Joint DOA and Multi-Pitch Estimation Using Block Sparsity

Kronvall, Ted; Adalbjörnsson, Stefan Ingi; Jakobsson, Andreas

Published in:
Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on

DOI:
10.1109/ICASSP.2014.6854344

2014

JOINT DOA AND MULTI-PITCH ESTIMATION USING BLOCK SPARSITY

Ted Kronvall, Stefan Ingi Adalbjörnsson, and Andreas Jakobsson

Centre for Mathematical Sciences, Lund University, Sweden.
email: {ted, sia, aj}@maths.lth.se

ABSTRACT

In this paper, we propose a novel method to estimate the fundamental frequencies and directions-of-arrival (DOA) of multi-pitch signals impinging on a sensor array. Formulating the estimation as a group sparse convex optimization problem, we use the alternating direction of multipliers method (ADMM) to estimate both temporal and spatial correlation of the array signal. By first jointly estimating both fundamental frequencies and time-of-arrivals (TOAs) for each sensor and sound source, we then form a non-linear least squares estimate to obtain the DOAs. Numerical simulations indicate the preferable performance of the proposed estimator as compared to current state-of-the-art methods.

Index Terms— multi-pitch estimation, group sparsity, convex optimization, ADMM, direction-of-arrival, time-of-arrival.

1. INTRODUCTION

Fundamental frequency estimation is a problem occurring in a wide range of applications, maybe most notably so in various forms of speech and audio processing (see e.g. [1] and the references therein). Much of the work within the area has focused on reliable estimation of signals containing only a single fundamental frequency, or pitch, such as [2], although there has, during recent years, also been efforts on estimating the fundamental frequencies of signals containing multiple pitches, see e.g., [3–5]. Notable contributions on exploiting or extracting spatial information of sound sources have also been developed, wherein an array of receiver sensors are used to separate the sources (see, e.g., [6–8]). However, typically existing methods for joint pitch and directions-of-arrival (DOA) estimation assumes strong *a priori* knowledge of the sound signals, such as knowledge of the number of sources, and the number of harmonics present in each pitch signal, and are also often restricted to only allow for single pitch sound sources. Model order information may be found using, for example, various forms of information criteria [9, 10], or by jointly estimating the pitch and model order using an optimal filtering approach reminiscent to the one proposed in [11]. In this work, we build on the recent block sparse multi-pitch estimator presented in [5], extending it to allow for multiple sensors and thereby allow for the estimation of both the spatial and temporal correlations. Thus, we reformulate the estimation problem using a sparse signal reconstruction framework, reminiscent to the one presented in [12], extending the signal model into a large and finely spaced dictionary by grouping together candidate pitches and their harmonics in blocks. By imposing block sparseness on the solution, i.e., by forcing most of the dictionary pitches to have zero magnitude, one may obtain a solution with as many pitches with non-zero magnitude as there are pitches in the received signal. We also show that the problem belongs to the class of convex minimization problems, and propose an efficient way to solve it using an alternating direction of multipliers method (ADMM) optimization procedure. Using the proposed method, one obtains a joint DOA and pitch estimate, freely allowing for multi-pitch sound sources without imposing any assumptions on the number of sources, pitches, or the number of harmonics for each pitch. Numerical simulations illustrate the preferable performance of the proposed algorithm as compared to the non-linear least squares (NLS), and subspace (Sub) approaches, presented in [6].

This work was supported in part by the Swedish Research Council and Carl Trygger's foundation.

