Oblique Matching Pursuit

Laura Rebollo-Neira
Aston University, Birmingham B4 7ET, UK

Abstract

A method for selecting a suitable subspace for discriminating signal components through an oblique projection is proposed. The selection criterion is based on the consistency principle introduced by M. Unser and A. Aldroubi and extended by Y. Elder. An effective implementation of this principle for the purpose of subspace selection is achieved by updating of the dual vectors yielding the corresponding oblique projector.

1 Introduction

Oblique projectors are of assistance to signal processing applications [1–7], in particular due to their ability to discriminate signal components lying in different subspaces. Thereby, as discussed in [1], oblique projectors are suitable for filtering structured noise. Let us suppose for instance that a given signal f, represented mathematically as an element of a vector space $\mathcal{H}$, is produced by the superposition of two phenomena, i.e. $f = f_1 + f_2$ where f_1 belongs to a subspace $\mathcal{S}_1 \subset \mathcal{H}$ and f_2 belongs to subspace $\mathcal{S}_2 \subset \mathcal{H}$. Provided that $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0\}$ we can obtain from f the component f_1 by an oblique projection onto S_1 along S_2, which maps f_2 to zero without altering f_1. The procedure is straightforward and effective if the corresponding subspaces S_1 and S_2, such that $S_1 \cap S_2 = \{0\}$, are known [1]. Nevertheless, this may not be always the case. In this letter we address the problem of selecting the appropriate subspace S_1, from the spanning set of a larger subspace, in order to fulfil the condition $S_1 \cap S_2 = \{0\}$ assuming that S_2 is known and fixed.

Given a signal, our strategy for the selection of the representation subspace is in the line of Matching Pursuit (MP) methodologies [8–12] and is made out of two ingredients i) the sampling/reconstruction *consistency requirement* introduced in [2] and extended in [6] ii) a recursive procedure for adapting the dual vectors giving rise to the corresponding oblique projector [13]. It will be shown here that the latter yields an effective implementation of a selection criterion that we base on the consistency principle.

The letter is organized as follows: Sec 2 introduces the general framework and discusses the ingredients of the approach. Namely, the consistency principle and the recursive updating of the measurement vectors for achieving the required oblique projection. The Oblique Matching Pursuit strategy is introduced in Sec 3. Its implementation is discussed in Sec 4 along with a numerical example. The conclusions are drawn in Sec 5.

2 The consistency principle and stepwise updating of measurement vectors

We represent a signal f as an element of an inner product space that without loss of generality is assumed to be finite dimensional. The square norm is computed as $||f||^2 = \langle f, f \rangle$, where the brackets denote the corresponding inner product and we define the inner product in such a way that if c is a complex number $\langle cf, g \rangle = c^* \langle f, g \rangle$, with c^* the complex conjugate of c. Measurements of a signal f (also called samples) will be represented as linear functionals. Thus a set of, say k, sampling vectors w_i^k, $i = 1, \ldots, k$ provides us with a set of k measurements on f given by the inner products $\langle w_i^k, f \rangle$, $i = 1, \ldots, k$. The superscript k is used to indicate that to reconstruct the signal we will need to modify the measurement vectors w_i^k if an additional measure is considered. From the sampling measurements we can construct an approximation f^k of f using a set of reconstruction vectors v_i, $i = 1, \ldots, k$. The *consistency principle* introduced in [2] states that the reconstruction f^k from $\langle w_i^k, f \rangle$, $i = 1, \ldots, k$ should be self-consistent in the sense that if the approximation is sampled with the same vectors the same samples should be obtained. In other words, a consistent reconstruction must satisfy:

$$\langle w_i^k, f^k \rangle = \langle w_i^k, f \rangle, \quad i = 1, \ldots, k.$$

