

APPLICATION OF THE GLOBAL MATCHED FILTER TO STAP DATA AN EFFICIENT ALGORITHMIC APPROACH

Sébastien MARIA, Jean-Jacques FUCHS

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

ABSTRACT

The Global Matched Filter (GMF) is a mean to solve simultaneously detection and estimation problems. It has been applied mainly in source localization and delay estimation problems where the amount of data to be handled is reasonably small. This is not the case in the Space Time Adaptive Processing (STAP) context where the amount of data is extremely large and constitutes probably the major challenge. We present a very efficient algorithm that allows to apply the GMF to problems having several thousands unknowns. This allows its application to the STAP data. We detail both the algorithm and the implementation of the GMF to the STAP data. Since the GMF is a high resolution technique it outperforms the algorithms that are traditionally used in this context.

1. INTRODUCTION

The Global Matched Filter (GMF) has been proposed in [1] as a mean to solve simultaneously detection and estimation problems. It can also be seen as allowing to solve linear inverse problems when it is known a priori that the input is a train of spikes. Similar approaches had been proposed somehow earlier [2, 3, 4]. It belongs to the global topic of sparse representation that has been active in the signal processing community [5] for almost ten years now. This topic has known a new start recently [6, 7] in applied mathematics where deterministic and much easier problems amenable to clear solutions were introduced. Simultaneously its connection with long standing problems in statistics such as Matching Pursuit [8], or step-wise regression [9] has been recognized and it is now an active domain of research in all these different communities [10, 11].

When applying these sparse representation techniques or similarly the GMF to practical situations one faces the computational complexity issue, hence the necessity to develop new algorithms allowing to handle larger problems in reasonable time. We develop here such an algorithm that allows to solve the GMF criterion for problems having up to 10000 unknowns in a few seconds on a standard desk computer. Similar algorithms have been proposed in the statistics literature under the name Least Angular Regression (LARS) [11, 12] and presented to the signal processing community in [10].

In order to demonstrate the performance of the algorithm, we apply it to the Space Time Adaptive Processing context [13]. We indicate how the problem of the detection of targets in the presence of clutter and noise can be handled by the GMF and solved in reasonable time.

In Section 2, we relate the GMF approach to more recent results in the sparse representations domain. In Section 3 we detail the efficient algorithm that solves the criterion used by the GMF. Then, in

Section 4, we adapt the basic STAP problem to the GMF technique and present simulation results that allow to assess the potentialities of the approach.

2. PRELIMINARIES

2.1. The Global Matched Filter

The Global Matched Filter (GMF) applies to any situation where an m -dimensional observation vector b can be represented as the sum of a small and unknown number of elements of a parameterized family in additive noise:

$$b = \sum_{i=1}^p a(\theta_i)x_i + e \quad (1)$$

where p is the unknown number of components to be estimated together with θ_i , the values of the scalar or vector of parameters and x_i the associated weights. The idea is then to replace the maximum likelihood approach which if e is Gaussian amounts to solve the non-linear least squares problem

$$\min_{\theta_i, x_i} \|b - \sum_{i=1}^p a(\theta_i)x_i\|_2^2$$

for different values of p , by a single convex program having a large number of unknowns as follows.

One uniformly discretizes the θ parameter values in its domain of interest and constructs with the n columns $a_j = a(\theta_j)$ -with θ_j on the regular grid- a matrix A having thus far more columns than rows ($n \gg m$). And, one solves the following quadratic program

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + h \|x\|_1, \quad h > 0 \quad (2)$$

for an adequately chosen parameter h . The ℓ_1 -norm penalty term ensures the sparsity of the solution. A physical interpretation can be given to the dual of this criterion [1]

