Subspace modelling for structured noise suppression

Zhiqiang Xu*
Institute of Computational Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100080, China

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

A. Plastino
IFLP-CCT-Conicet
Universidad National University La Plata
CC 727, 1900 La Plata, Argentina

January 16, 2010

Abstract

The problem of structured noise suppression is addressed by i)modelling the subspaces hosting the components of the signal conveying the information and ii)applying a nonlinear non-extensive technique for effecting the right separation. Although the approach is applicable to all situations satisfying the hypothesis of the proposed framework, this work is motivated by a particular scenario, namely, the cancellation of low frequency noise in broadband seismic signals.

1 Introduction

The problem of structured noise suppression concerns the elimination of signal components produced by phenomena interfering with the observations of interest. This problem can be addressed by linear techniques provided that the subspaces hosting the signal components are complementary and well separated [1–4]. More precisely, if a signal represented by the ket $|f\rangle$ is produced as the superposition of two components $|f_1\rangle \in \mathcal{S}_1$ and $|f_2\rangle \in \mathcal{S}_2$, provided

*This work was initiated during Z. Xu stay at the Mathematics Department of Aston University, UK

that $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0\}$, the components of the superposition $|f\rangle = |f_1\rangle + |f_2\rangle$ can be separated by an oblique projection. Even when this condition is theoretically fulfilled, if the subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ are not well separated, the concomitant linear problem for extracting one of the signal components may be ill posed, which causes the failure to correctly split the signal by a linear operation. Hence, nonlinear techniques for determining a subspace $\mathcal{V} \subset \mathcal{S}_1$, such that $|f_1\rangle \in \mathcal{V}$, and the projection onto $\mathcal{V}$ along $\mathcal{S}_2$ is well posed, have been considered [3–5]. In those publications the theoretically complementary subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ are assumed to be known. Nevertheless, the condition $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0\}$ is strong and the possibility of meeting it depends on the ability to generate the right model for the subspaces. Unfortunately, the modelling of the complementary subspaces by pure physical considerations is not always possible and one needs to relay on more general mathematical modelling.

Although the technique for subspace modelling we introduce here is applicable to different situations, the work is motivated by a particular problem relevant to the processing of seismic signals. In the nearshore these signals may be affected by a population of low-frequency waves called *infragravity waves* [6]. This type of noise may be also unavoidable in bottom broadband seismic observations [7]. The interested reader is refereed to [8] for explanations on how infragravity waves are generated. We restrict our consideration to the problem of reducing that type of structured low frequency noise from broadband seismic signals.

Our purpose is twofold. We aim at i)mathematically modelling the subspaces to represent the signal components ii)provide a sparse enough representation of the signals so as to make sure that the correct splitting can be realized.

Under the hypothesis that one of the signals components lies in the subspace of low frequency signals, we determine the subspace of the other component in an adaptive manner. We assume that such a component belongs to an unknown spline space and determine the knots characterising the space by taking into account the curvature points of the signal in hand. In that sense, the space is 'adapted' to the particular signal being analysed. In line with [4] we tackle the problem of finding the representation of this component through the minimisation of the q-norm-like quantity, which is closely related to the non-extensive entropy introduced as ingredient of a thermodynamic framework in the seminal paper by Tsallis [9] and ever since broadly applied in physics [9–16] and other disciplines [17].

The paper is organised as follows: In Section 2 we address the problem of subspaces modelling and discuss the nonlinear non-extensive technique yielding the right signal splitting. A numerical simulation concerning the filtering of low frequency noise from a seismic signal is presented in Section 3. The conclusions are drawn in Section 4.

2 Adaptive subspace modelling for structured noise filtering

As already mentioned, we are concerned with the problem of separating from a signal those components which are not relevant to the phenomenon of interest. For simplicity we consider that a signal $|f\rangle$ is the superposition of only two components and the goal is to find a suitable model for the subspaces hosting such component. Since our work is motivated by the specific problem of filtering low frequency noise from a seismic signal, we further assume that the subspace representing that type of structured noise is spanned by a few Fourier functions. In accordance with previous works, we denote such a subspace as $\mathcal{W}^\perp$ and consider it to be fixed. The composed signal is the superposition $|f\rangle = |f_\mathcal{V}\rangle + |f_{\mathcal{W}^\perp}\rangle$ with $|f_\mathcal{V}\rangle \in \mathcal{V}$ and $|f_{\mathcal{W}^\perp}\rangle \in \mathcal{W}^\perp$.