This requirement has been considered further in [6] where it is proved that: *if the reconstruction vectors v_i, $i = 1, \ldots, k$ span a subspace $\mathcal{V}_k$ and the sampling vectors w_i^k, $i = 1, \ldots, k$ span a subspace $\mathcal{W}_k$ such that its orthogonal complement $\mathcal{W}^\perp$ satisfies $\mathcal{V}_k \cap \mathcal{W}^\perp = \{0\}$, then f^k is a consistent reconstruction of f if and only if f^k is the oblique projection of f onto $\mathcal{V}_k$ along $\mathcal{W}^\perp$.* We represent the corresponding oblique projector as $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp}$. Hence, it is endowed with the following properties

$$\begin{aligned}
\hat{E}^2_{\mathcal{V}_k\mathcal{W}^\perp} &= \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} \\
\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} v &= v, \quad \text{for any} \quad v \in \mathcal{V}_k \\
\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} w &= 0, \quad \text{for any} \quad w \in \mathcal{W}^\perp.
\end{aligned}$$

Given the conditions of the above statement, the unique consistent approximation of f is therefore $f^k = \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f$. The oblique projector can be expressed as $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} = \sum_{i=1}^k v_i \langle w_i^k, \cdot \rangle$ where $\langle w_i^k, \cdot \rangle$ indicates that $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp}$ acts by performing inner products as in $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f = \sum_{i=1}^k v_i \langle w_i^k, f \rangle$. Explicit equations for updating an oblique projector when a new pair of reconstruction/measurement vectors is to be considered are given in [13]. As will be discussed in the next sections, for the purpose of this contribution we can restrict the measurement vectors to be lineally independent. Hence the vectors w_i^{k+1} yielding oblique projectors along $\mathcal{W}^\perp$ onto nested subspaces $\mathcal{V}_{k+1} = \mathcal{V}_k + v_{k+1} = \text{span}\{v_i\}_{i=1}^{k+1}$ can be inductively obtained as follows:

Construct vectors $u_i = v_i - \hat{P}_{\mathcal{W}^\perp} v_i$, with $\hat{P}_{\mathcal{W}^\perp}$ the orthogonal projector onto $\mathcal{W}^\perp$. From $w_1^1 = \frac{u_1}{||u_1||^2}$ every time a new vector is needed compute it, and update the previous ones, through the equations [13]:

$$w_i^{k+1} = w_i^k - w_{k+1}^{k+1} \langle u_{k+1}, w_i^k \rangle, \quad i = 1, \ldots, k \tag{1}$$

$$w_{k+1}^{k+1} = \frac{q_{k+1}}{||q_{k+1}||^2}, \quad q_{k+1} = u_{k+1} - \hat{P}_{\mathcal{W}_k} u_{k+1}, \tag{2}$$

where $\hat{P}_{\mathcal{W}_k}$ is the orthogonal projector onto $\mathcal{W}_k = \text{span}\{u_i\}_{i=1}^k$. It should be noticed that $\mathcal{V}_{k+1} + \mathcal{W}^\perp = \mathcal{W}_{k+1} \oplus \mathcal{W}^\perp$, with $\oplus$ indicating the orthogonal sum and $+$ the direct sum.

In the next section we introduce a method for stepwise selection of the measurement vectors aiming at finding a subspace $\mathcal{V}_k$ for reconstruction such that $\mathcal{V}_k \cap \mathcal{W}^\perp = \{0\}$. This property guarantees that the reconstructed signal has no component in $\mathcal{W}^\perp$, i.e. the reconstruction of the signal in $\mathcal{V}_k$ behaves like a filter of the component in $\mathcal{W}^\perp$.

3 Oblique Matching Pursuit (OBLMP)

Matching Pursuit strategies for signal representation evolve by stepwise selection of vectors, called atoms, which are drawn from a large set called a dictionary. Unless the dictionary is orthonormal, the seminal approach [8] does not yield a stepwise reconstruction of the orthogonal projection of the signal onto a selected subspace. A variation of this approach, called Orthogonal Matching Pursuit (OMP) does yield the orthogonal projection [9]. Such a reconstruction is therefore optimal in the sense of minimizing the norm of the approximation error. However, to render a matching pursuit strategy suitable for discriminating signals representing different phenomena, the approach needs to be generalized. In order to propose the Oblique Matching Pursuit (OBLMP) method addressing this problem we make the following assumptions.

