbf{1}_{L_1} \\ \mathbf{1}_{L_1}^H \mathbf{\Lambda}_1 \mathbf{A}^H(\tilde{\omega}_1, \tilde{\alpha}_1)\mathbf{R}^{-1}\mathbf{B}_{2,1}\mathbf{\Lambda}_1\mathbf{1}_{L_1} \end{bmatrix}\right\}, \end{aligned}$$

where $\mathbf{\Lambda}_1 = \left(\mathbf{A}^H(\tilde{\omega}_1, \tilde{\alpha}_1)\mathbf{R}^{-1}\mathbf{A}(\tilde{\omega}_1, \tilde{\alpha}_1)\right)^{-1}$, and $\mathbf{B}_{1,1} = \frac{\partial \mathbf{A}(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\omega}_1}$, $\mathbf{B}_{2,1} = \frac{\partial \mathbf{A}(\tilde{\omega}_1, \tilde{\alpha}_1)}{\partial \tilde{\alpha}_1}$. In practice, we substitute the covariance matrix estimate $\hat{\mathbf{R}}$ of (3) for the unknown $\mathbf{R}$, to accomplish the estimation with the optimal filtering method. All the above steps constitute the RIAA-OF estimator.

2.3. Multi-Pitch Estimation

Now, we proceed to deal with the multi-pitch case. The covariance matrix estimation in the multi-pitch scenario is the same as that in the single-pitch case. As for the joint parameter estimation, we solve the following optimization problem, and find the peaks as the estimates of DOA and pitch of the corresponding sources:

$$(\hat{\omega}_k, \hat{\alpha}_k) = \arg\min_{\tilde{\omega}_k, \tilde{\alpha}_k} \mathbf{1}_{L_k}^H \left[\mathbf{A}^H(\tilde{\omega}_k, \tilde{\alpha}_k)\hat{\mathbf{R}}^{-1}\mathbf{A}(\tilde{\omega}_k, \tilde{\alpha}_k)\right]^{-1}\mathbf{1}_{L_k},$$

where $\mathbf{A}(\tilde{\omega}_k, \tilde{\alpha}_k) = \left[\mathbf{a}(\tilde{\omega}_k, \tilde{\alpha}_k)\ \cdots\ \mathbf{a}(L_k\tilde{\omega}_k, L_k\tilde{\alpha}_k)\right], k = 1, 2, \cdots, K$.

3. SIMULATION RESULTS

In this section, we perform Monte Carlo simulations to evaluate the joint estimation accuracy of the RIAA-OF method for pitch and DOA. The estimation performance is evaluated using the mean square error (MSE): $\mathrm{MSE}_f = \sqrt{\frac{1}{S}\sum_{k=1}^{K}\sum_{s=1}^{S}(\hat{\omega}_k^{(s)} - \omega_k)^2}$ and $\mathrm{MSE}_\alpha = \sqrt{\frac{1}{S}\sum_{k=1}^{K}\sum_{s=1}^{S}(\hat{\alpha}_k^{(s)} - \alpha_k)^2}$, with ω_k, α_k and $\hat{\omega}_k^{(s)}, \hat{\alpha}_k^{(s)}$ being the true parameter values and their estimates, respectively, and S being the number of trials. We use the number of iterations as the stopping criterion as illustrated above. All the results provided are averages of 100 independent runs.

Firstly, we provide an example of single-pitch estimation. The harmonic signal consists of $L = 4$ sinusoids with pitch $\omega_1 = \sqrt{2} \cdot 0.5$ and DOA $\alpha_1 = \sqrt{2} \cdot 0.4$. The parameter setting is listed in Table 3. Fig. 1 shows the MSEs of pitch and DOA estimates by the RIAA-OF method and its non-relaxed version (which is termed as IAA-OF method here) as well as CRLB, with $N = 20, I = 10, K_1 = 10N, K_2 = 10I$. It is seen that the MSEs of the RIAA-OF estimates are close to CRLB in the whole SNR range, and there is about 3 dB gap from CRLB. Although the MSEs of the IAA-OF estimates are also close to CRLB in the middle SNR zone, their accuracy keeps nearly constant when the SNR becomes high. This is due to the relatively coarse estimation of the covariance matrix in the case of high SNR.