$\min_x \|Ax\|_2^2$ subject to $\|A^T(Ax - b)\|_\infty \leq h$
which is also equivalent to

$$\min_x \|x\|_1, \quad \text{subject to } \|Ax - b\|_2^2 \leq B \quad (3)$$

2.2. The sparse representation approach

In the recent literature [6, 14, 15] on sparse representations, the initial main concern was to obtain precise conditions under which it is possible to recover an exact sparse representation from the sole knowledge of the observation vector b . By assuming that $b = Ax_o$ with x_o a sparse vector having thus just a few non zero components, one seeks conditions under which x_o is the unique optimum of an easy to solve convex program. Indeed since seeking the sparsest representation of $b = Ax_o$ requires a combinatorial search over all

potential solutions that is infeasible in practice, one considers instead the so-called ℓ_1 -sparsity and solves

$$\min_x \|x\|_1 \quad \text{s.t.} : \quad b = Ax \quad (4)$$

with $\|x\|_1 = \sum_1^n |x_j|$, which can be transformed into a linear program. One then analyses under which condition on x_o the true exact representation is the unique solution to (4). The answer is now known but is extremely conservative and thus essentially useless in practice. The trend is now to analyze the same problem when additive noise is present, i.e., when $b = Ax_o + e$. Some results are available essentially when the noise vector is ℓ_2 -bounded [16, 17]. The idea is to replace the exact reconstruction constraint $Ax = b$ in (4) by an approximate reconstruction constraint $\|Ax - b\|_2^2 \leq B$ in (3) or equivalently to solve (2), the criterion used in the GMF.

3. THE OPTIMIZATION ALGORITHM

The problem (2), which is convex unconstrained but non-smooth, can be converted into a quadratic program [18] and thus solved using existing subroutines. Here we propose to describe an algorithm specifically designed to solve (2) that has a far lower computational cost.

3.1. Optimality conditions

Mainly to simplify the exposition we will assume in the sequel that (2) has a unique optimum, this is probably generically true in applications. To develop the algorithm we need to state the necessary and sufficient conditions (NSC) satisfied by the optimum of (2) in a precise way that is most easily obtained by introducing the sub-differential of $\|x\|_1$ [18], a set of vectors called the sub-gradients, denoted $\partial\|x\|_1$,

$$\begin{aligned} \partial\|x\|_1 &= \{u | u^T x = \|x\|_1, \|u\|_\infty \leq 1\} \\ &= \{u | u_i = \text{sign}(x_i) \text{ if } x_i \neq 0 \text{ and } |u_i| \leq 1 \text{ otherwise}\}. \end{aligned}$$

where x_i denotes the i -th component of x , and $\text{sign}(x_i) = \pm 1$ is well defined since $x_i \neq 0$. The following result then holds [15].

Theorem 1: The point x^* is a global minimum of (2) if and only if the vector zero is a sub-gradient of the criterion at x^* , i.e.:

$$\exists u \in \partial\|x\|_1 \text{ such that } A^T(Ax^* - b) + hu = 0. \quad (\text{NSC}) \quad \square$$

Relation (NSC) is an implicit equation that is difficult to solve. There is a unique couple $\{x^*, u\}$ where u is a sub-gradient of $\|x\|_1$ at x^* that satisfies this equation and only a lengthy combinatorial search could lead to it.

One can show [15, 1] that if $h \geq \|A^T b\|_\infty$, the unique optimum is at the origin, while otherwise the optimum of (2) is a function of h that has at most m non zero components. To learn more about x^* we split it into its non zero components we denote \bar{x}^* and its zero components we denote $\bar{\bar{x}}^*$, we accordingly split the associated sub-gradient u and the A matrix, one then has, e.g., $Ax^* = \bar{A}\bar{x}^*$. One can then split the n equations in (NSC) into two parts

$$\begin{aligned} \bar{A}^T(b - \bar{A}\bar{x}^*) &= h \text{sign}(\bar{x}^*), \\ \bar{\bar{A}}^T(b - \bar{A}\bar{x}^*) &= h\bar{\bar{u}}^*. \end{aligned} \quad (5)$$

