Estimating Multiple Pitches Using Block Sparsity

Adalbjörnsson, Stefan Ingi; Jakobsson, Andreas; Christensen, Mads Græsbøll

Published in:
2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

DOI (link to publication from Publisher):
10.1109/ICASSP.2013.6638861

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

ESTIMATING MULTIPLE PITCHES USING BLOCK SPARSITY

S. I. Adalbjörnsson, A. Jakobsson*, and M. G. Christensen*†

*Dept. of Mathematical Statistics, Lund University, Sweden
†Audio Analysis Lab, Dept. of Arch., Design & Media Technology, Aalborg University, Denmark

ABSTRACT

We study the problem of estimating the fundamental frequencies of a signal containing multiple harmonically related sinusoidal signals using a novel block sparsity representation of the signal model. An efficient algorithm for solving the resulting optimization is devised exploiting an alternating directions method of multipliers (ADMM) formulation of the problem. The superiority of the proposed method, as compared to earlier methods, is demonstrated using both simulated and measured audio signals.

***Index Terms*—** Multiple pitch signals, pitch estimation, block sparsity, order estimation.

1. INTRODUCTION

The problem of estimating the fundamental frequency, or *pitch*, of a periodic waveform occurs in various forms of applications, and has received notable interest over the recent years. For example, several speech and audio problems notably depend on the initial forming of an estimate of the pitch or pitches, including problems in source separation, enhancement, compression, and classification (see, e.g., [1, 2] and the references therein). Commonly, the pitch estimate from single source signals, i.e., signals containing only a single pitch, are formed using different kinds of similarity measures, such as the cross-correlation, cepstrum, or the average squared difference function (see also [1]). Generally, such techniques suffer from not yielding unique estimates even in the ideal case, even for a single source. Recent work has aimed at instead forming the pitch estimate using second order statistics [1–4]. For multi-source signals, containing several harmonically related signals, these methods estimate each of the present pitch signals separately, forming different forms of iterative estimation schemes, typically requiring a priori knowledge of both the number of sources and the model order of each of these sources. In this work, we examine a novel method for estimating the fundamental frequencies of a signal with multiple pitches, without assuming any prior knowledge of either the number of sources present or their number of harmonics. The proposed method, here termed the Pitch Estimation using Block Sparsity (PEBS) algorithm, introduces a block sparse formulation of the estimation problem, exploiting that the signal may be viewed as formed from a dictionary consisting of a set of blocks each containing a set of harmonically related signals, for each possible fundamental frequency. As the resulting convex optimization problem is computationally cumbersome, we also derive an efficient algorithm based on the alternating directions methods of multipliers (ADMM) technique (see, e.g., [5, 6]).

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

2. BLOCK SPARSE SIGNAL MODEL

Consider a complex-valued signal, $y(n)$, consisting of K harmonically related sources with fundamental frequencies ω_k, for $k = 1, \ldots, K$, such that [1]

$$y(n) = \sum_{k=1}^{K} \sum_{l=1}^{L_k} a_{k,l} e^{j\omega_k l n} + e(n) \tag{1}$$

for $n = 1, \ldots, N$, where $a_{k,l}$ and L_k denote the (complex-valued) amplitude of the lth harmonic of the kth source, and the number of harmonically related sinusoids for the kth source, respectively, and where $w(n)$ is an additive noise term, here assumed to be a circularly symmetric complex Gaussian process. It is worth noting that due to the restriction of the allowed frequency range, the number of harmonics are restricted as a function of the fundamental frequency, such that $L_k < \lfloor 2\pi/\omega_k \rfloor$, where $\lfloor \cdot \rfloor$ denotes the round-down to nearest integer operation. Introducing

$$\mathbf{y} = \begin{bmatrix} y(1) & \ldots & y(N) \end{bmatrix}^T \tag{2}$$

where $(\cdot)^T$ denotes the transpose, allows one to express (1) succinctly in matrix notation, such that the model is represented as a linear combination of columns, where each column corresponds to one sinusoidal component, i.e.,