The goal is to model the subspace $\mathcal{V}$ fulfilling the theoretical condition $\mathcal{V} \cap \mathcal{W}^{\perp} = \{0\}$, regardless of the fact that the two subspaces may be too close to each other for the signal separation to be obtained via a linear approach. We allow for this difficulty by introducing the additional hypothesis that $|f_{\mathcal{V}}\rangle$ is well represented in a subspace of $\mathcal{V}$, which is tantamount to assuming that $|f_{\mathcal{V}}\rangle$ has a 'sparse enough' representation in $\mathcal{V}$. We further assume that $|f_{\mathcal{V}}\rangle$ is well approximated in a dedicated spline space to be constructed as described in the next section.

2.1 Finding the appropriate spline space

Let us start by stating the few definitions on spline spaces which are needed for setting up our mathematical framework. For a complete treatment of splines we refer to [18–20].

Definition 1. *Given a finite closed interval* $[c, d]$ *we define a* partition *of* $[c, d]$ *as the finite set of points*

$$\Delta := \{x_i\}_{i=0}^{N+1}, N \in \mathbb{N}, \text{ such that } c = x_0 < x_1 < \cdots < x_N < x_{N+1} = d. \tag{1}$$

We further define N *subintervals* $I_i, i = 0, \ldots, N$ *as:* $I_i = [x_i, x_{i+1}), i = 0, \ldots, N-1$ *and* $I_N = [x_N, x_{N+1}]$.

Definition 2. *Let* Π_m *be the space of polynomials of degree smaller or equal to* $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. *Let* m *be a positive integer and define*

$$S_m(\Delta) = \{f \in C^{m-2}[c, d] \ : \ f|_{I_i} \in \Pi_{m-1}, i = 0, \ldots, N\}, \tag{2}$$

where $f|_{I_i}$ *indicates the restriction of the function* f *on the interval* I_i.

An *extended partition* with single inner knots associated with $S_m(\Delta)$ is a set $\tilde{\Delta} = \{y_i\}_{i=1}^{2m+N}$ such that

$$y_{m+i} = x_i, \ i = 1, \ldots, N, \ x_1 < \cdots < x_N$$

and the first and last m points $y_1 \leq \cdots \leq y_m \leq c, \quad d \leq y_{m+N+1} \leq \cdots \leq y_{2m+N}$ can be arbitrarily chosen. With each fixed extended partition $\tilde{\Delta}$ there is associated a unique B-spline basis for $S_m(\Delta)$, that we denote as $\{B_{m,j}\}_{j=1}^{m+N}$. The B-spline $B_{m,j}$ can be defined by the recursive formulae [18]:

$$\begin{aligned} B_{1,j}(x) &= \begin{cases} 1, & t_j \leq x < t_{j+1}, \\ 0, & \text{otherwise}, \end{cases} \\ B_{m,j}(x) &= \frac{x - y_j}{y_{j+m-1} - y_j} B_{m-1,j}(x) + \frac{y_{j+m} - x}{y_{j+m} - y_{j+1}} B_{m-1,j+1}(x). \end{aligned}$$

For each order, m, the corresponding spline space is determined by the number and position of the knots. In Fig 1 we show the B-spline basis for two different cubic spline spaces ($m = 4$) having the same number of knots but located at different positions.

In order to find the appropriate partition Δ giving rise to the appropriate spline space $S_m(\Delta)$ to represent a given signal f, we first determine the critical points of the signal's curvature, i.e., we find the set T defined as

$$T \ := \ \{t : \left(\frac{f''}{(1 + f'^2)^{3/2}} \right)' (t) = 0\}. \tag{3}$$

