SAMPLING THEORY IN SIGNAL AND IMAGE PROCESSING Vol. 1, No. 1, Jan. 2002, pp. 0-50

ISSN: 1530-6429

Comparison of numerical methods for the computation of energy spectra in 2D turbulence. Part II: Adaptative algorithms

Ch.H. Bruneau, P. Fischer, Z. Peter, A. Yger
Institut de Mathématiques de Bordeaux, Université Bordeaux 1
351 cours de la Libération
33405 Talence France
bruneau@math.u-bordeaux.fr

Abstract

The first part of this work was devoted to the comparison of direct numerical methods for the computation of energy spectra in 2D turbulence. Here such direct methods are mixed together and combined with adaptative algorithms such as matching pursuit or orthogonal matching pursuit. It appears curiously that the proper orthogonal decomposition basis is sometimes less adapted to the reconstruction process than cosine or wavelet packets dictionaries.

Key words and phrases : matching pursuit algorithms, proper orthogonal decomposition, wavelet and cosine packets.

2000 AMS Mathematics Subject Classification — 94A12, 62M10

1 Introduction

In the first part of this work [1] various direct methods were applied to compute energy spectra of 1D signals coming from direct numerical simulations of the 2D turbulent flow behind arrays of cylinders [2]. Some of these direct methods, namely the proper orthogonal decomposition (POD) and the cosine or wavelet packets decomposition provide dictionaries which are used here in adaptative matching pursuit (MP) or orthogonal matching pursuit (MPO) algorithms. After a brief presentation of these algorithms, numerical results are provided using first the POD modes dictionary, then wavelet or cosine packets dictionaries and finally mixed dictionaries to analyse the same signals than in the first part.

2 Matching pursuit algorithms

The matching pursuit (MP) algorithm, introduced by S. Mallat [4, 5], allows a clever decomposition of a given signal s into a linear combination of functions (also called atoms) which are selected from a redundant dictionary of signals with normalised energy equal to 1. These atoms are selected in order to fit in the best way the structure of the signal. The first selected atom d_1 is chosen so that the modulus $|\langle s\,,\,d_1\rangle|$ of its correlation coefficient with s is maximal ; then d_2 is chosen within the dictionary so that

$$\left|\Big\langle s - \langle s\,,\,d_1\rangle\, d_1\,,\,d_2\Big\rangle\right|$$

is maximal, and so on.
The sequence of coefficients is decreasing and indicates the order in which the corresponding atoms are selected. The same atom may be selected several times. Thus the final coefficient of an atom summarises its global contribution to the reconstruction, its order of appearance being forgotten. To overcome this inconvenient, the MP algorithm can be combined with the Gram-Schmidt process so that at any step, the partial reconstructed signal

$$\sum_{j=1}^{n} \alpha_k d_k$$

is orthogonal to the $(k+1)$-th atom d_{k+1} which is selected. When using this orthogonal version of the algorithm (MPO), one may take away from the dictionary any atom once it has been selected. In the MPO algorithm the coefficients give directly the whole contribution of the corresponding selected atoms. Besides, the number of iterations equals the number of atoms needed for a required ratio of reconstruction as a new atom is selected at each iteration. The weakness of that MPO algorithm lies on the fact that the Gram-Schmidt re-orthogonalization process might modify the coefficients and eventually erase some of them. Therefore, to keep track of the decomposition it is necessary for each selected atom to take into account its coefficient the first time the atom appears.
The MP algorithm is a nonlinear procedure which selects the components of the signal which are coherent respect to a given dictionary. It also provides a decomposition of the signal with such coherent structures as atoms. Whenever a segment of the signal is badly correlated with any element of the dictionary, it is cut into many pieces and thus its content is spread out. The MP algorithm can be compared to the wavelet packets method as follows [5] :

- on the one hand, wavelet packets decompositions are not well adapted to the analysis of highly non-stationary signals since entropy criteria exploit global properties of the signal. Indeed, wavelet packets correspond to precise frequency ranges which are directly connected with the selection tree of the algorithm. On the opposite, in the greedy approach of the MP algorithm, the choice of the atoms is completely free.