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{V}_k \mathbf{a}_k + \mathbf{e}_N \triangleq \mathbf{W}\mathbf{a} + \mathbf{e}_N \tag{3}$$

where $\mathbf{e}_N$ is a vector of noise terms constructed in the same manner as $\mathbf{y}$, and

$$\mathbf{W} = \left[\begin{array}{ccc} \mathbf{V}_1^T & \dots & \mathbf{V}_K^T \end{array} \right]^T \tag{4}$$

$$\mathbf{V}_k = \left[\begin{array}{cccc} \mathbf{z}_k & \mathbf{z}_k^2 & \dots & \mathbf{z}_k^{L_k} \end{array} \right]^T \tag{5}$$

$$\mathbf{a} = \left[\begin{array}{ccc} \mathbf{a}_1 & \dots & \mathbf{a}_K \end{array} \right]^T \tag{6}$$

$$\mathbf{a}_k = \left[\begin{array}{ccc} a_{k,1} & \dots & a_{k,L_k} \end{array} \right]^T \tag{7}$$

with the powers of $\mathbf{z}_k$ being evaluated element-wise, and

$$\mathbf{z}_k = \left[\begin{array}{ccc} e^{j\omega_k} & \dots & e^{j\omega_k N} \end{array} \right]^T \tag{8}$$

Reminiscent to the models considered for line spectra (see, e.g., [7–9]), we expand the matrix $\mathbf{W}$ to be formed instead over a (large) range of possible fundamental frequencies, ν_ℓ, for $\ell = 1, \dots, P$, where P denotes the total number of considered frequencies, such that the corresponding amplitude vector, $\mathbf{a}$, will have elements different from zero only for those frequencies actually coinciding with the frequencies in the signal. Thus, for the signal in (1), for each source in the signal, there will be a corresponding non-zero block in the amplitude vector, i.e., if the source has fundamental frequency ω_ℓ, the sub-block $\mathbf{a}_\ell$ will be non-zero. It should be noted that this formulation thus assumes that P is selected large enough so that the true frequencies lie at least close to the used grid over ν_ℓ, but practical experience shows that the developed algorithm is quite robust to this approximation (see also the related discussions in [8, 10]). Given the structure of (3), the resulting approximation of the signal is not only sparse, but also *block sparse*, and reminiscent of the block sparse formulations introduced in [11, 12], one may thus form an estimate of the present sources as

$$\underset{\mathbf{a}}{\text{minimize}} \ \frac{1}{2}||\mathbf{y} - \mathbf{W}\mathbf{a}||_2^2 + \lambda||\mathbf{a}||_1 + \alpha \sum_{k=1}^{P} ||\mathbf{a}_k||_2 \tag{9}$$

where $||\cdot||_p$ denotes the ℓ_p norm, and with α and λ being tuning parameters that control the relative importance of the sparsifying regularizers and the 2-norm fitting term, discussed further below. It should be stressed that the number of harmonics of each source, L_k, is not known, and to be able to use the presented sparse approximation model, one needs thus to set a maximum allowed number of harmonics for all sources, L_{max}. However, as the number of harmonics for each source is upper limited by the Nyquist frequency, the harmonically related sub-blocks, $\mathbf{V}_k$, will typically contain fewer harmonically related components for the higher frequencies, thereby causing different blocks to be of different sizes. Since some blocks will overlap in all elements except the first one, e.g., blocks corresponding to a particular fundamental frequency and the double of that frequency, the minimization in (9) will then in all cases prefer the block corresponding to the lower frequency. In order to resolve this problem, we introduce a normalization in the minimization, such that the blocks are given comparable weights, instead forming the minimization

$$\underset{\mathbf{a}}{\text{minimize}} \ \frac{1}{2}||\mathbf{y} - \mathbf{W}\mathbf{a}||_2^2 + \lambda||\mathbf{a}||_1 + \alpha \sum_{k=1}^{P} \sqrt{\Delta_k}||\mathbf{a}_k||_2 \tag{10}$$