- The subspace $\mathcal{W}^\perp$ in which the signal component to be filtered lays is known.
- The signal we wish to filter admits a unique decomposition $f = f_1 + f_2$, with $f_1 \in \mathcal{V}_k$ and $f_2 \in \mathcal{W}^\perp$. This is equivalent to assuming $f \in \mathcal{V}_k + \mathcal{W}^\perp$ with $\mathcal{V}_k \cap \mathcal{W}^\perp = \{0\}$.
- The subspace $\mathcal{V}_k$ can be spanned by vectors of the dictionary in hand.

As discussed in the previous section, the reconstruction that eliminates the signal component in $\mathcal{W}^\perp$ is $f^k = \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f$. Our goal is to construct the oblique projector by using the appropriate dictionary vectors. We know how to update $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp}$ to $\hat{E}_{\mathcal{V}_{k+1}\mathcal{W}^\perp}$ so as to account for the inclusion of an additional vector v_{k+1}. The question arises now as to how to select v_{k+1} giving rise to the right subspace. We answer this question by recourse to the consistency principle [2,6]. Considering that at iteration k the approximation f^k of f is $\hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f$, let us define the consistency error with regard to a new measurement w_{k+1}^{k+1} as $\Delta = |\langle w_{k+1}^{k+1}, f - \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f\rangle|$. Thus to construct the approximation $f^{k+1} = \hat{E}_{\mathcal{V}_{k+1}\mathcal{W}^\perp} f$ we propose to select the measurement vector w_{k+1}^{k+1} such that

$$w_{k+1}^{k+1} = \arg\max_{\ell \in \mathcal{J}} |\langle w_\ell^{k+1}, f - \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f\rangle|, \tag{3}$$

where $\mathcal{J}$ is the set of indices labeling the corresponding dictionary vectors not selected in the previous steps.

Proposition 1. *If vectors w_i^k, $i = 1 \ldots, k$ have been selected by criterion* (3) *and $|\langle w_{k+1}^{k+1}, f - \hat{E}_{\mathcal{V}_k\mathcal{W}^\perp} f\rangle| \neq 0$, the measurement vector w_{k+1}^{k+1} and the previously selected vectors w_i^k, $i = 1 \ldots, k$ are linearly independent.*

Proof. Assume that, on the contrary, $|\langle w_{k+1}^{k+1}, f - \hat{E}_{\mathcal{V}_k \mathcal{W}^\perp} f \rangle| \neq 0$ and there exists a set of numbers $\{a_i\}_{i=1}^k$ such that $w_{k+1}^{k+1} = \sum_{i=1}^k a_i w_i^k$. Since for the previously selected vectors the consistency condition holds, i.e. $\langle w_i^k, f \rangle = \langle w_i^k, \hat{E}_{\mathcal{V}_k \mathcal{W}^\perp} f \rangle$, $i = 1 \ldots, k$, we have

$$|\langle w_{k+1}^{k+1}, f - \hat{E}_{\mathcal{V}_k \mathcal{W}^\perp} f \rangle| = |\langle \sum_{i=1}^k a_i w_i^k, f - \hat{E}_{\mathcal{V}_k \mathcal{W}^\perp} f \rangle| = |\sum_{i=1}^k a_i^* (\langle w_i^k, f \rangle - \langle w_i^k, \hat{E}_{\mathcal{V}_k \mathcal{W}^\perp} f \rangle)| = 0.$$

This contradicts our assumption, which implies that $w_{k+1}^{k+1} \neq \sum_{i=1}^k a_i w_i^k$. □

Proposition 2. *All measurement vectors* w_ℓ^{k+1} *(C.f. eq. (3)) are orthogonal to the reconstruction vectors selected in previous iterations.*

Proof. Every w_ℓ^{k+1} is computed as in (2) and for $i = 1, \ldots, k$ it is true that $\langle q_\ell, v_i \rangle = \langle u_\ell, v_i \rangle - \langle \hat{P}_{\mathcal{W}_k} u_\ell, v_i \rangle = \langle u_{k+1}, u_i \rangle - \langle u_{k+1}, \hat{P}_{\mathcal{W}_k} v_i \rangle = \langle u_{k+1}, u_i \rangle - \langle u_{k+1}, u_i \rangle = 0.$ □

The last proposition allows us to re-state the OBLMP selection criterion (3) as

$$w_{k+1}^{k+1} = \arg\max_{\ell \in \mathcal{J}} |\langle w_\ell^{k+1}, f \rangle|. \tag{4}$$

