ISSN 1626-8334

CNRS - Université Pierre et Marie Curie - Université Versailles-Saint-Quentin
CEA - IRD - Ecole Normale Supérieure - Ecole Polytechnique

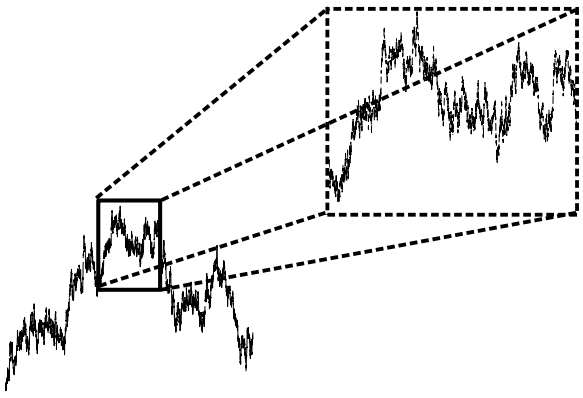
Institut Pierre Simon Laplace
des Sciences de l'Environnement Global

N . A . I
Notes des Activités Instrumentales
Instruments - Expériences - Observations

Compte-rendu de la IIIème journée statistique de l'IPSL (II, Ondelettes)

Marie Farge[1], Kai Schneider[2] and Patrice Abry[3]

[1] *Institut Pierre Simon Laplace / Laboratoire de Météorologie Dynamique, Paris, France*
[2] *CMI, Université de Provence, Marseille, France*
[3] *Laboratoire de Physique du CNRS, ENS, Lyon, France*



Avril 2002 / *April 2002* - **Note n° 20**

Notes des Activités Instrumentales de l'IPSL

http://www.ipsl.jussieu.fr/documentation/NAI/index.htm

• Directeur de la publication:	Jean JOUZEL
• Responsable éditorial:	Laurent MENUT [menut@lisa.univ-paris12.fr]
• Comité d'édition:	Hélène CHEPFER (LMD)
	Yves DANDONNEAU (LODyC)
	Cyrille FLAMANT (SA)
	Cyril MOULIN (LSCE)
	Alain PROTAT (CETP)
	Rémi ROCA (LMD)

Institut Pierre-Simon Laplace - *http://www.ipsl.jussieu.fr*

Université Pierre et Marie Curie	Université Versailles Saint Quentin
B 102 - T15 - E5	Collège Vauban
4, place Jussieu	47, boulevard Vauban
75252 Paris Cedex 05	78047 Guyancourt Cedex

CLIMSTAT GROUP	
Pascale Braconnot	Philippe Kechkut
Alain Chédin (co-chair)	Jean-Pierre Nadal
Monique Dechambre	Philippe Naveau (Project leader)
Laurence Eymard	Sylvie Thiria
Marie Farge	Olivier Talagrand
Claude Frankignoul (co-chair)	Pascal Yiou
Alain Hauchecorne	

Cette note fait suite à la première partie publiée dans les NAI en note n^{o}11, datée de Juillet 2001. Dans ce numéro, l'accent est porté sur les **ondelettes**.

L'emploi des ondelettes pour analyser des processus physiques ou climatiques n'est pas sans pièges pour le non-specialiste. Toutefois, la spécificité des donnees géophysiques (différentes échelles spatiales et temporelles), nécessite des outils d'analyse qui puissent tenir compte de tels changements d'échelle. Pour introduire les nombreuses méthodes liées aux ondelettes à la communaute climatique, des spécialistes de ce domaine en pleine evolution, Marie Farge, Kay Schneider et Patrice Abry nous proposent deux articles dans lesquels une introduction sur les ondelettes et les lois d'echelle et différentes applications liées à la turbulence sont présentées.

Ce numéro a été coordonné par Philippe Naveau.

Analyzing and compressing turbulent fields with wavelets

Marie Farge * & Kai Schneider †

May 27, 2002

Contents

*LMD-CNRS, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05
†CMI, Université de Provence, 39 rue Joliot–Curie, 13453 Marseille Cedex 13

1 Principle

The continuous wavelet transform has been discovered by Alex Grossmann and Jean Morlet who published the first paper on wavelets in 1984 [32]. This mathematical technique, based on group theory and square integrable representations, allows to decompose a signal, or a field, into space, scale and directions. The orthogonal wavelet transform has been discovered in 1986 by Pierre–Gilles Lemarié and Yves Meyer [39]. The development of wavelets has been interdisciplinary, with contributions coming from very different fields such as engineering (subband coding, quadrature mirror filters, time–frequency analysis), theoretical physics (coherent states of affine groups in quantum mechanics) and mathematics (Calderon–Zygmund operators, characterization of function spaces, harmonic analysis). Many reference textbooks are available, among them we recommand [42, 8, 41]. Meanwhile a large spectrum of applications has grown and is still developing, ranging from signal analysis and image processing to data compression. The first application of wavelets to analyze turbulent signals has been published in 1988 [19], and since then a long-term research program is developed for analyzing, computing and modelling turbulent flows with wavelets [16, 20, 21, 23].

Wavelets constitute basic building blocks of various function spaces out of which some can be used to contruct orthogonal bases. The starting point is a function $\psi(x)$, called mother wavelet, which is well-localized (it exhibits a fast decay for $|x|$ tending to infinity), is oscillating (ψ has at least a vanishing integral, or better the first m moments of ψ vanish), and is smooth (the Fourier transform of ψ exhibits fast decay in wavenumber space). The mother wavelet then generates a family of wavelets, $\psi_{l,x}(x') = l^{-1/2}\,\psi(\frac{x'-x}{l})$, by dilatation (or contraction) by the parameter $l > 0$ and translation by the parameter $x \in I\!R$, all wavelets being normalized in L^2-norm. In Fig.1, we show an example of such a family in the orthogonal case for discrete scales $l = 2^{-j}$ and positions $x = 2^{-j} i \Delta x$, with j the scale index for each octave, i the space index and Δx the grid size). The wavelet transform of a function f is then defined as a convolution of the analyzing wavelet with the signal $f(x) \in I\!R$, which gives the wavelet coefficients: $\tilde{f}(l,x) = \int f(x')\psi_{l,x}(x')dx'$. They measure the fluctuations of f around the point x and scale l. The function f can be reconstructed as a linear combination of wavelets $\psi_{l,x}(x')$ and wavelet coefficients $\tilde{f}(l,x)$: $f(x) = {C_\psi}^{-1} \int\int \tilde{f}(l,x)\psi_{l,x}(x') l^{-2}\, dl dx$, C_ψ being a constant which depends on the wavelet ψ. Let us mention that, due to the localization of wavelets in physical space, the behaviour of the signal at infinity does not play any role. Therefore the wavelet analysis and synthesis can be performed locally, in contrast to the Fourier transform where the nonlocal nature of the trigonometric functions does not allow to perform a local analysis.

In this course, we will first define the continuous and the orthogonal wavelet transforms, then recall classical statistical tools used to study turbulent flows and finally present several wavelet statistical tools, together with some associated intermittency measures.

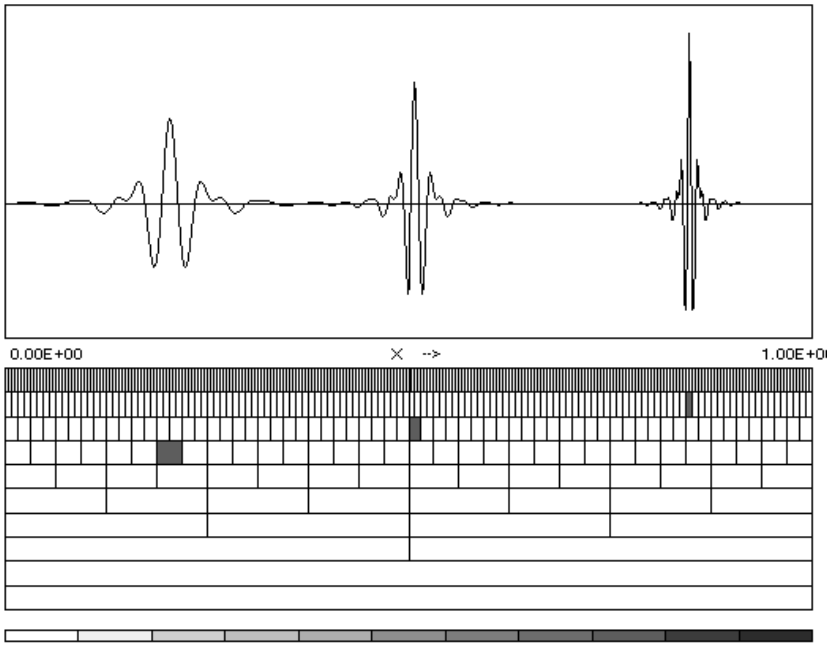


Figure 1: Example: Discrete wavelets $\psi_{j,i}(x) = l_0^{j/2}\,\psi(l_0^j x' - i x_0)$ with $l_0 = 2$ and $x_0 = 1$ and $l = l_0^j$, $x = k x_0 l_0^j$ for the case of orthogonal quintic spline wavelets $\psi_{5,6}(x')$, $\psi_{6,32}(x')$ $\psi_{7,108}(x')$.

2 The continuous wavelet transform

2.1 In one dimension

2.1.1 Analyzing wavelet

Starting point for the wavelet transform is a real or complex valued function $\psi(x)$, called wavelet, which has to fulfill the admissibility condition,

$$C_\psi = \int_0^\infty \left|\widehat{\psi}(k)\right|^2 \frac{dk}{|k|} \; < \; \infty \tag{1}$$

where

$$\widehat{\psi}(k) = \int_{-\infty}^{\infty} \psi(x) e^{-i2\pi kx} dx \tag{2}$$

denotes the Fourier transform. If ψ is integrable this implies that ψ has zero mean,

$$\int_{-\infty}^{\infty} \psi(x) dx \; = \; 0 \qquad \text{or} \quad \widehat{\psi}(k=0) \; = \; 0. \tag{3}$$

In practice however one also requires that the wavelet ψ should be well–localized in both physical and Fourier spaces, which implies smoothness. We also require that higher order moments of ψ vanish, *i.e.*

$$\int_{-\infty}^{\infty} x^m \psi(x) dx \; = \; 0 \qquad \text{for} \quad m = 0, M \tag{4}$$

which means that monomials up to degree M are exactly reproduced. In Fourier space this property is equivalent to

$$\frac{d^m}{dk^n} \, \widehat{\psi}(k)|_{k=0} \; = \; 0 \qquad \text{for} \quad m = 0, M \tag{5}$$

so that the Fourier transform of ψ decays smoothly at $k = 0$.

2.1.2 Analysis

From this function ψ, the so–called mother wavelet, we generate a family of continuously translated and dilated wavelets, normalized in L^2–norm

$$\psi_{l,x}(x') = l^{-1/2} \, \psi(\frac{x' - x}{l}) \qquad \text{for} \quad l > 0 \qquad \text{and} \quad x \in I\!R \tag{6}$$

where l denotes the scale dilation parameter, corresponding to the width of the wavelet and x the translation parameter, corresponding to the position of the wavelet.
In Fourier space this reads

$$\widehat{\psi}_{l,x}(k) = \sqrt{l} \, \widehat{\psi}(lk) \, e^{-i2\pi kx} \tag{7}$$

where the contraction with $1/l$ is reflected in a dilation with l and the translation with x implies a rotation in the complex plane.

The continuous wavelet transform of a signal $f \in L^2(\mathbb{R})$ is then defined as a convolution of f with the wavelet family $\psi_{l,x}$

$$\tilde{f}(l,x) = \int_{-\infty}^{\infty} f(x')\overline{\psi}_{l,x}(x')dx' \tag{8}$$

where $\overline{\psi}_{l,x}$ denotes in the case of complex valued wavelets the complex conjugate. Using Parseval's identity we also get

$$\tilde{f}(l,x) = \int_{-\infty}^{\infty} \hat{f}(k)\overline{\hat{\psi}_{l,x}}(k)dk \tag{9}$$

so that the wavelet transform may be interpretated as a frequency decomposition using band pass filters $\hat{\psi}_{l,x}$ centered at frequency $k = \frac{k_\psi}{l}$, where k_ψ denotes the center of the wavelet in Fourier space, and having variable width $\frac{\triangle k}{k}$, so for increasing scales the bandwidth is getting wider.

In Fig.2, we show an example of a continous wavelet analysis of an academic signal, which is the superposition of harmonic function.

2.1.3 Synthesis

The admissibility condition (1) of ψ implies the existence of a finite energy reproducing kernel, see *e.g.* [8], which is a necessary condition for being able to reconstruct a function from its wavelet coefficients. The signal can thus be reconstructed entirely from its wavelet coefficients,

$$f(x) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} \tilde{f}(l,x)\, \psi_{l,x}(x')\, \frac{dldx}{l^2} \tag{10}$$

which is the inverse wavelet transform.

2.1.4 Energy conservation

There also holds an energy conservation like for Fourier transforms, *i.e.* a Plancherel identity, which means that the total energy of a signal can be either calculated in physical space or in wavelet coefficient space,

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx \;=\; \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} |\tilde{f}(l,x)|^2\, \frac{dldx}{l^2} \tag{11}$$

This formula is also the starting point for the definition of wavelet spectra and scalogram (see 4.3).

2.2 In higher dimensions

2.2.1 Analyzing wavelet

The theory of the continuous wavelet transform can be generalized in several dimensions [47] using rotation in addition to dilatation and translation.

The family of functions $\psi_{l,\vec{x},\theta}$ is obtained from a single one ψ by dilatation with l^{-1}, by translation by $\vec{x}$ and by rotation of angle θ,

$$\psi_{l,\vec{x},\theta}(\vec{x}') \;=\; \frac{1}{l}\psi(R_\theta(\frac{\vec{x}'-\vec{x}}{l})) \tag{12}$$

where R_θ denotes a rotation matrix.
Note that the rotation matrix belongs to the group of rotations in $\mathbb{R}^n$ and depends on the $(n(n-1))/2$ Euler angles θ.

Analogously to the one-dimensional case the wavelet transform is invertible and isometric, provided that ψ fulfills the admissibility condition

$$C_\psi = \int\int_{\mathbb{R}^2} \left|\widehat{\psi}(\vec{k})\right|^2 \frac{d^2\vec{k}}{|\vec{k}|^2} \;<\; \infty. \tag{13}$$

In the following we restrict ourselves to isotropic real-valued wavelets, so there is no more a dependence on the angle θ. The wavelet coefficients can then be calculated using the formula

$$\tilde{f}(l,\vec{x}) \;=\; \int\int_{\mathbb{R}^2} \widehat{f}(\vec{k})\, l\overline{\widehat{\psi}}(l\vec{k})\, e^{i2\pi\vec{k}\cdot\vec{x}}\, d^2\vec{k} \tag{14}$$

where the Fourier transform of the wavelet ψ_l is essentially supported on an annulus a radius $|\vec{k}|$.

2.2.2 Analysis

For a two-dimensional function $f \in L^2(\mathbb{R}^2)$ (constructions for higher dimensions are analogous) we get

$$\tilde{f}(l,\vec{x},\theta) \;=\; \int\int_{\mathbb{R}^2} f(\vec{x}')\,\overline{\psi}_{l,\vec{x},\theta}(\vec{x}')\, d^2\vec{x}' \quad \text{for} \quad l>0,\;\; \theta\in[0,2\pi[,\quad \vec{x}\in\mathbb{R}^2. \tag{15}$$

2.2.3 Synthesis

The function $f \in L^2(\mathbb{R}^2)$ can be recontructed from its wavelet coefficients by the relation

$$f(\vec{x}') = \frac{1}{C_\psi} \int_0^\infty \int\int_{\mathbb{R}^2} \tilde{f}(l,\vec{x})\,\psi_l(\vec{x}'-\vec{x})\,\frac{d^2\vec{x}\,dl}{l^3}. \tag{16}$$

2.2.4 Energy conservation

The energy conservation reads

$$\int\int_{\mathbb{R}^2} |f(\vec{x})|^2 \, d\vec{x} \;=\; \frac{1}{C_\psi} \int_0^\infty \int\int_{\mathbb{R}^2} |\tilde{f}(l,\vec{x})|^2 \, \frac{dl \, d^2\vec{x}}{l^3} \tag{17}$$

2.3 Algorithm

To illustrate the practical implementation of the continuous wavelet transform we consider a one-dimensional signal $f(x)$ sampled on a regular grid with $N = 2^J$ points, *i.e.* the given data are $f(2^{-J}i)$ for $i = 0, ..., 2^J - 1$. We assume the signal to be periodic and compute the wavelet coefficients by means of the Fast Fourier Transform (FFT).

The large scale corresponds to the domain size, which is by construction equal to 1 and the smallest scale of the discretization corresponds to $\frac{2}{N} = 2^{1-J}$. The range of scales is discretized logarithmically, such as

$$l_j \;=\; l_0^{-j} \qquad j \geq 0 \quad . \tag{18}$$

The choise of l_0 is determined to ensure a given precision for the reconstruction of the signal by a discretized version of formula (10). For instance, we choose $l_0 = 2^{1/4}$ (i.e. four voices per octave) for Hermitte wavelets, which are derivatives of Gaussians, to insure that the scale sampling is sufficient to have enough precision for the reconstruction of f [8].

To compute $\tilde{f}(l,x)$ we discretize the formula (9). First we compute the disrete Fourier transform of the signal samples. Then we multiply it with the wavelet filter in Fourier space, and we subsequently obtain at each scale l_j, for $j = 0$ to $J-1$, and at all positions $x_i = \frac{i}{N} for i = 0, N-1$ the wavelet coefficients by executing an inverse FFT

$$\tilde{f}(l_j, x_i) \;=\; \sum_{k=-N/2}^{N/2-1} \widehat{f}_k \, \sqrt{l_j} \, \overline{\widehat{\psi(l_j k)}} e^{+i2\pi k i/N} \tag{19}$$

where $\widehat{f}_k$ denotes the discrete Fourier transform of the samples $f\left(\frac{n}{N}\right)$,

$$\widehat{f}_k \;=\; \frac{1}{N} \sum_{n=0}^{N-1} f\left(\frac{n}{N}\right) \, e^{-i2\pi nk/N} \tag{20}$$

Due to periodicity of the signal f, no boundary effects are introduced by using the FFT. The complexity of the algorithm is of order $N \log N$ at each scale due to the use of the FFT.

The above algorithm can be applied analogously in the two-dimensional case using the tensor product discretization together with 2D FFTs.

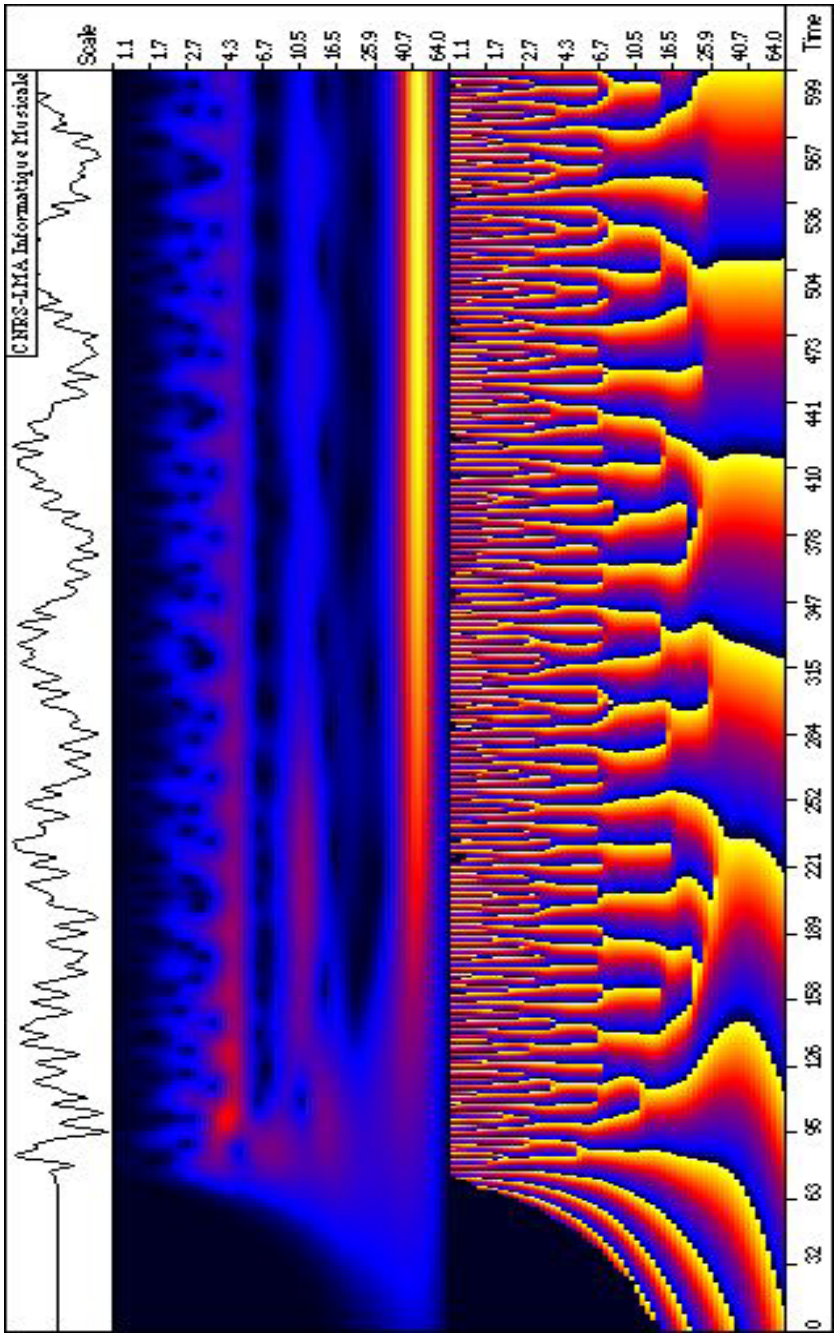


Figure 2: Example of a one-dimensional continuous wavelet analysis. Top: The signal to be analyzed. Middle: the modulus of its wavelet coefficients. Bottom: the corresponding phase.