2. THE PITCH-DOA SIGNAL MODEL

Consider a number of acoustic signals impinging on an array of sensors, such that $\mathbf{y}(t) = \begin{bmatrix} y_0(t) & \ldots & y_{M-1}(t) \end{bmatrix}^T$, for $t = 1, \ldots, N$, where $(\cdot)^T$ denotes the transpose, and $y_m(t)$ the response of sensor m at time t. Assuming that the sound sources consist of K complex-valued harmonic pitch signals corrupted by interference and noise, the m:th sensor response may be well modeled as (see also [6])

$$y_m(t) = \sum_{k=1}^{K} \sum_{\ell=1}^{L_k} |a_{k,\ell,m}| e^{j(\omega_k \ell (t+\tau_{k,m}) + \phi_{k,\ell})} + e_m(t) \quad (1)$$

where $a_{k,l,m}$ is the complex-valued amplitude of the ℓ:th harmonic of the kth pitch as measured at sensor m, whereas L_k, ω_k, and $\phi_{k,\ell}$ are the number of harmonics, the pitch, and the phase of the ℓ:th harmonic, for the kth signal source. It should be noted that the different pitch signals may originate from the

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

same direction, thus forming a multi-pitch signal. As an example, this could occur if the sound source is a musical chord. Furthermore, let $e_m(t)$ denote the noise term at the m:th sensor, and let $\tau_{k,m}$ represent the translated (with respect to some reference point) time-of-arrival (TOA) at sensor m for sound source k. In the case of a uniform linear array (ULA), the TOA is related to the DOA as

$$\tau_{k,m} = (m-1)d\sin(\theta_k)c^{-1} \tag{2}$$

with d, c, and θ denoting the uniform distance between sensors, the wave propagation velocity, and the DOA, respectively [9]. Collecting the snapshots of the array signal, $\mathbf{y}(t) \in \mathbb{C}^M$, we introduce

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}(1) & \dots & \mathbf{y}(N) \end{bmatrix}^T \tag{3}$$

allowing (1) to be expressed in an extended form as

$$\mathbf{Y} = \sum_{p=1}^{P} \mathbf{W}_p \mathbf{A}_p + \mathbf{E} = \mathcal{W}\mathcal{A} + \mathbf{E} \tag{4}$$

where $\mathbf{E}$ denotes the combined noise term constructed in the same manner as $\mathbf{Y}$, and

$$\mathcal{W} = \begin{bmatrix} \mathbf{W}_1 & \dots & \mathbf{W}_P \end{bmatrix} \tag{5}$$

$$\mathbf{W}_p = \begin{bmatrix} \mathbf{w}_p & \dots & \mathbf{w}_p^{Q_p} \end{bmatrix} \tag{6}$$

$$\mathbf{w}_p = \begin{bmatrix} e^{j\omega_p} & \dots & e^{j\omega_p N} \end{bmatrix}^T \tag{7}$$

$$\mathcal{A} = \begin{bmatrix} \mathbf{A}_1^T & \dots & \mathbf{A}_P^T \end{bmatrix}^T \tag{8}$$

$$\mathbf{A}_p = \begin{bmatrix} \mathbf{a}_{p,1} & \dots & \mathbf{a}_{p,Q_p} \end{bmatrix}^T \tag{9}$$

$$\mathbf{a}_{p,q} = \begin{bmatrix} a_{p,q,1} & \dots & a_{p,q,M} \end{bmatrix}^T \tag{10}$$

Here, the resulting dictionary matrix, $\mathcal{W}$, is formed as a column-wise stack of a large number of Fourier matrices, $\mathbf{W}_p$ for $p = 1, \dots, P$, each containing the Fourier vectors, $\mathbf{w}_p$, for a candidate pitch ω_p and its (nominal) Q_p harmonics $q = 1, \dots, Q_k$. The dictionary matrix is thus formed such that all the candidate pitches and harmonics that coincided with the spectral components in (1) have a non-zero amplitudes, whereas all other candidate elements have zero amplitude. To allow for a sparse representation, the dictionary is expanded to include all feasible candidate elements, with the candidate pitches being assumed to be chosen to be so numerous and finely spaced that K of them may coincide with the K pitches in the signal. For these pitches, the amplitude matrix $\mathcal{A}$ will thus have non-zero blocks $\mathbf{A}_p$, and every such block will have non-zero array amplitude vectors $\mathbf{a}_{p,q} \in \mathbb{C}^{1\times M}$, for $q = 1, \dots, Q_p$. All other blocks of $\mathcal{A}$ are thus zero, making the amplitude matrix block-sparse. Note that the considered number of harmonics, Q_p, for each pitch, ω_p, needs to be chosen such that $Q_p \geq L_p, \forall p$, although being limited above by the Nyquist criterion, i.e., $\omega_p Q_p < \pi f_s$, with f_s denoting the sampling frequency. Using (4), the non-linear model in (1) is thus reformulated into a sparse structured linear system of sinusoids with respect to the complex-valued amplitudes. As the amplitudes $a_{p,q,m}$ estimate the product of the signal magnitude, the initial phase, and the TOA, one may thus, by imposing a sparse assumption on $\mathcal{A}$, form a joint estimation of both the pitch signal and the TOA.