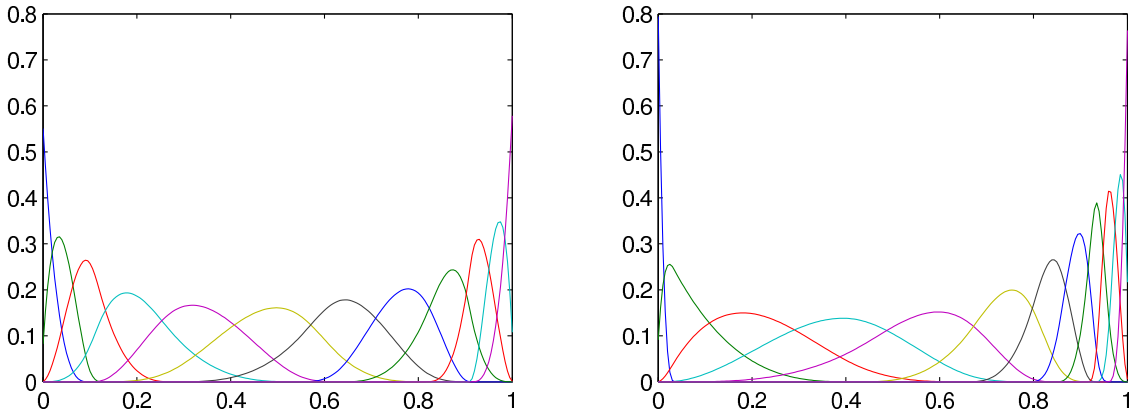

Figure 1: Left graph: Cubic B-spline basis functions for an arbitrary partition. Right graph: Same description as in the left graph but for a different partition with the same number of knots

The entries in T are chosen as the initial knots of Δ. Extra knots are obtained by subdivisions between consecutive knots in T so as to generate a partition Δ with the desired number of knots. An algorithm for implementing this procedure on a signal given as a discrete piece of data is outlined in [21]. We would like to be able to use this procedure on the signal we need to represent, namely the component $|f_{\mathcal{V}}\rangle$, but, of course we do not have access to this signal; our goal is to find it! Thus, in line with [3,4] we proceed as explained below.

Since in our framework $\mathcal{W}^{\perp}$ is fixed and known, we can construct the orthogonal projector onto $\mathcal{W} = (\mathcal{W}^{\perp})^{\perp}$ that we denote $\hat{P}_{\mathcal{W}}$. Assuming now that $\mathcal{V} = S_m(\Delta)$, for some order m and some partition Δ, we can use B-splines to span the space so that for $|f_{\mathcal{V}}\rangle \in S_m(\Delta)$ we have

$$f_{\mathcal{V}}(x) = \langle x|f_{\mathcal{V}}\rangle = \sum_{i=1}^{M} c_i \langle x|B_{i,m}\rangle = \sum_{i=1}^{M} c_i B_{i,m}(x). \tag{4}$$

Hence, by applying the projector $\hat{P}_{\mathcal{W}}$ on both sides of (4) we further have

$$|f_{\mathcal{W}}\rangle = \sum_{i=1}^{M} c_i |u_i\rangle, \quad \text{where} \quad |u_i\rangle = \hat{P}_{\mathcal{W}}|B_{i,m}\rangle, \quad \text{and} \quad |f_{\mathcal{W}}\rangle = \hat{P}_{\mathcal{W}}|f_{\mathcal{V}}\rangle. \tag{5}$$

Denoting by $\hat{I}_{\mathcal{S}}$ the identity operator in $\mathcal{S} = \mathcal{V} + \mathcal{W}^{\perp}$, the projector $\hat{P}_{\mathcal{W}}$ is obtainable from the relation $\hat{P}_{\mathcal{W}} = \hat{I}_{\mathcal{S}} - \hat{P}_{\mathcal{W}^{\perp}}$. Therefore the component $|f_{\mathcal{W}}\rangle$ is available and can be used to determine the knots of the spline space to represent it.

Let us suppose then that through the curvature function (3) we obtain a suitable partition for the space to represent $|f_{\mathcal{W}}\rangle$. Now, in order to obtain the component $|f_{\mathcal{V}}\rangle$ from $|f\rangle$ in the most usual case involving subspaces $\mathcal{V}$ and $\mathcal{W}^{\perp}$ close to each other, we need to find the representation of $|f_{\mathcal{W}}\rangle$ in a subspace of $\mathcal{W} = \text{span}\{|u_i\rangle\}_{i=1}^{M}$. The approach for achieving such an aim is discussed in the next section.