3 The orthogonal wavelet transform

3.1 In one dimension

3.1.1 Orthogonal scaling functions

The discrete wavelet transform relies on the concept of multi-resolution analysis (MRA) which is a sequence of imbedded subspaces V_j verifying

$$V_j \subset V_{j+1} \qquad \forall j \in \mathbb{Z} \tag{21}$$

$$\overline{\bigcup}_{j\in\mathbb{Z}} V_j = L^2(\mathbb{R}) \tag{22}$$

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\} \tag{23}$$

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1} \tag{24}$$

A scaling function $\phi(x)$ is required to exist. Its translates generate a basis in each V_j, *i.e.*

$$V_j = \overline{span}\{\phi_{ji}\}_{i\in\mathbb{Z}} \quad . \tag{25}$$

where

$$\phi_{ji}(x) = 2^{j/2}\phi(2^j x - i) \qquad j, i \in \mathbb{Z} \quad . \tag{26}$$

This basis is orthonormal, so that

$$< \phi_{ji}, \phi_{jk} >_{\mathbb{R}} = \delta_{ik} \tag{27}$$

with $< f, g >_{\mathbb{R}} = \int_{-\infty}^{+\infty} f(x)\,\overline{g}(x)\,dx$ being the inner product in $L^2(\mathbb{R})$.

The main issue of the wavelet approach now is to work with the orthogonal complement spaces W_j defined by

$$V_{j+1} = V_j \oplus W_j \tag{28}$$

3.1.2 Orthogonal wavelets

Based on the function $\phi(x)$ one can find a function $\psi(x)$, the so-called mother wavelet. Their translates and dilates constitute a orthonormal bases of the spaces W_j,

$$W_j = \overline{span}\{\psi_{ji}\}_{i\in\mathbb{Z}} \tag{29}$$

where

$$\psi_{ji}(x) = 2^{j/2}\psi(2^j x - i) \qquad j, i \in \mathbb{Z} \tag{30}$$

use in 6.1 for analyzing turbulent signals.

Each function $f \in L^2(\mathbb{R})$ can now be expressed as

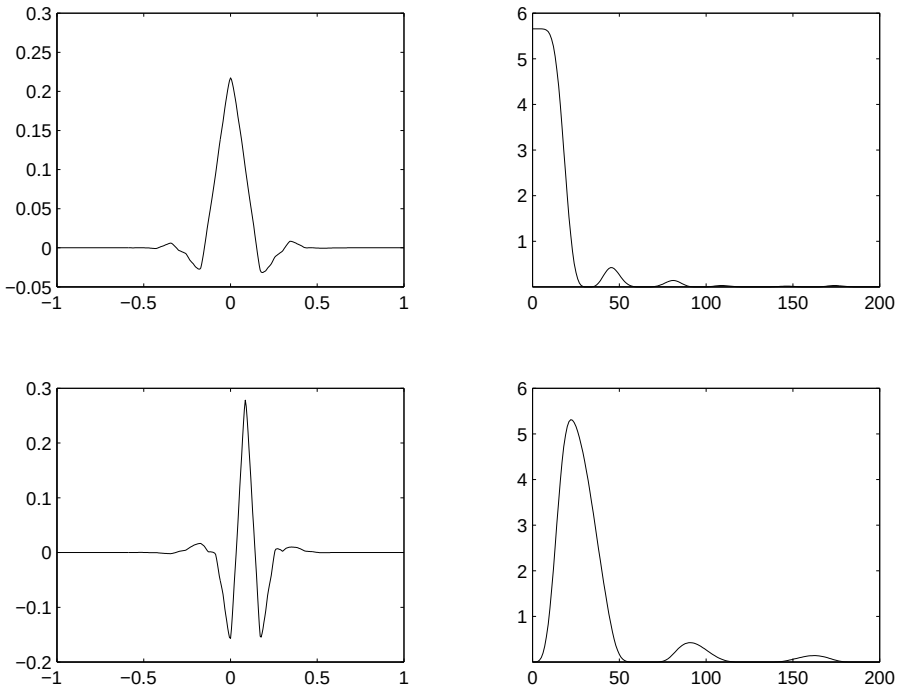


Figure 3: Coiflet 12. Top: Scaling function ϕ. Bottom: Corresponding wavelet ψ. (left: in physical space, right: in Fourier space)

$$f(x) = \sum_{i \in \mathbb{Z}} \overline{f}_{j_0 i}\, \phi_{j_0 i}(x) \quad + \quad \sum_{j=j_0}^{\infty} \sum_{i \in \mathbb{Z}} \tilde{f}_{ji}\, \psi_{ji}(x) \tag{31}$$

where

$$\overline{f}_{ji} = < f, \phi_{ji} >_{\mathbb{R}} \qquad \tilde{f}_{ji} = < f, \psi_{ji} >_{\mathbb{R}} \tag{32}$$

In numerical applications the sums in (31) have to be truncated which corresponds to the projection of f onto a subspace of $V_J \subset L^2(\mathbb{R})$. The decomposition (31) is orthogonal, as, by construction,

$$< \psi_{ji}, \psi_{lk} >_{\mathbb{R}} = \delta_{jl}\ \delta_{ik} \tag{33}$$

$$< \psi_{ji}, \phi_{lk} >_{\mathbb{R}} = 0 \qquad j \geq l \tag{34}$$

in addition to (27).

3.1.3 Relation between the regularity of a function and its wavelet coefficients

There is a relation between the local or global regularity of a function and the decay of its wavelet coefficients. The global regularity directly determines the

error being made when the wavelet sum is truncated at some scale. Depending on the type of norm and whether global or local characterization is concerned, various relations of this kind have been developed, see *e.g.* [42, 8, 41] for an overview.

As an example we consider the case of an α-Lipschitz function, with $\alpha \geq 1$ [33]. Suppose $f \in L^2(\mathbb{R})$, then for $[a,b] \subset \mathbb{R}$ the function f is α-Lipschitz for any $x_0 \in [a,b]$, *i.e.* $|f(x_0+h) - f(x_0)| \leq C|h|^\alpha$, if and only if there exists a constant A such that $| < f, \psi_{ji} > | \leq A 2^{-j\alpha - \frac{1}{2}}$ for any (j,i) with $\frac{i}{2^j} \in]a,b[$. This shows the relation between the local regularity of a function and the decay of its wavelet coefficients in scale.

The adaptive discretization we have proposed [21, 22] to compute turbulent flows is based on taking into account spatially varying regularity of the solution through a cut off of its wavelet series, where the strongest wavelet coefficients are retained. This guarantees that regions where the fields are less regular (quasi-singulatities) are well sampled.

3.2 In higher dimensions

This section consists of an extension of the previously presented one-dimensional construction to higher dimensions. For simplicity, we will consider only the two-dimensional case, since higher dimensions can be treated analogously. We start with a brief description of the construction principle and then turn in more detail to the two-dimensional case with periodicity, which is relevant for the subsequent applications.

3.2.1 Tensor product construction

Having developed a one-dimensional orthonormal basis ψ_{ji} of $L^2(\mathbb{R})$ one would like to use these functions as building blocks in higher dimensions. One way of doing so is to take the tensor product of two one-dimensional bases [8] and to define

$$\psi_{j_x,j_y,i_x,i_y}(x,y) \;=\; \psi_{j_x,i_x}(x)\; \psi_{j_y,i_y}(y) \tag{35}$$

The resulting functions constitute an orthonormal wavelet basis of $L^2(\mathbb{R}^2)$. Each function $f \in L^2(\mathbb{R}^2)$ can then be developed into

$$f(x,y) \;=\; \sum_{j_x,i_x} \sum_{j_y,i_y} \tilde{f}_{j_x,j_y,i_x,i_y}\; \psi_{j_x,j_y,i_x,i_y}(x,y) \tag{36}$$

with $\tilde{f}_{j_x,j_y,i_x,i_y} =< f, \psi_{j_x,j_y,i_x,i_y} >$.
Remark:
In this basis the two variables x and y are dilated separately. Therefore the functions ψ_{j_x,j_y} involve two scales, $l_x = 2^{j_x}$ and $l_y = 2^{j_y}$, and each of the functions is essentially supported on a rectangle with side lengths l_x and l_y. Hence the decomposition is often called rectangular wavelet decomposition. This is closely related to the standard form of operators using the nomenclature of Beylkin [5]. From the algorithmic viewpoint, this is equivalent to apply the

one-dimensional wavelet transform to the rows and the columns of a matrix representing an operator or a two–dimensional function. For some applications such a basis is advantageous, for others not. For example in turbulence the notion of scale has an important meaning and one would like to have a unique scale assigned to each basis function, which is not the case with the tensor product construction. This is why we prefer the MRA construction to analyze and compute turbulent flows.

3.2.2 2D multi-resolution analysis

A suitable concept which fulfills the above requirement of having a unique scale is the construction of a truly two-dimensional MRA of $L^2(\mathbb{R}^2)$. It can be obtained through the tensor product of two one-dimensional MRA's of $L^2(\mathbb{R})$ [42]. More precisely one defines the spaces $\mathbf{V}_j, j \in \mathbb{Z}$ by

$$\mathbf{V}_j = V_j \otimes V_j \tag{37}$$

and $\mathbf{V}_j = \overline{span}\,\{\phi_{j,i_x,i_y}(x,y) \;=\; \phi_{j,i_x}(x)\;\phi_{j,i_y}(y)\;, i_x, i_y \in \mathbb{Z}\}$ fulfilling analogous properties as in the one–dimensional case (3.1) - (3.4).
Likewise, we define the complement space $\mathbf{W}_j$ to be the orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j+1}$, *i.e.*

$$\begin{aligned}
\mathbf{V}_{j+1} &= V_{j+1} \otimes V_{j+1} \;=\; (V_j \oplus W_j) \otimes (V_j \oplus W_j) && (38)\\
&= V_j \otimes V_j \;\oplus\; ((W_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes W_j)) && (39)\\
&= \mathbf{V}_j \;\oplus\; \mathbf{W}_j && (40)
\end{aligned}$$

It follows that the orthogonal complement $\mathbf{W}_j = \mathbf{V}_{j+1} \ominus \mathbf{V}_j$ consists of three different types of functions and is generated by three different wavelets

$$\psi^{\mu}_{j,i_x,i_y}(x,y) \;=\; \begin{cases} \psi_{j,i_x}(x)\;\phi_{j,i_y}(y) & ;\; \mu = 1 \\ \phi_{j,i_x}(x)\;\psi_{j,i_y}(y) & ;\; \mu = 2 \\ \psi_{j,i_x}(x)\;\psi_{j,i_y}(y) & ;\; \mu = 3 \quad . \end{cases} \tag{41}$$

Observe that here the scale parameter j simultaneously controlls the dilatation in x and in y. We recall that in d dimensions this construction yields $2^d - 1$ types of wavelets spanning $\mathbf{W}_j$.
Using (41) each function $f \in L^2(\mathbb{R}^2)$ can be developed into an MRA basis as

$$f(x,y) \;=\; \sum_{j} \sum_{i_x,i_y} \sum_{\mu=1,2,3} \tilde{f}^{\mu}_{j_x,j_y,i_x,i_y}\;\psi^{\mu}_{j,i_x,i_y}(x,y) \tag{42}$$

with $\tilde{f}^{\mu}_{j_x,j_y,i_x,i_y} =< f, \psi^{\mu}_{j,i_x,i_y} >$. The wavelets ψ^{μ}_{j,i_x,i_y} are the basis functions of the so-called square wavelet decomposition. The algorithmic structure of the one-dimensional transforms carries over to the two-dimensional case by simple tensorisation, *i.e.* applying the filters at each decomposition step to rows and columns. Applying this kind of transform to matrices representing operators (differential, integral, integro-differential) leads to the non-standard form in the terminology of Beylkin [5].

Remark:
The described two-dimensional wavelets and scaling functions are separable. This advantage facilitates the generation of a multidimensional MRA from several one-dimensional MRA's. However the main drawback of this construction is that three wavelets are needed to span the orthogonal complement space $\mathbf{W}_j$ in two dimensions and seven in three dimensions. Another property should be mentioned. By construction the wavelets are anisotropic, *i.e.* horizontal, diagonal and vertical directions are preferred. This could be an advantage in digital signal processing to recognize corners and edges.

3.2.3 Periodic 2D multi-resolution analysis

Using the tensor product construction of two-dimensional wavelets on the real line and the periodization technique, see *e.g.* [49], we now recall the essential features of periodic two-dimensional wavelets of $L^2(\mathbb{T}^2)$. For notational ease we drop from now on the tilde introduced to distinguish the periodic wavelets from those on the real line. In the latter applications the periodic basis is used throughout unless otherwise explicitly stated.
A two-dimensional MRA of $L^2(\mathbb{T}^2)$ is a sequence of embedded subspaces $\mathbf{V}_j \subset \mathbf{V}_{j+1}$, $j \in \mathbb{N}_0$. It can be obtained through the tensor product of two one–dimensional MRA's of $L^2(\mathbb{T})$ [42]. This induces a decomposition of $L^2(\mathbb{T}^2)$ into mutually orthogonal hierarchical subspaces

$$L^2(\mathbb{T}^2) = \mathbf{V}_0 \oplus_{j\geq 0} \mathbf{W}_j \tag{43}$$

The space $\mathbf{V}_j$ is generated by the bivariate scaling functions

$$\mathbf{V}_j = \overline{span}\,\{\phi_{j,i_x,i_y}(x,y) \;=\; \phi_{j,i_x}(x)\;\phi_{j,i_y}(y)\}_{i_x,i_y=0,\ldots,2^j-1} \tag{44}$$

and the orthogonal complement space $\mathbf{W}_j \;=\; \mathbf{V}_{j+1} \ominus \mathbf{V}_j\,, j \geq 0$ by three different wavelets

$$\mathbf{W}_j = \overline{span}\,\{\psi^{\mu}_{j,i_x,i_y}(x,y)\}_{i_x,i_y=0,\ldots,2^j-1,\mu=1,2,3} \tag{45}$$

with ψ^{μ}_{j,i_x,i_y} defined as in (41) using the periodic analogons.
Correspondingly, any function $f \in L^2(\mathbb{T}^2)$ which is at least continuous can be projected onto $\mathbf{V}_J$ by collocation

$$f_J(x,y) \;=\; \sum_{i_x=0}^{2^J-1} \sum_{i_y=0}^{2^J-1} \; f\left(\frac{i_x}{2^J}, \frac{i_y}{2^J}\right)\, S_{J,i_x,i_y}(x,y) \tag{46}$$

using the two-dimensional cardinal Lagrange function

$$S_{j,i_x,i_y}(x,y) = S_{j,i_x}(x) S_{j,i_y}(y) \quad . \tag{47}$$

It can then be expressed as

$$f_J(x,y) \;=\; \bar{f}_{0,0,0}\;\phi_{0,0,0}(x,y) \quad + \quad \sum_{j=0}^{J-1} \sum_{i_x=0}^{2^j-1} \sum_{i_y=0}^{2^j-1} \sum_{\mu=1}^{3} \; \tilde{f}^{\mu}_{j,i_x,i_y}\;\psi^{\mu}_{j,i_x,i_y}(x,y) \tag{48}$$

with scaling coefficients

$$\bar{f}_{0,0,0} = \int_{\mathbb{T}^2} f(x,y)\,dx\,dy, \tag{49}$$

and wavelet coefficients

$$f^{\mu}_{j,i_x,i_y} = \langle f, \psi^{\mu}_{j,i_x,i_y} \rangle \tag{50}$$

using that $\phi_{0,0,0} = 1$.

Representing a function in terms of wavelet coefficients has the following advantages. Smooth functions yield rapid decay of the coefficients in scale (depending on the number of vanishing moments of ψ_{ji}). At locations where u develops a singularity or an 'almost singularity' only local coefficients have to be retained (depending on the decay of ψ_{ji} in space). Second, all employed basis functions are mutually orthogonal, a property which is the keystone of the algorithms.

3.3 Algorithm

In case of a regular sampling and periodic functions, computations can be done in physical space by periodizing the required filters (defined below) using Mallat's algorithm [41]. For long filters it is more economical, however, to use fast convolution in Fourier space employing FFT. The algorithm is based on the application of three discrete filters. The scaling coefficients $\bar{f}_{J,k}$ are computed by application of the interpolation filter I

$$I^J_n = \langle S_{J,n}, \phi_{J,0} \rangle\ , \quad \widehat{I^J} = 2^{-3J/2}\,\widehat{I}(\frac{k}{2^j})\ , \quad \widehat{I}(k) = \widehat{S}(k)/\widehat{\phi}(k) \tag{51}$$

to the sampled values $\{f(\frac{i}{2^J})\}_i$.

The scaling filter H and the wavelet filter G

$$H^j_n = \langle \phi_{j,n}, \phi_{j-1,0} \rangle\ , \qquad G^j_n = \langle \phi_{j,n}, \psi_{j-1,0} \rangle \tag{52}$$

are classically used for computing the wavelet transform. They can be obtained in physical space for compactly supported bases and in Fourier space through

$$\widehat{H}(k) = \widehat{\phi}(2k)/\widehat{\phi}(k) \quad , \qquad \widehat{G}(k) = \widehat{\psi}(2k)/\widehat{\phi}(k) \tag{53}$$

The algorithm then reads

step 0. FFT of the values $\{f_i\}_{i=0,\ldots,2^J-1}$ at the points $\{x_i = \frac{i}{2^J}\}_{i=0,\ldots,2^J-1}$ to the
Fourier coefficients $\{\widehat{f}_k\}_{k=0,\ldots,2^J-1}$.

step 1. Interpolation using the Lagrange function $S_J(x)$ of the space V_J by computation in Fourier space : application of $\widehat{I_J}$ gives $(\widehat{\overline{f}_J})_k$, $k = 0, \ldots, 2^J - 1$.

step 2. Application of Filters G and H in Fourier space (* indicating double length sequences)

$$(\widehat{\overline{f}^*_{J-1}})_k = \overline{\hat{H}_k}\,(\widehat{\overline{f}_J})_k \qquad k = 0, \ldots, 2^J - 1 \tag{54}$$

$$(\widehat{\tilde{f}^*_{J-1}})_k = \overline{\hat{G}_k}\,(\widehat{\tilde{f}_J})_k \qquad k = 0, \ldots, 2^J - 1 \tag{55}$$

step 3. Instead of setting

$$\overline{f}_{J-1,i} = \overline{f}^*_{J-1,2i} \qquad \tilde{f}_{J-1,i} = \tilde{f}^*_{J-1,2i} \qquad i = 0, \ldots, 2^{J-1} - 1 \tag{56}$$

in physical space, downsampling can be done directly in Fourier space through

$$(\widehat{\overline{f}_{J-1}})_k = (\widehat{\overline{f}^*_{J-1}})_k + (\widehat{\overline{f}^*_{J-1}})_{k+2^{J-1}} \qquad k = 0, \ldots, 2^{J-1} \tag{57}$$

$$(\widehat{\tilde{f}_{J-1}})_k = (\widehat{\tilde{f}^*_{J-1}})_k + (\widehat{\tilde{f}^*_{J-1}})_{k+2^{J-1}} \qquad k = 0, \ldots, 2^{J-1} \tag{58}$$

step 4. Inverse FFT of lenght 2^{J-1} to get $\{\tilde{f}_{J-1,i}\}_{i=0,\ldots,2^{J-1}-1}$.

iterate steps 2 to 4 replacing J by $j = J-1, \ldots, 0$. Observe that in the last step 2^{-1} is replaced by 0 and $(\widehat{\overline{f}_0})_0 = \overline{f}_{0,0}$.

The use of (57), (58) instead of (56) leads to a speed up by a factor 6 with respect to extracting the coefficients in physical space. The inverse transform is obtained by executing the above steps in reversed order omitting the conjugate complex in (54), (55) and replacing step 3 with upsampling in Fourier space.

4 Statistical tools

4.1 Methodology

Turbulence research is based on observations, laboratory and numerical experiments. Typical quantities measured to characterize turbulent flows are scalar fields (temperature, concentration, pressure, etc), vector fields (velocity, vorticity, etc), tensor fields (stress, strain, etc).

4.1.1 Laboratory experiments

Observations and laboratory experiments (*e.g.* in wind tunnels or water tank) are done using flow visualisations and time measurements performed in few points of the flow (*e.g.* hot-wire anemometry, laser velocimetry). Flow visualizations give mostly qualitative information. Time measurements give quantitative information by accumulating well-sampled and well-converged time statistics, although only at very few spatial locations. By checking Taylor's hypothesis, namely that the time fluctuations are small compared to the mean flow velocity, one assumes that the time statistics can be identified with the space statistics. This allows to compare observations or laboratory measurements with the predictions of the statistical theory of turbulence, and also with the statistics obtained from numerical experiments.