To get the first part we use the fact $\bar{\bar{u}}^* = \text{sign}(\bar{x}^*)$. From the first equation we now deduce an expression for \bar{x}^* which we write $\bar{x}(h)$ to emphasize the dependency on h , and substituting \bar{x}^* in the second part yields an explicit expression for $\bar{\bar{u}}(h)$:

$$\begin{aligned} \bar{x}(h) &= \bar{A}^+ b - h(\bar{A}^T \bar{A})^{-1} \text{sign}(\bar{x}(h)), \\ \bar{\bar{u}}(h) &= \frac{1}{h} \bar{\bar{A}}^T b^\perp + \bar{\bar{A}}^T d, \end{aligned} \quad (6)$$

where $b^\perp = (I - \bar{A}\bar{A}^{+T})b$ is the projection of b on the subspace orthogonal to the range of \bar{A} and $d = \bar{A}^{+T} \text{sign}(\bar{x}(h))$ with \bar{A}^{+T} the transpose of the pseudo-inverse of \bar{A} .

Remember that the equation (NSC) is useless as such. If a solution i.e., a couple $\{x^*, u\}$ or more precisely $\{x(h), u(h)\}$ is known, for a given and fixed value of h , it can however be used to obtain the expressions in (6) that are valid within an interval around the initial h . This means that the expressions of $\bar{x}(h)$ and $\bar{\bar{u}}(h)$ in (6) when complemented by $\bar{\bar{x}}(h) = 0$ and $\bar{\bar{u}}(h) = \text{sign}(\bar{x}(h))$ yield a couple $\{x(h), u(h)\}$ that is optimal not just for the initial value of h but over a whole interval. Indeed it is valid as long as no component in $\bar{x}(h)$ becomes zero (which generally occurs when one increases h) or a component in $\bar{\bar{u}}(h)$ becomes equal to ± 1 (which generally occurs when one decreases h).

Since moreover the boundaries of the so-defined interval in h are easy to obtain and the modification to be introduced in the equations (5) or equivalently (6) to cross a boundary are simple as well, we are ready to write the algorithm generating the optimal $x(h)$ as h varies, provided we know how to initialize this procedure. But this also is trivial since we know that for $h \geq \|A^T b\|_\infty$, the optimum is at the origin and that for a slightly smaller value of h , within the first interval, the index and the value of the single non zero component in $x(h)$ are easy computable. These observations have first been done in [12] in a slightly different form.

3.2. The algorithm

We construct the optimum $x(h)$ of (2) or more precisely its non-zero components $\bar{x}(h)$ as h decreases, i.e., we construct a sequence of adjacent intervals $[h_{\text{inf}}^{(k)}, h_{\text{sup}}^{(k)}]$ with $h_{\text{sup}}^{(k)} = h_{\text{inf}}^{(k-1)}$ for increasing k within which $\bar{x}(h)$ and $\bar{\bar{u}}(h)$ are given by the expressions in (6) with $\bar{A} = \bar{A}^{(k)}$, $\text{sign}(\bar{x}(h)) = s^{(k)}$ and $d^{(k)} = \bar{A}^{(k)+T} s^{(k)}$. One stops when, say, h^* the value of h for which the optimum is sought belongs to the current interval. We describe below how to find the boundaries of the intervals and how to reactivate the different quantities, $n^{(k)}$ is the number of columns in $\bar{A}^{(k)}$.