Proposition 1 ensures that, for a given tolerance $\delta > 0$, by stopping the selection process when the condition $\arg\max_{\ell \in \mathcal{J}} |\langle w_\ell^{k+1}, f \rangle| < \delta$ is reached, the method only selects linearly independent measurement vectors. Assuming that at iteration $k+1$ the selected indices are $\ell_1, \ldots, \ell_{k+1}$, the signal reconstruction is given as

$$f^{k+1} = \hat{E}_{\mathcal{V}_{k+1} \mathcal{W}^\perp} f = \sum_{i=1}^{k+1} \langle w_i^{k+1}, f \rangle v_{\ell_i} = \sum_{i=1}^{k+1} c_i^{k+1} v_{\ell_i}. \tag{5}$$

The coefficients in the last equation can be updated at each iteration according to (1) and (2), i.e.,

$$c_{k+1}^{k+1} = \langle w_{k+1}^{k+1}, f \rangle \tag{6}$$

$$c_i^{k+1} = c_i^k - c_{k+1}^{k+1} \langle w_i^k, u_{k+1} \rangle, \quad i = 1, \ldots, k. \tag{7}$$

It is appropriate to point out that these equations, as well as (1) and (2), have the identical form of the equations to modify the dual vectors and the coefficients in the Optimized Orthogonal Matching Pursuit Approach (OOMP) [10]. However, now the equations involve vectors of different nature yielding therefore a different approach. OOMP updating arises as the particular case, corresponding to $u_i \equiv v_i$, for which $\hat{E}_{\mathcal{V}_{k+1}\mathcal{W}^\perp} \equiv \hat{P}_{\mathcal{V}_{k+1}}$. Nevertheless, since the criterion for the selection process we have adopted here does not necessarily minimize the norm of the residual error, OOMP is not a truly particular case of the new approach. On the contrary, we are introducing an alternative selection criterion based on the consistency principle, which could also be considered for producing yet one more variation of OMP.

4 Implementation details and numerical example

In consistence with the hypothesis itemized in Sec. 3 we consider that the subspace $\mathcal{W}^\perp$ is given, i.e. $\{\eta_i\}_{i=1}^n$ such that $\mathcal{W}^\perp = \text{span}\{\eta_i\}_{i=1}^n$ is known. For constructing $\hat{P}_{\mathcal{W}^\perp}$ there are a number of possibilities. In the example we present here the set $\{\eta_i\}_{i=1}^n$ is linearly dependent and we have used the technique for dictionary redundancy elimination proposed in [14]. MATLAB code for its implementation is available at [15]. The method produces a set of orthonormal vectors $\{\psi_i\}_{i=1}^m$, $m \leq n$ that we use to construct $\hat{P}_{\mathcal{W}^\perp} = \sum_{i=1}^m \psi_i \langle \psi_i, \cdot \rangle$.

Given a dictionary $\{v_\ell\}_{\ell \in \mathcal{J}}$ we proceed to compute vectors $\{u_\ell\}_{\ell \in \mathcal{J}}$ as $u_\ell = v_\ell - \sum_{n=1}^m \psi_n \langle \psi_n, v_\ell \rangle$. Except for the selection criterion the next steps parallel those for the implementation of OOMP but considering the dictionary $\{u_\ell\}_{\ell \in \mathcal{J}}$. A routine for implementation of OOMP based on Modified Gram Smidth orthogonalization with re-orthogonalization is also available at [15]. With very minor changes that routine can be used for the implantation of OBLMP. The algorithm is described below.