4.1.2 Numerical experiments

Numerical experiments are based on the accepted assumption that Navier–Stokes are the fundamental equations of fluid dynamics, whatever the flow regime. The fully-developed turbulent regime is reached when the nonlinear advective term dominates the linear dissipative term (limit kinematic viscosity $\nu \to 0$ or Reynolds number $Re \propto 1/\nu \to \infty$). In this highly nonlinear regime, Navier–Stokes solutions can only be computed by numerical approximation. The computation predicts the time evolution of one flow realization only. Statistical analyses are performed afterwards in three different ways:

- by computing spatial statistics of instanteneous turbulent fields, which is valid only if the computational domain is much larger than the integral scale where turbulence is produced,
- by computing time statistics of long flow history, which is valid only if the time evolution is much larger than the eddy turn over time characteristic of turbulent flow instabilities,
- by running a large number of numerical simulations for the same parameters and flow configuration, but with different initial conditions. The ensemble averages are computed afterwards. This procedure requires a number of independant realisations sufficient to ensure the stationarity of the probability distribution (PDF).

4.1.3 Averaging procedure

Turbulent flows are characterized by their unpredictability, namely each flow realization is different, although the statistics are reproducible as long as the flow configuration and parameters are the same. This is the reason why turbulence models predict only statistical quantities.

Another essential characteristic of turbulent flows is their intermittency, *i.e.* the fact that we observe in each flow realization well localized strong events (bursts). This intermittent behaviour is not very pronounced in the velocity field, but becomes dominant when one considers the velocity gradients or the vorticity field. They are characterized by non-Gaussian PDFs, whose tails correspond to the intermittent bursts. The flow intermittency comes from the nonlinear dynamics of turbulent flows which, for incompressible fluids, tend to form well localized coherent vortices (vortex spots in two dimensions and vortex tubes in three dimensions) which wander around in a chaotic way resulting from their mutual interactions.

A crucial difficulty in turbulence modelling is to define averages able to take into account intermittency. The L^2-norm averages, *i.e.* $(\int |f(x)|^2 dx)^{1/2})$, classically used in turbulence (*e.g.* two-point correlations, second order structure functions, spectra) are 'blind' to intermittency, because the well localized strong events responsible for intermittency are too rare to affect the L^2-norm, since their weight remains negligible in the integral.

To define averages able to take into account intermittency there are two possible strategies:

- consider L^p-norms, *i.e.* $((\int |f(x)|^p dx)^{1/p})$ with p large enough to have the values of the PDF tails contributing significantly to the integral,
- extract the rare (intermittent) events, responsible for the heavy tails of the PDF, from the dense (non-intermittent) events, which contribute only to the center of the PDF, and perform classical L^2-norm averages for the dense events only.

The second approach corresponds to conditional averages and requires a criterium to separate the rare events from the dense events. As we have assumed that the intermittency of turbulent flows is due to the presence of coherent vortices, responsible for the rare events, we first need to identify them in order to extract them.

As we have shown [19], [17] the coherent vortices can be characterized by the fact that they correspond to the strongest wavelet coefficients of the vorticity field. Based on this property we have defined a procedure to extract them [20], [21], which consists in retaining only those vorticity wavelet coefficients $\tilde{\omega}$ which are larger than a threshold value $\tilde{\omega}_T = (2Z \log_e N)^{1/2}$, with Z the total enstrophy and N the resolution (*i.e.* the number of grid points or wavelet coefficients). We then verify that the PDF of the discarded coefficients, *i.e.* those smaller than the treshold $\tilde{\omega}_T$ which correspond to the dense non-intermittent events, is Gaussian.

4.2 Classical tools

4.2.1 Probability distribution function (PDF)

To motivate the introduction of a probability space $(\Xi, \mathcal{F}, \mathcal{P})$ we consider:

- the set of all possible configurations of the flow, *i.e.*, the phase space of the Navier–Stokes equations, denoted by Ξ,
- the set $\mathcal{F}$ of all experiments with a definite outcome which have been performed, called the flow realisations,
- a probability measure $\mathcal{P}$ for the set of experiments $\mathcal{F}$, such that $\mathcal{P}(O) = \{0\}$ and $\mathcal{P}(\Xi) = \{1\}$, which assigns a probability to each experiment which has been performed.

We consider a stationary, homogeneous and isotropic random signal $f(\xi, x, t) \in \mathcal{F}$, with $\xi \in \Xi$, $x \in I\!R$ and $t \in I\!R_0^+$. For fixed ξ the function $f(x,t)$ is called a realization of the random field or a sample, *e.g.*, one component of the velocity field.

Using the probability measure $\mathcal{P}$ we define the distribution function $F(g) = \mathcal{P}(-\infty < f \leq g,\ f \in \Xi)$ which measures the probability of f having a value less or equal to g.

4.2.2 Radon-Nikodyn's theorem

If the probability measure $\mathcal{P}$ is absolutely continuous, there exists a probability density function (PDF) p of $\mathcal{P}$ such that $p(f) = \frac{d\mathcal{P}}{df}$, which corresponds to the derivative $\frac{dF}{dg}$ of the distribution F, *i.e.*, $p(f)df = \mathcal{P}(f < g \leq f + df,\ g \in \Xi)$. The PDF is normalized such that $\int_{I\!R}\ p(f)\, df = 1$.

4.2.3 Definition of the joint probability

Let f and g be two random fields, one can define the joint probability

$$F(f,g) = \mathcal{P}(-\infty < f' \leq f\,,\ -\infty < g' \leq g,\ [f', g'] \in \Xi).$$

The corresponding joint probability density function is given by $p(f,g) = p(f)p(g) - p(f \cap g)$ (Bayes' theorem). If f and g are independent and identically distributed (i.i.d.), then $p(f,g) = p(f)p(g)$.

4.2.4 Statistical moments

The p–th order moments of the random field f are defined as

$$M_p(f) = \langle f^p \rangle = \int\ f^p\, p(f)\, df \tag{59}$$

If $f \in \Xi$ is ergodic, the p–th order moments can also be expressed as space average

$$M_p(f) = \int (f(\vec{x}))^p \, d^n\vec{x}, \tag{60}$$

where the integral is defined as

$$\int = \lim_{L\to\infty} \frac{1}{L^n} \int_0^L \int_0^L \int_0^L \dots , \tag{61}$$

L being the size of the domain and n the space dimension.
$M_p(f)$ corresponds to the p–th power of the L^p–norm of f.

Remark:
Note that a stationary random process is ergodic, if the ensemble average is equal to the time average of one realization. This means that if one waits long enough, the ergodic process visits all states of the phase space, and the time it spends in each state is proportional to the probability of occurence of this state. Similarly, an homogeneous random process is ergodic, if the ensemble average is equal to the space average of one realization. This means that if one considers all space points, the ergodic process accesses to all states of the phase space and then both averages are equivalent.

Ratios of moments are defined, such as

$$Q_{p,q}(f) = \frac{M_p(f)}{(M_q(f))^{p/q}} \tag{62}$$

Classically one chooses $q = 2$, which leads to define statistical quantitities such as:

- skewness $S = Q_{3,2}(f)$,
- flatness $F = Q_{4,2}(f)$,
- hyperskewness $S_h = Q_{5,2}(f)$
- hyperflatness $F_h = Q_{6,2}(f)$.

4.2.5 Structure functions

The p–th order structure function of a random scalar field f is defined as

$$S_p(\vec{l}) \; = \; \int (f(\vec{x}+\vec{l}) - f(\vec{x}))^p \, d^n\vec{x} \; . \tag{63}$$

The p–th order structure function can be computed from the PDF of the increments of f, defined as $\Delta(l) = f(\vec{x}+\vec{l}) - f(\vec{x})$, since it is the p–th moment of their PDF

$$S_p(\vec{l}) \; = \; \int \Delta(\vec{l}) \, p(\Delta(\vec{l})) \, d(\Delta(\vec{l})) \; . \tag{64}$$

Remark:
The drawback of higher-order statistics, however, is that the number of data points required for an accurate estimation increases with order p. For instance, the number of points required for moments of order 12 is about 10^9, and thus estimation of high-order moments quickly becomes impracticable.

4.2.6 Autocorrelation function

The autocorrelation function of the random scalar field $f \in L^2(\mathbb{R}^2)$ is defined as

$$R(\vec{l}) = \int f(\vec{x})\, f(\vec{x}+\vec{l})\, d^n\vec{x} \tag{65}$$

and for vector fields $\vec{f} \in L^2(\mathbb{R}^2)$ we get the two–point correlation tensor

$$R_{ij}(\vec{l}) = \int f_i(\vec{x})\, f_j(\vec{x}+\vec{l})\, d^n\vec{x} \tag{66}$$

Note that in turbulence the above quantity computed for the velocity field is called the Reynolds stress tensor, and plays a key role in turbulence modelling.

4.2.7 Fourier spectrum

The spectrum of the random scalar field f is the Fourier transform of its auto-correlation function:

$$\Phi(\vec{k}) = \frac{1}{(2\pi)^n} \int R(\vec{l})\, e^{-i\vec{k}\cdot\vec{l}}\, d^n\vec{l} \tag{67}$$

Similarly for vector fields we obtain

$$\Phi_{ij}(\vec{k}) = \frac{1}{(2\pi)^n} \int R_{ij}(\vec{l})\, e^{-i\vec{k}\cdot\vec{l}}\, d^n\vec{l} \tag{68}$$

One can integrate $\Phi(\vec{k})$ on shells of radius $k = |\vec{k}|$ which gives the one-dimensional spectrum

$$E(k) \;=\; \int \Phi(\vec{k})\, k^{n-1}\, d\theta\ . \tag{69}$$

4.2.8 Wiener-Khinchin's theorem

For a function $R(l) \in L^1(\mathbb{R})$, to be the correlation function of a homogeneous field $f(x)$ which satisfies the condition $S_2(l) \to 0$ for $l \to 0$, it is necessary and sufficient that it has a representation of the form $R(l) = \int_{\mathbb{R}} E(k)\, e^{ikl}\, dk$ where $E(k) \geq 0$ is the spectral density of the random variable $f(x)$.

Remark:
If $R(l)$ is not in $L^1(\mathbb{R})$, then Wiener-Khinchin holds in a distributional sense only, *e.g.* for a Gaussian white noise $R(l) = \delta_{l,0}$, therefore it is not in L^1 and $E(k) = 1$.

The spectrum $E(k)$, the second order structure function $S_2(l)$ and the autocorrelation function $R(l)$ fulfill the following relations:

$$R(l) \;=\; \int f(x+l)f(x)\,dx \;=\; 2\int_0^\infty \cos(2\pi kl)E(k)dk \tag{70}$$

and hence we get

$$S_2(l) = \int |f(x+l)-f(x)|^2\,dx\;. = 2R(0)-2R(l) = 2\int_0^\infty (1-\cos(2\pi kl))E(k)dk \tag{71}$$

Remark:
The above relation illustrates that the structure function corresponds to a high pass filtered spectrum although the corresponding filter is not very selective. We will propose wavelet tools to improve the filter selectivity.

4.3 Tools based on the continuous wavelet transform

4.3.1 Local wavelet spectrum

If the signals are stationary (for time signals) or homogeneous (for space signals), one classically considers the modulus of their Fourier transform (*i.e.* their energy spectrum) and thus the phase information is lost. This is a major limitation of the classical way of analyzing complex signals since their localized temporal or spatial structures are thus neglected. This local information can be observed in each realization but not within the Fourier coefficients since it is hidden among the phases of all Fourier coefficients. For statistically non-stationary or inhomogeneous signals the classical energy spectrum, based on the Fourier transform, is no more appropriate.

The wavelet transform extends the concept of energy spectrum so that one can define a local energy spectrum $\tilde{E}(x,k)$ using the wavelet transform (which, as we have seen conserves, the L^2–norm of a function), such that

$$\tilde{E}(k,x) = \frac{1}{C_\psi k_\psi}\left|\tilde{f}\left(\frac{k_\psi}{k},x\right)\right|^2 \quad \text{for} \quad k \geq 0 \tag{72}$$

where k_ψ is the peak wavenumber of the analyzing wavelet ψ and C_ψ as defined in (1). By measuring $\tilde{E}(k,x)$ at different instants or positions in the signal one estimates what parts in the signal contribute most to the global Fourier energy spectrum and might suggest a way to decompose the signal into different contributions. For example, if one considers turbulent signals, one can determine the type of energy spectrum contributed by coherent structures, such as isolated vortices, and the type of energy spectrum contributed by the unorganized background flow.

4.3.2 Global wavelet spectrum

Although the wavelet transform analyses the flow into wavelets rather than complex exponentials, one shows [51] that the mean wavelet energy spectrum

converges to the Fourier energy spectrum provided the analyzing wavelets have enough cancellations. More precisely the mean wavelet spectrum

$$\tilde{E}(k) = \int_0^{+\infty} \tilde{E}(k,x)dx \tag{73}$$

gives the correct Fourier exponent for a power-law Fourier energy spectrum $E(k) \propto k^{-\beta}$ if the analysing wavelet has at least $n > (\beta - 1)/2$ vanishing moments. This condition is the same as that for detecting singularities derived in the previous section since $\beta = 1 + 2\alpha$ for isolated cusps. Thus, the steeper the energy spectrum, the more vanishing moments the analyzing wavelet should have. The inertial range in turbulence has a power-law form. The ability to correctly characterize power-law energy spectra is therefore a very important property of the wavelet transform (which is related to its ability to detect and characterize singularities).

4.3.3 Relation with the Fourier spectrum

The mean wavelet energy spectrum $\tilde{E}(k)$ is a smoothed version of the Fourier energy spectrum $E(k)$. This can be seen from the following relation between the two spectra

$$\tilde{E}(k) = \frac{1}{C_\psi k_\psi} \int_0^{+\infty} E(k') \left| \hat{\psi}\left(\frac{k_0 k'}{k}\right) \right|^2 dk' \tag{74}$$

which shows that the mean wavelet spectrum is an average of the Fourier spectrum weighted by the square of the Fourier transform of the analysing wavelet shifted at wavenumber k. Note that the larger k is, the larger the averaging interval, because wavelets are bandpass filters with $\frac{\Delta k}{k}$ constant. This property of the mean wavelet energy spectrum is particularly useful for turbulent flows. Indeed, the Fourier energy spectrum of a single realization of a turbulent flow is too oscillating to be able to clearly detects a slope, but it is no more the case for the mean wavelet energy spectrum which is much smoother.

The real-valued Mexican hat wavelet

$$\hat{\psi}(k) = k^2 \exp(-k^2/2) \tag{75}$$

has only two vanishing moments and thus can correctly measure energy spectrum exponents up to $\beta < 5$.

In the case of the complex-valued Morlet wavelet, only the zeroth order moment is zero,

$$\begin{aligned} \hat{\psi}(k) &= \frac{1}{2\pi} \exp(-(k - k_\psi)^2/2) \quad \text{for} \quad k > 0 \\ \hat{\psi}(k) &= 0 \quad \text{for} \quad k \leq 0 \end{aligned} \tag{76}$$

is zero, but the higher n^{th} order moments are very small ($\propto k_\psi^n \exp(-k_\psi^2/2)$) provided that k_ψ is sufficiently large. Therefore the Morlet wavelet transform

gives accurate estimates of the power–law exponent of the energy spectrum at least for approximately $\beta < 7$ (if $k_\psi = 6$).

There is also a family of wavelets [51] with an infinite number of cancellations

$$\hat{\psi}_n(k) = \alpha_n \exp\left(-\frac{1}{2}\left(k^2 + \frac{1}{k^{2n}}\right)\right), \quad n \geq 1, \tag{77}$$

where α_n is chosen for normalization. The wavelets defined in (77) can therefore correctly measure any power–law energy spectrum. Furthermore, these wavelets can detect the difference between a power–law energy spectrum and a Gaussian energy spectrum ($E(k) \propto \exp(-(k/k_0)^2)$). For instance in turbulence it is important to determine at what wavenumber the power–law energy spectrum becomes exponential since this wavenumber defines the end of the inertial range of turbulence and the begining of the dissipative range.

4.4 Tools based on the orthogonal wavelet transform

4.4.1 Motivation

In this section we describe some statistical tools based on the orthogonal wavelet transform. The wavelet approach avoids the limitations of structure functions and allows moment ratios to be defined as a function of scale. We present them considering, as example, a one-dimensional scalar field $f(x) \in L^2(\mathbb{R}^2)$ which has vanishing mean and is periodic (the extension to higher dimensions and vector fields is straightforward). Hence we employ a 1D periodic multiresolution analysis (MRA) [8, 16] and develop the signal f, sampled on $N = 2^J$ points, as an orthonormal wavelet series from the largest scale $l_{max} = 2^0$ to the smallest scale $l_{min} = 2^{-J}$:

$$f(x) = \sum_{j=0}^{J-1} \sum_{i=0}^{2^j-1} \tilde{f}_{j,i}\, \psi_{j,i}(x), \tag{78}$$

where $\psi_{j,i}$ is the 2π-periodic wavelet. Due to orthogonality the coefficients are given by $\tilde{f}_{j,i} = \langle f, \psi_{j,i} \rangle$ where $\langle \cdot, \cdot \rangle$ denotes the L^2–inner product. The wavelet coefficients contain information on scale, position and direction (in higher dimensions) information.

4.4.2 Local wavelet spectrum

We define the scale distribution of energy, also called scalogram, as

$$E_j = \sum_{i=0}^{2^j-1} |\tilde{f}_{j,i}|^2. \tag{79}$$

To be able to relate the scale distribution to the Fourier spectrum, we introduce the mean wavenumber k_ψ of the wavelet ψ, defined by

$$k_\psi = \frac{\int_0^\infty k|\widehat{\psi}(k)|dk}{\int_0^\infty |\widehat{\psi}(k)|dk}. \tag{80}$$

Thus each scale 2^{-j} of the wavelet ψ_j is inversely proportional to the mean wavenumber $k_j = k_\psi 2^j$. The discrete local wavelet spectrum [10, 51] is then defined as

$$\widetilde{E}(k_j, x_i) = |\tilde{f}_{j,i}|^2 \frac{2^j}{k_j}, \tag{81}$$

where $\Delta k_j = \sqrt{k_j k_{j+1}} - \sqrt{k_j k_{j-1}}$, is the mean wavenumber.

By measuring $\widetilde{E}(k_j, x_i)$ at different positions x_i in a turbulent flow one can study how the energy spectrum depends on local flow conditions and estimate the contribution to the overall Fourier energy spectrum of different components of the flow. For example, one can determine the scaling of the energy spectrum contributed by coherent structures, such as isolated vortices, and the scaling of the energy spectrum contributed by the unorganized part of the flow.

The spatial variability of the local energy spectrum $\widetilde{E}(k_j, x_i)$ measures the flow's intermittency.

4.4.3 Global wavelet spectrum

The global spectral behaviour of f is studied by summing the local energy spectrum over all positions,

$$\widetilde{E}(k_j) = \sum_{i=0}^{2^j - 1} \widetilde{E}(k_j, x_i). \tag{82}$$

The relationship between the global wavelet spectrum $\widetilde{E}(k_j)$ and the usual Fourier energy spectrum $E(k)$ is described in the following section.

4.4.4 Relation between wavelet and Fourier spectra

First note that due to the orthogonality of the wavelet decomposition, the total energy is preserved and we have $E = \sum_j E_j$. The global wavelet spectrum is related with the Fourier energy spectrum according to [16, 51]

$$\widetilde{E}(k) = \frac{1}{C_\psi k_\psi} \int_0^\infty E(k') |\widehat{\psi}(\frac{k_\psi k'}{k})|^2 dk'. \tag{83}$$

where

$$C_\psi = \int_0^\infty \frac{|\widehat{\psi}(k)|^2}{k}\, dk. \tag{84}$$

The wavelet spectrum is therefore a smoothed Fourier spectrum weighted by the modulus of the Fourier transform of the analyzing wavelet [51]. Note that as the wavenumber increases the smoothing interval becomes larger [16]. A sufficient condition guaranteeing that the global wavelet spectrum is able to detect the same power-law behaviour $k^{-\alpha}$ as the Fourier spectrum is that ψ has enough vanishing moments [51], *i.e.*

$$\int_{-\infty}^{+\infty} x^n \psi(x)\, dx = 0 \quad \text{for} \quad 0 \le n \le \frac{\alpha - 1}{2}. \tag{85}$$

If this condition is not satisfied the global wavelet spectrum saturates at the critical cancellation order n. In this case it only shows a power-law behaviour with a slope not steeper than $\alpha = 2(n+1)$. Since large α corresponds to smooth functions, as the function analyzed is smoother we require $\psi(x)$ to have more vanishing moments in order to correctly detect the signal's spectrum. If the wavelet does not have enough zero moments we simply measure the wavelet's own spectral scaling!