DOA	l	Frequency	Amplitude	Initial Phase
	1	$\sqrt{2} \cdot 0.5$	2.0	1
$\sqrt{2} \cdot 0.4$	2	$\sqrt{2} \cdot 1.0$	1.5	2
	3	$\sqrt{2} \cdot 1.5$	2.5	3
	4	$\sqrt{2} \cdot 2.0$	4.0	4

Table 3. Simulation setting of single-pitch estimation

The next example is about multi-pitch estimation. The harmonic signal consists of $K = 2$ pitches, each with $L_k = 2$ tones. The parameter setting is listed in Table 4, and Fig. 2 shows the MSE results with $N = 20, I = 10, K_1 = 10N, K_2 = 10I$. We can also see that the MSEs of the RIAA-OF method keep about 5 dB gap from CRLB in the whole SNR range. Still, the performance of the IAA-OF method becomes worse than that of the RIAA-OF method when the SNR is high.

k	DOA	l_k	Frequency	Amplitude	Initial Phase
1	$\sqrt{2} \cdot 0.4$	1	$\sqrt{2} \cdot 0.5$	2.0	1
		2	$\sqrt{2} \cdot 1.0$	1.0	2
2	$\sqrt{2} \cdot 0.6$	1	$\sqrt{2} \cdot 0.7$	2.0	3
		2	$\sqrt{2} \cdot 1.4$	1.0	4

Table 4. Simulation setting of two-pitch estimation

4. DISCUSSION

The work presented here is focused on the joint estimation of pitch and DOA for multi-channel harmonic sinusoidal signal, whose parameter estimation is conducted with the optimal filtering method, while whose covariance matrix estimation is based on the IAA technique. Thus, our joint estimation needs no user-defined parameter. In addition, to keep the covariance matrix estimation reliability, the relaxation technique is utilized. The work of [8] is based on

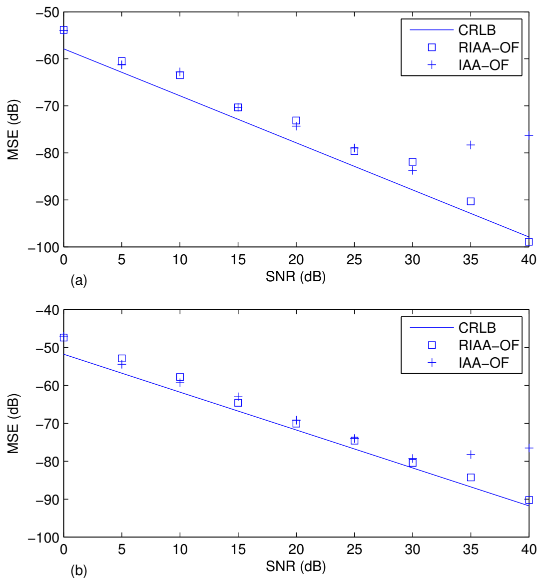

Fig. 1. MSEs of single-pitch estimation versus SNR for: (a) Pitch and (b) DOA.

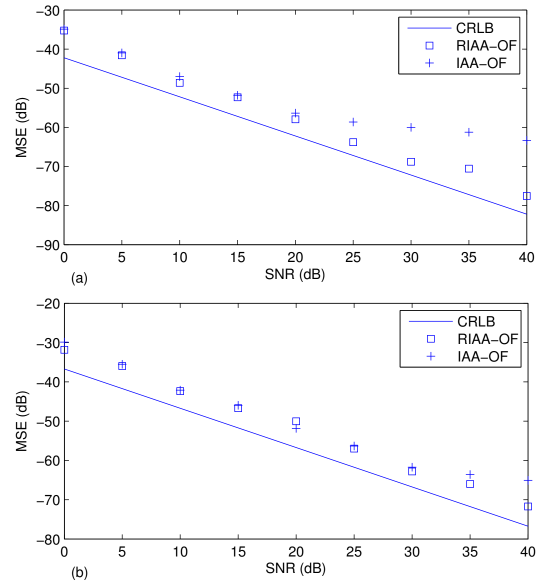

Fig. 2. MSEs of two-pitch estimation versus SNR for: (a) Pitch and (b) DOA.

the NLS method, and encounters decoupling difficulty when dealing with multi-pitch estimation [9]. The work by Zhou $et\ al.$ [14] is based on two-stage estimation instead of joint estimation, which encounters difficulty in differentiating the overlapping harmonics. The work by Jensen $et\ al.$ [10] utilizes the optimal filtering method, while it estimates the covariance matrix by data partitioning and averaging in the time domain. Consequently, the user-defined parameter such as the size of the covariance matrix should be chosen to achieve accurate estimation performance properly.