Initialization

- If $h^* > \|A^T b\|_\infty$, $x^* = 0$, stop the algorithm,
- else
 - * $i^{(1)} = \arg \max_i (|A^T b|_i)$,
 - * $\bar{A}^{(1)} = [a_{i(1)}]$, $\bar{\bar{A}}^{(1)} = A \setminus a_{i(1)}$, $n^{(1)} = n - 1$,
 - * $s^{(1)} = \text{sign}((A^T b)_{i(1)})$,
 - * $h_{\text{sup}}^{(1)} = \|A^T b\|_\infty$.

k -step of the algorithm

Search for $h_{\text{inf}}^{(k)} = h_{\text{sup}}^{(k+1)}$:

- **Case 1:** a component in $\bar{\bar{u}}$ becomes equal to ± 1 with associated value $h_1^{(k)}$ of h :

For $j = 1, 2, \dots, n^{(k)}$ and $\epsilon = +1$ or -1 , compute:

$$\text{Val}1_j(\epsilon) = \epsilon \frac{(\bar{A}^{(k)T} b^\perp)_j}{1 - \epsilon (\bar{A}^{(k)T} d^{(k)})_j},$$

$$\begin{cases} (ind_1^{(k)}, \epsilon_1^{(k)}) &= \arg \max_{j, \epsilon} (\text{Val}1_j(\epsilon)) \\ h_1^{(k)} &= \text{Val}1_{ind_1^{(k)}}(\epsilon_1^{(k)}). \end{cases}$$

- **Case 2:** a component in \bar{x} becomes equal to 0 with associated value $h_2^{(k)}$. For $j = 1, 2, \dots, n-n^{(k)}$:

$$\begin{aligned} Val2_j &= \frac{(\bar{A}^{(k)+b})_j}{((\bar{A}^{(k)T}\bar{A}^{(k)})^{-1}s^{(k)})_j}, \\ \begin{cases} ind_2^{(k)} &= arg \max_j (Val2_j) \\ h_2^{(k)} &= Val2_{ind_2^{(k)}}. \end{cases} \end{aligned}$$

- If $(h_{inf}^{(k)} > h_{sup}^{(k)} \text{ or } h_{inf}^{(k)} < 0)$, take $h_{inf}^{(k)} = 0$.

- else $h_{inf}^{(k)} = \max(h_1^{(k)}, h_2^{(k)})$.

Updating :

- If $h^* \in]h_{inf}^{(k)}, h_{sup}^{(k)}[$, we are in the desired interval:

$$x^* = \bar{A}^{(k)+b} - h^*(\bar{A}^{(k)T}\bar{A}^{(k)})^{-1}s^{(k)}, \text{ and stop.}$$

- else update $\bar{A}^{(k+1)}, \bar{A}^{(k+1)}$ and $s^{(k+1)}$:

- if case 1 above is valid: remove in $\bar{A}^{(k)}$ the $ind_1^{(k)}$ -th column, add it in $\bar{A}^{(k)}$ and $s^{(k+1)} = [s^{(k)T}, \epsilon_1^{(k)}]^T$, $n^{(k+1)} = n^{(k)} - 1$,
- if case 2 above is valid: remove in $\bar{A}^{(k)}$ the $ind_2^{(k)}$ -th column and the corresponding $ind_2^{(k)}$ -th component in $s^{(k)}$, then add this column in $\bar{A}^{(k+1)}$, $n^{(k+1)} = n^{(k)} + 1$,
- then $h_{sup}^{(k+1)} = h_{inf}^{(k)}$, $k = k + 1$ and return.

4. THE APPLICATION TO STAP

4.1. The STAP context

STAP is an increasingly popular signal processing approach for detecting slow moving targets in the presence of clutter and noise. STAP consists in simultaneously processing spatial and temporal data. The basic theory of STAP is by now well developed [13, 19, 20, 21]. An advanced airborne radar uses a N_s -element uniform linear array with interelement spacing d taken equal to half-wavelength. It transmits a N_p -pulse waveform at pulse repetition interval T . For each pulse, the radar return from scatterers is associated to a given range from the transmitter. The time slot corresponding to a given range is commonly referred as a ‘‘range bin’’. Data are collected over P range bins. The received data from the j -th range bin can be viewed as a $N_s N_p$ -dimensional complex vector called ‘‘snapshot’’. When a target is present in the j -th range bin the associated snapshot s can be written as:

$$s = \alpha s_t(\tilde{\phi}, \tilde{f}) + s_{int}, \quad (7)$$

where $\alpha, \tilde{\phi}, \tilde{f}$ are the target amplitude, normalized angle and normalized Doppler frequency respectively. Here $s_t(\tilde{\phi}, \tilde{f})$ represents the spatio-temporal steering vector. It corresponds to:

$$s_t(\tilde{\phi}, \tilde{f}) = u(\tilde{f}) \otimes v(\tilde{\phi}),$$

where \otimes represents the Kronecker product and

$$\begin{aligned} u(\tilde{f}) &= [1 \ e^{-j2\pi\tilde{f}} \ \dots \ e^{-j2\pi(N_p-1)\tilde{f}}]^T, \\ v(\tilde{\phi}) &= [1 \ e^{-j2\pi\tilde{\phi}} \ \dots \ e^{-j2\pi(N_s-1)\tilde{\phi}}]^T, \end{aligned}$$

are the N_p -dimensional spatial steering vector and N_s -dimensional temporal steering vector. Moreover the interferences are represented by s_{int} , they are modelled as:

$$s_{int} = s_c + n = \sum_k \alpha_k s_{tk}(\tilde{\phi}_k, \tilde{f}_k) + n,$$

where s_c is the clutter component and $n \sim \mathcal{N}(0, \sigma^2 I)$. Indeed, the clutter perturbation present in the j -th range bin corresponds to the sum of many targets with a complex Gaussian amplitude α_k . The principal characteristic that distinguishes clutter from potential moving targets is the dependence of the clutter normalized Doppler frequency on the normalized angle:

$$\tilde{f}_k = \frac{2v_{airborne}T}{d} \tilde{\phi}_k = \beta \tilde{\phi}_k.$$

In the following, we will consider that the parameter β is adjusted to be equal to 1. As the principal goal of STAP is to separate targets and clutter at each range bin, the sensor outputs are linearly combined in order to maximize the output-signal to clutter-plus-noise ratio. The well-known $N_s N_p$ -dimensional optimum weight vector [22] is

$$w = R_{int}^{-1} s_t(\tilde{\phi}, \tilde{f}),$$

where $R_{int} = E(s_{int} s_{int}^H)$ is the covariance matrix of the interference (clutter plus noise). To apply this weight vector, we need an estimate of the R_{int} matrix and so one usually uses the $2N_s N_p$ adjacent range bins to perform this estimation ([13]). Since this is unfeasible in practice, many sub-optimal methods have been developed. Below we apply the GMF technique to this problem, it requires no R_{int} -estimate and works on a single snapshot.

4.2. The GMF in the STAP context

In this part we describe the different steps which permit to apply GMF to (7). At first, we apply to the complex s a set of standard beamformers tuned to m equispaced spatio-temporal points. This transforms the $N_s N_p$ -dimensional vector s into a m -dimensional real vector. More precisely, a typical beamformer-output evaluated at \tilde{f}_i and $\tilde{\phi}_i$ and applied to the snapshot is:

$$b_i = |s_{ti}(\tilde{\phi}_i, \tilde{f}_i)^H s|^2 / (N_s N_p), \text{ for } i = 1..m$$

This allows to obtain the m -dimensional real vector b in (1),(2) which constitutes the input-data to the GMF. In [1], one proposes to take m equal to the number of degrees of freedom (d.o.f) of the covariance matrix of the snapshot. In our case, this gives $m = 2N_s N_p - 1$ which is also the number of real d.o.f. in s . To construct the column vectors of the A matrix in (2), all values in frequency and azimuth which are liable to characterize a target are discretized at equispaced spatio-temporal points. At each of these, say, n points, one creates a basis vector by transforming the associated elementary $N_s N_s$ -dimensional steering vector by the same set of m beamformers described above and used to build the observation vector b .