4.4.5 Wavelet intermittency measures

Intermittency is defined as localized bursts of high frequency activity. This means that intermittency is a phenomenon localized in both physical space and spectral space, and thus a suitable basis for representing intermittency should reflect this dual localization. The Fourier basis is perfectly localized in spectral space, but completely delocalized in physical space. Therefore when a turbulence signal is filtered using a high-pass Fourier transform and then reconstructed in physical space, e.g. to calculate the flatness, some spatial information is lost. This leads to smoothing of strong gradients and spurious oscillations in the background, which come from the fact that the modulus and phase of the discarded high wavenumber Fourier modes have been lost. The spatial errors introduced by such a filtering lead to errors in estimating the flatness, and hence intermittency, of the signal.

When a quantity (*e.g.* velocity derivative) is intermittent it contains rare but strong events (*i.e.* bursts of intense activity), which correspond to large deviations reflected in the 'heavy tails' of the PDF. Second-order statistics (e.g. energy spectrum, second-order structure function) are relatively insensitive to such rare events because their time or space support is very small and thus do not dominate the integral. However, these events become increasingly important for higher-order statistics, and finally come to dominate. High-order statistics therefore characterize intermittency. Of course, intermittency is not essential for all problems: second-order statistics will suffice to measure dispersion (dominated by energy-containing scales), but not to calculate drag (dominated by vorticity production in thin boundary layers).

The goal of this paragraph is to point out the limitations of classical measures of intermittency, and to present a unified set of wavelet-based alternatives (many of which have been introduced separately elsewhere). We show how the classical measures can be thought of as a special case of wavelet filtering using an extremely non-smooth wavelet defined as the difference of two Diracs (DOD). It is this lack of regularity of the underlying wavelet that limits the usefulness of classical measures for sufficiently smooth signals. We propose wavelet-based diagnostics that overcome these limitations, and produce accurate results whatever the signal to be analysed.

To measure intermittency we use the space-scale information contained in the wavelet coefficients to define scale-dependent moments and moment ratios. Useful diagnostics to quantify the intermittency of a field f are the moments of

its wavelet coefficients at different scales j [54],

$$M_{p,j}(f) = 2^{-j} \sum_{i=0}^{2^j-1} |\tilde{f}_{j,i}|^p. \tag{86}$$

Note that $E_j = 2^j M_{2,j}$.

The sparsity of the wavelet coefficients at each scale is a measure of intermittency, and it can be quantified using ratios of moments at different scales,

$$Q_{p,q,j}(f) = \frac{M_{p,j}(f)}{(M_{q,j}(f))^{p/q}} \quad , \tag{87}$$

which may be interpreted as quotient norms computed in two different functional spaces, L^p- and L^q-spaces. Classically, one chooses $q = 2$ to define typical statistical quantities as a function of scale. Recall that for $p = 4$ we obtain the scale dependent flatness $F_j = Q_{4,2,j}$. It is equal to 3 for a Gaussian white noise at all scales j, which proves that this signal is not intermittent. The scale dependent skewness, hyperflatness and hyperskewness are obtained for $p = 3, 5$ and 6, respectively. For intermittent signals $Q_{p,q,j}$ increases with j.

4.4.6 Relation to structure functions

In this section we link the scale dependent moments of wavelet coefficients, structure functions and Besov norms (which are typically used in nonlinear approximation theory [9]). In the case of second order statistics, we show that global wavelet spectra correspond to second order structure functions. Furthermore, we give a rigorous bound for the maximum exponent of the structure functions and propose a way to overcome this limitation.

The increments of a signal, also called the modulus of continuity, can be seen as its wavelet coefficients using the DOD (difference of Diracs) wavelet

$$\psi^{DOD}(x) = \delta(x+1) - \delta(x). \tag{88}$$

We thus obtain

$$f(x+l) - f(x) = \tilde{f}_{x,l} = \langle f, \psi^{DOD}_{x,l} \rangle \tag{89}$$

with $\psi_{x,l}(y) = 1/l[\delta((y-x)/l + 1) - \delta((y-x)/l)]$. Note that the wavelet is normalized with respect to the L^1 norm. The p-th order structure function $S_p(l)$ therefore corresponds to the p-th order moment of the wavelet coefficients at scale l,

$$S_p(l) = \int (\tilde{f}_{x,l})^p dx. \tag{90}$$

As the DOD wavelet has only one vanishing moment (its mean), the exponent of the p-th order structure function in the case of a self-similar behaviour is limited by p, *i.e.* if $S_p(l) \propto l^{\zeta(p)}$ then $\zeta(p) < p$. This saturation was first observed by [2] for DNS of two-dimensional flows, but without proposing an explanation. To be able to detect larger exponents one has to use increments with a larger

stencil, or wavelets with more vanishing moments, *i.e.* $\int x^n \psi(x) dx = 0$ for $n = 0, 1, \ldots, M-1$. This will become clearer below in the context of Besov regularity of functions.

We now concentrate on the case $p = 2$, *i.e.* the energy norm. Equation (83) gives the relation between the global wavelet spectrum $\widetilde{E}(k)$ and the Fourier spectrum $E(k)$ for an arbitrary wavelet ψ. For the DOD wavelet we find, since $\hat{\psi}^{DOD}(k) = e^{ik} - 1 = e^{ik/2}(e^{ik/2} - e^{ik/2})$ and hence $|\hat{\psi}^{DOD}(k)|^2 = 2(1 - \cos k)$, that

$$\widetilde{E}(k) = \frac{1}{C_\psi k} \int_0^\infty E(k')(2 - 2\cos(\frac{k_\psi k'}{k}))\, dk'. \tag{91}$$

Setting $l = k_\psi / k$ and comparing with (71) we see that the wavelet spectrum corresponds to the second order structure function, *i.e.*

$$\widetilde{E}(k) = \frac{1}{C_\psi k} S_2(l). \tag{92}$$

The above results show that, if the Fourier spectrum behaves like $k^{-\alpha}$ for $k \to \infty$, $\widetilde{E}(k) \propto k^{-\alpha}$ if $\alpha < 2M + 1$, where M is the number of vanishing moments of the wavelets. Consequently we find for $S_2(l)$ that $S_2(l) \propto l^{\zeta(p)} = \left(\frac{k_\psi}{k}\right)^{\zeta(p)}$ for $l \to 0$ if $\zeta(2) \leq 2M$. In the present case we have $M = 1$, *i.e.* the second order structure function can only detect slopes smaller than 2, corresponding to an energy spectrum with slopes shallower than -3. Thus we find that the usual structure function gives spurious results for sufficiently smooth signals. In the appendix we generalize the relation between structure functions and wavelet coefficients in the context of Besov spaces.

5 Extraction of coherent vortices

5.1 Methodology

To study turbulent signals we now propose to separate the rare and extreme events from the dense and weaker events, and then calculate the statistics for each independently. A major difficulty in turbulence research is that there is no clear scale separation between the two kinds of events. This lack of a 'spectral gap' excludes Fourier filtering for disentangling these two behaviours. Since the rare events are well localized in physical space, one might try to use an on-off filter defined in physical space to extract them. However, this approach changes the spectral properties by introducing spurious discontinuities (adding an artificial k^{-2} component to the energy spectrum). To avoid these problems we use the wavelet representation, which combines both physical space and spectral space localizations (bounded from below by Heisenberg's uncertainty principle). In turbulence the relevant rare events are the coherent vortices and the dense events correspond to the residual background flow. In [21] we have proposed a nonlinear wavelet filtering of the wavelet coefficients of vorticity to separate the coherent vortices from the residual background flow. We now detail the different steps of this procedure.

5.2 Wavelet decomposition of vorticity

We describe the wavelet algorithm to extract coherent vortices out of turbulent flows and consider as example the 3D case (for the 2D case refer the reader to [21]). We consider the vorticity field $\vec{\omega}(\vec{x}) = \nabla \times \vec{v}$, computed at resolution $N = 2^{3J}$, N being the number of grid points and J the number of octaves. Each component is developed into an orthogonal wavelet series from the largest scale $l_{max} = 2^0$ to the smallest scale $l_{min} = 2^{J-1}$ using a 3D multi-resolution analysis (MRA) [8, 16]:

$$\omega(\vec{x}) \;=\; \bar{\omega}_{0,0,0}\;\phi_{0,0,0}(\vec{x}) \;+\; \sum_{j=0}^{J-1}\sum_{i_x=0}^{2^j-1}\sum_{i_y=0}^{2^j-1}\sum_{i_z=0}^{2^j-1}\sum_{\mu=1}^{2^n-1} \tilde{\omega}^{\mu}_{j,i_x,i_y,i_z}\;\psi^{\mu}_{j,i_x,i_y,i_z}(\vec{x}), \tag{93}$$

with $\quad \phi_{j,i_x,i_y i,i_z}(\vec{x}) \;=\; \phi_{j,i_x}(x)\;\phi_{j,i_y}(y)\;\phi_{j,i_z}(z)$, and

$$\psi^{\mu}_{j,i_x,i_y,i_z}(\vec{x}) \;=\; \begin{cases} \psi_{j,i_x}(x)\;\phi_{j,i_y}(y)\;\phi_{j,i_z}(z) & ;\;\mu=1 \quad , \\ \phi_{j,i_x}(x)\;\psi_{j,i_y}(y)\;\phi_{j,i_z}(z) & ;\;\mu=2 \quad , \\ \phi_{j,i_x}(x)\;\phi_{j,i_y}(y)\;\psi_{j,i_z}(z) & ;\;\mu=3 \quad , \\ \psi_{j,i_x}(x)\;\phi_{j,i_y}(y)\;\psi_{j,i_z}(z) & ;\;\mu=4 \quad , \\ \psi_{j,i_x}(x)\;\psi_{j,i_y}(y)\;\phi_{j,i_z}(z) & ;\;\mu=5 \quad , \\ \phi_{j,i_x}(x)\;\psi_{j,i_y}(y)\;\psi_{j,i_z}(z) & ;\;\mu=6 \quad , \\ \psi_{j,i_x}(x)\;\psi_{j,i_y}(y)\;\psi_{j,i_z}(z) & ;\;\mu=7 \quad , \end{cases} \tag{94}$$

where $\phi_{j,i}$ and $\psi_{j,i}$ are the one-dimensional scaling function and the corresponding wavelet, respectively. Due to orthogonality, the scaling coefficients

are given by $\bar{\omega}_{0,0,0} = \langle \omega \, , \, \phi_{0,0,0} \rangle$ and the wavelet coefficients are given by $\tilde{\omega}^{\mu}_{j,i_x,i_y,i_z} = \langle \omega \, , \, \psi^{\mu}_{j,i_x,i_y,i_z} \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product.

5.3 Nonlinear thresholding

We then split the vorticity field into $\vec{\omega}_C(\vec{x})$ and $\vec{\omega}_I(\vec{x})$ by applying a nonlinear thresholding to the wavelet coefficients. The threshold is defined as $\epsilon = (4/nZ \ln N)^{1/2}$, where n is the space dimension which is here $n = 3$. It only depends on the total enstrophy Z and on the number of grid points N without any adjustable parameters. The choice of this threshold is based on theorems by Donoho and Johnstone [11, 12] proving optimality of the wavelet representation to denoise signals in presence of Gaussian white noise, since this wavelet-based estimator minimizes the maximal L^2-error for functions with inhomogeneous regularity.

5.4 Vorticity and velocity reconstruction

The coherent vorticity field $\vec{\omega}_C$ is reconstructed from the wavelet coefficients whose modulus is larger than ϵ and the incoherent vorticity field $\vec{\omega}_I$ from the wavelet coefficients whose modulus is smaller or equal to ϵ. The two fields thus obtained, $\vec{\omega}_C$ and $\vec{\omega}_I$, are orthogonal, which ensures a separation of the total enstrophy into $Z = Z_C + Z_I$ because the interaction term $\langle \vec{\omega}_C \, , \vec{\omega}_I \rangle$ vanishes. We then use Biot–Savart's relation $\vec{v} = \nabla \times (\nabla^{-2} \vec{\omega})$ to reconstruct the coherent velocity $\vec{v}_C$ and the incoherent velocity $\vec{v}_I$ for the coherent and incoherent vorticities respectively.

6 Application to 2D turbulence

6.1 Wavelet analysis

In the following examples of wavelet applications to study 2D and 3D turbulence, we use Coiflets 12 [8], *i.e.* wavelets with $M = 4$ vanishing moments and a filter length of $3M = 12$ (see Fig. 3 bottom). The advantage of the Coiflets is that they are almost symmetric and that the corresponding scaling functions (see Fig. 3 top) have also M vanishing moments.

We apply the wavelet diagnostics presented in 4.4 to analyze a freely decaying two-dimensional turbulent flow [21]. It has been computed by DNS at resolution $N = 256^2$ using a pseudo-spectral method. The spatial domain is periodic and the initial state is a random vorticity field with a Gaussian PDF.

The Navier–Stokes equations rapidly organize the initial homogeneous random vorticity field into isolated coherent vortices which contain most of the enstrophy Z (figure 4).

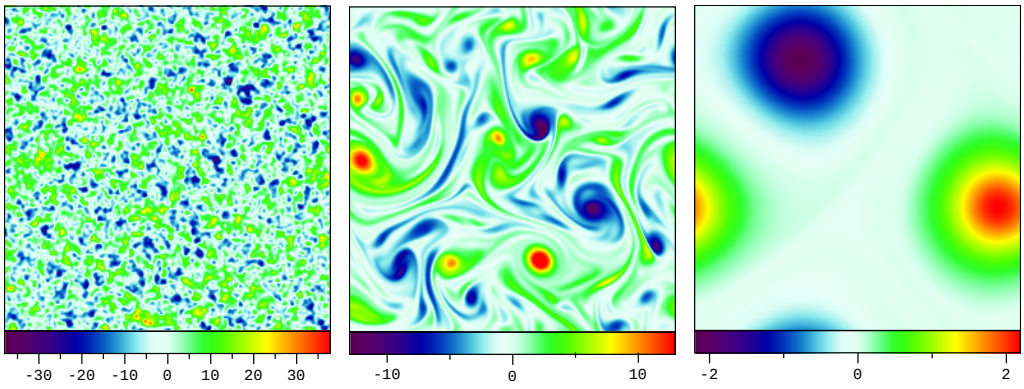


Figure 4: Vorticity field at $t = 0, 4, 100$. Note the emergence of coherent vortices from an initially homogeneous vorticity distribution.

As shown in figure 5, the initial Gaussian distribution evolves via a stretched exponential to a quasi-stationary final state approximating a Cauchy distribution. This Cauchy distribution is similar to the prediction of [44] for the PDF of the derivative of the velocity field of a system of point vortices. For such a distribution the variance and all higher order moments diverge, showing that the Navier–Stokes equation can generate extremely non-Gaussian distributions with coherent vortices. This evolution of the PDF from Gaussian to Cauchy can be explained dynamically. Due to the Biot–Savart law the flow organizes itself around initial extreme values of the vorticity. The gradients formed between the coherent vortices by this process tend to dissipate weak vorticity and therefore isolate the vortices. The coherent vortices then merge, which results in further dissipation of weaker vorticity. As a result, the strong values of the vorticity

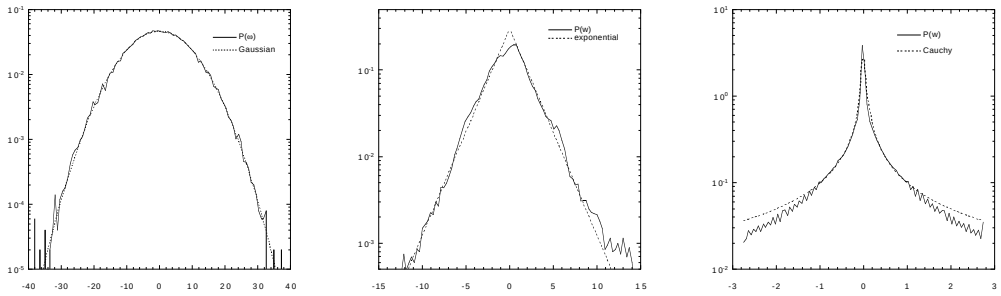


Figure 5: PDF of vorticity field at $t = 0, 4, 100$. The PDF is initially Gaussian and changes to a Cauchy distribution via a stretched exponential. The dotted curves show the ideal distribution (Gaussian, stretched exponential or Cauchy).

decay more slowly than the weak values, which results in a steepening of the PDF. Note that, as is the case for three-dimensional turbulence, the velocity remains Gaussian for all times.

In figure 6(a), we show the time evolution of the energy spectrum at initial, intermediate ($t = 0, 4\tau$) and late ($t = 100\tau$), the time unit being the eddy-turnover time $\tau = Z^{-1/2}$. It follows a power law in the inertial range, namely from $k = 1$ up to the dissipative wavenumbers larger than $k = 64$, where the slope changes from -3 at $t = 0$, to -4 at $t = 4$ and finally -6 at $t = 100$. These negative slopes reveal the long range dependence of the energy spectrum which increases in time, *i.e.* the velocity field becomes increasingly correlated and smooth. Note that the statistical theory of two–dimensional homogeneous turbulence [38] predicts a k^{-3} power-law behaviour. The steepening of the slope we observe as time evolves is attributed to the intermittency resulting from the emergence of coherent vortices [43].

In figure 6 (b) we show the scale dependence of enstrophy (82) at early, intermediate and late times. The scale of maximum enstrophy increases from 2^{-5} at $t = 0$ to 2^{-2} at $t = 100$. Therefore the correlation scale of the vorticity field increases in time, which is due to the formation and subsequent merging of coherent vortices, as illustrated in figure 4.

The scale dependent flatness of vorticity F_j is shown in figure 6(c). It evolves from Gaussian (*i.e.* $F_j \approx 3$ for all j) at $t = 0$ to non-Gaussian (characterized by the fact that F_j strongly increases with j) as time increases.

In the following we focus on the instant $t = 4$ (figure 4b), which is typical of the regime where coherent vortices have already formed and are interacting strongly. To study the dynamics of this flow regime we analyze the vorticity field ω, the linear dissipation term $L = \nu\nabla^2\omega$ (where ν denotes kinematic viscosity) and the nonlinear advection term $N = -\vec{u}\cdot\nabla\omega$ of the governing vorticity transport equation

$$\partial_t\omega = L + N \tag{95}$$

at time $t = 4$. We recall that, since $\nabla\cdot\vec{u} = 0$, the velocity can be reconstructed

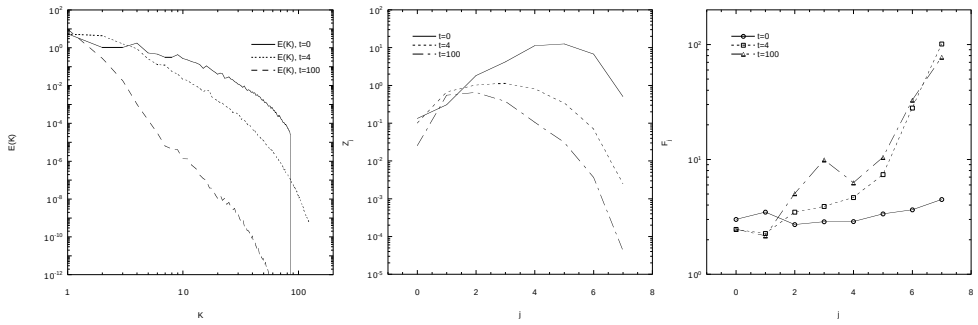


Figure 6: (a) Energy spectra $E(k)$. (b) Scale distribution of enstrophy Z_j. (c) Scale dependent flatness F_j at $t = 0, 4, 100$ for the decaying case.

from the vorticity by $\vec{u} = \nabla^{\perp}\nabla^{-2}\omega$, where $\nabla^{\perp} = (-\partial_y, \partial_x)$ and ∇^{-2} denotes the Green's function of the Laplacian (from Biot–Savart's relation). We plot vorticity, dissipation and advection at $t = 4$ in both physical and wavelet space in figure 7. Figure 7b shows that dissipation is localized in the sheared regions between interacting vortices. The advection term (figure 7c) is also well-localized in sheet-like regions. The wavelet coefficients of the three fields have similar intermittent structure. Note that the wavelet coefficients become increasingly intermittent at smaller scales (figure 7d, e, f). It is interesting to note that the wavelet coefficients that are active for vorticity are also active for dissipation and advection, *i.e.* the same wavelet coefficients represent all three quantities.

The wavelet coefficients $\tilde{\omega}, \tilde{L}$ and $\tilde{N}$ reveal that vorticity, dissipation and advection are strongly intermittent, *i.e.* for these 3 fields the spatial support decreases with the scale, likewise their wavelet coefficients become sparser when scale decreases. This intermittency is quantified by computing the scale dependent flatness F_j (cf. figure 8c). The moments M_p and the flatness F_j strongly increase with p and j, respectively, with the same scaling law for the three fields ω, L and N. This confirms the fact that they have the same type of intermittency.

In figure 8a we display the scale distribution (in L^2-norm) of vorticity, dissipation and advection. They all are multiscale but have different distributions: vorticity is most active around scale $2^{-2.5}$, dissipation around scale 2^{-6} and advection around scale 2^{-5}. The fact that dissipation has a maximum at small scales agrees with the classical phenomenology. However, its multiscale distribution contradicts the assumption of a non-dissipative inertial range (assumed by the statistical theory of turbulence), but agrees with the hypothesis of progressive dissipation throughout the inertial range as proposed by [6], and [25], for three-dimensional turbulence.

