

ON THE USE OF THE GLOBAL MATCHED FILTER FOR DOA ESTIMATION IN THE PRESENCE OF CORRELATED WAVEFORMS.

Jean-Jacques FUCHS

IRISA/INRIA/Université de Rennes I
Campus de Beaulieu - 35042 Rennes Cedex - France
fuchs@irisa.fr

ABSTRACT

The correlation between waveforms arriving at an array of sensors, which is generally a consequence of multipath propagation, can drastically degrade the performance of source localization and detection schemes. Most approaches proposed so far to handle this difficulty, consist in performing some preprocessing aiming to decorrelate the signals and have limited effects. They are moreover restricted to specific array geometries. We present a technique that works for any geometry and that is able to handle coherent signals as well. It simply relies on no prior assumption on spatial stationarity.

Index Terms— DOA estimation, correlated waveforms, global matched filter, generalized maximum likelihood ratio

1. INTRODUCTION

Correlation between waveforms generally appears in the case of multipath propagation and can severely degrade the performance of an antenna array system. Most source localization techniques, including those that are eigen structure-based such as MUSIC [1, 2], encounter difficulties only when the signals are perfectly correlated or coherent. But in practice, however, significant degradations arise when the signals are highly correlated.

Though of high practical importance, the highly correlated or the fully correlated (coherent) case did not receive the attention it deserves [3] and in general only some preprocessing schemes whose aim is to decorrelate the signals are proposed to improve the performance or reduce the degradations of techniques that are build and developed for uncorrelated sources. A preprocessing technique known as spatial smoothing has been proposed in [4] and further investigated in [5, 6]. The main drawback of this approach, besides being only applicable to regular array geometries, is the reduction of the effective array aperture and hence lower resolution and accuracy. Another technique known as redundancy averaging [7, 8] is known to induce bias in the bearing estimates.

In the sequel we consider a technique that works for any geometry and that is able to handle correlated as well as coherent waveforms. It simply relies on no prior assumption.

2. THE MODEL

We consider a uniform linear array of M omnidirectional sensors separated by one half wave-length at the temporal frequency of interest, but the reader can easily check that the method extends to any geometry and this choice is simply done to fix ideas and ease the exposition. There are $m < M$ potentially correlated plane waves that impinge on the array and for simplicity we assume that the array and sources are coplanar. We denote $d(\theta)$ the steering vector associated with a bearing θ with respect to broadside.

The observed vector of sensor-outputs, i.e. snapshot, at time k is then modeled as

$$X_k = D(\theta)s_k + \mathbf{n}_k$$

where $D(\theta) = [d(\theta_1) \ d(\theta_2) \ \dots \ d(\theta_m)] \in C^{M \times m}$, $s_k \in C^{m \times 1}$ are the Gaussian signal vector at time k and $\mathbf{n}_k \in C^{M \times 1}$ is Gaussian spatially white noise with power σ_0^2 . More precisely $s_k \sim CN(m, 0, P)$ and $\mathbf{n}_k \sim CN(M, 0, \sigma_0^2 I_M)$ where $CN(M, 0, R)$ denotes a complex (circular) Gaussian distribution of dimension M , mean 0 and covariance matrix R . It follows that the snapshots X_k are themselves $CN(M, 0, R)$ with covariance matrix

$$R = D(\theta)PD(\theta) + \sigma_0^2 I_M. \quad (1)$$

If the m impinging paths are spatially uncorrelated, their covariance matrix P is diagonal with diagonal components σ_p^2 and we define the signal to noise ratio (SNR) of the p -th source to be equal to σ_p^2/σ_0^2 . If the paths are spatially correlated, P is no longer diagonal and in the case of two correlated paths, the covariance matrix P becomes

$$P = \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho_{1,2} e^{j\varphi_{1,2}} \\ \sigma_1 \sigma_2 \rho_{1,2} e^{-j\varphi_{1,2}} & \sigma_2^2 \end{bmatrix} \quad (2)$$

with $\rho_{1,2} \leq 1$ the modulus of the correlation between the two signals. When $\rho_{1,2} = 1$ they are fully correlated or coherent and the rank of P becomes equal to one.

Note that spatial correlation is induced by temporal correlation between the signals carried by the waveforms. Typically a temporally correlated signal emitted by a source arrives at the array through different paths, with different delays and the temporal correlation induces spatial correlation.