2.2 Determination of the signal representation through a nonlinear non-extensive approach

At this point we can assume that we know the spanning set for $\mathcal{W}$ so that henceforth the problem is reduced to finding the representation (5) with sparse coefficients. For this we could

apply the approach proposed in [4], which entails to use the normal equations

$$\langle u_n | f_{\mathcal{W}} \rangle = \sum_{i=1}^{M} c_i \langle u_n | u_i \rangle, \quad n = 1 \ldots, M \tag{6}$$

as constraints for the minimisation of the q-norm-like quantity $\sum_{i=1}^{M} |c_i|^q$, $0 < q \leq 1$, but incorporating the equations in a stepwise manner. However, in the case motivating this work the number of necessary constraints is large enough to make the whole process slow. Hence rather than using the approach proposed in [4] we take an alternative route and apply a regularised version of the FOCUSS algorithm, which implies to minimise the functional

$$\mathcal{L} = \sum_{i=1}^{M} |c_i|^q + \lambda \| |f_{\mathcal{W}}\rangle - \sum_{i=1}^{M} c_i |u_i\rangle \|^2,$$

where λ is a regularisation parameter. The algorithm for implementing the approach is based on re-weighted least squares and is given in [22]. It comprises the following simple steps:

1) For each fixed q set a value for λ and a value for the initial vector $|c^o\rangle = \sum_{i=1}^{M} c_i^o |i\rangle$.

2) At each iteration, say iteration k, define the operators

$$\hat{U} = \sum_{i=1}^{M} \hat{P}_{\mathcal{W}} |B_{i,m}\rangle\langle i| = \sum_{i=1}^{M} |u_i\rangle\langle i|, \quad \hat{\Pi}_k = \sum_{i=1}^{M} |i\rangle |c_i^{k-1}|^{1-\frac{q}{2}} \langle i|, \quad \hat{U}_k = \hat{U}\hat{\Pi}_k,$$

where the representation $\langle x | B_{i,m} \rangle$, $i = 1 \ldots, M$ of the kets $|B_{i,m}\rangle$, $i = 1 \ldots, M$ are B-spline functions $\langle x | B_{i,m} \rangle = B_{i,m}(x)$ of order m.

3) Compute $|c^k\rangle$ as

$$|c^k\rangle = \hat{\Pi}_k (\hat{U}_k^* \hat{U}_k + \lambda \hat{I})^{-1} |f_{\mathcal{W}}\rangle,$$

where $\hat{U}_k^*$ indicates the adjoint of $\hat{U}_k$ and $\hat{I}$ the identity operator.

4) Given a small δ, while $\| |c^k\rangle - |c^{k-1}\rangle \| > \delta$ repeat 2) and 3).

When for the given δ the stopping condition in 4) has been reached, say at iteration K, set $c_i = c_i^K$, $i = 1, \ldots, M$ and compute the required component $|f_{\mathcal{V}}\rangle$ as

$$|f_{\mathcal{V}}\rangle = \sum_{i=1}^{M} c_i |B_{i,m}\rangle.$$

Remark 1. *For $q < 1$ the optimisation of the functional $\mathcal{L}$ is a non convex problem yielding, typically, a large number of local minima. However, it is shown in [22] that the above algorithm reduces the cost function $\mathcal{L}$ at each step and the stable fixed points of the algorithm are sparse. This is a desired feature for finding the solution of our problem.*

3 Application to filtering of structured low frequency noise from a seismic signal

We apply here the proposed approach for filtering low frequency noise from a seismic signal. As already mentioned, a common interference with broadband seismic signals is produced by long waves, generated by known or unknown sources, called infragravity waves [6]. This interference is referred to as low frequency noise as it falls in a frequency range of up to 0.05 Hz. Thus, the model for the subspace of that type of structured noise, on a signal given by $L = 403$ samples, is

$$\mathcal{W}^{\perp} = \text{span}\{e^{\imath\frac{2\pi n(i-1)}{L}},\, i = 1, \ldots, L\}_{n=-21}^{21}. \tag{7}$$

The particular realization of the noise we have simulated is plotted in the top left graph of Fig 2 (signal f_i^n, $i = 1, \ldots, L$). However, it is appropriate to recall that the success of the approach does not depend on the actual form of the noise (as long as it belongs to the subspace given in (7)) because the approach guarantees the suppression of the whole subspace $\mathcal{W}^{\perp}$.