Now we take the vorticity field at $t = 4$ and randomize the phase of its Fourier coefficients, in order to construct a fractal field corresponding to a Gaussian

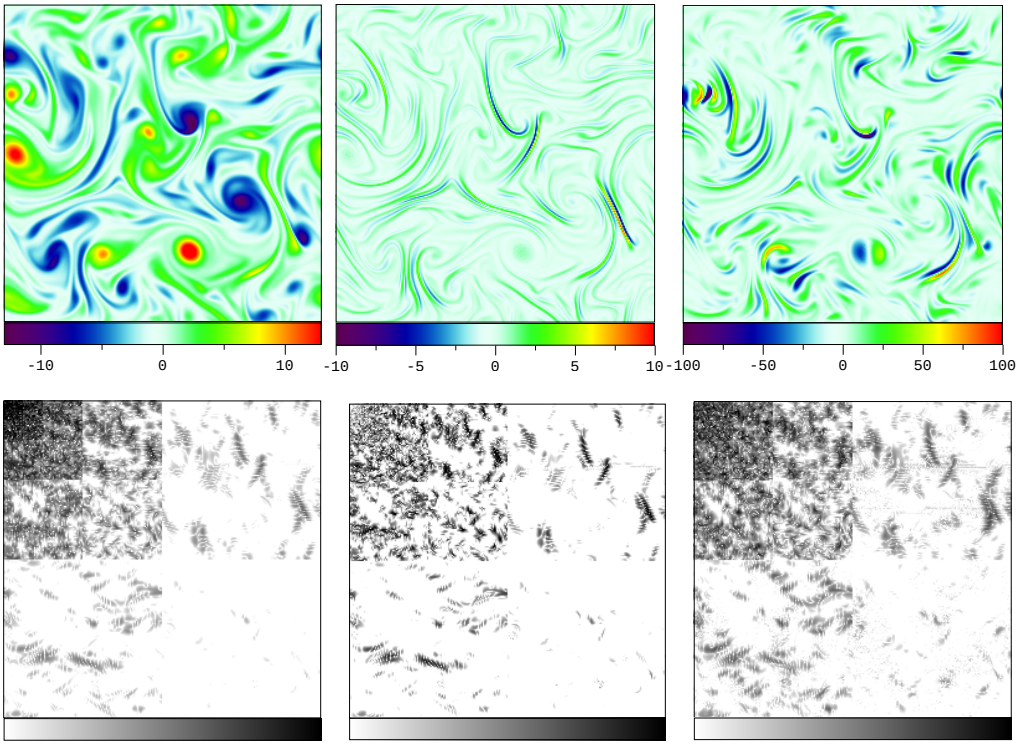


Figure 7: Top: vorticity field ω, dissipation term $-\nu\nabla^2\omega$ and advection term $u \cdot \nabla\omega$ at $t = 4$. Bottom: corresponding wavelet coefficients $\tilde{\omega}$, $-\nu\widetilde{\nabla^2\omega}$ and $\widetilde{\vec{u} \cdot \nabla\omega}$.

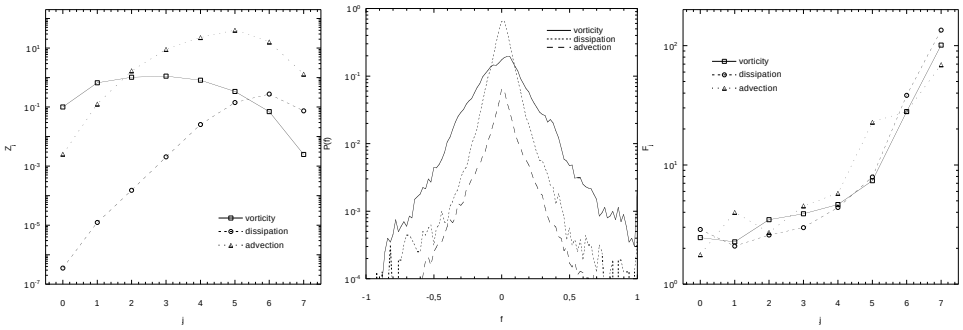


Figure 8: Scale distribution of enstrophy Z_j, normalized PDFs $p(f_n)$ for $f_n = \frac{f}{||f||_\infty}$ and scale-dependent flatness F_j for vorticity, dissipation and advection terms at $t = 4$.

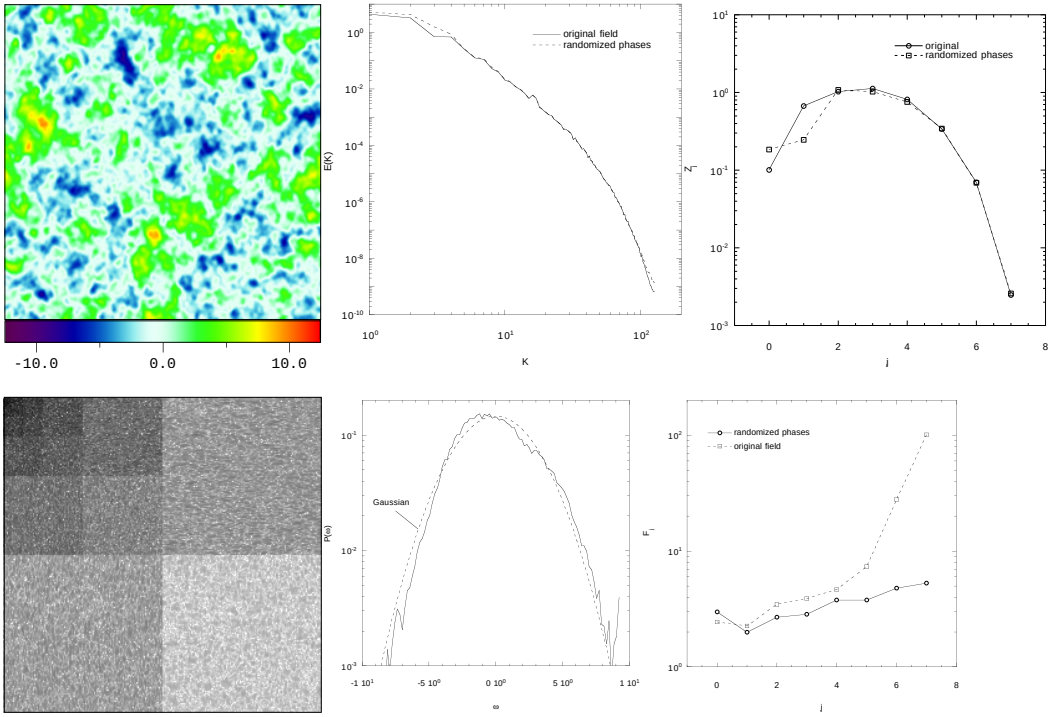


Figure 9: Top: vorticity with randomized phases, ω_{rand} at $t = 4$. Corresponding energy spectra $E(k)$ and scale distribution of enstrophy Z_j. Bottom: corresponding wavelet coefficients $\tilde{\omega}_{rand}$, PDF of vorticity ω_{rand} and flatness F_j.

PDF, with the same spatial correlation as the original vorticity field. The phase-scrambled vorticity field ω_r is defined as

$$\omega_r(\vec{x}) = \sum_{\vec{k}} |\widehat{\omega}(\vec{k})|\, e^{i2\pi\theta} e^{i2\pi\vec{k}\cdot\vec{x}}, \tag{96}$$

where $\widehat{\omega}(\vec{k}) = \int \omega(\vec{x}) exp(-i2\pi\vec{k}\cdot\vec{x}) d\vec{x}$ denotes the Fourier transform of the original vorticity field and θ are uniformly distributed random numbers, *i.e.* $\theta \in U(0,1)$. Although the resulting field has the same spectral behaviour (cf. figure 9b) and the same scale distribution (cf. figure 9c) as vorticity ω, it has neither coherent vortices in physical space (figure 9a), nor intermittency, *i.e.* no sparsity of its wavelet coefficients (figure 9d). The phase-scrambled field also has no significant increase of flatness F_j with scale (figure 9f) and the PDF is Gaussian with flatness $F = 2.9$ (cf. figure 9e) compared with $F = 5.7$ and stretched exponential PDF for the original field (figure 5b). We have thus shown that a fractal field with the same long-range dependence as a turbulent field (*i.e.* same energy spectrum), is not necessarily intermittent. This also demonstrates that intermittency in turbulence is due the presence of coherent vortices.

To illustrate the relation between structure functions and scale dependent moments of the wavelet coefficients we consider a typical statistically stationary two-dimensional turbulent flow field at resolution $N = 256^2$, extensively studied in [21]. In fig. 10 (top, left) we plotted its energy and enstrophy spectra exhibiting a k^{-5} and a k^{-3} power law behaviour, respectively. Figure 10 (top, right) shows the enstrophy Fourier spectrum together with the global wavelet spectrum using quintic spline wavelets. We find perfect agreement between both Fourier and wavelet spectra as predicted by the theory since the wavelet used has 5 vanishing moments. Furthermore we plot the global wavelet spectrum, plus its standard deviation at the different scales, to illustrate the fact that the fluctuations of the spectrum in physical space increase with the wavenumber. These diagnostics all indicate the presence of intermittency.

In figure 10 (middle) we plot the longitudinal structure functions $S_{||,p}(l)$ of the velocity for $p = 1$ to 6. On figure 10 (middle, left) we use the DOD wavelet, *i.e.* the classical structure function, and on figure 10 (middle, right) we use quintic spline wavelets. For the classical structure functions (figure 10 (middle, left)) we observe that the slope is limited by the number of vanishing moments of the wavelet. For example at $p = 2$ we observe that $S_2(l) \propto l^2$, whereas one should find for $S_2(l) \propto l^4$ since $E(k) \propto k^{-5}$. This limitation is due to the fact that the DOD wavelet has vanishing mean only, and therefore the structure functions show the scaling of the wavelet ($\zeta(p) = p$) and not that of the velocity! Using the quintic spline wavelets instead of the DOD wavelets, we find the correct slope of l^4 for S_2 since the quintic spline wavelets have 5 vanishing moments. For the higher order structure functions we find $\zeta(p) \approx 2p$, which is the expected property of the signal.

Finally, in figure 10 (bottom) we plot the structure functions versus the third order structure function, as used in the ESS approach. In both cases we observe that the functions are less curved than without using ESS. Using

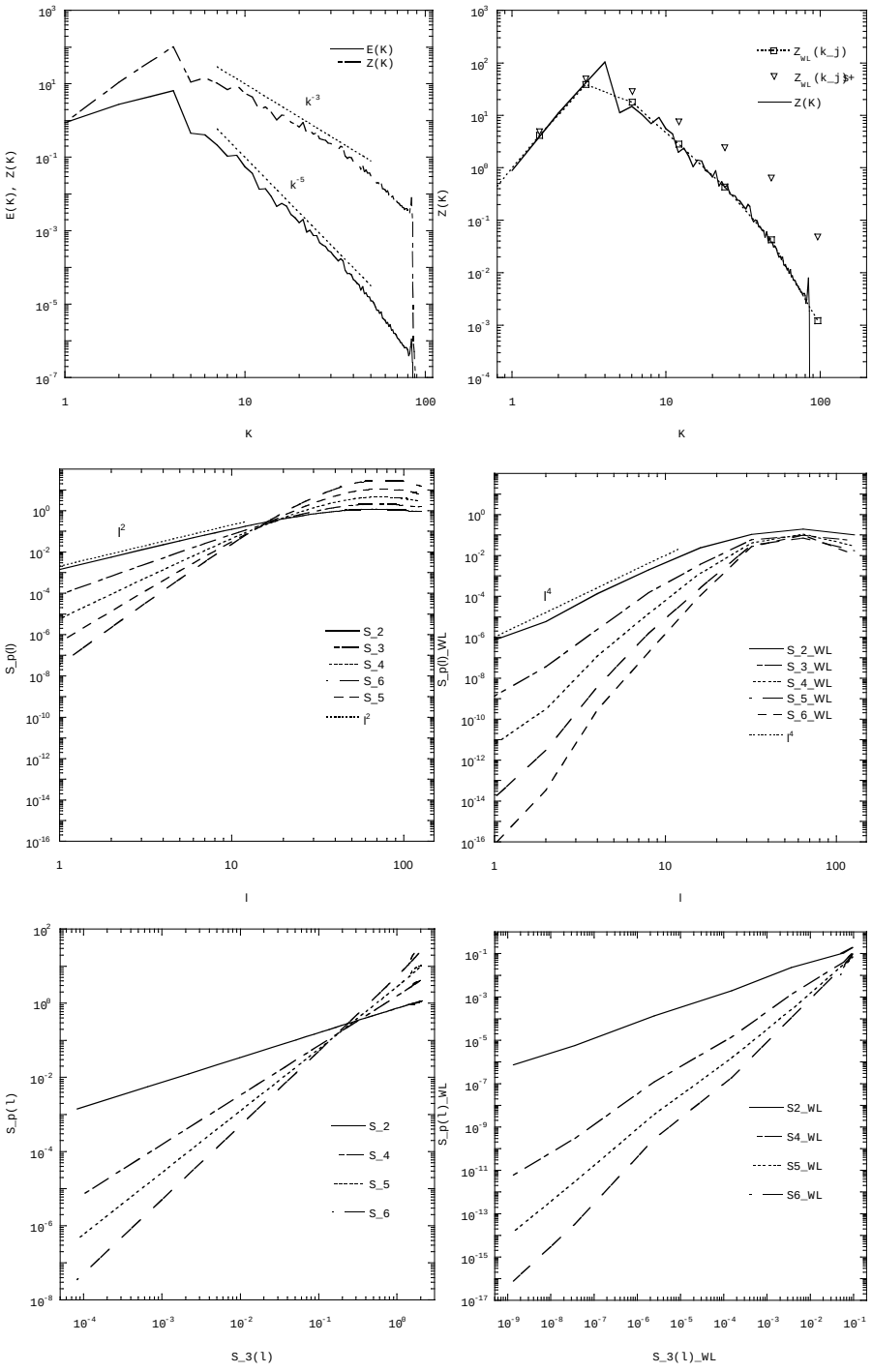


Figure 10: Top: Isotropic Fourier energy and enstrophy spectra (left). Fourier and global wavelet enstrophy spectra and the standard deviation of the wavelet spectrum in physical space. Middle: Classical longitudinal structure functions $S_{||,p}(l)$ for $p = 2, ..., 6$ of velocity (left) and corresponding wavelet based longitudinal structure functions (right), both averaged over 256 lines. Bottom: Structure functions versus third order structure function, left classical case, right wavelet case.

classical structure functions (figure 10 bottom, left) we find slopes of $\frac{2}{3}, \frac{4}{3}, \frac{5}{3}$ and 2 for $p = 2, 4, 5$ and 6 respectively. When we use the wavelet structure function based on quintic splines, we find the same slopes of$\frac{2}{3}, \frac{4}{3}, \frac{5}{3}$ and 2 for $p = 2, 4, 5$ and 6 respectively. This is due to the fact the we plot the structure functions versus the third order structure function and hence only information about the relative slope is obtained.

The above results show clearly that the slope of the classical structure functions is limited by the regularity of the underlying wavelet. The scaling behaviour of smoother fields can only be detected using structure functions based on wavelets with a sufficient number of vanishing moments. We have also shown that the ESS approach may be misleading as it only yields information about the relative slopes. These might be the same, even if the slope of the original structure functions are wrong.

6.2 Extraction of coherent vortices

As example, we apply the CVS filter to DNS data of a two-dimensional homogeneous and isotropic turbulenct flow, which was computed using a pseudo-spectral scheme with resolution $N = 512^2$ and without hyperdissipation. One realization of the vorticity field is shown on Figure 11a.(left) and most of the enstrophy is concentrated in the coherent vortices one observe. We observe that the coherent vortices are extracted by retaining only 2% of the N wavelet modes, which contain 99.99% of the total energy and 99.01% of the total enstrophy (Figure 11a., middle). They have the same velocity and vorticity PDFs (Figure 11c.), the same energy spectrum in the inertial range (Figure 11d., left) and the same cross PDF between vorticity and streamfunction characteristic of coherence (Figure 11b.) as the total flow.

The remaining incoherent background flow is made of vorticity filaments densely distributed in space (Figure 11d., right), deprived of coherence (Figure 11b., right), which have Gaussian PDFs for both velocity (Figure 11c., left) and vorticity (Figure 11c., right), and present a k^{+1} scaling of the energy spectrum characteristic of enstrophy equipartition in two dimensions (Figure 11d., left). This gives evidence that the incoherent background flow has reached a statistical equilibrium which corresponds to turbulent dissipation.

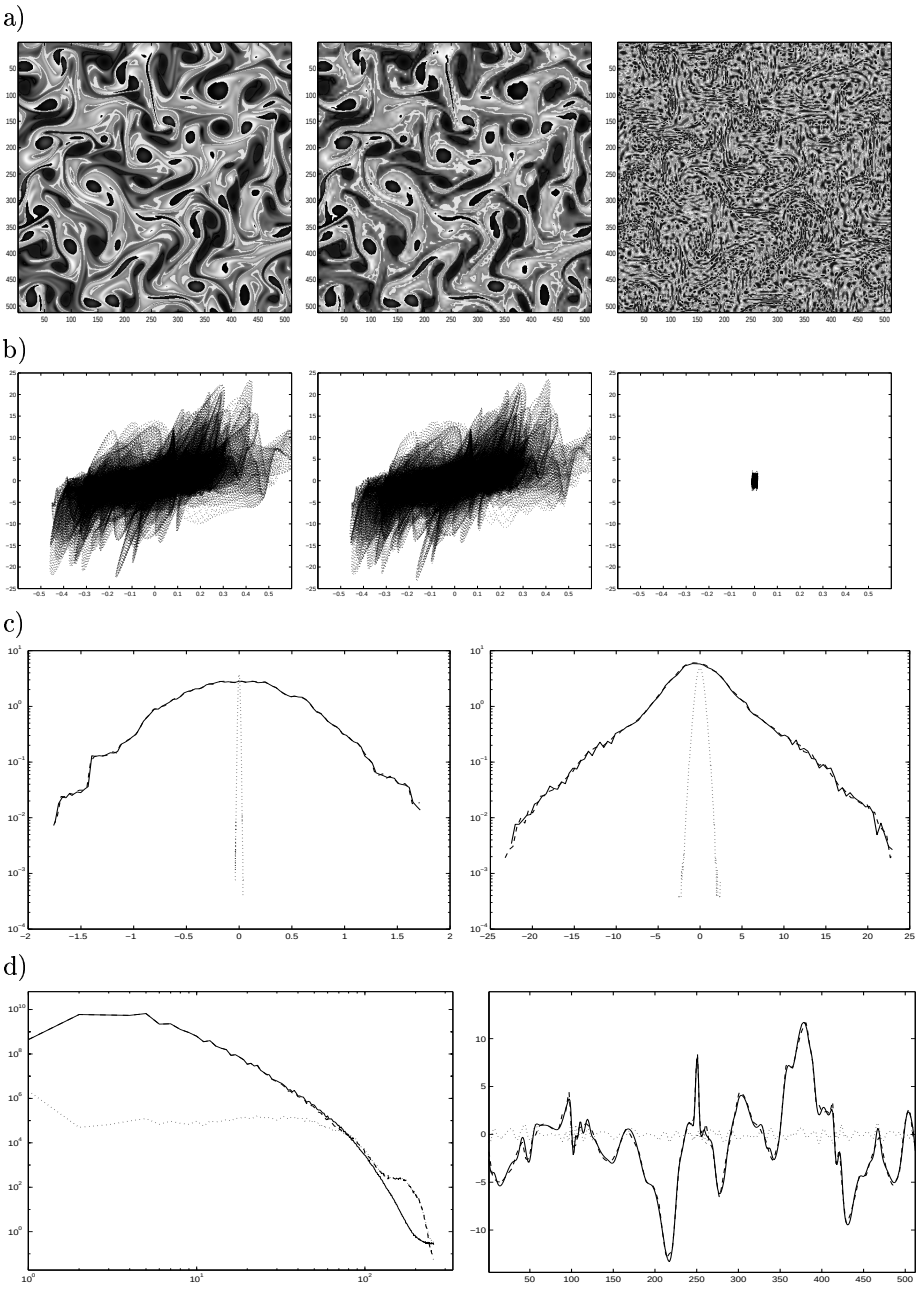


Figure 11: CVS filtering of a two-dimensional turbulent flow computed at $N = 512^2$: a) total vorticity (left), coherent vorticity (middle) which corresponds to $2\%N$ modes and contains 99% of both energy and enstrophy, incoherent vorticity (right) which corresponds to $98\%N$ modes and contains less than 1% of both energy and enstrophy, b) corresponding coherence scatter plots c) PDF of velocity (left), PDF of vorticity (right), d) energy spectrum (left), cuts of vorticity (right). Solid lines for the total, dashed lines for the coherent and dotted lines for the incoherent contributions.

7 Application to a 3D turbulence

7.1 Extraction of coherent vortices

We apply the CVS decomposition to 3D forced homogeneous isotropic turbulence, computed by DNS (Direct Numerical Simulation) at microscale Reynolds number $R_\lambda = 168$ with resolution $N = 256^3$ [34].

In Fig. 13 we plot the modulus of the vorticity fluctuations of the total flow field, considering only a 64^3 subcube. We observe that the field contains well defined vortex tubes, as previously observed in laboratory and numerical experiments [13, 59].