- on the other hand, the best basis algorithm is more efficient when applied to stationary signals.

Therefore, as soon as the dictionary contains enough irregular functions, the MP algorithm should give better results than wavelet packets decompositions since it isolates non-stationary components.

3 Matching pursuit with POD modes dictionaries

Given a signal s, a natural dictionary $\mathcal{V}$ for the MP algorithm could be the POD modes $V_1, ..., V_q$ computed from a set of snapshots corresponding to successive overlapping or not segments partitioning the signal. The more redundant the segmentation of s with snapshots is, the more efficient the MP algorithm realised with corresponding POD modes will be. Let us take for example the four dictionaries corresponding to the segmentation of the signals of the pressure p, the two components of the velocity u and v, and the vorticity ω with 39 non-overlapping segments. The MP algorithm is applied to each of the four signals p, u, v, ω with the four dictionaries and the crossed reconstruction ratio after 25 iterations of the algorithm is given on Table 1. Of course the best results are obtained on the diagonal as the used dictionary corresponds to the signal. Moreover, due to the nature of the signal which is very different for the vorticity as we have seen in [1], the crossed results with the vorticity are the worst. We see also that after only 25 iterations we get more than 99% of the L^2 norm whereas the 33 first POD modes were necessary to get the same amount of the norm with the POD decomposition [1].

Now, let $\mathcal{D}$ be the dictionary with the 39 POD modes constructed from the 39 non overlapping snapshots of the vorticity signal. The MPO algorithm is applied to the family of 305 segments of length 1024 which are obtained by translation of 128 points. On Figure 1, is plotted the number of times a given POD mode V_j is selected as the first atom d_1. We can see that the first POD mode V_1 is chosen 31 times as d_1, the second mode 25 times, and so on. The first four modes are more often selected whereas the last ones are sometimes not even selected once. This is in accordance with the construction of the POD modes although the modes have been constructed only with 39 snapshots. On Figure 2(a) are plotted the absolute values of the MPO coefficients (from dark grey for the biggest ones to light grey for the smallest ones) obtained when decomposing each of the 305 segments. Let us point out that the clear vertical strips show the intermittencies of the signal. In addition Figure 2(b) shows that in these strips, taking away the 39 snapshots, the reconstruction rate is much lower.

Finally in this subsection, we would like to point out how the POD modes can be very well adapted to a segment of the original signal since the first coefficient of the MP algorithm may induce a reconstruction up to 58%. This is better than what can be obtained for some segments with the whole reconstruction as we have seen for S_2 in the previous section. Indeed, if we carry out a point by point translation of a window of 1024 points over the entire original vorticity signal of

39936 points, one will have a total of 38913 windows. Then, we proceed the MP algorithm on each window and pick up the segment for which the reconstruction after the first iteration is the best. On the Figure 3 one can see the correspondence between the shape of this original segment and its reconstruction.

	u	v	p	ω
u POD set	99.73%	96.18%	92.42%	79.47%
v POD set	94.81%	99.87%	93.40%	80.40%
p POD set	93.11%	96.35%	99.81%	83.24%
ω POD set	78.20%	81.36%	80.19%	99.44%

Table 1: Crossed reconstruction ratio for various physical quantities after 25 iterations of the matching pursuit algorithm with POD dictionary