The seismic signal f_i^s, $i = 1, \ldots, L$ shown in the top right graph of Fig 2 is a segment of the seismic signal which is provided by the WaveLab802 Toolbox (it is acknowledged there that such a signal is distributed throughout the seismic industry as a test dataset). The left bottom graph is the superposition of the signals in the top graphs, i.e. $f_i = f_i^n + f_i^s$, $i = 1, \ldots, L$. We first subtract from f the component in $\mathcal{W}^{\perp}$ to obtain $f_{\mathcal{W}} = f - \hat{P}_{\mathcal{W}^{\perp}} f$ and use this signal to find the points in the set (3). Then we subdivide uniformly those points to obtain 341 nonuniform knots defining the signal space. Using these knots we construct the nonuniform B-spline basis for the space (MATLAB codes for the implementation of both steps are available from [23]).

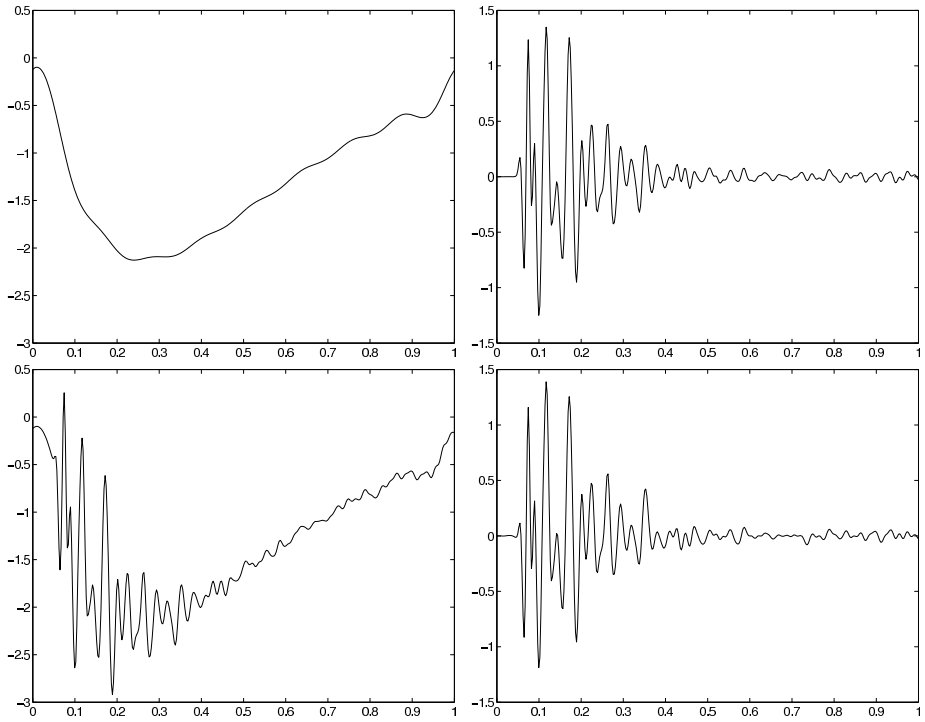

Figure 2: Top left graph: Simulated low frequency noise (f^n). Top right graph: Piece of seismic signal (f^s). Left bottom graph: Signal plus noise ($f = f^s + f^n$). Right bottom graph: Approximation f^q recovered from the left graph by applying the proposed approach for $q = 0.4$ and $\lambda = 5 \times 10^{-8}$.

The methodology discussed in the previous section requires to fix the values for q and λ. To allow for good resolution the regularisation parameter λ should be given a small value. The right bottom graph of Fig 2 depicts the filtered signal arising by applying the regularised FOCUSS method for $\lambda = 5 \times 10^{-8}$ and $q = 0.4$. The left graph of Fig 3 plots, $|f^q - f^s|$, the absolute value of the difference between the approximation f^q (for $q = 0.4, \lambda = 5 \times 10^{-8}$) and the true signal f^s. For the sake of comparison, in the right graph we have plotted $|f^f - f^s|$, where f^f is the approximation arising by filtering with Fast Fourier Transform (FFT). For this approximation we simply take the FFT of f, eliminate the frequencies components in (7), and apply the inverse transform to obtain f^f. The comparison shows that the proposed approach may perform considerably better than standard FFT filtering.