It may nevertheless happen that, even in that case, the snapshots X_k are temporally uncorrelated. If this is the case all the information in the, say N , snapshots X_k is contained in their sample covariance matrix

$$\hat{R} = \frac{1}{N} \sum_1^N X_k X_k^H.$$

One can then build a DOA estimation scheme using \hat{R} without loosing any information. This is the case of MUSIC, for instance. In the presence of correlation, some schemes (for instance, sub-aperture or redundancy averaging) have been proposed to somehow decorrelate the signals but their efficiency is quite limited and, in any case, they cannot handle the case of coherent paths.

3. THE PROPOSED TECHNIQUE

3.1. The basic version

We propose a high resolution DOA estimation scheme that works whatever the correlation between the sources and even somehow detects the number m of paths that are present. It is a sparse representations technique, we called the Global Matched Filter (GMF) in [9, 10], that can also be seen as a model-fitting approach or an inverse-problem solver. It works whenever one wants to decompose a signal/vector into the sum of a small number of vectors belonging to a parametrized family of vectors [9] (manifold). In an uncorrelated sources context, we proposed to apply it to a set of beamformer outputs and it generally outperforms MUSIC (at a higher computational cost though) and attains performances close to the Cramer-Rao bounds [10].

In the present correlated or coherent plane waves context, we propose to apply it to the first column of \hat{R} , we denote \mathbf{y} in the sequel. If the sources are uncorrelated, R is Toeplitz and the first column of \hat{R} after averaging of the diagonals is a sufficient statistic whose covariance matrix can be evaluated. If the sources are partially or fully correlated, R is no longer Toeplitz and averaging the diagonals of \hat{R} though being one of the tricks (redundancy averaging) proposed to decorrelate the sources has no justification, and we thus stay with \mathbf{y} being the plain first column of \hat{R} .

For two correlated plane waves in white noise (1,2), the first column of R , we denote \mathbf{Y} , is given by

$$\mathbf{Y} = (\sigma_1^2 + \sigma_1\sigma_2\rho_{1,2}e^{j\varphi_{1,2}})d(\theta_1) + (\sigma_2^2 + \sigma_1\sigma_2\rho_{1,2}e^{-j\varphi_{1,2}})d(\theta_2) + \sigma_0^2\mathbf{e}_1 \quad (3)$$

with \mathbf{e}_1 the first column of the identity matrix. One can check that quite generally for any number of paths uncorrelated or correlated partially or fully, \mathbf{Y} can be decomposed uniquely into the sum of a weighted sum of the sought-for steering vectors and the \mathbf{e}_1 vector. The same holds then approximately for the first column of \hat{R} , we will denote \mathbf{y} , i.e., \mathbf{y} has

a sparse representation on the redundant basis formed of \mathbf{e}_1 and a large number of equispaced steering vectors with bearings on a regular grid. Typically if one allows for some residual error between \mathbf{y} and its sparse reconstruction an existing steering vector $d(\theta_p)$, whose bearing θ_p does not belong to the regular grid, will be approximated by the two neighboring bearings on the grid.

If A denotes the $M \times K$ matrix whose columns are the $K-1$ equispaced steering vectors plus the column $\sqrt{M}\mathbf{e}_1$ (so that all the columns of A have the same euclidean norm: \sqrt{M}), the GMF, in its most basic form, amounts to solve

$$\min_z \frac{1}{2} \|Az - \mathbf{y}\|_2^2 + h \|z\|_1, \quad h > 0 \quad (4)$$

with $z \in C^K$, $\|z\|_1 = \sum |z_k|$, $\|z\|_2^2 = z^H z = \sum z_k \bar{z}_k$ and h a positive real to be fixed.