Algorithm 1 The ADMM algorithm

1: Initiate $\mathbf{z} = \mathbf{z}_0, \mathbf{u} = \mathbf{u}_0$, and $k = 0$
2: **repeat**
3: $\quad \mathbf{z}_{k+1} = \underset{\mathbf{z}}{\operatorname{argmin}}\, f(\mathbf{z}) + \frac{\mu}{2}||\mathbf{z} - \mathbf{u}_k - \mathbf{d}_k||_2^2$
4: $\quad \mathbf{u}_{k+1} = \underset{\mathbf{u}}{\operatorname{argmin}}\, g(\mathbf{u}) + \frac{\mu}{2}||\mathbf{z_{k+1}} - \mathbf{u} - \mathbf{d}_k||_2^2$
5: $\quad \mathbf{d}_{k+1} = \mathbf{d}_k - (\mathbf{z_{k+1}} - \mathbf{u_{k+1}})$
6: $\quad k \leftarrow k+1$
7: **until** convergence

3. MULTI-PITCH ESTIMATION USING ADMM

We proceed to express the estimation of the pitches and the TOAs using a group sparse minimization. Generalizing the sparse multi-pitch scheme presented in [5] to multiple sensors, the minimization with respect to $\mathcal{A}$ is formed as

$$\underset{\mathcal{A}}{\text{minimize}} \Bigg\{ \frac{1}{2} \|\mathbf{Y} - \mathcal{W}\mathcal{A}\|_2^2 + \lambda\gamma \sum_{p=1}^{P}\sum_{q=1}^{Q_p} \|\mathbf{a}_{p,q}\|_2 + \lambda(1-\gamma)\sum_{p=1}^{P} \|\mathbf{A}_p\|_{\mathcal{F}} \Bigg\} \tag{11}$$

where two different kinds of group sparsities are imposed. Firstly, it should be noted that there is no particular reason for the sensor amplitude vector $\mathbf{a}_{k,l}$ to be sparse, i.e., if one sensor receives a certain frequency component, the remaining sensors will also receives it. However, the set of 2-norms for these vectors, which is in the second entry of the minimization, should be sparse, and ideally only have as many non-zero elements as there are sinusoids in the signal, i.e., $\sum_{k=1}^{K} L_k$. The third entry makes the solution block sparse over the candidate pitches, penalizing the number of pitches with non-zero magnitude in the signal, ideally making them as many as there are pitches in the signal, i.e., K. The user parameters λ and γ weights the fit of the solution to its sparsity and the priority between the vector and the matrix sparsity, respectively. In the case of different noise variance at each sensor in the array, the 2-norm in the first entry of the minimization criterion may be replaced with a weighted 2-norm, that is $\|\cdot\|_{\hat{\mathbf{R}}}^2$, where $\hat{\mathbf{R}} \in \mathbb{R}^{M\times M}$ denotes an estimate of the spatial noise covariance matrix. The minimization in (11) may be solved via one of the freely available interior point based solvers, such as SeDuMi [13] and SDPT3 [14], although such solvers will typically scale poorly both with increasing data