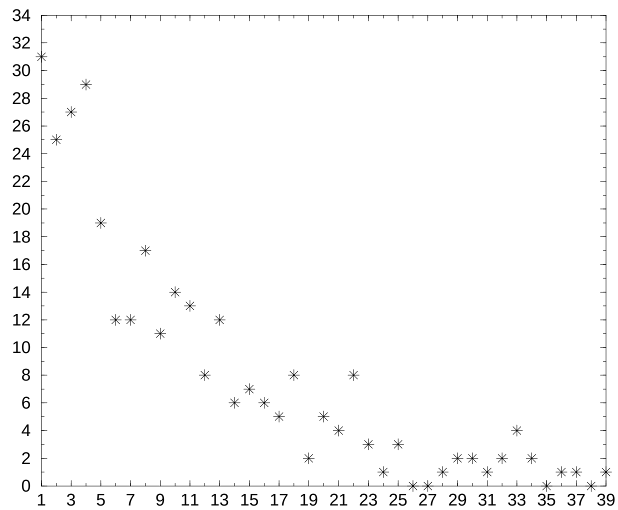
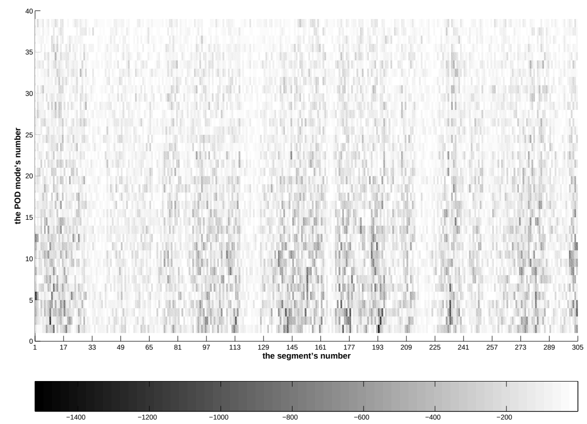


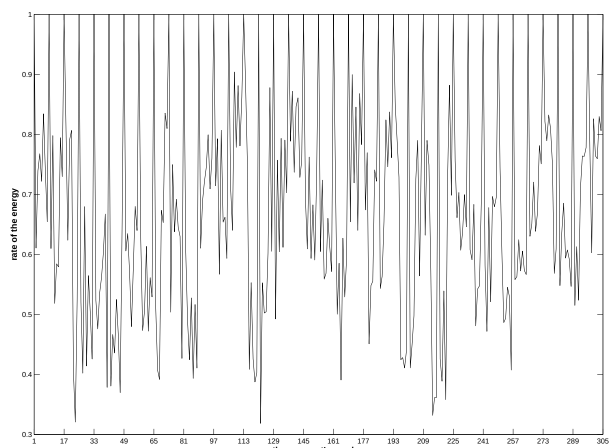
Figure 1: Number of times a POD mode is chosen as d_1

4 Matching Pursuit with wavelet packets or cosine packets dictionaries

Wavelet packets and cosine packets described in [1] can also be used as dictionaries as implemented in [3]. We decide to decompose both segments S_1 and S_2 of the vorticity signal with such dictionaries in order to see how many atoms are necessary to get a reconstruction up to 99% of the L^2 norm. The Table 2 indicates the number of atoms required for four different wavelet packets basis and cosine packets. Once again, less atoms are required to approximate segment

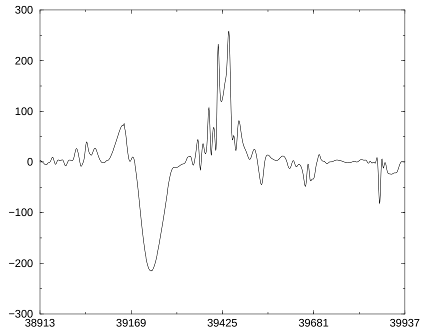


(a) *MPO coefficients intensity*

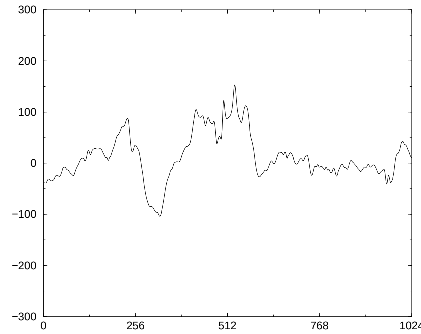


(b) *Reconstruction rate*

Figure 2: Influence of the position of the segments on the MPO decomposition with the POD modes dictionary