3.2. A more elaborate version

To improve the performance of the method, one can take into account the statistical properties of the observations in \mathbf{y} . Under the current assumptions, $N\hat{R}$ follows a complex Wishart distribution $CW(N, M, R)$. One can also observe that \hat{R} is an estimate of the covariance matrix of $X_k \in CN(M, 0, R)$ and that $N\hat{R}$ is then a sample of $CW(N, M, R)$. From the first of these two observations it is easy to deduce the statistical properties of the components of \hat{R} , with $\tilde{r}_{i,j} = \hat{r}_{i,j} - r_{i,j}$, one has

$$E(\tilde{r}_{i,j} \tilde{r}_{m,n}) = \frac{1}{N} r_{i,n} r_{m,j} \quad (5)$$

where N is the number of snapshots. It then follows easily from this formula that Σ , the covariance matrix of \mathbf{y} the first column of \hat{R} , an estimate of \mathbf{Y} the first column of R , is

$$\Sigma = \frac{1}{N} r_{1,1} R. \quad (6)$$

It is then natural to premultiply both A and \mathbf{y} in (4) by $\Sigma^{-\frac{1}{2}}$ to whiten the observations. Since R and thus Σ are not known, in practice we replace R by its estimate \hat{R} , the components in the resulting observation vector, say, $\mathbf{y}_w = \hat{\Sigma}^{-\frac{1}{2}}\mathbf{y}$ are then uncorrelated and of unit variance, at least asymptotically in N .

The more elaborate version which should somehow be close to a maximum likelihood approach applied to the observations in \mathbf{y} amounts to replace (4) by

$$\min_z \frac{1}{2} \|\bar{A}_w z - \mathbf{y}_w\|_2^2 + h \|z\|_1, \quad h > 0 \quad (7)$$

with $\mathbf{y}_w = \hat{\Sigma}^{-\frac{1}{2}}\mathbf{y}$, $A_w = \hat{\Sigma}^{-\frac{1}{2}}A$ and \bar{A}_w represents the matrix A_w with columns normalized to one in ℓ_2 norm.

The number of waveforms present, their *amplitudes* and their bearings are then deduced from the optimum z of (4) or (7) as explained in the following section.

4. IMPLEMENTATION ISSUES

For a well chosen h , the optimal z has typically about $2m + 1$ nonzero components, i.e., a pair of neighboring nonzero components for each of the waveforms that are present and an additional component to model the noise contribution. The estimate of the number of waveforms present is then given by the number of clusters of nonzero components, the *amplitudes* is estimated by the sum of the weights in its associated cluster and the bearing estimate is obtained by linear interpolation of the associated *indices*. From (3) it follows that there exists no simple relation between the *amplitudes* of the steering vectors and the sources powers, we will not investigate this point further.

4.1. Setting the threshold

We consider the elaborate version of the proposed technique. The value of h will have a major influence upon the detection part of the whole procedure. To make this more apparent we introduce the dual of the criterion (7) which is

$$\min_z \|\bar{A}_w z\|_2 \quad \text{under} \quad \|\bar{A}_w^H (\bar{A}_w z - \mathbf{y}_w)\|_\infty \leq h \quad (8)$$

where A^H is the conjugate transpose of A and $\|z\|_\infty$ denotes the ℓ_∞ -norm of z , $\max_k |z_k|$. This dual which is equivalent to (7), is important because it allows to understand the role played by h [10]. One can already observe from the constraints in (8) that, as h increases, so does the size of the admissible domain. More precisely, for h larger than $\|\bar{A}_w^H \mathbf{y}_w\|_\infty$, the point $z = 0$ becomes admissible and is then the optimum. Roughly speaking, the larger h , the sparser the optimal z and vice versa.

If the value of h is too large, the procedure may not detect weak paths. On the other hand, if h is too small, there may appear many false alarms, i.e., the procedure will detect and identify paths that do not exist. We will thus aim to set h to achieve a given probability of false alarms.

To do so we observe that in the ideal case where the optimal z and associated $\bar{A}_w z$ represents exactly the waveforms and the noise contribution, i.e., when $\bar{A}_w z = \mathbf{Y}_w$, the whitened residual vector, say $r = \bar{A}_w z - \mathbf{y}_w = \mathbf{Y}_w - \mathbf{y}_w$ is, asymptotically in N , $N(0, I)$ (see section 3.2). It follows that, in this ideal situation, the vector $\bar{A}_w^H r$ in the dual constraint is a vector filled with K scalar (dependent) complex random Gaussian variables build upon the M independent scalar standard Gaussian random variables present in r . More precisely, each component in $\bar{A}_w^H r$ is a zero mean random variable whose real and imaginary parts are zero mean correlated Gaussian random variables whose variances sum to one.

In a detection context, one would like to choose h so that the probability for the maximum of the K random variable in $\bar{A}_w^H r$, with $r \in N(0, I_M)$, to exceed h in modulus be p , the chosen probability of false alarm. This is quite difficult in this realization-dependent whitening context and would require further investigations.