where Δ_k denotes the number of harmonics in block k given as $\min(L_{\text{max}}, L_k)$. Considering that the signals of interested are only approximately sparse in $\mathbf{W}$, and as the columns of $\mathbf{W}$ are correlated, one cannot expect the resulting (block) pseudo spectral solution, formed over the peaks of the 2-norm of the estimated amplitudes, $||\mathbf{a}_k||_2$, to have exactly as many non-zero blocks as there are sources present in the signal. In order to determine the number of sources present, we therefore form a BIC-style criterion, where (cf. [8, 13])

$$\text{BIC}_k = 2N \ln(\hat{\sigma}_k^2) + (5H_k + 1)\ln(N) \tag{11}$$

with $\hat{\sigma}_k^2$ dentoing the variance of the estimation residual when subtracting the H_k harmonics corresponding to the k dominant peaks of the pseudo spectrum, such that the number of sources are selected as

$$\hat{K} = \underset{k \in [1, K_{\text{max}}]}{\text{argmin}} \ \text{BIC}_k \tag{12}$$

where K_{max} denotes the number of peaks present in the pseudo-spectra.

Algorithm 1 The general ADMM algorithm

Initiate $\mathbf{z} = \mathbf{z}_0, \mathbf{u} = \mathbf{u}_0$, and $k = 0$
repeat
  $\mathbf{z}_{k+1} = \underset{\mathbf{z}}{\text{argmin}} \ f_1(\mathbf{z}) + \frac{\mu}{2}||\mathbf{G}\mathbf{z} - \mathbf{u}_k - \mathbf{d}_k||_2^2$
  $\mathbf{u}_{k+1} = \underset{\mathbf{u}}{\text{argmin}} \ f_2(\mathbf{u}) + \frac{\mu}{2}||\mathbf{G}\mathbf{z}_{\mathbf{k+1}} - \mathbf{u} - \mathbf{d}_k||_2^2$
  $\mathbf{d}_{k+1} = \mathbf{d}_k - \mathbf{G}\mathbf{z}_{\mathbf{k+1}} - \mathbf{u}_{\mathbf{k+1}}$
  $k \leftarrow k + 1$
until convergence

3. AN EFFICIENT ADMM IMPLEMENTATION

The convex minimization in (10) can be readily implement using various freely available solver, such as CVX [14, 15], SeDuMi [16] and SDPT3 [17], although such solvers will in many cases be too computationally intensive to be useful for larger data sets or in situations where one has real-time limitation on the computation time. In order to form a more efficient implementation, we therefore reformulate the minimization in (10) using the ADMM approach (see, e.g., [5]), which solves general convex optimization problems of the form

$$\underset{\mathbf{z}}{\text{minimize}} \quad f_1(\mathbf{z}) + f_2(\mathbf{G}\mathbf{z}) \tag{13}$$

where $\mathbf{z} \in \mathbf{R}^p$ is the optimization variable, $f_1(\cdot)$ and $f_2(\cdot)$ are convex functions, and $\mathbf{G} \in \mathbf{R}^{N \times p}$ is a known matrix.