One further normalizes (to one in Euclidean norm) all these vectors to construct the $m \times n$ real matrix A . The number n of columns is fixed by the discretization steps in azimuth and Doppler frequency. It has to be chosen so that it does not prevent the approach to attain the Cramer Rao bounds. For standard values Signal to Noise Ratios (SNR), one expects to improve the standard beamformer resolution by a factor 2 or 3. For this to be easily achievable, the normalized frequency domain and normalized azimuth are typically discretized respectively in $9N_s$ and $9N_p$ below. The matrix A has thus $n \simeq 81N_s N_p$ columns and $m = 2N_s N_p - 1$ rows.

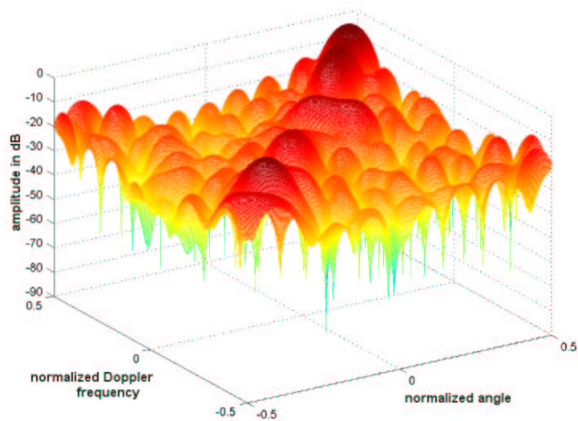
Note that since clutter can be modelled as a sum of stationary targets, they will be identified by the GMF just like the targets of interest. Since GMF is a high-resolution technique it may detect slowly moving targets undetectable by standard techniques. Its ability to separate target and clutter is of course a function of SNR, CNR (Clutter to Noise Ratio), N_s and N_p .

4.3. Application of the algorithm

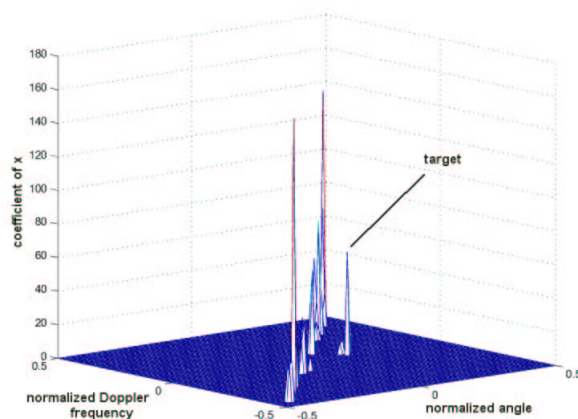
Simulations have been done for $N_s = N_p = 10$. The normalized azimuth and normalized Doppler frequency both belong to $[-0.5, 0.5]$. To build b the input-vector, we discretize both spaces with a step of 0.07 that leads to $m = 225 \simeq 2N_s N_p - 1$ beamformers. To build the n columns of the A matrix we take a step equal to 0.01 in azimuth and frequency, this gives $n = 10201$.

We simulate a scenario with one target at $(\tilde{f} = 0, \tilde{\phi} = 0.15)$ in presence of clutter with $SNR = CNR = 20dB$. The parameter h is fixed to .4, due to limited space we do not specify how to tune this parameter, one can actually take any value between, say, .2 and .8 and still get the same type of outputs. The algorithm described in section 4 takes about 10 seconds on standard desk computer to solve (2).

In figure 1(a) and 1(b) we compare the classical 2D-IFFT method to the GMF approach respectively. Target and clutter are separated by GMF but not really by the 2D-IFFT method. If the target is taken closer to the clutter, GMF will still localize it but additional spurious peaks will appear around it making a true detection more questionable.