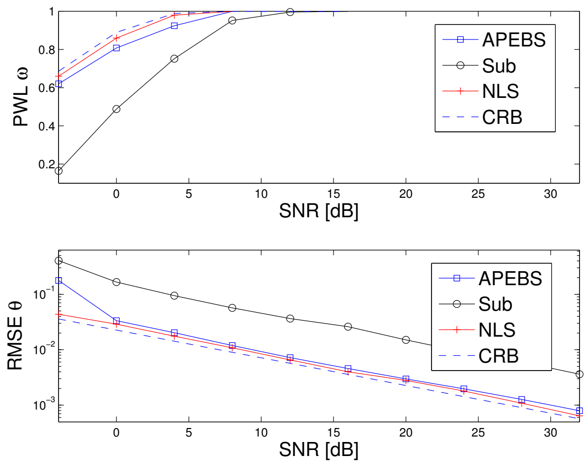

Fig. 1. The PWL and RMSE for a single-pitch signal as compared with the optimal performance of an estimator reaching the CRB.

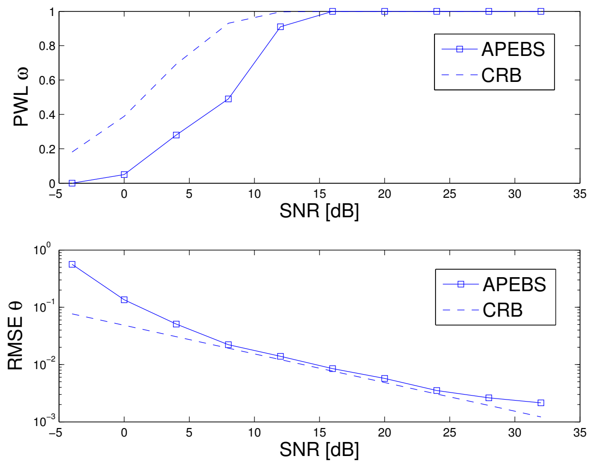

Fig. 2. The PWL and RMSE for a multi-pitch signal with two pitches, as compared to the corresponding CRB.

length, the use of a finer grid for the fundamental frequency, and the number of sensors. As a result, such a solution may in many cases be too computationally intensive to be practically useful. In order to form a more efficient implementation, the minimization in (11) is therefore reformulated using a novel ADMM formulation. For completeness and to introduce our notation, we here include an outline of the main steps involved: consider the convex optimization problem

$$\min_{\mathbf{z}} \quad f(\mathbf{z}) + g(\mathbf{z}) \tag{12}$$

where $\mathbf{z} \in \mathbb{R}^p$ is the optimization variable, with $f(\cdot)$ and $g(\cdot)$ being convex functions. If one introduces an auxiliary variable, $\mathbf{u}$, then (12) may be equivalently be expressed as

$$\min_{\mathbf{z},\mathbf{u}} \quad f(\mathbf{z}) + g(\mathbf{u}) \quad \text{subj. to} \quad \mathbf{z} - \mathbf{u} = \mathbf{0} \tag{13}$$

since at any feasible point $\mathbf{z} = \mathbf{u}$. Under the assumption that there is no duality gap, which is true for (11), one may solve the optimization problem via the dual function defined as the infimum of the augmented Lagrangian with respect to $\mathbf{x}$ and $\mathbf{z}$ [15], i.e.,

$$L_\mu(\mathbf{z},\mathbf{u},\mathbf{d}) = f(\mathbf{z}) + g(\mathbf{u}) + \mathbf{d}^T(\mathbf{z}-\mathbf{u}) + \frac{\mu}{2}||\mathbf{z}-\mathbf{u}||_2^2$$