Algorithm 2 Block Sparse Pitch Estimation via ADMM

1: Initiate $\mathbf{z} = \mathbf{z}_0, \mathbf{u} = \mathbf{u}_0$ and $k := 0$
2: **repeat**
3: $\quad \mathbf{z}_{k+1} = \left[\mathbf{A}^H\mathbf{A} + 2\mathbf{I}\right]^{-1}\left(\mathbf{A}^H\xi_k^{(1)} + \xi_k^{(2)} + \xi_k^{(3)}\right)$
4: $\quad \mathbf{u}_{k+1}^{(1)} = \frac{\mathbf{y}-\mu\left(\mathbf{A}\mathbf{z}_{k+1}-\mathbf{d}_k^{(1)}\right)}{1+\mu}$
5: $\quad \mathbf{u}_{k+1}^{(2)} = \Psi\left(\mathbf{z}_{k+1} - \mathbf{d}_k^{(2)}, \frac{\lambda}{\mu}\right)$
6: $\quad$ **for** $m = 1, \ldots, P$ **do**
7: $\quad\quad \mathbf{u}_{k+1}^{(3,m)} = \Psi^V\left(\mathbf{z}_{k+1}^{(m)} - \mathbf{d}_k^{(3,m)}, \frac{\alpha_m}{\mu}\right)$
8: $\quad$ **end for**
9: $\quad \mathbf{d}_{k+1} = \mathbf{d}_k - (\mathbf{G}\mathbf{z}_{k+1} - u_{k+1})$
10: $\quad k \leftarrow k+1$
11: **until** convergence

This is accomplished by introducing the auxiliary variable, $\mathbf{u}$, allowing (13) to be decomposed into two, hopefully simpler, optimization problems, i.e.,

$$\begin{aligned} \underset{\mathbf{z},\,\mathbf{u}}{\text{minimize}} \quad & f_1(\mathbf{z}) + f_2(\mathbf{u}) + \frac{\mu}{2}||\mathbf{G}\mathbf{z} - \mathbf{u}||_2^2 \\ \text{subject to} \quad & \mathbf{G}\mathbf{z} - \mathbf{u} = \mathbf{0} \end{aligned} \tag{14}$$

where μ is a tuning variable which needs to be set by the user, further discussed below. Note that the added quadratic term will not affect the optimal value since at any feasible point it will be equal to zero. Solving (14) by alternatively maximizing over $\mathbf{z}$ and $\mathbf{u}$ leads to the general ADMM algorithm outlined as Algorithm 1. Clearly, this reformulation is only relevant when the optimizations in steps 3 and 4 can be carried out easily as compared to the original problem. One possibility to reformulate (10) in this fashion would be to choose $f_1(\cdot)$ as the 2-norm fitting term and $f_2(\cdot)$ as the sum of the sparse regularization term, although this would then lead to a difficult optimization problem in step 4. Herein, we instead exploit the recent idea introduced in [6], where by a clever choice of functions the $f_1(\cdot)$ and $f_2(\cdot)$, one can extend (13) to a minimization of a sum of B convex functions, i.e.,

$$\underset{\mathbf{z}}{\text{minimize}} \quad \sum_{k=1}^{B} g_k(\mathbf{H}\mathbf{z}) \tag{15}$$

where B is the number of functions, $\mathbf{H}_k \in \mathbf{R}^{N\times p}$ are known matrices, and $g_k(\cdot)$ convex functions. This is accomplished by setting $f_1(\mathbf{z}) = 0$, and

$$f_2(\mathbf{G}\mathbf{u}) = \sum_{k=1}^{B} g_k(\mathbf{G}\mathbf{u}) = \sum_{k=1}^{B} g_k(\mathbf{H}_k\mathbf{u}^{(k)}) \tag{16}$$

where

$$\mathbf{G} = \left[\begin{array}{ccc} \mathbf{H}_1^T & \ldots & \mathbf{H}_K^T \end{array}\right]^T \tag{17}$$

$$\mathbf{u} = \left[\begin{array}{ccc} (\mathbf{u}^{(1)})^T & \ldots & (\mathbf{u}^{(K)})^T \end{array}\right]^T \tag{18}$$

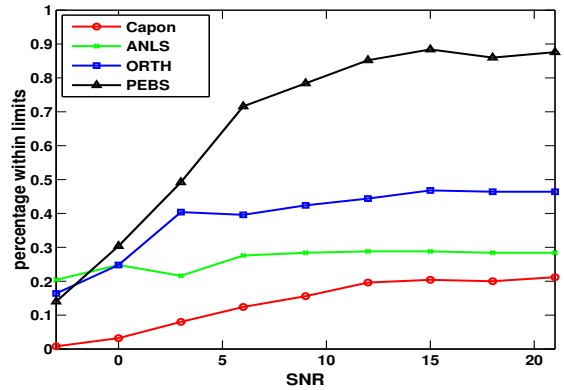

Fig. 1. Ratio of estimated pitches where both fundamental frequencies lie at most 0.001 from the ground truth, plotted as a function of SNR for $\alpha = \lambda = 0.5\chi$ and $\chi = 0.2$.