In this paper, the RIAA-OF method for joint estimation of the fundamental frequency and DOA of the multi-channel harmonic sinusoidal signal is proposed. In our approach, the covariance matrix is estimated based on the IAA and relaxation techniques. The parameters of interest are estimated with the optimal filtering method and the covariance matrix estimate, which needs no user-defined parameter. Due to the use of relaxation, the covariance matrix estimation keeps its performance when the sources are off-grid. Simulation results show that the RIAA-OF method bears better accuracy compared with its non-relaxed version for the off-grid sources. And its performance is close to CRLB for the single-pitch and two-pitch estimation. Moreover, the RIAA-OF method keeps its good performance in the existence of the difference in noise level. Further works include dealing with higher-dimensional and adaptive pitch-DOA estimation problems, detection of harmonics, and their application to speech and audio signal processing.

5. REFERENCES

[1] J.R. Jensen and M.G. Christensen, "Joint spatio-temporal filtering methods for DOA and fundamental frequency estimation," submitted for publication.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, Mar. 1986.

[3] R. Roy, T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984-995, Jul. 1989.

[4] D.H. Johnson and D.E. Dudgeon, *Array Signal Processing : Concepts and Techniques*, Englewood Cliffs, N.J. : Prentice Hall, 1993.

[5] M.G. Christensen, A. Jakobsson and S.H. Jensen, "Joint high-resolution fundamental frequency and order estimation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 5, pp. 1635-1644, Jul. 2007.

[6] M.G. Christensen, P. Stoica, A. Jakobsson and S.H. Jensen, "Multi-pitch estimation," *Signal Processing*, vol. 88, no. 4, pp. 972-983, Apr. 2008.

[7] M.G. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, [San Rafael, Calif.] : Morgan and Claypool Publishers, 2009.

[8] J.R. Jensen, M.G. Christensen and S.H. Jensen, "Statistically efficient joint DOA and pitch estimation," submitted for publication.

[9] Z. Zhou, H.C. So and F.K.W. Chan, "Optimally weighted music algorithm for frequency estimation of real harmonic sinusoids," in *Proc. ICASSP*, pp. 3537-3540, 25-30, Mar. 2012, Kyoto, Japan.

[10] J.R. Jensen, M.G. Christensen and S.H. Jensen, “Joint DOA and fundamental frequency estimation methods based on 2-D filtering,” in *Proc. EUSIPCO*, vol. 2010, pp. 2091-2095, 23-27, Aug. 2010, Aalborg, Denmark.

[11] J.X. Zhang, M.G. Christensen, S.H. Jensen and M. Moonen, “Joint DOA and multi-pitch estimation based on subspace techniques,” *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1-11, Jan. 2012.

[12] M. Xue, L. Xu and J. Li, “IAA spectral estimation: Fast implementation using the Gohberg-Semencul factorization,” *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3251-3261, Jul. 2011.

[13] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Upper Saddle River, N.J. : Prentice-Hall, 2005.

[14] Z.Zhou, M.G. Christensen and H.C. So, “Two stage DOA and fundamental frequency estimation based on subspace techniques,” in *Proc. ICSP*, Oct. 2012, Beijing, China.

[15] J. Benesty, J. Chen and Y. Huang, *Microphone Array Signal Processing*, Berlin : Springer, 2008.

[16] T. Yardibi, J. Li, P. Stoica, M. Xue and A.B. Baggeroer, “Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 425-443, Jan. 2010.

[17] J. Li and P. Stoica, “An adaptive filtering approach to spectral estimation and SAR imaging,” *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1469-1484, Jun. 1996.

[18] J. Li, D. Zheng and P. Stoica, “Angle and waveform estimation via RELAX,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 1077-1087, Jul. 1997.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*, New York : Cambridge University Press, 2004.

[20] A. Jennings and J.J. McKeown, *Matrix Computation*, Chichester, West Sussex : J. Wiley, 1992.