(a) *the selected segment*



(b) *the 58% reconstruction after the first iteration of the MP algorithm*

Figure 3: The part of the original vorticity signal which gives the best correlation with a POD dictionary atom

S_1 than segment S_2. In addition, we see that the number of atoms is lower than the number of atoms which is involved in the split and merge algorithm based on the minimisation of the entropy given on Table 3 (quoted from part I). This is due to the fact that in the MP algorithm there is no selection tree to explore the correlation of the signal with the dictionary atoms.

Then, we perform the same experiment on the whole signal s and compute the slopes of the energy and enstrophy cascades by basic Fourier method for the reconstructed signals. The results are reported in Table 4. The computed slope of the inverse cascade remains in the same range whatever the dictionary is and fits the slope computed in the previous sections with other methods. On the contrary, the results for the enstrophy cascade are very different from one basis to another as the slope varies from -4.36 to -3.09.

Basis type	# elements for S_1	# elements for S_2
Haar	33	106
Daubechies6	20	77
Coiflet2	13	76
Symmlet8	18	68
Cosine	23	72

Table 2: Number of atoms needed to reconstruct S_1 and S_2 up to 99% of the L^2 norm with the matching pursuit algorithm

Basis type	# elements for S_1	# elements for S_2
Haar	47	143
Daubechies6	22	121
Coiflet2	24	101
Symmlet8	21	100
Cosine	33	156

Table 3: Number of elements necessary to reconstruct signals S_1 and S_2 up to 99% of the L^2 norm with the best basis algorithm

Basis type	# elements	enstrophy cascade	inverse cascade
Haar	1150	-3.09	-1.82
Daubechies6	639	-3.60	-1.76
Coiflet2	594	-4.06	-1.83
Symmlet8	573	-4.36	-1.86
Cosine	701	-3.50	-1.78

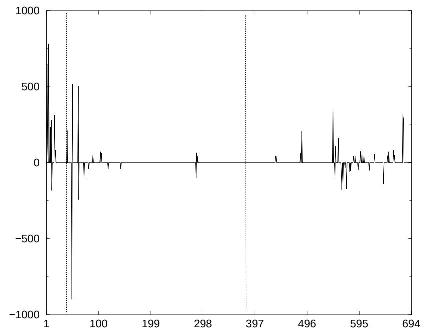
Table 4: Number of atoms needed to reconstruct the whole signal up to 99% of the L^2 norm with the matching pursuit algorithm and cascades slopes of the reconstructed signal