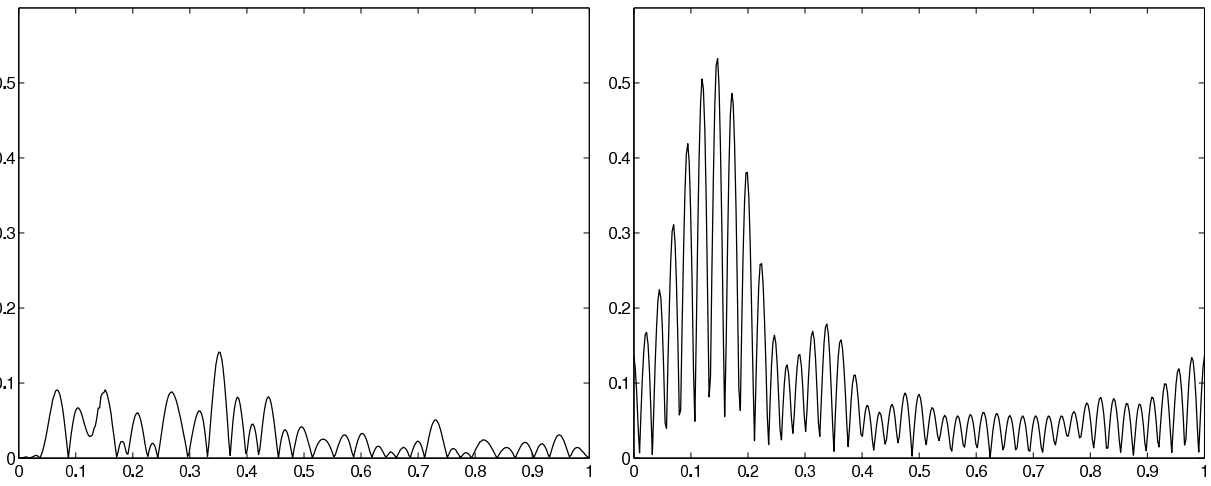

Figure 3: Left graph: Absolute value of the difference between the true seismic signal f^s and the approximation f^q obtained by the proposed approach with $q = 0.4$ and $\lambda = 5 \times 10^{-8}$. Right graph: Absolute value of the difference between the true seismic signal f^s and the approximation f^f obtained by filtering frequencies in the FFT of f.

In order to analyse the dependence of the solution on the parameters q and λ we have considered the range $5 \times 10^{-9} \leq \lambda \leq 5 \times 10^{-6}$ and let the parameter q vary in the interval $(0, 1]$ with step 0.001. For each λ we calculated the error's norm $\epsilon(q) = \|f^q - f^s\|$. The plot of this error, against q, is depicted by the three lines of Fig 4, each of which corresponds to a different value of λ. The two close lines correspond to $\lambda = 5 \times 10^{-9}$ and $\lambda = 5 \times 10^{-8}$ whilst the broken line corresponds to $\lambda = 5 \times 10^{-6}$. For producing these graphs we have varied the parameter λ with step 10^{-9} and use the solution for each λ to initialise the recursive procedure yielding the solution for the next value of λ. By this process we have ensured to fall in equivalent local minima and that is the reason why the curves of Fig 4 for the smallest and largest value of λ are still similar. By initialising the regularised FOCUSS algorithm in a different way different curves were obtained. However, for small λ values all the solutions were more satisfactory than those obtained by filtering with FFT. As a final remark it may be worth stressing that, as opposed to the non-extensive q-entropy [9, 10], the q-norm-like quantity for $q = 1$ (yet leading to a convex optimisation problem) is non-extensive.

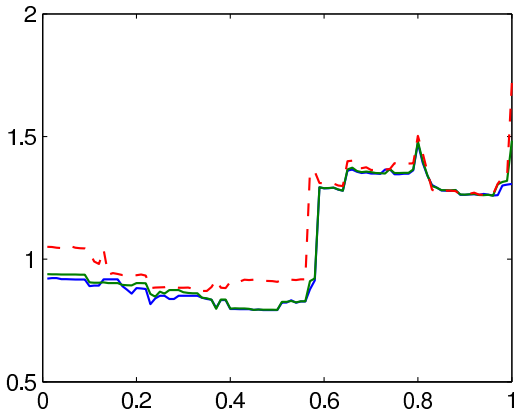

Figure 4: Norm of the approximation error $\epsilon(q) = ||f^s - f^q||$ as a function of the parameter q for $\lambda = 5 \times 10^{-9}$, $\lambda = 5 \times 10^{-8}$ (close lines) and $\lambda = 5 \times 10^{-6}$ (broken line).