Starting by assigning $\gamma_\ell = u_\ell$, $\ell \in \mathcal{J}$, at the first step we select the index ℓ_1 corresponding to the index for which $\langle \gamma_\ell, f \rangle / ||\gamma_\ell||^2$ is maximal and set $q_1 = \gamma_{\ell_1}/||\gamma_{\ell_1}||$, $w_1^1 = q_1/\gamma_{\ell_1}$ and $c_1^1 = \langle w_1^1, f \rangle$. The index set $\mathcal{J}$ is changed to $\mathcal{J} = \mathcal{J} \setminus \ell_1$. At step $k+1$ the sequence γ_ℓ, $\ell \in \mathcal{J}$ (at this stage $\mathcal{J}$ is the subset of indices not selected in the previous k steps) is orthogonalized with respect to q_k as: $\gamma_\ell = \gamma_\ell - q_k \langle q_k, \gamma_\ell \rangle$ and, if necessary, reorthogonalized with respect to $q_1, \ldots, q_k$ i.e., $\gamma_\ell = \gamma_\ell - \sum_{j=1}^k q_j \langle q_j, \gamma_\ell \rangle$. After selecting the index ℓ_{k+1} as the maximizer of $\langle \gamma_\ell, f \rangle / ||\gamma_\ell||^2$ we set $q_{k+1} = \gamma_{\ell_{k+1}}/||\gamma_{\ell_{k+1}}||$, $w_{k+1}^{k+1} = q_{k+1}/\gamma_{\ell_{k+1}}$ and $c_{k+1}^{k+1} = \langle w_{k+1}^{k+1}, f \rangle$ and compute $\{w_i^{k+1}\}_{i=1}^k$ according to (1) and $\{c_i^{k+1}\}_{i=1}^k$ according to (7). For a given tolerance parameter δ the algorithm is to be stopped when $\langle \gamma_\ell, f \rangle / ||\gamma_\ell||^2 < \delta$ for all $\ell \in \mathcal{J}$. The reconstructed signal is then obtained as in (5).

We illustrate now the proposed method and its motivation by the following example: Suppose that from the two signals of Figure 1 we wish to eliminate the corresponding backgrounds. The backgrounds are different, but both are known to belong to the subspace spanned by the set of functions $\eta_i(x) = (x + 0.5)^{-0.03i}$, $i = 1, \ldots, 50$. This set is highly redundant. A good representation of the span can be achieved by just five linearly independent functions. Actually to avoid possible bad conditioning we considered only three orthonormal functions for constructing $\hat{P}_{\mathcal{W}^\perp}$. As verified a posteriori, that was enough for the backgrounds we were dealing with. In regard to the signal space we considered the cardinal cubic spline space, with distance 0.065 between consecutive knots, spanned by the corresponding B-spline basis on the interval $[0, 4]$. Since in both cases the signal space is also suitable for representing the background, the oblique projection onto the whole space does not yield the desired signal splitting. The failed attempt to separate the signal components in the left graph of Figure 1 is displayed by the broken line in the left graph of Figure 2. On the contrary, by applying the OBLMP approach, we could pick from the whole basis some elements spanning a subspace such that in the intersection with $\mathcal{W}^\perp$ there are only vectors of very small norm. Hence, as depicted in the right graph of Figure 2, the signal discrimination is successful. The approximation coincides in the scale of the figure with the true signal. The equivalent results, but concerning the signal in the right graph of Figure 1, are shown in Figure 3. The continuous lines depict the target signal and the broken lines the corresponding approaches.

5 Conclusions

A method, termed OBLMP, which allows for the selection of a suitable subspace for representing one of the signal components, and leaving aside other components of different nature, has been proposed. The approach evolves by stepwise selection of the subspace. The selection criterion is based on the consistency requirement introduced in [2] and extended in [6]. An effective implementation is achieved by stepwise updating of the measurement vectors yielding the appropriate oblique projector [13]. With regard to implementation and complexity OBLMP is equivalent to the OOMP approach [10, 12].