5 Matching pursuit with mixed dictionaries

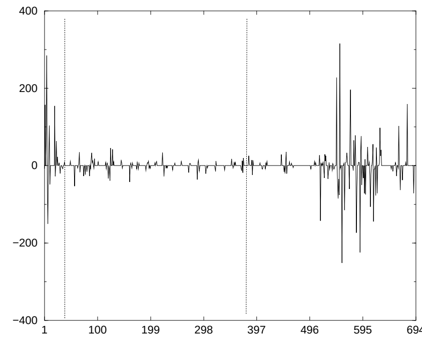
It was shown in the previous sections that POD modes are not enough to reconstruct the two segments S_1 and S_2 up to 99% of the L^2 norm whereas this was possible with wavelet packets or cosine packets dictionaries. Therefore it is interesting to see what happens when dealing with a mixed POD modes and packets dictionary in order to compare the contribution of the atoms in both classes to the reconstruction of the segments. The new dictionary $\mathcal{D}_1$ is made of the 39 POD modes and approximately 341 cosine packets and 314 wavelet packets chosen in such a way that they represent properly the different scale and frequency ranges. Using the matching pursuit algorithm it is then possible to determine the importance of each atom of that mixed dictionary in the reconstruction. On Figure 4 are represented the coefficients of the selected atoms, necessary to reconstruct the two segments with more than 99% of the L^2 norm. The reconstruction needs respectively 54 atoms for S_1 and 203 atoms for S_2 ; the bigger number for S_2 could be expected from the previous numerical experiments. We see that in both cases the atoms are selected among the three sub-dictionaries separated on the figure by dotted vertical lines and their repartition is directly linked to the shape of the original segments. We can notice that the values of the coefficients are bigger for S_1 than for S_2 so that the selected atoms better correlate with the first than with the second segment. This explains why we need much more atoms to reconstruct S_2. Although the difference between the two segments is clear, it is not so easy to compare the number of selected atoms with those obtained in Table 2. On the one hand, the values presented in that table are the results of the matching pursuit performed on the entire packet dictionary while $\mathcal{D}_1$ is not so rich. On the other hand, the mixing of POD modes and packets which are not correlated in the same way to the original segments induces extra corrections of the algorithm to smooth the difference. Besides it is interesting to remark that, although the POD sub-dictionary contains few atoms, its weight in the reconstruction process is significant for segments such as S_1 and S_2 and not only for the snapshots. For S_1 which is smoother, the algorithm chooses at first cosine packets and POD modes whereas for S_2 which contains more high frequencies the algorithm selects first wavelet packets and POD modes.

Even if each snapshot can be exactly reconstructed using all the available POD modes, it is also possible to use on it the MP or MPO algorithms with $\mathcal{D}_1$ dictionary. The results on Figure 5 show that for almost half of them the first selected atom is not a POD mode as it is located above the dotted horizontal line which indicates the limit of the first sub-dictionary. Even when POD modes are selected as first atoms, some atoms are chosen later in the other sub-dictionaries. For example for the first snapshot, the MP algorithm selects 10 POD modes at the beginning and then chooses a Symmlet8 packet as 11^{th} atom as shown on Figure 6(a). To understand this choice, we plot on Figure 6(b) the graph of the remainder after 10 iterations of the MP algorithm and the selected wavelet atom times its coefficient. We see that this atom approximates precisely the L^2 norm but ignores the high frequencies. Nevertheless, it has been selected instead of the other POD modes.

Now, we construct two new reduced dictionaries $\mathcal{D}_{21}$ and $\mathcal{D}_{22}$ respectively adapted to S_1 and S_2 by adding to the POD modes the sub-dictionaries consisting of the 23 or 72 cosine packets and the 18 or 68 Symmlet8 wavelet packets which have been selected in the decomposition of S_1 or S_2 (Table 2). Applying the MPO algorithm, we see on Figure 7 that this time we get the reconstruction with only 20 atoms without any POD modes for S_1 and 81 atoms with only 5 POD modes with small coefficients for S_2. So, it appears that for segments which do not correspond to the snapshots the adapted packets are better correlated than the POD modes. Finally, the POD modes are not so relevant although they are constructed from the snapshots of the original signal. This is due to the fact that they represent a mean behaviour and not the instantaneous behaviour. All these numerical tests show that the choice of the dictionary plays a major role.



(a) *for S1*



(b) *for S2*

Figure 4: The coefficients of the atoms, selected by the matching pursuit algorithm with the mixed dictionary $\mathcal{D}_1$, used to reconstruct the two segments with more than 99% of the L^2 norm