(a) 2D-IFFT processing



(b) GMF

Fig. 1. Detection of one target in clutter plus noise environment

5. CONCLUSION

We have applied the GMF [1] to the basic STAP problem. The major difficulty was to overcome the computational complexity problem. To do so we have presented an algorithm that allows to apply GMF to problems having several thousands unknowns. Since GMF identifies both the targets and a model of the clutter it has the advantage to be able to work on a single snapshot without prior estimation of the interference covariance matrix.

6. REFERENCES

- [1] 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, p. 702–709, avr. 2001.
- [2] B.M. Bell and T.E. Ewart. Separating multipaths by global optimization of a multidimensional matched filter. *IEEE-T-ASSP*, 37:378–391, mar. 1986.
- [3] M.D. Sacchi, T.J. Ulrych and C.J. Walker. Interpolation and extrapolation using a high-resolution discrete Fourier transform. *IEEE Trans on S.P.*, 46, 1: 31–38, Jan. 1998.
- [4] S. Alliney and S.A. Ruzinsky. An algorithm for the minimization of mixed ℓ_1 and ℓ_2 norms with application to Bayesian estimation. *IEEE Trans. on S.P.*, 42, 3: 618–627, Mar. 1994.
- [5] B.D. Rao. Signal Processing with the sparsness constraint. *IEEE ICASSP*, I, 369–372, Seattle, 1998.
- [6] D.L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. on I.T.*, 47, 11, 2845–2862, Nov. 2001.
- [7] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24, 21, 227–234, april 1995.
- [8] S. Mallat and Z. Zhang. “Matching pursuit with time-frequency dictionaries”. *IEEE Trans. on S.P.*, vol. 41, 12, december 1993.
- [9] A. J. Miller, “Subset Selection in Regression” *Chapman and Hall*, 2002.
- [10] D.M. Malioutov, M. Cetin and A.S. Willsky. “Homotopy continuation for sparse signal representation” *IEEE ICASSP*, V, 733–737, Philadelphia,, 2005.
- [11] B. Efron, T. Hastie, I. Johnstone and R. Tibshirani, “Least angle regression,” *Annals of Statistics*, vol. 32, pp. 407–499, Apr. 2004.
- [12] M. R. Osborne, B. Presnell, and B. A. Turlach, A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, 20, 3, 389–403, 2000.
- [13] J.Ward, “Space-Time Adaptive Processing for Airborne Radar”. Lincoln Laboratory Technical Report 1015, Dec. 1994.
- [14] R. Gribonval and M. Nielsen. Sparse representations in unions of bases. *IEEE Trans. on I.T.* 49, 12, 3320–3325, Dec. 2003.
- [15] J.J. Fuchs. More on sparse representations in arbitrary bases. *IEEE Trans. on I.T.* 50, 6, 1341–1344, June 2004
- [16] D.L. Donoho, M. Elad and V. Temlyakov, “Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise,” submitted to *IEEE Trans. on I.T.*, Feb. 2004.
- [17] J.J. Fuchs. Recovery of exact sparse representations in the presence of bounded noise *IEEE Trans. on I.T.* 51, 10, 3601–3608, Oct. 2005.
- [18] R. Fletcher. Practical methods of optimization. *John Wiley and Sons*, 1987.
- [19] R. Klemm, “Principles of space-time adaptive processing”. *London: Institution of Electrical Engineers*, 2002.
- [20] R. Klemm, “Introduction to space-time adaptive processing”. *Electron., Commun. Eng. J.*, 11, 1, 5–12, Feb. 1999.
- [21] G. W. Titi and D. F. Marshall, “The ARPA/NAVY mountaintop program: Adaptive signal processing for airborne early warning radar”. *IEEE ICASSP*, II, 1165–1168, Atlanta, 1996.
- [22] L.E. Brennan, J.D. Mallett and I.S. Reed, “Adaptive Arrays in Airborne MTI Radar”. *IEEE Trans. on Antennas and Propagation* Vol. AP-24, 607–615, Sept. 1976.