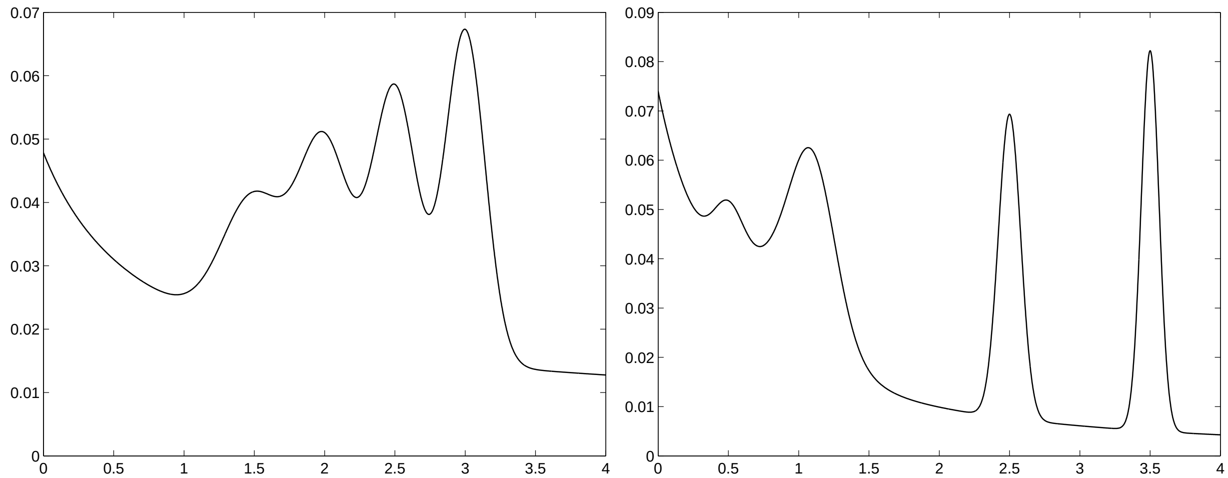

Figure 1: Two different signals (left and right graphs) superpose on two different backgrounds belonging to the given subspace $\mathcal{W}^{\perp} = \mathrm{span}\{(x+1)^{-0.03i}\}_{i=1}^{50}$.

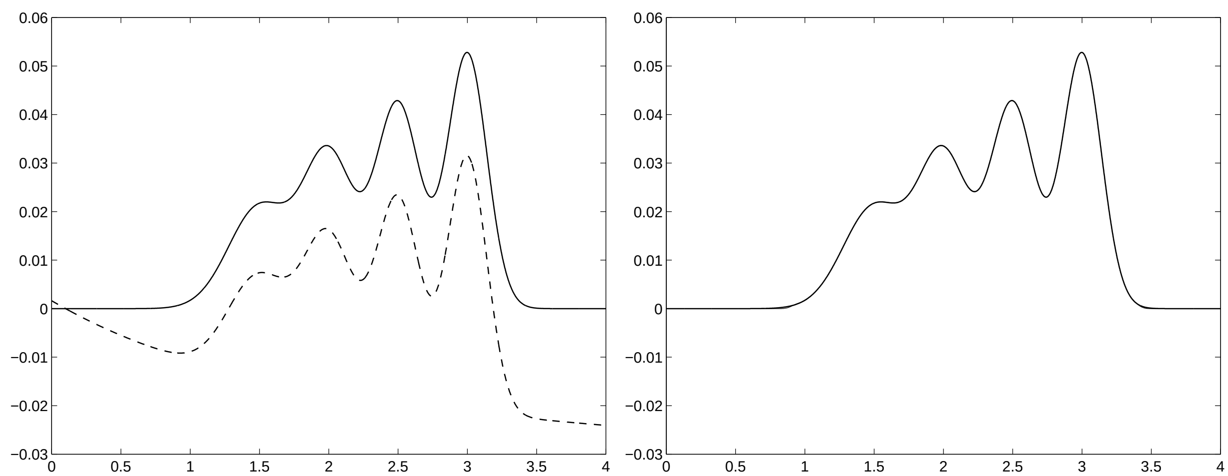

Figure 2: The continuous line in both graphs represents the signal to be separated from the background. The broken line in the left graph depicts the result of applying the oblique projection onto the subspace spanned by the whole B-spline basis on $[0, 4]$. The broken line in the right graph, coinciding with the continuous one, depicts the output of the proposed OBLMP approach.

Since the subspace selection is performed by picking a single vector at each step, there is no guarantee that the required signal splitting will always be achieved. The success should depend on the nature of the signal components and the dictionaries spanning the subspaces for representing them. We hope that the results presented in this letter will stimulate further analysis of the proposed approach.