To nevertheless fix ideas, we propose to take

$$h \simeq \frac{1}{2} \sqrt{\frac{1}{2}} \sqrt{t} \quad \text{with} \quad Pr(\chi_2^2 > t) \simeq \frac{p}{M},$$

which would be close to the optimal value if the real and imaginary parts of the K complex random variables in $\bar{A}_w^H r$ had identical variance (equal to $1/2$) and were uncorrelated. In the simulations, we perform in Section 5, we take $p = .1$, i.e., a 10 percent probability of false alarm in the formula and, with $M = 10$, we get $h = 1.07$.

4.2. Getting the estimates

Having obtained the optimum of (4) or (7), denoted z below, we now explain how to deduce from z the estimates of the amplitudes, we will denote A_j , and bearings, θ_j .

We first consider the case $m = 1$, where there is a single source/path present in the observations. For a well chosen h , the optimal z has then just two non-zero complex weights. A non-zero component associated with the noise contribution, i.e., column $\sqrt{M} \mathbf{e}_1$ and another one associated with the column in A with bearing closest to the true bearing. The amplitude and bearing estimates are then simply this non-zero component (modified to take into account the normalization step in the elaborate version) and the bearing associated with the corresponding column.

The so-obtained bearing estimate being on the discretization grid may not be the best possible. A better choice of h (slightly smaller than the previous one) might yield a optimal z having three non-zero weights. Besides the 'noise' component, there would appear a couple of neighboring non-zero weights that, by simple linear interpolation, would yield a bearing estimate that does not lie on the discretization grid and may thus well be closer to the true bearing. In case, there are more than 2 contiguous nonzero components in z (which actually never happens) one similarly estimates the bearing by linear interpolation and the amplitude by simply adding the different weights.

If there are $m > 1$ paths present in \mathbf{y} , one expects the optimal z to have, in addition to the noise component, m disjoint couples of non-zero weights and the same procedure applies to each couple. This, of course, corresponds to the ideal situation where the proposed algorithm does a perfect detection job.

4.3. Detecting the number of paths

An estimate of m , the number of paths present in \mathbf{y} , is given by the number of (significant) clusters of non-zero components in the optimal z . A cluster being typically a couple of neighboring non-zero weights. For a given scenario, the detection performance of the proposed approach depends mainly upon the choice of h in the criterion, see Section 4.1 above. While for easy scenarios this choice is quite robust, this is no longer true for difficult scenarios. Globally if h is fixed to get a probability of false alarms equal to p , the procedure

will detect an additional (false) bearing in p percent of the realizations. This additional path will have a small amplitude estimate and may thus appear suspicious.

The choice of the number K of columns in A and the associated discretization step in bearing or better, for a uniform linear array, in spatial frequency $\delta f = 1/(K - 1)$ is also important in this context. If K is too large, it will essentially increase the computation time and lead to a poorly conditioned optimization problem. Remember though that the optimization problems (4) or (7) are convex and have thus generically a unique optimum.

If K is too small, it may happen that the two clusters associated with paths having close bearings merge which results in the non-separation of the two paths. This situation can and must be avoided by adequately choosing K . Indeed knowing roughly the resolution limit, say Δf , associated with the array, we propose to choose $K - 1$ so that $\Delta f = 10/(K - 1)$. As an example, for an array with M equispaced sensors and signal to noise ratios around 0 dB, the resolution limit (Rayleigh limit) is known to be about $\Delta f = 1/M$ so that one will typically choose $K - 1 = 10M$. This choice allows for about 2 zero weights between two paths that are close in bearings but nevertheless potentially separable and thus guarantees that two disjoint clusters will be obtained in case the two paths are detected.

4.4. Unbiasing the estimates, link with a detection test

Since even for $\mathbf{y} = \mathbf{Y}$, the optimal z will never be such that $Az = \mathbf{y}$, for $h > 0$, it is obvious that the presence of the term $h\|z\|_1$ in (4) or (7) induces bias into the estimates.

This bias while concerning quite directly the amplitude estimates, also affects the other parameters. It is thus important to remove the bias when the amplitude estimates are of interest as such. This is not the case here since the amplitude associated with a path is a function of the power of the associated source but also of the correlation with other paths, see (3). We will thus not elaborate on the way this bias can be removed, though we shall see in Section 5 that it is only non negligible on the bearing estimates. One of the ways being to use the outcome of the algorithm as input to a Maximum Likelihood algorithm.