thereby allowing step 4 in Algorithm 1 to be decomposed into B independent optimization problems. Rewriting (10) on the form in (15), noting that for this case, $B = 3$, and

$$f_2(\mathbf{G}\mathbf{u}) = \frac{1}{2}||\mathbf{u}^{(1)} - \mathbf{y}|| + \lambda||\mathbf{u}^{(2)}||_1 + \alpha\sum_{k=1}^{P}\sqrt{\Delta_k}||\mathbf{u}_k^{(3,k)}||_2$$

where $\mathbf{G} = \left[\begin{array}{ccc} \mathbf{A}^T & \mathbf{I} & \mathbf{I} \end{array}\right]^T$, and

$$\mathbf{u} = \left[\begin{array}{ccc} (\mathbf{u}^{(1)})^T & (\mathbf{u}^{(2)})^T & (\mathbf{u}^{(3)})^T \end{array}\right]^T$$
$$\mathbf{u}^{(3)} = \left[\begin{array}{ccc} (\mathbf{u}^{(3,1)})^T & \ldots & (\mathbf{u}^{(3,P)})^T \end{array}\right]^T$$

This implies that step 3 in Algorithm 1 can be solved as

$$\begin{aligned} \mathbf{z}_{k+1} &= \underset{\mathbf{z}}{\text{argmin}}\,||\mathbf{G}\mathbf{z} - \mathbf{u}_k - \mathbf{d}_k||_2^2 \\ &= \left[\mathbf{A}^H\mathbf{A} + 2\mathbf{I}\right]^{-1}\left(\mathbf{A}^H\xi_k^{(1)} + \xi_k^{(2)} + \xi_k^{(3)}\right) \end{aligned}$$

where $\mathbf{d}$ is decomposed in the same manner as $\mathbf{u}$, and

$$\xi_k^{(m)} \triangleq \mathbf{u}_k^{(m)} - \mathbf{d}_k^{(m)}$$

for $m = 1, 2, 3$. Step 4 in Algorithm 1 thereby decomposes into three different and decoupled optimization problems; firstly, for the first block,

$$\begin{aligned} \mathbf{u}_{k+1}^{(1)} &= \underset{\mathbf{u}}{\text{argmin}}\,||\mathbf{u} - \mathbf{y}||_2^2 + \frac{\mu}{2}||\mathbf{A}\mathbf{z}_{\mathbf{k+1}} - \mathbf{u} - \mathbf{d}_k^{(1)}||_2^2 \\ &= \frac{\mathbf{y} - \mu\left(\mathbf{A}\mathbf{z}_{k+1} - \mathbf{d}_k^{(1)}\right)}{1+\mu} \end{aligned}$$

Secondly, the second block decouples into LP one-dimensional optimizations, whose solution yields

$$\begin{aligned} \mathbf{u}_{k+1}^{(2)} &= \underset{\mathbf{u}}{\text{argmin}}\,\lambda||\mathbf{u}||_1 + \frac{\mu}{2}||\mathbf{A}\mathbf{z}_{\mathbf{k+1}} - \mathbf{u} - \mathbf{d}_k^{(2)}||_2^2 \\ &= \Psi\left(\mathbf{z}_{k+1} - \mathbf{d}_k^{(2)}, \frac{\lambda}{\mu}\right) \odot \left(\mathbf{z}_{k+1} - \mathbf{d}_k^{(2)}\right) \end{aligned} \tag{19}$$