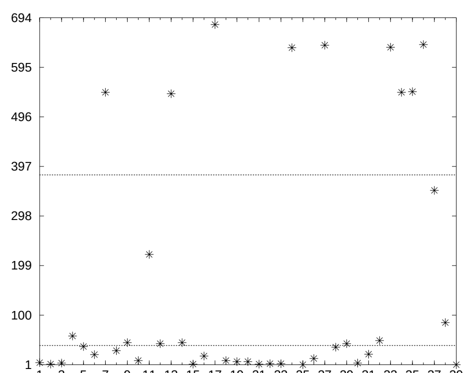
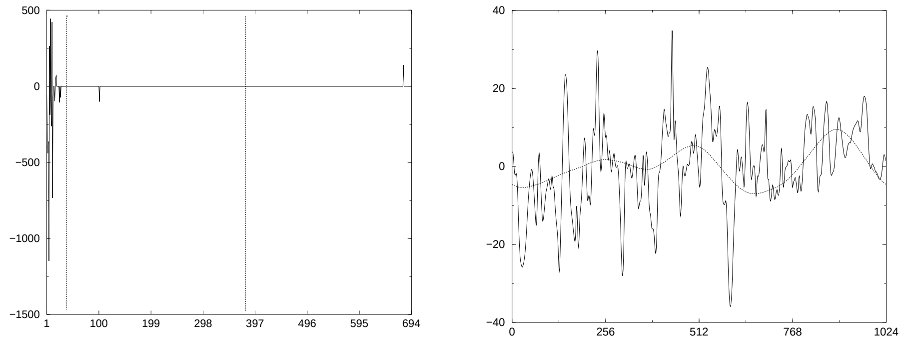


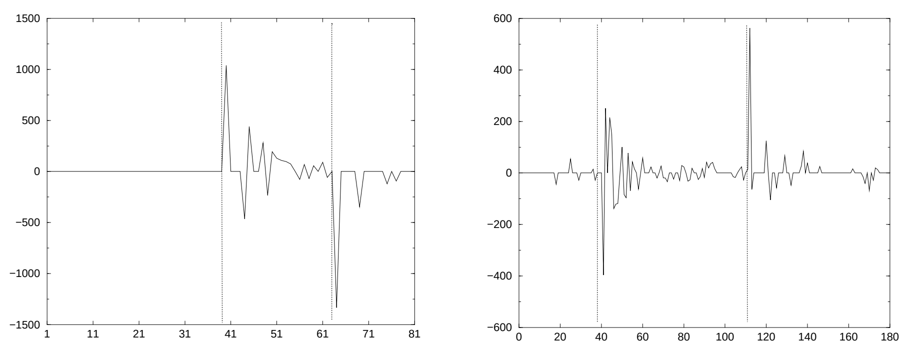
Figure 5: Index of the first selected atom of $\mathcal{D}_1$ for each snapshot



(a) *Coefficients of the selected atoms*

(b) *Comparison between the remainder after* 10 *iterations and the contribution of the* 11^{th} *atom*

Figure 6: Reconstruction of the first snapshot by the matching pursuit algorithm with the mixed dictionary



(a) *For segment* S_1

(b) *For segment* S_2

Figure 7: Coefficients of the selected atoms for the reconstruction by the matching pursuit algorithm with adapted dictionaries

6 Conclusions

Several dictionaries are constructed with the POD modes and with cosine or wavelet packets. Compared to the best basis algorithm developped in part I, the matching pursuit is less efficient for regular parts of the signal but is very well adapted to highly oscillating segments. Surprisingly, when using mixed dictionaries involving POD modes, wavelet and cosine packets, the POD modes are less selected than the packets and even some segments, which have not been chosen as snapshots, can be reconstructed without using POD modes.

References

[1] C.H. Bruneau, P. Fischer, Z. Peter, A. Yger, *Comparison of numerical methods for the computation of energy spectra in 2D turbulence. Part I: Direct methods*, Sampling Theory Signal Image Process. J. , 2005.

[2] C.H. Bruneau, O. Greffier, H. Kellay, *Numerical study of grid turbulence in two dimensions and comparison with experiments on turbulent soap films*, Phys. Rev. E **60** R1162, 1999.

[3] D. Donoho, *Wavelab 802* software, http://www-stat.stanford.edu/~wavelab

[4] S. Mallat, *A wavelet tour of signal processing*, Academic press, 1999.

[5] S. Mallat, Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. Signal Process. **41** $n°$12 3397-3415, 1993.