In fact, see Section 4.1 above, h is chosen to achieve a given probability of false alarm, i.e., to decide wrongly that a path is present (that hypothesis H_1 is true) while there is none (hypothesis H_o is true). Usually deciding between H_1 (a single path is present) and H_o (the observations are just noise) is performed by comparing a quantity (the likelihood ratio or here the generalized likelihood ratio) to a threshold. In the GMF, the bias induced by the presence of h in the criterion translates this threshold to zero. One thus decides H_1 , i.e., that a source is present, if a component in the optimal z (besides the noise component) is nonzero, and H_o otherwise.

Indeed one can show that, while the matched filter (MF) is equivalent to the likelihood ratio (LR), the global (or generalized) matched filter (GMF) is equivalent to the generalized likelihood ratio (GLR), hence its name !

4.5. Summary of the algorithm.

Before presenting some simulation results in the next section, we summarize the algorithm we propose. For a given array, we build the A matrix of dimension (M, K) with $K - 1$ equal to roughly $10M$ and column vectors the equispaced steering vectors and a noise vector, see Section 3.1. With the N observed snapshots we build an estimate \hat{R} of their covariance matrix and denote \mathbf{y} its first column vector. We then solve (4) or (7) with h as defined in Section 4.1, using the subroutines SeDuMi [11] which is a second order cone programming library, to get the optimum.

For a quick and transitory test purpose, to solve, say, (4), one can implement the recursion

$$z \rightarrow z^+ = |Z|(A|Z|A^H + hI)^{-1}\mathbf{y},$$

with $|Z| = \text{diag}(|z|)$ and initial condition $z_o = A^+\mathbf{y}$ which has been proved to converge to the optimum [12].

We then deduce the estimates from the optimum z as described in Section 4.2.

To implement the more elaborate version associated with criterion (7) one further has to whiten A and \mathbf{y} and (re-) normalize the columns of the whitened A -matrix.

5. SIMULATION RESULTS.

To evaluate the performance of the proposed GMF approach, we compare its elaborate version to MUSIC after redundancy averaging (denoted MUSIC-R). We do not consider MUSIC after spatial smoothing since it does not separate the 3 paths even in the absence of correlation. The antenna is linear and has $M=10$ sensors. We consider 3 paths all at 0db (i.e. $\sigma_i^2 = \sigma_o^2 = 1$ in (1,2) and bearings -5, 0 and 15 degrees with respect to broadside. There are $N=100$ snapshots in all simulations. In the GMF criterion (7), we take $h = 1.07$ as recommended in Section 4.1 and the columns in A are steering vectors that are equispaced in spatial frequency with a step $1/160$. The same set of steering vectors is used in MUSIC-R where we further assume to know a priori the true number of paths present.

The basic MUSIC algorithm does not systematically separate the two paths that are close even if all the sources are uncorrelated, hence the above remark concerning the spatially smoothed MUSIC which has a lower resolution.

Below we keep all these features and in the different sets of simulations we only vary the correlations between the 3 paths denoted path 1, 2 and 3 with respective bearings -5, 0 and 15 degrees. We only consider fully correlated or coherent paths. More precisely if paths 1 and 2 are said to be coherent, we take $\rho_{1,2} = 1$ and $\varphi_{1,2} = 0$ in (2).

We take in turn paths 1 and 2 coherent, see Table 1, paths 1 and 3 coherent, Table 2, paths 2 and 3 coherent, Table 3 and in Table 4, all three paths are coherent. MUSIC-R locates systematically the three paths in the scenarios associated with Table 1 and 2 but fails to do so systematically in the scenarios associated with Table 3 and 4, despite the fact that we assume

the number of paths to be known to the algorithm. The proposed method GMF separates systematically the 3 paths in all 4 scenarios without prior knowledge.

Estimates of the mean, standard deviation averaged over 1000 independent realizations for MUSIC after redundancy averaging and the elaborate version of GMF.

	MUSIC-R		GMF	
	mean	std. dev.	mean	std. dev.
$\theta_1 = -5$	-4.746	2.50	-5.072	.593
$\theta_1 = 0$.442	.817	-1.145	.371
$\theta_1 = 15$	15.085	.252	15.522	.3131

Table 1: Two coherent paths at -5 and 0 and a third path at 15 degrees, equal SNR's= 0dB, 100 snapshots.