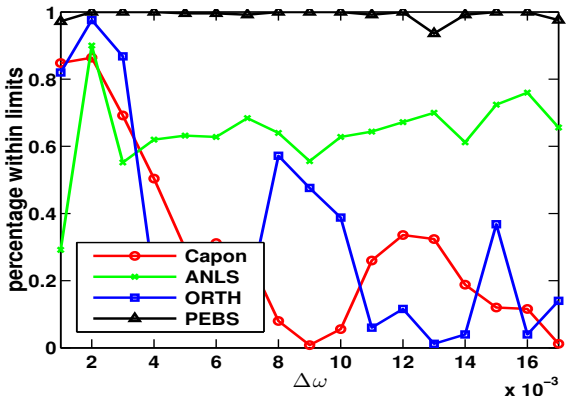

Fig. 2. Ratio of estimated pitches where both fundamental frequencies lie at most 0.001 from the ground truth, plotted as a function of $\Delta\omega$, for $\alpha = \lambda = 0.5\chi$ and $\chi = 0.2$.

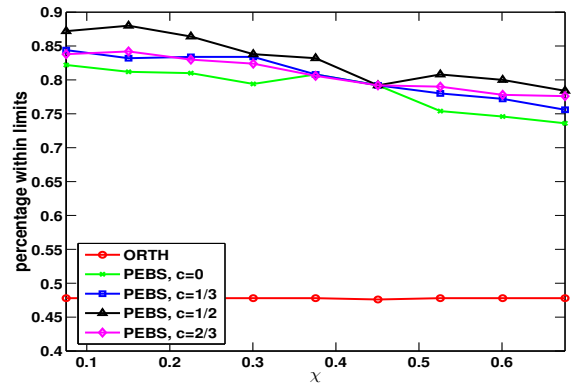

Fig. 3. Ratio of estimated pitches where both fundamental frequencies lie at most 0.001 from the ground truth, plotted as a function of χ for $\alpha = c\chi$, $\lambda = (1-c)\chi$, for $c \in \{0,\ 1/2,\ 1/3,\ 2/3\}$.

where $\odot$ denotes the Hadamard product and Ψ is a shrinkage function, defined as

$$\Psi(\mathbf{a}, \gamma) = \frac{\max(|\mathbf{a}| - \gamma, 0)}{\max(|\mathbf{a}| - \gamma, 0) + \gamma} \tag{20}$$

with both the max function and the absolute value being interpreted element-wise. Finally, the third block can be found to decouple into P optimization problems, each of the form

$$\begin{aligned}\mathbf{u}_{k+1}^{(3,m)} &= \underset{\mathbf{u}}{\operatorname{argmin}}\, \alpha_m ||\mathbf{u}||_1 + \frac{\mu}{2}||\mathbf{z}_{k+1}^{(m)} - \mathbf{u} - \mathbf{d}_k^{(3,m)}||_2^2 \\ &= \Psi\left(||\mathbf{z}_{k+1}^{(m)} - \mathbf{d}_k^{(3,m)}||_2, \frac{\alpha_m}{\mu}\right)\left(\mathbf{z}_{k+1}^{(m)} - \mathbf{d}_k^{(3,m)}\right)\end{aligned}$$

The resulting ADMM algorithm for the block sparse pitch estimation problemis summarized in Algorithm 2.