After decomposing the vorticity field into an orthogonal wavelet series, we calculate the square of each wavelet amplitude, which corresponds to the enstrophy retained in that mode. Subsequently, we sort them by decreasing order of magnitude and compute their partial sum to obtain the compression curve of the wavelet basis. In Fig. 12 we plot the percentage of retained enstrophy versus the fraction of retained wavelet modes. This curve shows that very few wavelet modes contain most of the enstrophy and that, above 10% of the modes, it saturates rapidly. This saturation corresponds to a quasi-equipartition of the enstrophy which is characteristic of random fields. On Fig. 12 we indicate by a star the threshold ϵ we use (5.3), which retains 3% of the wavelet coefficients and 79% of the enstrophy. The coherent vorticity $\vec{\omega}_C$ is then reconstructed from the retained wavelet coefficients and the incoherent vorticity $\vec{\omega}_I$ is the remainder.

We find that only 3% wavelet modes correspond to the coherent flow, which retains 98.9% of the energy and 79.1% of the enstrophy, while the remaining 97.1% incoherent modes have only 0.5% of the energy and 21% of the enstrophy.

In Fig. 14 we display the modulus of the coherent (left) and incoherent (right) vorticity fields. Note that the values of the vorticity isosurfaces are the same for the total and the coherent fields while they have been reduced by a factor 2 for the incoherent field since its amplitude is much smaller. In the

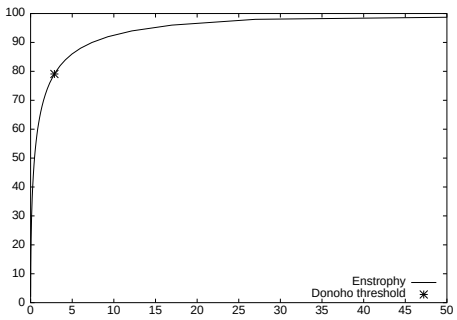


Figure 12: Compression curve: % retained enstrophy versus % number of retained wavelets. The star corresponds to the Donoho-Johnstone threshold.

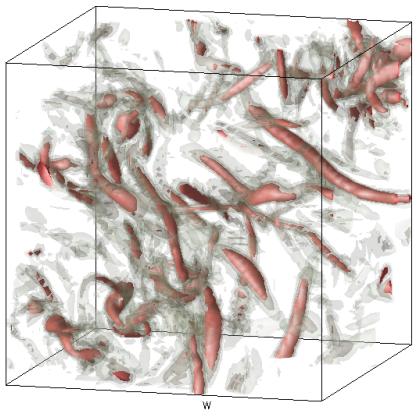


Figure 13: Total field: vorticity modulus (isosurfaces $|\vec{\omega}| = 3\sigma, 4\sigma, 5\sigma$, σ being the variance)

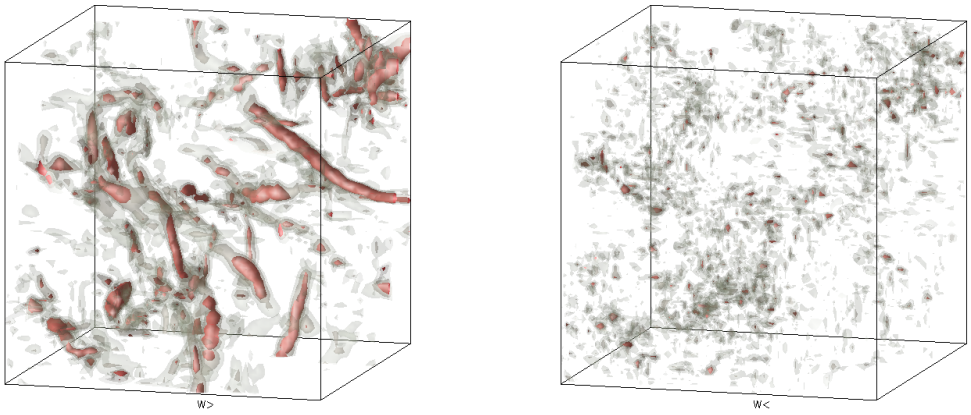


Figure 14: CVS decomposition: coherent (left) and incoherent (right) contributions (isosurfaces $|\vec{\omega}| = 3\sigma, 4\sigma, 5\sigma$ and $3/2\sigma, 2\sigma, 5/2\sigma$, respectively)

coherent vorticity (Fig. 14, left) we recognize the same vortex tubes as those present in the total field. In contrast, the incoherent vorticity (Fig. 14, right) is structureless and does not exhibit any organized structures. Hence, the CVS decomposition retains all the vortex tubes in the coherent part, whatever the scale where they are active, and disentangle the intermittent contribution of the vortices from the non intermittent background flow.

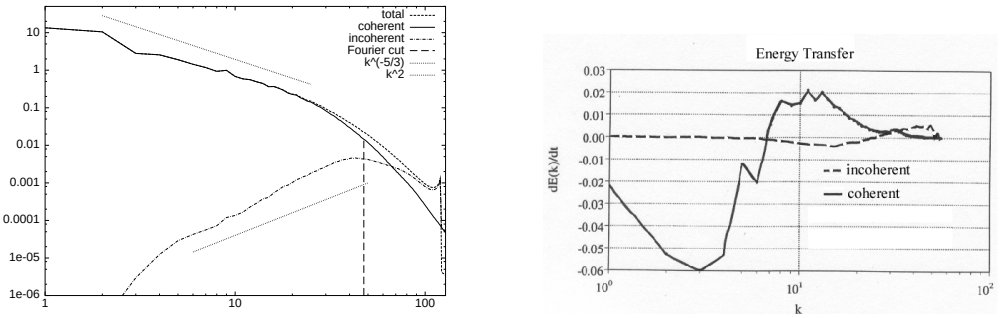


Figure 15: Energy spectra $E(k)$ (left) and energy transfers $d_t E(k)$ (right)

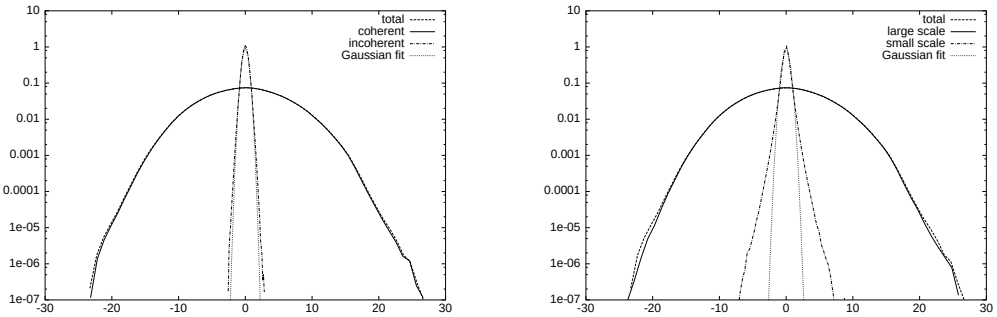


Figure 16: PDF of velocity: CVS (left) and POD (right) decompositions

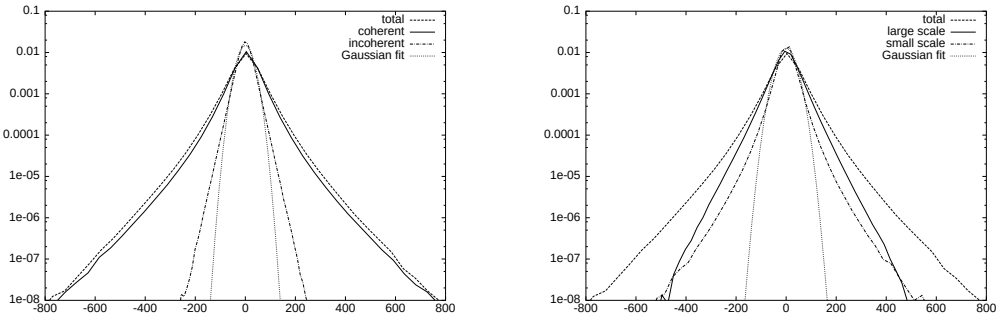


Figure 17: PDF of vorticity: CVS (left) and POD (right) decompositions

The energy spectra for the total, coherent and incoherent velocity fields, computed using the Biot–Savart law from the corresponding vorticity fields, are plotted in Fig. 15 (left). The spectrum of the coherent velocity is identical to the spectrum of the total velocity in the inertial range, which means that the vortex tubes are responsible for the $k^{-5/3}$ scaling there. Only in the dissipative range does the coherent spectrum decay more rapidly than the incoherent one, since

total energy is mostly converted into incoherent energy there and dissipated. The incoherent contribution exhibits a scaling close to k^2 which corresponds to energy equipartition. The incoherent velocity field is therefore decorrelated, which is consistent with the fact that the incoherent vorticity is structureless (see Fig. 14 (right)).

To check the dynamical behavior of the coherent and incoherent contributions we computed their energy transfer in wavenumber space. Fig. 15 (right) shows that the coherent flow is responsible for most of the energy transfer, giving an energy cascade from large to small scales which almost vanishes in the viscous range. In contrast, the incoherent flow does not contribute to the energy transfer in the inertial range, but dominates in the dissipative range. From these observations, we propose the following scenario for the turbulent cascade: the coherent energy injected into the large scales is nonlinearly transferred towards the small scales by nonlinear interactions between the vortex tubes, which produces incoherent energy. The incoherent energy becomes dominant in the smallest scales where it is dissipated. We conjecture that, on the contrary, the incoherent background flow does not transfer energy into the coherent flow as it is structureless and well decorrelated.

Fig. 16 (left) shows the PDF of the velocity in semilogarithmic coordinates. The coherent velocity has the same Gaussian distribution as the total velocity. The PDF of the incoherent velocity is also Gaussian but its variance is reduced by a factor 13. In contrast to the velocity, the PDF of vorticity (Fig. 17, left) is a stretched exponential with significant tails. The coherent vorticity has the same stretched exponential behaviour as the total vorticity including the tails, while the incoherent vorticity has an exponential behaviour with much weaker tails.

Since the CVS filtering is based on wavelet denoising, without any dynamical assumption or pattern recognition procedure, we now check *a posteriori* that we have actually extracted the vortex tubes from the background flow. The coherent vortex tubes can be described as local steady solutions of Euler equations which correspond to regions where there is a depletion of nonlinearity, which happens when the vorticity and velocity vectors are aligned. This situation maximizes the flow helicity $H = \vec{v} \cdot \vec{\omega}$ and corresponds to flow Beltramization [45]. To study this tendency towards alignment of the vorticity $\vec{\omega}$ and the velocity $\vec{v}$, we plot in Fig. 18 the PDF of the relative helicity $h = \frac{\vec{\omega} \cdot \vec{v}}{|\vec{\omega}||\vec{v}|}$. We observe that the coherent flow has the same tendency towards Beltramization as the total flow, which is characterized by the two maxima encountered in both PDFs for $h = -1$ (alignment) and $+1$ (anti-alignment). In contrast, the incoherent flow is more evenly distributed with a maximum at $h = 0$, which indicates a tendency towards a local two-dimensionalization, since the probability that the vortex stretching term $\vec{\omega} \cdot \nabla \vec{v}$ vanishes is large. This observation, together with the evidence for strong dissipation in the incoherent contribution (see transfers in Fig. 15), agrees with a remark of Moffatt: *Euler flows contain blobs of maximal helicity (positive or negative) which may be interpeted as 'coherent structures',*

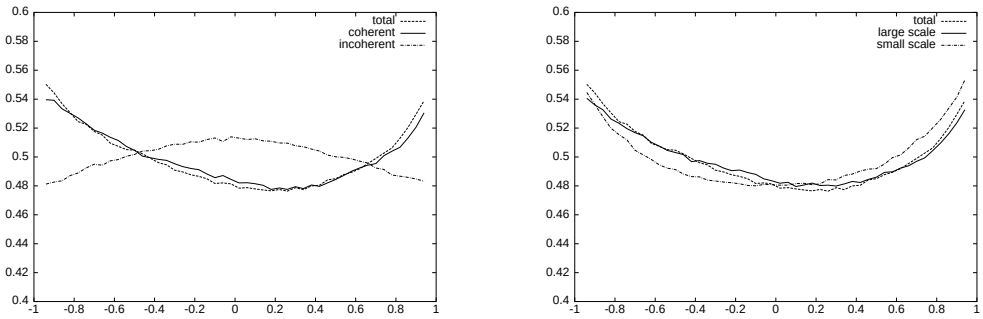


Figure 18: PDF of cosine of the relative helicity: CVS (left) and POD (right) decompositions

separated by regular surfaces on which vortex sheets, the site of strong viscous dissipation, may be located [45]. Following this picture the coherent vorticity corresponds to the coherent structures, which tend to maximize helicity where vorticity and velocity vectors tend to align with each other, while the incoherent vorticity corresponds to quasi-2D foliated regions which tend to maximize dissipation.

The results discussed here confirm those we obtained for the CVS decomposition of 3D forced homogeneous isotropic turbulent flow computed by DNS at microscale Reynolds number $R_\lambda = 150$ [59] with resolution $N = 240^3$ [14].

7.2 Comparison between CVS and POD decompositions

The procedure CVS decomposition uses to separate turbulent flows into organized and random fluctuations differs from the POD (Proper Orthogonal Decomposition). POD, also called PCA (Principal Component Analysis), or EOF (Empirical Orthogonal Functions) or Karhunnen-Loève decomposition, computes the auto-correlation tensor of an ensemble of realizations, then diagonalizes it and retains only those eigenmodes corresponding to the $N_>$ largest eigenvalues. This yields the best basis for the ensemble of realizations with respect to the L^2-norm. Note that the computational cost of POD decomposition scales as N^3, while it scales as N for the CVS decomposition.

In the POD decomposition the retained modes are defined *a priori* for all realizations. In contrast, CVS performs the separation *a priori* from each flow realization by selecting a given set of basis functions, the orthogonal wavelets, those having the strongest coefficients. Hence, the selection procedure is nonlinear, as the retained basis functions depend on each flow realization. From statistical point of view the CVS method is based on a Bayesian approach while POD is based on a non-Bayesian (also called frequentist) approach. For the time integration, the CVS *a priori* retains the wavelets whose coefficients are larger than the threshold ϵ at time step t and prepares some neighbour wavelets,

in space and scale, to compute the flow evolution from time step t to $t+1$ [22]. This selection of the active wavelets is nonlinear because it depends on the flow realization at time t and on the direction of the enstrophy transfers in wavelet space evaluated at the previous time step.

For a stationary homogeneous isotropic turbulent flow, such as the one studied here, the POD yields the Fourier basis, since the correlation tensor is translation invariant. In order to perform comparison between POD and CVS decompositions, we project the same vorticity field as chosen before on a Fourier basis and split the flow into low and high wavenumber contributions. Note that for this linear separation it doesn't matter whether we decompose the vorticity or the velocity fields, as the Fourier basis diagonalizes the curl operator. To get the same compression ratio as CVS, *i.e.*, 3% of the modes retained, the cut-off wavenumber is $k_c = 48$ (see Fig. 15, left). This is a particular case of LES filtering, with a Fourier low-pass filter, in which case the 97% high wavenumber modes are the LES subgrid scale modes.

In Fig. 19 we plot the modulus of vorticity for the POD decomposition. In the low wavenumber modes (left) we observe some vortex tubes. If we compare them with those retained in the CVS coherent vorticity (Fig. 14, left) we find that only a subset of the vortex tubes is extracted and that their structure is smoothed, due to the low-pass filtering. Consequently the small scale contributions of the vortex tubes are removed and the high wavenumber modes to be modelled (Fig. 19, right) exhibit organized structures similar to those found in the total vorticity field (Fig. 13).

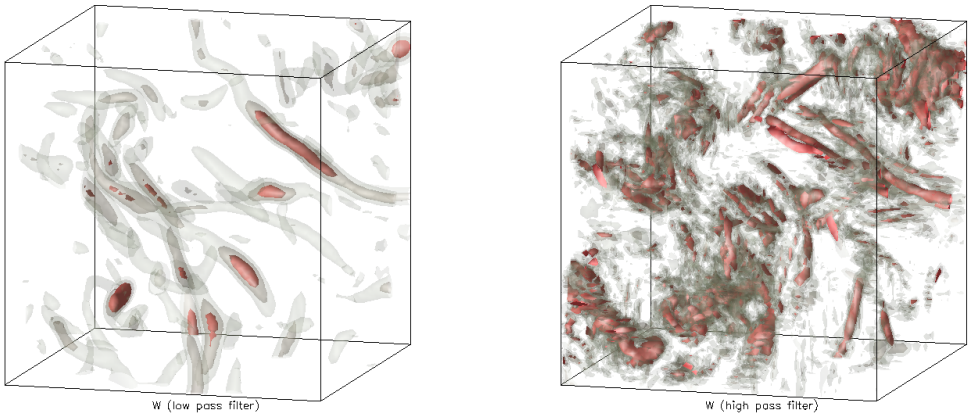


Figure 19: POD decomposition: low wavenumber (left) and high wavenumber (right) contributions (isosurfaces $|\vec{\omega}| = 3\sigma, 4\sigma, 5\sigma$ and $3/2\sigma, 2\sigma, 5/2\sigma$, respectively)

The POD retains 99.2% of the energy, while CVS retains only 98.9%. On

the other hand, CVS retains 79.1% of the enstrophy, while POD retains only 70.6%Z. The skewness of velocity and vorticity is negligible, a property preserved by both the CVS and POD decompositions.

In Fig. 17 the vorticity PDFs show that both the large and small scale contributions, have strong variance, with the peak of the small scale PDF being slightly larger than that of the large scale PDF. It is important to note that the vorticity PDFs are interchanged compared to the CVS decomposition. The vorticity PDF of the large scales retained by POD is exponential (with flatness 6.1), while it is stretched exponential (with flatness 9.6) for the discarded small scales. The vorticity PDF of the coherent modes retained by CVS is stretched exponential (with flatness 9.6), while it is exponential (with flatness 4.8) for the discarded incoherent vorticity.

Moreover, the velocity PDFs for POD (Fig. 16, right) show that, although the large scale contribution is gaussian (with flatness 2.9), this is not the case for the small scale contribution which maintains a stretched exponential behaviour as for the vorticity PDF (with flatness 6.8). In contrast, the velocity PDFs for CVS (16, left) are gaussian for both the coherent and the incoherent modes. This non-Gaussian behavior of both the vorticity and velocity PDFs of the POD/LES small scales, may make modeling of its effect on the resolved large scales difficult. This difficulty is much less acute with the CVS decomposition, since the PDFs of the incoherent contribution is Gaussian for the velocity (16, left) and exponential for the vorticity (17, left).

Concerning the alignement properties between vorticity and velocity we found that both the large and small scale contributions have the same PDF of $\cos\alpha$ as the total flow (18, right). This is further evidence that coherent vortex tubes are present in both components since, in contrast to CVS, the POD decomposition does not separate different topological behaviors.

Acknowledgments

We thank Nicholas Kevlahan, Alexandre Azzalini and Giulio Pellegrino for their kind help, and Maurice Meneguzzi for providing us the 3D vorticity field we have analyzed.

Appendix

Structure functions and Besoc norms

The relation between structure functions and wavelet coefficients can be generalized by using appropriate function spaces. For this we introduce Besov spaces [56, 9], which can be characterized using wavelet coefficients and are related to structure functions [50].

For $q < \infty$ we define the Besov space

$$\mathcal{B}^s_{p,q} = \left\{ f \in L^p(I\!R) \ ; \ l^{-s}(\int |f(x+l) - f(x)|^p dx)^{1/p} \in L^q(I\!R^*_+, \frac{dl}{l}) \right\} \tag{97}$$

with $0 < s < 1$, $p, q \geq 1$.

This means that $f \in \mathcal{B}^s_{p,q}$ if and only if $f \in L^p$, and

$$\left(\int_0^{+\infty} l^{-sq} (\int |f(x+l) - f(x)|^p dx)^{p/q} \ \frac{dl}{l} \right)^{\frac{1}{q}} < \infty \ . \tag{98}$$

Using the p-th order structure function $S_p(l)$ this is equivalent to

$$\left(\int_0^{+\infty} l^{-sq} \ S_p(l)^{\frac{q}{p}} \frac{dl}{l} \right)^{\frac{1}{q}} < \infty. \tag{99}$$

This means that the p-th order structure function is related to Besov norms via the modulus of continuity.

The corresponding norm is given by

$$||f||_{\mathcal{B}^s_{p,q}} = ||f||_{L^p} + |f|_{\mathcal{B}^s_{p,q}} \tag{100}$$

where the semi-norm $|f|_{\mathcal{B}^s_{p,q}}$ is defined as:

$$|f|_{\mathcal{B}^s_{p,q}} = \left(\int_0^{+\infty} \left[l^{-s} \ (\int |f(x+l) - f(x)|^p dx)^{1/p} \right]^q \ \frac{dl}{l} \right)^{\frac{1}{q}} \tag{101}$$

$$= \left(\int_0^{+\infty} l^{-sq} \ S_p(l)^{\frac{q}{p}} \ \frac{dl}{l} \right)^{\frac{1}{q}} \tag{102}$$

$$\approx \left(\int_0^{+\infty} \left[a^{-\alpha} \ ||\widetilde{f}(a, \cdot)||_{L^p} \right]^q \ \frac{da}{a} \right)^{\frac{1}{q}}. \tag{103}$$

This shows that the Besov norms ($q < \infty$) are intimately related to the structure functions $S_p(l)$ and the wavelet coefficients $\widetilde{f}(a, \cdot) = \langle f(x), \psi_{a,b}(x) \rangle$.