The ADMM does this by iteratively maximizing the dual function such that at step $k+1$, one minimizes the Lagrangian for the one of the variables while holding the other fixed at its most recent value, i.e.,

$$\mathbf{z}_{k+1} = \arg\min_{\mathbf{z}} L_\mu(\mathbf{z},\mathbf{u}_k,\mathbf{d}_k) \tag{14}$$

$$\mathbf{u}_{k+1} = \arg\min_{\mathbf{u}} L_\mu(\mathbf{z}_{k+1},\mathbf{u}_k,\mathbf{d}_k) \tag{15}$$

and finally updating the dual variable by taking a gradient ascent step to maximize the dual function, resulting in

$$\tilde{\mathbf{d}}_{k+1} = \tilde{\mathbf{d}}_k - \mu(\mathbf{z}_{k+1} - \tilde{\mathbf{d}}_{k+1}) \tag{16}$$

where μ is the dual variable step size (see [15] for further details). The general ADMM steps are outlined in Algorithm 1, using the scaled version of the dual variable $\mathbf{d}_k = \tilde{\mathbf{d}}/\mu$, which is more convenient for implementation. The ADMM is as most useful when the optimizations in steps 3 and 4 of Algorithm 1 can be carried out more efficient than the original problem. For (11), note that the matrix variable is only of notational convenience and that the criterion may, using the *vec* operation, be rewritten equivalently as the norm of an affine function of the vector variable plus a sum of ℓ_2 norms of different partitions of the vector variable. Thus, convexity of the criterion follows by the convexity of norms and the composition rules for convex functions [16]. Defining

$$f(\mathcal{Z}) = \frac{1}{2}\left\|\mathbf{Y} - \mathcal{W}\mathcal{Z}\right\|_2^2 \tag{17}$$

$$g(\mathcal{U}) = \lambda\gamma\sum_{p=1}^{P}\sum_{q=1}^{Q_p}\left\|\mathbf{u}_{p,q}\right\|_2 + \lambda(1-\gamma)\sum_{p=1}^{P}\left\|\mathbf{U}_p\right\|_{\mathcal{F}} \tag{18}$$

with $\mathcal{Z}$, $\mathcal{U}$, and $\mathcal{D}$ being defined similarly to $\mathcal{A}$ in (8)-(10), leads to a quadratic problem in step 3 in Algorithm 1, with closed form solution given by

$$\mathcal{Z}_{k+1} = \left(\mu I + \mathcal{W}^H\mathcal{W}\right)^{-1}\left(\mu\left(\mathcal{U}_k - \mathcal{D}_k\right) + \mathcal{W}^H\mathbf{Y}\right)$$

whereas in step 4, by solving the sub-differential equations, one obtains $\mathbf{U}_{k+1} = \mathcal{S}\left(\mathcal{S}'\left(\mathbf{Z}_k - \mathbf{D}_k, \lambda\gamma\right), \lambda(1-\gamma)\right)$, where $\mathcal{S}(\mathbf{X},\xi) = \mathbf{X}\left(1 - \xi/||X||_F\right)^+$ for a matrix $\mathbf{X}$ and postive scalar ξ, with $(\cdot)^+$ denoting the identity function for

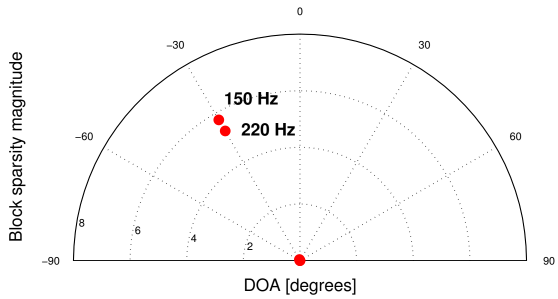

Fig. 3. Polar plot for a multi-pitch signal. The magnitude scale indicates the power of the sound source, formed as the sum of all magnitudes in the pitch. Every dot indicates the direction and power of a candidate pitch in the dictionary.