	MUSIC-R		GMF	
	mean	std. dev.	mean	std. dev.
$\theta_1 = -5$	-5.150	.554	-5.817	.770
$\theta_1 = 0$.273	.665	-1.001	.396
$\theta_1 = 15$	15.248	.152	15.501	.270

Table 2: Two coherent paths at -5 and 15 and a third path at 0 degrees, equal SNR's= 0dB, 100 snapshots.

	MUSIC-R		GMF	
	mean	std. dev.	mean	std. dev.
$\theta_1 = -5$	-6.639	.645
$\theta_1 = 0$	-1.864	.342
$\theta_1 = 15$	15.856	.209

Table 3: Two coherent paths at 0 and 15 and a third path at -5 degrees, equal SNR's= 0dB, 100 snapshots.

	MUSIC-R		GMF	
	mean	std. dev.	mean	std. dev.
$\theta_1 = -5$	-7.115	.746
$\theta_1 = 0$	-1.771	.617
$\theta_1 = 15$	16.100	2.349

Table 4: Three coherent paths at -5, 0 and 15 degrees, equal SNR's= 0dB, 100 snapshots.

6. CONCLUDING REMARKS.

We have proposed a technique that allows to detect and locate paths that are correlated or even coherent. It simply makes no assumption on the spatial stationarity or on the uncorrelatedness of the signals. As opposed to most proposed approaches that rely on some pre-processing of the observations which is only feasible for very specific array geometries, it works for any geometry and consists in applying the Global Matched Filter (GMF) [10] to the first column of the estimated covariance matrix of the snapshots. We also indicate that, from a detection point of view, it has some interesting

properties that makes it comparable to a generalized likelihood ratio test. Some simulations are proposed to demonstrate its performance. In its present form it presents some inherent bias that can however be removed by using our estimates as initial conditions of, for instance, a maximum likelihood algorithm. This last point requires however further investigations.

7. REFERENCES

- [1] G. Bienvenu and L. Kopp, "Source power estimation method associated with high resolution bearing estimation" Proceedings of the *IEEE ICASSP* Conference, Atlanta, p. 153-156, 1981.
- [2] R.O. Schmidt. "Multiple emitter location and signal parameter estimation," Proceedings of the *RADC Spectral Est. Workshop*, p. 243-258, 1979.
- [3] D. Rahamim, J. Tabrikian and R. Shavit, "Source localization using vector sensor array in a multipath environment," *IEEE Trans. on S.P.*, 52, 11, 33096-3103, 2004.
- [4] J.E. Evans, J.R. Johnson and D.F. Sun, "High resolution angular spectrum estimation techniques," Proceedings of the *1st ASSP Workshop Spectral Estimation*, Hamilton, ON, p. 134-139, 1981.
- [5] J.S. Thompson, P.M. Grant and B. Mulgrew, "Performance of spatial smoothing algorithms for correlated sources," *IEEE Trans. on Signal Processing*, 44, p. 1040-1046, Apr. 1996.
- [6] J. Li, "Improved angular resolution for spatial smoothing techniques," *IEEE Trans. on Signal Processing*, 40, p. 3078-3081, Dec. 1992.
- [7] D.A. Linebarger and D.H. Johnson, "The effect of spatial averaging on spatial correlation matrices in the presence of coherent signals" *IEEE Trans. on ASSP*, 38, p. 880-884, May 1990.
- [8] L.C. Godara, "Beamforming in the presence of correlated arrivals using structured correlation matrix" *IEEE Trans. ASSP*, 38, p. 1-15, Jan. 1990.
- [9] J.J. Fuchs, "Detection and estimation of superimposed signals", Proceedings of the *IEEE ICASSP* Conference, Seattle, 4, pp. 1649-1652, 1998.
- [10] J.J. Fuchs, "On the application of the global matched filter to DOA estimation with uniform circular arrays," *IEEE Trans. on S.P.*, vol. 49, 4, pp. 702-709, Apr. 2001.
- [11] SeDuMi, <http://sedumi.mcmaster.ca>
- [12] J.J. Fuchs, "Convergence of a Sparse Representations Algorithm Applicable to Real or Complex Data", *IEEE Journ. of STSP* 1, 598-605, Dec. 2007.