In the case $q = \infty$, we obtain

$$\mathcal{B}^s_{p,\infty} = \left\{ f \in L^p(I\!R) \ ; \ l^{-s}(\int |f(x+l) - f(x)|^p dx)^{1/p} \in L^\infty(I\!R^*_+) \right\} \tag{104}$$

and for the semi-norm we get

$$|f|_{\mathcal{B}^s_{p,\infty}} = ||l^{-s}(\int |f(x+l) - f(x)|^p dx)^{1/p}||_\infty = ||l^{-s} S_p(l)^{1/p}||_\infty. \qquad (105)$$

In the case of a self-similar behaviour of the type

$$S_p(l) \sim l^{\xi(p)} \qquad (106)$$

it follows that

$$f \in \mathcal{B}^{\xi(p)/p}_{p,\infty}. \qquad (107)$$

This implies that $\xi(p) < p$ as we restricted ourself to the case $s < 1$. To overcome this limitation the Besov spaces can be generalized for $s > 1$, where s is not an integer. For $s > 1$, s not an integer, we decompose s into $s = [s] + \sigma$, ($[s]$ being the integer part of s) and we introduce

$$\mathcal{B}^s_{p,q}(\mathbb{R}) = \left\{ f \in L^p(\mathbb{R}) \ ; \ f^{(m)} \in \mathcal{B}^\sigma_{p,q}(\mathbb{R}) \, , \, 0 \leq m \leq [s] \right\} \qquad (108)$$

where $f^{(m)}$ denotes the m-th derivative of f. The corresponding norm is defined as

$$||f||_{\mathcal{B}^s_{p,q}} = \sum_{m=0}^{[s]} ||f^{(m)}||_{\mathcal{B}^\sigma_{p,q}}. \qquad (109)$$

In order to have norm equivalence with the wavelet coefficients the wavelet ψ has to have at least $[s] + 1$ vanishing moments.

Let us mention that in the case where s is an integer, the modulus of continuity should be modified to $||f(x+l) - 2f(x) - f(x-l)||_{L^p}$ [56]. Note that this second-order stencil is no longer equivalent to the structure function. Finally, when $p = q = 2$ we obtain the Sobolev space H^s and for $p = q = \infty$ the Hölder space C^s [56].

To summarize, structure functions of order p correspond to Besov norms of functions which can be characterized by means of weighted sums of wavelet coefficients due to norm equivalences. This remark completes the link between structure functions, wavelet coefficients and Besov norms. It also suggests that the limitations of classical structure functions may be overcome by using structure functions based on wavelets with more vanishing moments than the DOD wavelet.

Extended self-similarity and quotients of Besov norms

To extend the scaling behaviour of structure functions one typically uses rescaled structure functions [4], *i.e.* one considers ratios of structure functions of different order $S_p(l)/S_q(l)$. One then studies the scaling behaviour of the p-th order structure function as a function of the q-th order structure function, *i.e.*

$$S_p(l) \propto S_q(l)^{\zeta(p,q)} \qquad (110)$$

with $S_q(l) \propto l^{\zeta(q)}$. Typically, $q = 3$ due to the fact that $S_3(l)$ is known exactly from the Karman–Howarth equation [26]. This approach is called extended self-similarity (ESS), and greatly increases the range over which one observes a well-defined power-law, even at moderate Reynolds numbers.

Because the scale-dependent moments of the wavelet coefficients are equivalent to the structure functions using the L^1 normalization of the wavelets, *i.e.* $S_p^{WL}(2^{-j}) = 2^{jp/2} M_{p,j}(f) = 2^{jp/2}/2^j \sum_{i=0}^{2^j-1} |\tilde{f}_{j,i}|^p$, the ratios of the moments at different scales $Q_{p,q,j}(f)$ (87) correspond to a generalized extended self similarity in wavelet space. This allows us to detect self-similar behaviour of functions with steeper slopes: if ψ has m vanishing moments then $\zeta(p)$ is bounded from above by mp.

The relations summarized in this section have been presented for the one-dimensional case only, but they can be generalized easily to higher dimensions using tensor product constructions of wavelets [8] and Besov spaces in $\mathbb{R}^n$ [56].

References

[1] A. Babiano, C. Basdevant, B. Legras and R. Sadourny. Vorticity and passive–scalar dynamics in two–dimensional turbulence. *J. Fluid Mech.*, **183**, 379–397 (1987).

[2] A. Babiano, C. Basdevant and R. Sadourny. Structure functions and dispersion laws in two–dimensional turbulence. *J. Atm. Sci.***42**(9), 941–949 (1985).

[3] G. K. Batchelor and A. A. Townsend. The nature of turbulent motion at large wave-number. *Proc. Roy. Soc.* A **199**, 238–255 (1949).

[4] R. Benzi, S. Ciliberto, R. Tripicione, C. Baudet and F. Massaioli. Extended self similarity in turbulen flows. *Phys. Rev. E* **48**, 29 (1993).

[5] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. *Comm. Pure Appl. Math.*, **44**, 141 (1991).

[6] B. Castaing. Consequence d'un principe d'extremum en turbulence. *J. Physique* **59**, 147–156 (1989).

[7] P. Chaisnais, P. Abry and J. Pinton. Intermittency and coherent structures in a swirling flow: A wavelet analysis of joint pressure and velocity measurements. *Phys. Fluids* **11**(11), 3524–3539 (1999).

[8] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[9] R. DeVore. Nonlinear approximation. *Acta Numerica* **8**, Cambridge University Press (1999).

[10] M. Do-Khac, C. Basdevant, V. Perrier and K. Dang-Tran. Wavelet analysis of 2d turbulent flows. *Physica D*, **76**, 252–277 (1994).

[11] D. Donoho. Wavelet shrinkage and Wavelet-Vaguelette Decomposition. *Progress in Wavelet Analysis and Applications* (eds. Y. Meyer and S. Roques) Editions Frontière, 109–128, 1992.

[12] D. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard. Wavelet shrinkage: Asymptopia? *J. Royal Stat. Soc., Ser. B* **57**, 301–369 (1995).

[13] S. Douady, Y. Couder and M. E. Brachet. Direct observation of the intermittency of intense vorticity filaments in turbulence. *Phys. Rev. Lett.* **67**, 983–990 (1991).

[14] M. Farge, G. Pellegrino and K. Schneider. Coherent Vortex Extraction in 3D Turbulent Flows using orthogonal wavelets. *Phys. Rev. Lett.*, **87**(5), 054501-1–054501-4 (2001).

[15] M. Farge, K. Schneider and N. Kevlahan. Non-Gaussianity and Coherent Vortex Simulation for two-dimensional turbulence using an adaptive orthonormal wavelet basis. *Phys. Fluids*, **11**(8), 2187–2201 (1999).

[16] M. Farge. Wavelet Transforms and their Applications to Turbulence. *Ann. Rev. of Fluid Mech.*, **24**, 395–457 (1992).

[17] M. Farge, M. Holschneider and J. F. Colonna. Wavelet analysis of coherent structures in two-dimensional turbulent flows. *Topological Fluid Mechanics* (eds. H. K. Moffatt and A. Tsinober) Cambridge University Press, 765–776 (1989).

[18] M. Farge and T. Philipovitch. Coherent structure analysis and extraction using wavelets. *Progress in Wavelet Analysis and Applications* (ed. Y. Meyer and S. Roques) Editions Frontières, 477–481 (1993).

[19] M. Farge and G. Rabreau. Transformée en ondelettes pour detecter et analyser les structures coherents dans les ecoulements turbulents bidimensionnels. *C. R. Acad. Sci. Paris Série II* **307**, 1479 (1988).

[20] M. Farge, N. Kevlahan, V. Perrier and E. Goirand. Wavelets and Turbulence. *IEEE Proceedings*, **84**, 4, 639–669 (1996).

[21] M. Farge, N. Kevlahan, V. Perrier and K. Schneider. Turbulence analysis, modelling and computing using wavelets. *Wavelets in Physics* (Ed. J.C. van den Berg), Cambridge University Press, 117–200 (1999).

[22] M. Farge and K. Schneider. Coherent Vortex Simulation (CVS), a semi-deterministic turbulence model using wavelets. *Flow, Turb. Combust.*, **66**, 393–426 (2001).

[23] M. Farge and K. Schneider. Analyzing and computing turbulent flows using wavelets. *Cours des Houches LXXIV: New Trends in Turbulence* (Ed. M. Lesieur, A. Yaglom and F. David) Springer, 393–426 (2002).

[24] J. H. Ferziger. Large Eddy Simulation. *Simulation and Modeling of Turbulent Flows.* ICASE Series in Computational Science and Engineering, eds. T. B. Gatski, M. Y. Hussaini and J. L. Lumley,109 (1992).

[25] U. Frisch and M. Vergassola. A prediction of the multifractal model: the intermediate dissipation range. *Europhys. Lett.* **14**, 439–444 (1991).

[26] U. Frisch. *Turbulence. The legacy of A.N. Kolmogorov.* Cambridge University Press (1995).

[27] J. Fröhlich and K. Schneider. An adaptive wavelet Galerkin algorithm for one– and two–dimensional flame computation. *Eur. J. Mech.* B, **13**, 439–471 (1994).

[28] J. Fröhlich and K. Schneider. A fast algorithm for lacunary wavelet bases related to the solution of PDE's. *C. R. Math. Rep. Acad. Sci. Canada* **17**(6), 283 (1995).

[29] J. Fröhlich and K. Schneider. Numerical simulation of decaying turbulence in an adaptive wavelet basis. *Appl. Comput. Harm. Anal.* **3**, 393–397 (1996).

[30] J. Fröhlich and K. Schneider. Computation of Decaying Turbulence in an Adaptive Wavelet Basis. *Physica D*, **134** 337–361 (1999).

[31] J. Fröhlich and K. Schneider. An Adaptive Wavelet–Vaguelette Algorithm for the Solution of PDE's, *J. Comput. Phys.*, **130**, 174–190 (1997).

[32] A. Grossmann and J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, **15**, 723–736 (1984).

[33] M. Holschneider and Ph. Tchamitchiam. Regularité locale de la fonction non-differentiable de Riemann, *Les ondelettes en 1989*, Lecture Notes in Mathematics, ed. P.G. Lemarié, Springer (1989).

[34] J. Jimenez J. and A. A. Wray, The structure of intense vorticity in isotropic turbulence, *J. Fluid Mech.*, **255**, 65–90 (1993).

[35] D. A. Kennedy and S. Corrsin. Spectral flatness factor and 'intermittency' in turbulence and in non-linear noise. *J. Fluid Mech.* **10**, 366–370 (1961).

[36] A. N. Kolmogorov. The local structure of turbulence in incompressible fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR.* **30**(4), 301–305 (1941).

[37] A. N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.* **13**, 82–85 (1962).

[38] R. H. Kraichnan. Inertial ranges in two-dimensional turbulence. *Phys. Fluids* **10**, 1417–1423 (1967).

[39] P. G. Lemarié and Y. Meyer. Ondelettes et bases Hilbertiennes. *Rev. Mat. IberoAm.*, **2**, 1–18 (1986).

[40] C. C. Lin. On Taylor's hypothesis and the acceleration terms in the Navier-Stokes equations. *Quart. Appl. Math.* **10**(4), 295–306 (1953).

[41] S. Mallat. *A wavelet tour to signal processing.* Academic Press (1997).

[42] Y. Meyer, *Ondelettes et Opérateurs I/II* Hermann, Paris (1990).

[43] J. C. McWilliams, The emergence of isolated coherent vortices in turbulent flows. *J. Fluid Mech.* ,**146**, 21–43 (1984).

[44] I. A. Min, I. Mezić and A. Leonard. Lévy stable distributions for velocity and velocity difference in systems of vortex elements. *Phys. Fluids* **8**(5), 1169–1180 (1996).

[45] H. K. Moffatt, Magnetostatic equilibria and analogous Euler flows of arbitrary complex topology, *J. Fluid Mech.*, **150**, 359–378 (1985).

[46] D. Montgomery, W. H. Matthaeus, W. T. Stribling, D. Martinez and S. Oughton. Relaxation in two dimensions and the "sinh-Poisson" equation. *Phys. Fluids* A, **4** (1992).

[47] R. Murenzi. Wavelet transforms associated to the n-dimensional Euclidean Group with dilatations. *PhD thesis*, UCL, Louvain-la-Neuve (1989).

[48] A. M. Obukhov. On the distribution of energy in the spectrum of turbulent flow. *Dokl. Akad. Nauk SSSR*, **32**(1), 22–24 (1941).

[49] V. Perrier and C. Basdevant. Periodical wavelet analysis. Theory and algorithms. *Rech. Aérosp.*, **3** (1989).

[50] V. Perrier and C. Basdevant. Besov norms in terms of the continuous wavelet transform. Applications to structure functions. *Math. Mod. Meth. Appl. Sci.*, **6** (1996).

[51] V. Perrier, T. Philipovitch and C. Basdevant. Wavelet spectra compared to Fourier spectra. *J. Math. Phys.*, **36**(3), 1506–1519 (1995).

[52] M. Rogers and R. Moser. Direct simulation of a self-similar turbulent mixing layer. *Phys. Fluids*, **6**(2), 903–923 (1994).

[53] V. A. Sandborn. Measurements of intermittency of turbulent motion in a boundary layer. *J. Fluid Mech*, **6**, 221–240 (1959).

[54] K. Schneider and M. Farge. Wavelet approach for modelling and computing turbulence. *Advances in turbulence modelling*. Von Karman Institute for Fluid Dynamics Lecture Series, **5** (1998).

[55] K. Schneider, N. Kevlahan and M. Farge. Comparison of an adaptive wavelet method and nonlinearly filtered pseudo-spectral methods for two-dimensional turbulence. *Theoret. Comput. Fluid Dyn.*, **9** (3/4), 191–206 (1997).

[56] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press (1970).

[57] G. I. Taylor. The spectrum of turbulence. *Proc. Roy. Soc. Lond.* A **164**, 476–490 (1938).

[58] A. A. Townsend. Local isotropy in the turbulent wake of a cylinder. *Austr. J. Sci. Res.* **1**, 161 (1948).

[59] A. Vincent, and M. Meneguzzi. The spatial structure and statistical properties of homogeneous turbulence. *J. Fluid Mech.*,**225**, 1–20 (1991).

Scaling Phenomena and Wavelets: a Review.

Patrice Abry*

CNRS, UMR 5672, laboratoire de Physique,
Ecole Normale Supérieure de Lyon,
46, allée d'Italie, F-69364, Lyon cedex 7, France.
Tél : 33 4 72 72 84 93, fax : 33 4 72 72 80 80,
e-mail : Patrice.Abry@ens-lyon.fr, http://www.ens-lyon.fr/∼pabry

Abstract: This note proposes first an introductory walk through the notions related to scaling phenomena and intuitions behind are gathered to formulate a tentative definition. Second, it introduces the mathematical model of self-similar processes with stationary increments, understood as the canonical reference for scaling. Then, it shows how and why the wavelet transform constitutes a powerful and relevant tool for the analysis (detection, identification, estimation) of self-similarity. It is finally explained why self-similarity is a too restrictive model to account for the large variety of scaling encountered in empirical data, and a review of various models related to scaling and of their interrelations is proposed (long range dependence, (multi)fractal processes, cascade processes). A set of *Matlab* routines has been developed to implement the wavelet-based analysis for scaling described here. It is available on the WEB page `www.ens-lyon.fr/∼pabry`.

Key-words: Scaling, Scaling Phenomena, Wavelet, Self-Similarity, Long-Range Dependence, Fractal, Multifractal, Infinitely Divisible Cascades, Estimation Procedure, Matlab Routines.

Résumé : Cette note propose d'abord une promenade introductive aux phénomènes d'invariance d'échelle et aux intuitions associées, pour permettre la formulation d'une définition. Elle définit, ensuite, avec détails, les processus auto-similaires à accroissements stationnaires, envisagés comme modèle mathématique de référence pour les lois d'échelle. Elle montre, dans un troisième temps, comment et pourquoi l'analyse en ondelettes constitue un outil efficace et pertinent pour l'analyse (détection, identification, estimation) de l'autosimilarité. Elle explique finalement en quoi l'autosimilarité constitue un modèle trop restrictif pour rendre compte de la grande variété des lois d'échelle rencontrées dans les données réelles et un panorama des déclinaisons de ce modèle et de leurs interrelations est présenté (dépendance à longue portée, processus (multi-)fractal, processus de cascade). Un ensemble de routines *Matlab* a été développé pour implanter les analyses en ondelettes des lois d'échelle décrites ici. Ces routines sont disponibles sur la page WEB `www.ens-lyon.fr/∼pabry`.

Mots-clé : Lois d'échelle, Invariance d'échelle, Ondelettes, Auto-similarité, Dépendance à longue portée, Fractal, Multifractal, Cascade infiniment divisible, Estimation, Routine Matlab.

*This note reviews works done in past years and still in progress, realized in collaboration with D. Veitch, EMUlab, Univ. of Melbourne, Australia, and with Pierre Chainais, Patrick Flandrin, ENS Lyon, France. These works have been partially supported by the CNRS grant TL99035, Programme Télécommunications and French MENRT grant ACI jeune Chercheur 2329.

1 Motivation - Introduction

Scaling phenomena. Power laws, scaling laws, scaling phenomena or, simply, scaling, recently became a very fashionable topic. Indeed, scaling behaviors were observed or studied or used as description paradigms in a large collection of research works covering a wide variety of different domains or applications. It is worth noting that those applications may be related to natural phenomena as well as to human activity. For the first category, one can, for instance, mention hydrology [8] with the study of variabilities of water levels in rivers, hydrodynamic with the study of developed turbulence [19, 15], statistical physics with the study of systems having long range interactions [35], microelectronics with $1/f$ noises in semi-conductors [7, 18], geophysics and fault repartitions or geological layers [34, 33], biology and physiology [41] with human rhythms variabilities, heart beat [36] or gait [21] for instance. For the second category, one can find human geography and population repartition in cities or continents [14], information flows on network and mainly computer network teletraffic [27], stock market volatilities or currency change rates fluctuations [24, 40]. The notion of scaling, however, remains defined poorly or in a loose way and may be related to a variety of different properties of a system or a process. A possible common tentative definition for scaling can be formulated through a negative statement: there is no characteristic scale (of time or space...) in the studied system or process. In other words, this is no longer possible to identify any scale that plays a privileged role compared to others, or equivalently, all scales play identical roles and are of equal importance in the dynamics of the analyzed system or process. Scaling, therefore, correspond to situations where the *whole* and a *subpart* of it can not be (statistically) distinguished one from the other. This is commonly associated to the picture of geometric fractal object, obtained from the iteration of an identical construction procedure.

Self-Similarity. From a data analysis (or signal processing) point of view, scaling in time series implies that the usual intuitive search techniques for characteristic scales are to abandoned and replaced instead by new ones aiming at evidencing relations, mechanisms between scales or involving a wide range of scales. This also means to abandon the use of models relying on the existence and definition of a characteristic scale (e.g, Markow chains, Poisson models, models with exponential autocorrelation functions,...). The canonical reference mathematical model to analyze scaling is that of self-similarity, and more particularly, the popular fractional Brownian motion. This will be introduced with details in Section 2.1.

Wavelets. The practical use and analysis of self-similar processes present however two major difficulties: they are non stationary and are characterized by long range dependence or long term correlations or long memory. Such statistical features turn the analysis of self-similar processes into a uneasy and non standard task. It has been shown recently in a collection of papers [1, 2, 37] that wavelet transforms constitute an ideal tool for the analysis of scaling. Wavelet analysis will be introduced in Section 2.2. More precisely, it has been shown [2] that wavelets can be considered as "matched" to self-similar processes in the sense that wavelet coefficients exactly reproduce, from scale to scale, the self-replicating statistical structure of such processes. This will be made explicit in Section 2.3. Section 2.4 explains how those statistical properties of the wavelet coefficients are to be used to design tools for the analysis (detection , identification and estimation) of scaling phenomena.

Beyond self-similarity. Self-similarity is a mathematically well-behaved model but it suffers from many limitations and can, obviously, not account for the large variety of scaling existing amongst actual empirical data. Section 3 will therefore allow a larger part to other models also of interest to describe scaling, such as long range dependence, fractal sample path, multifractal processes, multiplicative processes... and will underline their mutual interrelations, common denominators and differences.

Note. This note only intends to give a synthetic overview, without detailed proof, of the relations between scaling and wavelets. For full details, the interested reader is referred to the

tutorial chapter [2].

2 Self-Similarity and wavelets

2.1 Self-Similarity

Self-similar processes. A process X is said to be statistically self-similar, with self-similarity parameter $H > 0$, if [32]:

$$\forall c > 0, \{c^H X(t/c), t \in \mathbb{R}\} \stackrel{d}{=} \{X(t), t \in \mathbb{R}\} \tag{1}$$

where $\stackrel{d}{=}$ means equality of all the finite dimensional distributions. This essentially means that the sample paths (t, X) of the process $X(t)$ and those $(t/c, c^H X)$ of the process $c^H X(t/c)$ are statistically indistinguishable. In other words, the process X is statistically similar to any of its dilated versions. Therefore, no characteristic scale of time can be identified on these processes, self-similarity is hence a model for scaling behavior. This is illustrated on Figure 1.