positive finite values and zero otherwise, and $\mathcal{S}'$ is defined similarly but operating on each row of $\mathbf{X}$ separately. The resulting algorithm yields a computationally efficient algorithm for finding the $\mathcal{A}$ minimizing (11). As the estimate of $\mathcal{A}$ will inherently contain estimates of the TOAs, this enables a whole range of post-processing steps to, for instance, estimate position, track, and calibrate the sources and/or sensors. Here, we limit our attention to estimating the DOAs in the case of a ULA, and to that end, consider $\hat{\mathcal{A}}$ as the solution obtained from minimizing (11). Each non-zero $\hat{\mathbf{A}}_p$ thus corresponds to the element-wise product of the amplitudes, phases, and TOA for the p:th pitch, ω_p, and its harmonics in the signal model, i.e.,

$$\hat{a}_{p,q,m} = |a_{p,q,m}| e^{j\left(\omega_p d \sin(\theta_p) c^{-1} q(m-1) + \phi_q\right)} + \epsilon_{q,m} \quad (19)$$

where $\epsilon_{q,m}$ is the (q,m):th element of $\mathbf{E}$. Then, the $Q_p M$ amplitudes in $\hat{\mathbf{A}}_p$ will only depend on a single DOA, θ_p, and the initial phases ϕ_q, for $q = 1, \ldots, Q_p$. Introducing $\eta = \omega_p d \sin(\theta_p) c^{-1}$, one may view this as a frequency estimation problem for irregularly sampled data, and find an estimate of η by solving Figure 2

$$\min_{\eta, \boldsymbol{\phi}, \mathbf{P}_p} \sum_{q=1}^{Q_p} \sum_{m=1}^{M} \left| \hat{a}_{p,q,m} - \rho_{q,m} e^{j(\eta q m + \phi_q)} \right|^2 \quad (20)$$

where $\boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_{Q_p}\right]$, and with $\mathbf{P}_p$ formed reminiscent to $\mathbf{A}_p$, containing the elements $\{\rho_{q,m}\}$, for $q = 1, \ldots, Q_p$ and $m = 0, \ldots, M-1$. It is worth noting that the the minimization may thus be viewed as a generalization of the time-varying amplitude modulation problem examined in [9, 17], to the case of several realizations of the same signal, sampled at irregular time points, and with a different initial phase for each realization. Reminiscent to the solution presented in [9, p. 186], one may thus form a closed form solution for $\phi_1, \ldots, \phi_{Q_p}$ and $\mathbf{P}_p$ as

$$\hat{\eta} = \arg\max_{\eta} \sum_{q=1}^{Q_p} \left| \sum_{m=1}^{M} (\hat{a}_{p,q,m})^2 e^{-j 2 \eta q m} \right| \quad (21)$$

This is a one-dimensional optimization problem easily solvable using a grid search on $\omega_p d \sin\left([-\pi/2, \pi/2]\right) c^{-1}$, from which the DOA may then be found as $\hat{\theta}_p = \arcsin\left(d^{-1} \omega_p^{-1} \hat{\eta} c\right)$.

4. NUMERICAL RESULTS

We proceed to evaluate the performance of the presented method using synthetic audio signals. Figure 1 shows the percentage within limits (PWL), defined as the ratio of pitch estimates within a limit of ± 0.1 Hz from the true pitch, and the root mean square error (RMSE) of the DOA, defined as

$$\mathrm{RMSE}_\theta = \sqrt{\frac{1}{nK} \sum_{k=1}^{K} \sum_{i=1}^{n} \left(\hat{\theta}_{k,i} - \theta_k\right)^2} \quad (22)$$