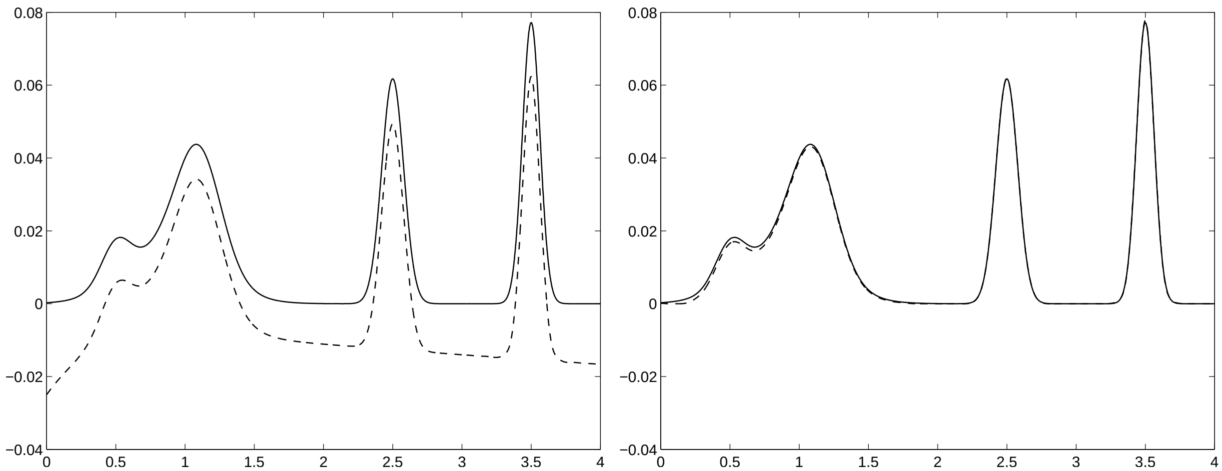

Figure 3: Same description as in Figure 2 but corresponding to the signal in the right graph of Figure 1

References

[1] R. T. Behrens, L. L. Scharf, Signal processing applications of oblique projection operators, IEEE Tran. Signal Processing, 42 (1994), 1413–1424.

[2] M. Unser, A Aldroubi, A general sampling theory from nonideal acquisition devices, IEEE Tran. Signal Processing, 42, no. 11 (1994), 2915–2925.

[3] P. L. Ainsley, Observations on oblique projectors and pseudoinverses, IEEE Tran. Signal Processing, 45, no 7, (1997), 1886–1889.

[4] T. Blu, M Unser, Quantitative Fourier analysis of approximation techniques. I Interpolations and projections, IEEE Tran. Signal Processing, 47 (1999), 2783–2795.

[5] Xiang Yu, Lang Tong, Joint channel and symbol estimation by oblique projections, IEEE Tran. Signal Processing, 49, No 12 (2001), 3074–3083.

[6] Y. C. Eldar, Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors, J. Fourier Anal. Appl, 9 no 1, (2003), 77–96.

[7] Chun Yi Peng, Xianda Zhang, On recursive oblique projectors, IEEE Signal Processing Letters, Vol. 12, no. 6 (2005), 433–436.

[8] S. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, 41, no 12, (1993), 3397–3415.

[9] Y.C. Pati and R. Rezaiifar and P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, Proceedings of the 27th Annual Asilomar Conference in Signals, System and Computers, 1, (1993), 40–44.

[10] L. Rebollo-Neira, D. Lowe, Optimized orthogonal matching pursuit approach, IEEE Signal Proc. Let., 9, no. 4, (2002), 137–140.

[11] J. A. Tropp, Greed is good: algorithmic results for sparse approximation, IEEE Transactions on Information Theory, 50, no. 10, (2004), 2231–2242.

[12] M. Andrle, L. Rebollo-Neira, A swapping-based refinement of orthogonal matching pursuit strategies, Signal Processing, 86, no. 3, (2006), 480-495.

[13] L. Rebollo-Neira, Constructive updating/downdating of oblique projectors, http://arxiv.org/abs/math.FA/0610094, (2006).

[14] L. Rebollo-Neira, Dictionary redundancy elimination, IEE Proceedings – Vision, Image and Signal Processing, 151, 1, (2004), 31–34.

[15] M. Andrle, L. Rebollo-Neira, "Biorthogonal techniques for optimal signal representation", http://www.ncrg.aston.ac.uk/Projects/BiOrthog, (2003).