4 Conclusions

The problem of structured noise suppression has been considered by modelling the subspaces of the signal components and applying a nonlinear non-extensive technique for separating them. The work was motivated by the problem of filtering infragravity waves from broadband seismic signals. For this, the noise subspace was modelled using low frequency Fourier functions and that of the other component by a dedicated spline space (adapted to the signal in hand). A simulation involving a piece of seismic signal distributed by the seismic industry and noise up to 0.05Hz has produced encouraging results in comparison with those arising by standard Fourier Transform filtering.

Acknowledgements

Support from the Engineering and Physical Sciences Research Council (EPSRC), UK, grant EP/D06263/1, is acknowledged.

References

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

[2] L. Rebollo-Neira, Constructive updating/downdating of oblique projectors: a generalization of the Gram–Schmidt process, Journal of Physics A: Mathematical and Theoretical 40 (2007) 6381–6394.

[3] L. Rebollo-Neira, Measurements design and phenomena discrimination, J. Phys. A: Math. Theor. 42 (2009) 165210.

[4] L. Rebollo-Neira, A. Plastino, Nonlinear non-extensive approach for identification of structured information, Physica A 388 (2009) 4703–4712.

[5] L. Rebollo-Neira, Oblique matching pursuit, IEEE Signal Processing Letters 14 (10) (2007) 703–706.

[6] R. D. Kosian, N. V. B. Pykhov, B. L. Edge, Coastal processes in tideless seas, ASCE Publications, 2000.

[7] D. Dolenc, B. Romanowicz, B. Uhrhammer, P. McGill, D. Neuhauser, D. Stakes, Identifying and removing noise from the Monterey ocean bottom broadband seismic station (MOBB) data, Geochem. Geophys. Geosyst., 8, (2007) Q02005, doi:10.1029/2006GC001403.

[8] W. C. Crawford, S. C. Webb, Identifying and removing tilt noise from low-frequency (< 0.1 Hz) seafloor vertical seismic data, Bull. Seism. Soc. Am., 90, 952-963, 2000.

[9] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys., 52, (1988) 479.

[10] C. Tsallis, Introduction to nonextensive statistical mechanics, Springer-Verlag, NY, (2009).

[11] A. R. Plastino, A. Plastino, Tsallis Stellar Polytropes and Tsallis' entropy, Physics Letters A, 174 (1993) 834–386.

[12] A. R. Plastino, A. Plastino, Tsallis Entropy, Erhenfest Theorem and Information Theory, Physics Letters A, 177 (1993) 177–179.

[13] S. Abe, Y. Okamoto, Nonextensive Statistical Mechanics and Its Applications Series: Lecture Notes in Physics , Vol. 560 Springer-Verlag, NY (2001).

[14] A. Taruya, M. Sakagami, Gravothermal catastrophe and Tsallis generalized entropy of self-gravitating systems, Physica A, 307, 1-2 (2002) 185–206. AS01

[15] J. Andrade Jr., M.P. Almeida, A.A. Moreira, G.A. Farias, Extended phase-space dynamics for the generalized nonextensive thermostatistics, Phys. Rev. E 65, (2002) 036121

[16] G. Adesso, A. Serafini, F. Illuminati, Extremal entanglement and mixedness in continuos variable systems, Phys. Rev. A 70 (2004) 022318.

[17] M. Gell-Mann, C. Tsallis, (Editors) Nonextensive entropy Interdisciplinary Application (Santa Fe Institute Studies on the Sciences of Complexity) Oxford University Press, USA (2004).

[18] L. Schumaker, Spline Functions: Basic Theory, Wiley, New York, 1981.

[19] C. K. Chui, Multivariate splines, SIAM, Philadelphia, 1988.

[20] Carl De Boor, A Practical Guide to Splines, Springer, New York, 2001.

[21] L. Rebollo-Neira, Z. Xu, Adaptive non-uniform B-spline dictionaries on a compact interval, arXiv:0908.0691v1 [math.FA]

[22] B.D. Rao, K. Engan, S. F. Cotter, J. Palmer, K. Kreutz-Delgado, Subset selection in noise based on diversity measure minimization, IEEE Transactions on Signal Processing, 51, 3 (2003) 760– 770, 10.1109/TSP.2002.808076.

[23] http://www.ncrg.aston.ac.uk/Projects/HNLApprox