4. NUMERICAL RESULTS

We proceed to examine the robustness and performance of the proposed PEBS estimator, using both simulated and real audio signals, comparing with the Capon, ANLS, and ORTH algorithms [1]. Initially, examining simulated signals, we form a signal consisting of two sources with the fundamental frequencies, ω_1 and ω_2, drawn uniformly on $[0.02\pi, 0.2\pi]$ in such a way that the minimum difference between the frequencies is at least 1/30 of the frequency range. The performance of the estimates for the different algorithms are computed using 250 Monte-Carlo simulations, wherein the number of harmonics are selected uniformly over $[3, \min(\text{floor}(\pi/\omega_i), 10)]$ in each simulation, in order to ensure that all frequencies are below the Nyquist limit, and with amplitudes drawn as $a_{i,k} \sim \mathbf{N}(0,1)$, except for the amplitude of the fundamental frequency which is set to one to avoid ambiguity in the true fundamental frequency. The signal to noise ratio (SNR), defined as $10log_{10}(||\mathbf{y}||_2/||\mathbf{w}||_2)$, is set to 15. To ensure the best possible performance, the reference methods are allowed perfect a priori knowledge of both the number of present sources and their respective number of harmonics, whereas the PEBS estimator is only given that the maximum number of harmonics for any present source is 10. Figure 1 shows the percentage of estimates where the estimated pitches are ± 0.001 from the true value, for varying SNR, clearly showing the preferable performance of the proposed PEBS algorithm. To examine the effects of closely spaced fundamental frequencies, we next consider the pitches $\omega_1 = 0.06\pi$ and $\omega_2 = \omega_1 + \Delta\omega$. Here, to clarify the effects of the source separation, $L_1 = 7$ and $L_2 = 5$, $\alpha_{k,l} = 1$, $\forall k, l$, although with random phase. Figure 2 shows the resulting performance as a function of $\Delta\omega$, again confirming the preferable performance of the proposed estimator. In the above simulations, we used $\alpha = c\chi$, $\lambda = (1-c)\chi$, for $c = 0.5$ and $\chi = 0.2$. We proceed to examine the robustness to the selection of these user parameters. Figure 3 illustrates the resulting performance as a function of χ for different values of c, for SNR=15. To increase clarity, the results are here only compared to the ORTH estimator, which exhibited the best performance of all the reference methods. As shown in the figure, the performance is quite insensitive to the choice of the user parameters, although their relative ratio, typically estimated using cross validation, does make some difference in performance. Finally, we examine real audio using the Sound Quality Assessment Material recordings for subjective tests (SQAM)[1], where we chose the file `21.flac`, which is a sequence of tones from a trumpet. All the discussed estimators were able to estimate the fundamental frequency for each considered tone, but after mixing two tones, equating their power, only the PEBS and ORTH algorithms managed to accurately estimate the pitch frequencies.

[1] Available at `http://tech.ebu.ch/publications/sqamcd`

5. REFERENCES

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

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

[3] Z. Zhou, H. C. So, and F. K. W. Chan, "Optimally weighted music algorithm for frequency estimation of real harmonic sinusoids," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, March 25-30 2012.

[4] J. X. Zhang, M. G. Christensen, S. H. Jensen, and M. Moonen, "A robust and computationally efficient subspace-based fundamental frequency estimator," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 3, pp. 487 –497, march 2010.

[5] 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.

[6] M. A. T. Figueiredo and J. M. Bioucas-Dias, "Algorithms for imaging inverse problems under sparsity regularization," in *Cognitive Information Processing (CIP), 2012 3rd International Workshop on*, may 2012, pp. 1 – 6.

[7] 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.

[8] P. Stoica and P. Babu, "Sparse Estimation of Spectral Lines: Grid Selection Problems and Their Solutions," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 962–967, Feb. 2012.

[9] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 101 – 111, jan. 2003.

[10] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182 –2195, may 2011.

[11] M. Yuan and Y. Lin, "Model Selection and Estimation in Regression with Grouped Variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[12] J. Friedman, T. Hastie, and R. Tibshirani, "A Note on the Group Lasso and a Sparse Group Lasso," 2010.

[13] P. Stoica and Y. Selén, "Model-order selection — A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, July 2004.

[14] CVX Research, Inc., "CVX: Matlab software for disciplined convex programming, version 2.0 beta," `http://cvxr.com/cvx`, Sept. 2012.

[15] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008, http://stanford.edu/∼boyd/graph_dcp.html.

[16] 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.

[17] 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.

[18] F. Huang and T. Lee, "Pitch estimation in noisy speech using accumulated peak spectrum and sparse estimation technique," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 99 –109, jan. 2013.