where n is the number of Monte Carlo (MC) simulation estimates and K is the number of pitches in the signal, for the resulting estimates. For comparison, we use the Cramér-Rao lower bound (CRB), the NLS estimator, and the Sub approach[1] (see [6] for further details on these methods). These results have been obtained using = 250 MC simulations of a single pitch signal with $\omega_1 = 220$ Hz and $L_1 = 7$ harmonics, impinging from $\theta_1 = -30°$, where both the NLS and the Sub estimators have been allowed perfect a priori knowledge of both the number of sources and their number of harmonics, whereas the proposed method, here termed the Array DOA and Pitch Estimation using Block Sparsity (APEBS), is allowed no such knowledge. As is clear from the figures, the APEBS method offers a preferable performance as compared to the Sub estimator, and only marginally worse than the NLS estimator, in spite of the latter being allowed perfect model orders information. Here, the number of sensors in the array was $M = 5$ and 20 ms of data sampled at $f_s = 8820$ Hz, i.e., $N = 176$ samples, were used. Further, we have $c = 324.3$ m/s and $d = c/f_s \approx 0.037$ m. We proceed to consider the case of multi-pitch signals impinging on the array. Measuring as in the single-pitch case, we now form a multi-pitch signal with two pitches and fundamental frequencies $[150,\ 220]$ Hz containing $[6,\ 7]$ harmonics, coming from $\theta_1 = -30°$. Figure 3 shows the an estimate of the parameters at SNR = 20 dB. The sparsity is clearly shown here as most dots, i.e., candidate pitches, are zero. Figure 2 shows the RMSE and PWL estimates, as obtained using 250 Monte Carlo simulations, clearly showing that the APEBS estimator is able to reach close to optimal performance also in this case. Here, no comparison is made with the NLS and Sub estimators of [6] as these are restricted to the single-pitch case.

[1]The authors would like to express their gratitude to the authors of [6, 18] for sharing their Matlab implementations with us.

5. REFERENCES

[1] M. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Morgan & Claypool, 2009.

[2] L. Rabiner, "On the use of autocorrelation analysis for pitch detection," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 25, no. 1, pp. 24–33, 1977.

[3] E. Benetos and S. Dixon, "Joint Multi-Pitch Detection Using Harmonic Envelope Estimation for Polyphonic Music Transcription," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 6, pp. 1111–1123, Oct. 2011.

[4] C. Lee, Y. Yang, and H. H. Chen, "Multipitch Estimation of Piano Music by Exemplar-Based Sparse Representation," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 608–618, 2012.

[5] S. I. Adalbjörnsson, A. Jakobsson, and M. G. Christensen, "Estimating Multiple Pitches Using Block Sparsity," in *38th IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing*, 2013.

[6] J.R. Jensen, M.G. Christensen, and S.H. Jensen, "Nonlinear Least Squares Methods for Joint DOA and Pitch Estimation," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 923–933, 2013.

[7] S. Gerlach, S. Goetze, J. Bitzer, and S. Doclo, "Evaluation of joint position-pitch estimation algorithm for localising multiple speakers in adverse acoustical environments," in *Proc. German Annual Conference on Acoustics (DAGA)*, Düsseldorf, Germany, 2011, vol. Mar. 2011, pp. 633–634.

[8] J. X. Zhang, M. G. Christensen, S. H. Jensen, and M. Moonen, "Joint DOA and Multi-pitch Estimation Based on Subspace Techniques," *EURASIP J. on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–11, 2012.

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

[10] M. G. Christensen, A. Jakobsson, and S. H. Jensen, "Sinusoidal Order Estimation using Angles between Subspaces," *EURASIP Journal on Advances in Signal Processing*, 2009, Article ID 948756, 11 pages.

[11] M. G. Christensen, J. .L Højvang, A. Jakobsson, and S. H. Jensen, "Joint fundamental frequency and order estimation using optimal filtering," *EURASIP Journal on Advances in Signal Processing*, vol. 13, pp. 1–18, 2011.

[12] J. J. Fuchs, "On the Use of Sparse Representations in the Identification of Line Spectra," in *17th World Congress IFAC*, Seoul,, jul 2008, pp. 10225–10229.

[13] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11-12, pp. 625–653, August 1999.

[14] R. H. Tutuncu, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming Ser. B*, vol. 95, pp. 189–217, 2003.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[17] O. Besson and P. Stoica, "Exponential signals with time-varying amplitude: parameter estimation via polar decomposition," *Signal Processing*, vol. 66, pp. 27–43, 1998.

[18] S. Karimian-Azari, J. Jensen, and C. Christensen, "Fast Joint DOA and Pitch Estimation Using a Broadband MVDR Beamformers," in *Proceedings of the European Signal Processing Conference (EUSIPCO)*, 2013.