A major consequence of self-similarity also lies in the fact that the moments of the process, when they exist (we assume in this section that they all exist), behave as power laws with respect to time,

$$\mathbb{E}|X(t)|^q = \mathbb{E}|X(1)|^q |t|^{qH}, \tag{2}$$

whose exponents are all controlled by the self-similarity parameter H. Besides the connection between scaling and power law, those relations also show that self-similar processes are non stationary.

Self-similar processes with stationary increments. Because it is always difficult to use practically on actual data non stationary models, one often restricts the class of self-similar processes to that of self-similar processes with stationary increments. A process is said to posses stationary increments if the finite dimensional distributions of $\{X(\tau, t) \equiv X_\tau(t) \equiv X(t+\tau) - X(t), t \in \mathbb{R}\}$ do not depend on t. Self-similarity and stationary increments together impose that $0 < H < 1$ and the increments of a self-similar processes with stationary increments satisfy the following similarity relation:

$$\forall c > 0, \{c^H X(\tau/c, t/c), t \in \mathbb{R}\} \stackrel{d}{=} \{X(\tau, t), t \in \mathbb{R}\}. \tag{3}$$

Moreover, the covariance function of the process X (with zero mean and $X(0) \equiv 0$) necessarily takes the following form:

$$\mathbb{E}X(t)X(s) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right). \tag{4}$$

with $\sigma^2 := \mathbb{E}|X(1)|^2$ and the correlation function of the increment process reads:

$$\begin{aligned} \mathbb{E}X_\tau(t)X_\tau(t+s) &= \frac{\sigma^2}{2}\left(|\tau+s|^{2H} + |\tau-s|^{2H} - 2|s|^{2H}\right) \\ &= \frac{\sigma^2}{2}|s|^{2H}\left(\left|1+\frac{\tau}{s}\right|^{2H} + \left|1-\frac{\tau}{s}\right|^{2H} - 2\right), \end{aligned} \tag{5}$$

Long range dependence or long term correlations or long memory. From this relation, one can infer the asymptotic behavior of the covariance function on the increment process in the limit of large s (i.e., $s \to +\infty, s \gg \tau$):

$$\mathbb{E}X_\tau(t)X_\tau(t+s) \sim \frac{\sigma^2}{2} 2H(2H-1)\tau^2 s^{2(H-1)}. \tag{6}$$

This asymptotic power law decrease is to be compared to the exponential one encountered in more usual processes (like Markov processes). An exponential decrease, by definition, implies a

characteristic time while a power law behavior does not, meeting again the intuition of the absence of characteristic scale of time. Moreover when $H > 1/2$, the autocovariance function decreases so slowly that its sum diverges, i.e., for $A > 0$,

$$\int_A^\infty \mathbb{E} X_\tau(t) X_\tau(t+s) ds = \infty.$$

This phenomenon is referred to as long range dependence, or long term correlations, and implies that the correlation between two samples of the process cannot be neglected without missing something crucial in the analysis of the process, no matter how far apart one from the other they are. Long range dependence among its increment process constitute a major difficulty in the analysis of a self-similar process. Independently from self-similarity, long range dependence, in itself, is a model for scaling observed in the limit of large scales (see Section 3 for details).

Local regularity of the process. Exploring the other limit, that of small scales, i.e., $s \ll \tau, s \to 0$, one obtains that the autocovariance function of X behaves as:

$$\mathbb{E} X_\tau(t) X_\tau(t+s) \sim \sigma^2 |\tau|^{2H} (1 - |\tau|^{-2H} |s|^{2H}).$$

Such a behavior indicates that the local regularity of each sample path of the process X is constant along time and controlled by the parameter H, since $1/2 < H < 1$, this means that the sample paths are everywhere continuous but nowhere differentiable. Local regularity is usually measured in terms of Hölder exponent. Processes with sample paths characterized by constant hölder exponent, are often referred to as *monofractal* processes, which constitute therefore a model for scaling observed in the limit of small scales. Local regularity and Hölder exponent will not be further detailed, see e.g., [28] for an introduction.

Fractional Brownian motion. The only Gaussian (zero-mean, $X(0) \equiv 0$) self similar process with stationary increments is the so-called *fractional Brownian motion* [22], which, in numerous situations, is quoted as the canonical reference for scaling. All its scaling properties, self-similarity, long-range dependence, fractal sample paths, are controll ed by the single parameter H, therefore used in applications to describe, sometimes confusingly, either global scaling properties or asymptotic ones...

2.2 Wavelet analysis

Continuous wavelet transform. The wavelet coefficients of the so called continuous wavelet transform (CWT) [12, 25] are defined as the results of comparisons, by means of inner products, between the process to be analyzed X and a family of functions, the wavelets $\psi_{a,t}$:

$$T_X(a,t) = \langle X, \psi_{a,t} \rangle, \quad (a,t) \in (\mathbb{R}^+, \mathbb{R}). \tag{7}$$

The wavelets are dilated and translated templates of a reference pattern ψ called the mother wavelet:

$$\psi_{a,t}(u) = \frac{1}{|a|} \psi\left(\frac{u-t}{a}\right). \tag{8}$$

Figure 2 shows dilated templates of a single mother wavelet. Note that some definitions prefer a $1/\sqrt{a}$ instead of $1/a$ normalization term, mainly because it ensures energy preservation. For the analysis of scaling, however, the choice $1/a$ is more convenient.

The function ψ is usually required to be bounded and to have time and frequency supports that are either bounded or decrease very fast, jointly in both domains, time and frequency. To ensure that the wavelet transform is invertible, ψ moreover has to be a zero-mean function:

$$\int_{\mathbb{R}} \psi(t) dt = 0. \tag{9}$$

Hence, the function ψ has to oscillate and to exist on reduced time support. this is therefore a small wave, a *wavelet*.

This is because the mother wavelet ψ has joint localization in time and frequency that the wavelet coefficients can be given the signification of frequency content of the data at a given time or joint time-frequency content of the information in X.

The mother-wavelet is characterized by its number N of vanishing (or zero) moments: the smallest integer such that $N \geq 1$ and:

$$\forall m \in \{0, \ldots, N-1\}, \quad \int_{\mathbb{R}} t^m \, \psi(t) dt = 0, \; \int_{\mathbb{R}} t^N \, \psi(t) dt \neq 0. \tag{10}$$

This means that, for a mother-wavelet with N vanishing moments, the wavelet coefficients of a polynomial of degree $P < N$ are strictly zero. More generally, it means that the wavelet coefficients $T_X(a,t)$ of a process X at time t are only sensitive to the part of the local behavior of X which is more irregular than that of a polynomial of degree N. In other words, the higher N, the less wavelet coefficients are sensitive to regular behavior. The number of vanishing moments of the mother wavelets also controls the behaviour of its Fourier transform at the origin:

$$|\Psi(\nu)| \sim |\nu|^N, \; |\nu| \to 0. \tag{11}$$

Figure 3 shows examples of wavelet with different vanishing moments.

Discrete wavelet transform. One also defines the coefficients of the discrete wavelet transform (DWT) as a discrete subset of the $T_x(a,t)$:

$$d_X(j,k) = T_X(a = 2^j, t = 2^j k) = \langle X, \psi_{j,k} \rangle, \;\; (j,k) \in (\mathbb{Z}^+, \mathbb{Z}), \tag{12}$$

where $\psi_{j,k}(u) = 2^{-j} \psi(2^{-j} u - k)$. This discrete subset of points is usually called the dyadic grid of the discrete wavelet transform and this definition also usually implies that the mother wavelet is constructed through a multiresolution analysis [12, 25]. The major interest in the use of the discrete wavelet transform lies in the facts that the $\{\psi_{j,k}, (j,k) \in (\mathbb{Z}^+, \mathbb{Z})\}$ form (possibly orthonormal) basis of $L^2(\mathbb{R})$ (so that the DWT is a non redundant representation of X) and that the $d_X(j,k)$ can be computed with fast pyramidal recursive algorithm whose computational costs is of the order of that of an fast Fourier transform. In all the methods and algorithms proposed here, we always use the DWT. For further details on wavelet transforms, the reader is referred to [12, 24].

Summary. All what the reader unfamiliar with wavelets needs to have in mind to follow the remainder of this text is that the relevance of the wavelet transform for the analysis of self-similar processes relies on two ingredients:

I1) the wavelet basis is designed from a dilation operator,

$$\psi_{a,t}(u) = \frac{1}{|a|} \psi(\frac{u-t}{a}),$$

I2) the mother wavelet is characterized by a strictly positive integer N, its number of vanishing moments,

$$\forall m \in \{0, \ldots, N-1\}, \quad \int_{\mathbb{R}} t^m \, \psi(t) dt = 0, \; \int_{\mathbb{R}} t^N \, \psi(t) dt \neq 0.$$

2.3 Self-Similarity and wavelets: theory

Let us assume that X is a self-similar process with stationary increments and let us examine the statistical properties of its wavelet coefficients. The results are given with no proof, the interested reader is referred to [2] for detailed demonstrations and original references.

P1 Self-Similarity: The wavelet coefficients reproduce, in an exact manner, the self-similarity of the process, through the relations:

$$\{2^{-jH} d_X(j,k), k \in \mathbb{Z}\} \stackrel{d}{=} \{d_X(0,k), k \in \mathbb{Z}\}. \tag{13}$$

$$\begin{aligned} \forall c > 0, \{c^H T_X(a/c, t/c), t \in \mathbb{R}, a \in \mathbb{R}^+\} \stackrel{d}{=} \\ \{T_X(a,t), t \in \mathbb{R}, a \in \mathbb{R}^+\}. \end{aligned} \tag{14}$$

These two relations result, fundamentally, from the fact that the wavelets are designed using a dilation operator: $\psi_a(t) = \psi(t/a)/a$. It is, moreover, interesting to note that the last relation has strong and obvious analogies to that satisfied by the increments (cf. equation (3)).

P2 Non-Stationarity: Though self similar processes are non stationary, their $\{d_X(j,k), k \in \mathbb{Z}\}$ form stationary sequences at each octave j. Identically, their $\{T_X(a,t), t \in \mathbb{R}\}$ form stationary processes at each scale a. This is again analogous to the increments and this is deeply connected to the fact that $N \geq 1$.

P3 Long Range Dependence: It can be shown that the covariance function of any two wavelet coefficients on the dyadic grid can be asymptotically bounded as, $|2^j k - 2^{j'} k'| \to +\infty$,

$$|\text{Cov } d_x(j,k), d_x(j',k')| \leq C|2^j k - 2^{j'} k'|^{-2(N-H)}. \tag{15}$$

which shows the key role played by the number of vanishing moments N. Increasing N allows to increase the rate of decrease of the covariance function and therefore to reduce as much as desired the range of the correlations amongst the wavelet coefficients. More precisely, it can be shown that, in the Gaussian case, if $N > H + 1/2$, the long range dependence that exists amongst the increments of X is turned into short range dependence.

P3ID Idealization: The "decorrelation effect", i.e., the reduction of the range of dependence of the wavelet coefficients under the increase of N, is idealized as follows :
any two wavelet coefficients of X, on the dyadic grid $\{d_X(j,k), k \in \mathbb{Z}, j \in \mathbb{Z}^+\}$, can be regarded as independent one from the other. This idealization is used to provide approximated but analytical studies of the performance of the estimators proposed below.

2.4 Self-Similarity and wavelets: application

Intuition: Because of their being non stationary, the study of the self-similar processes with stationary increments is usually performed through the analysis of their increments and more specifically through the use of the fundamental relation:

$$\mathbb{E}|X(t+\tau) - X(t)|^2 = \sigma^2 |\tau|^{2H}. \tag{16}$$

The existence of long term correlations amongst increments, however, substantially increase difficulties in the practical use of this formula.

The wavelet based analysis of self-similarity proposes to replace increments with wavelet coefficients. The intuitive ideas behind this are the following. First note the following identity between the increments of X and its wavelet coefficients obtained with a particular choice for the mother wavelet $\psi(t) = \delta(t + \tau_0) - \delta(t)$:

$$X(t+\tau) - X(t) \equiv T_x(a,t), \text{ with } \tau = a\tau_0.$$

Increments can therefore be thought of as a specific example of wavelet coefficients obtained from a particular mother wavelet, which has a poor spectral localization and only one vanishing moment (i.e., $N = 1$). One can naturally think of using a wavelet with better joint time and frequency localizations and higher numbers of vanishing moment, resulting in better statistical properties. Second, let us note, as already pointed out above, that the wavelet coefficients exactly reproduce

self-similarity (cf. equation **P1**), i.e., the scaling property. But from properties **P2** and **P3**, they form, at each scale 2^j, stationary sequences with short range and weak statistical dependence, they do not suffer any more from long range dependence (on condition that N is high enough). They are therefore statistically better behaved than increments and offer a versatile and convenient tool for the analysis of self-similarity.

Log-scale Diagram. More precisely from properties **P1** and **P2**, one can show that the variance of wavelet coefficients of self-similar processes with stationary increments behaves as power laws of the scales:

$$\mathbb{E}|d_x(j,k)|^2 = C2^{2jH}. \tag{17}$$

This last relation is highly reminiscent of equation (16) above for the increments and suggests that the analysis of self-similarity can be performed by the search of straight lines in a $\log_2 2^j = j$ versus $\log_2 \mathbb{E}|d_x(j,k)|^2$ plot. Practically, however, this implies to estimate the quantity $\log_2 \mathbb{E}|d_x(j,k)|^2$ from a single realization of finite length of the process. The properties **P2** (stationarity) and **P3** (weak statistical dependences) naturally lead to propose the following estimator:

$$Y_j = \log_2 \left(\frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^2 \right) ,$$

where n_j is the number of wavelet coefficients available at octave j. Let n denote the number of samples of the analyzed process X, neglecting the practical border effects resulting from the computation of the wavelet coefficients, the n_js behave as $n_j = 2^{-j}n$. We therefore propose to study self-similarity by plotting and analyzing *Logscale Diagrams*, i.e., plots $\log_2 2^j = j$ versus $Y_j = \log_2 \left(\frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^2 \right)$. Straight lines will evidence the existence of self-similarity and the measurement of their slopes will allow for an estimation of the parameter H. Figure 4 proposes an example of logscale diagram

Estimation Issues. Precisely, the estimator $\hat{H}$ for H is defined through the linear fit:

$$\hat{H} = \sum_j w_j Y_j / 2, \tag{18}$$

where $\sum_j$ runs over the range $\{j_1, \ldots, j_2\}$ of octaves where the linear fit is to be performed, this range is to be chosen a priori. The w_j satisfy the usual relations,

$$\begin{array}{lcl} \sum_j j w_j & = & 1 \\ \sum_j w_j & = & 0 \\ w_j & = & (1/a_j)(S_0 j - S_1)/(S_0 S_2 - S_1^2) \\ S_m & = & \sum_{j=j_1}^{j_2} a_j^{-1} j^m (m = 0, 1, 2), \end{array} \tag{19}$$

the a_js being arbitrary numbers.

The statistical performance of this estimator have been studied in detail in [1, 37, 2]. For Gaussian processes, analytical calculations relying on the idealization of exact independence of the wavelet coefficients **P3ID**, show that a residual bias can be determined and therefore subtracted to produce an unbiased estimator. The variance of the estimator asymptotically decreases as the inverse of the analyzed number of samples: $(\log_2 e)^2 \sum_j w_j^2 2^j/(2n)$. Numerical simulations show that the actual performance are very close to those idealized ones, even for non Gaussian processes. This approximate, however realistic, estimation of the variance of $\hat{H}$ enables us to have confidence intervals on the estimation of the parameter H.

The choice of the weights w_j remains to be specified through the choice of the a_js. It is well-known that the variance of the linear fit is minimal on condition that the a_j match the covariance structure of the Y_js. Assuming the idealization **P3ID**, one obtains that the Y_js are independent and that $\mathrm{Var}\, Y_j = 2\log_2(e))^2/n_j$. The choice $a_j = \mathrm{cste}\ n_j^{-1}$, therefore, ensures that the quantities

$(\sum_j w_j^2/n_j)$ and Var $\hat{H}$ takes close to minimal values. The estimator $\hat{H}$ is systematically implemented with this choice.

Additional properties. Thanks to its number of vanishing moments, the wavelet-based analysis of self-similarity, moreover, benefits from robustness against non-stationarities. For instance, if deterministic smooth trends (like a linear trend or an oscillating trend) are superimposed to self-similar processes, this may significantly complicate the detection and analysis of self-similarity. The comparison of wavelet-based analysis performed using mother-wavelet with different number of vanishing moments will allow to detect those trends and perform relevant analysis of self-similarity (see [1] or [2] for details).

From another perspective, non-stationarities and scaling may have practical effects and consequences that are practically very close and similar so that it may be difficult to distinguish one phenomena from the other. We have also shown that the wavelet framework offers a convenient way to design a statistical test allowing to discriminate actual scaling from some non stationary effects [39].

Finally, other interesting features of this wavelet based analysis lies in the facts that it is simple both conceptually and practically (DWT plus linear fits) and that it has a low computational cost thanks to the recursive pyramidal algorithm underlying the discrete wavelet transform. This is of importance when dealing with large sample of data, as is often the case when dealing with scaling and allowed us to propose real time on line algorithm for the analysis of scaling [30].

3 Beyond self-similarity

Practical limitations. Self-similar processes with stationary increments, and more specifically the fractional Brownian motion, are very attractive models to describe scaling because they are mathematically well-defined and well-documented and they moreover fulfil the intuition of scaling phenomena in a very satisfactory manner. Their major quality is simplicity: all the features of the process (large scales scaling and long range dependence as well as small scales scaling and local regularity) are entirely controlled by the single parameter H. Their major drawback is ... simplicity. It is very unlikely that the various types of scaling encountered in the many different applications where they occur can all be described by a unique model with single parameter. More precisely, exact self-similarity implies a number of specific properties and significant departures from those properties can be observed in the analysis of actual empirical data: *i*) moments of different orders may have scaling exponents that are not controlled by a single parameter, or more simply some moments may not present scaling or, even more simply, may not exist at all; *ii*) when scaling are observed, they may not exist over the whole range of scales as in the self-similar case, but only in a given range of scales, or only asymptotically in the limit of large scales or in the limit of small scales; *iii*) power-law behavior of the moments may not exist despite scaling behavior. In this section, we will explore those variations and describe some related mathematical models.

Beyond finite variance. In the previous section, we assumed that all the moments of the process X existed. However, one may encounter situations where scaling and self similarity are valid but where the variance of the process, for instance, and therefore all higher moments, are infinite. For those situations, the model of Gaussian self-similar processes as well as the analysis presented above and based on the variance of the wavelet coefficients no longer work. Such situations can be modelled using α-stable self-similar processes.

A usual definition of stable processes is through their integral representation [32]:

$$X(t) = \int_{\mathbb{R}} f(t,u) M(du)\,, \tag{20}$$

where $M(du)$ is an α stable measure ($0 < \alpha \leq 2$) and $f(t,u)$ is an integration kernel. This definition means that the process X can be read as a weighted sum of independent stable random

variables. Technically, this implies that, for a fixed t, $X(t)$ is a stable random variable, M therefore controls the marginals of the process, or in other words, its static properties. The weighting function $f(t,u)$ controls the statistical dependences of x and hence its joint statistics, or in other words, its dynamic properties. For well chosen forms of the kernel, [13, 26, 32], x is a self-similar process with stationary increments. A wavelet-based analysis of α-stable self-similar processes can be conducted but relies on the quantity, $\mathbb{E}\log_2|d_X(j,k)|$, instead of $\mathbb{E}|d_X(j,k)|^2$. This is described in details in [6, 3].

Beyond scaling over all scales: long range dependence or local regularity. A major limitation of the self-similarity model lies in the fact that the scaling behavior holds for all the scales, i.e.,

$$\mathbb{E}|d_X(j,k)|^2 \equiv C2^{2jH}, \forall j.$$

Practically, scaling may exist for the second order statistics (namely the variance) of the process, but may be only observed in a given, large but finite, range of scales, or in the asymptotic limits of small or large scales, rather than in the whole range of scales.

For instance, if the observation corresponds to the weakened version of the relation above,

$$\mathbb{E}|d_x(j,k)|^2 \simeq C2^{2jH}, j \to +\infty,$$

this is the signature of a scaling behavior that exists for the largest scales of the process only. This, actually, tells us that the data are not self-similar but rather present some long term correlations properties and can therefore be modelled as a stationary long range dependent process instead of a self-similar one [37, 2].

When scaling hold in the limit of small scales,

$$\mathbb{E}|d_x(j,k)|^2 \simeq C2^{2jH}, j \to -\infty,$$

this means that the data are not self similar but rather that their sample paths are characterized by a local regularity h controlled by H and that remains constant along time. This therefore betrays a small scale property.

For those two situations, long range dependence or local regularity, the analysis and estimation of the exponent can be performed with the logscale diagram, as in the self-similar case, except that linear fits are to be performed in the limit of large or small scales, respectively.

Beyond second-order statistics - Multifractals. Another limitation of self-similar processes is that scaling for all the moments are controlled by the single parameter H:

$$\mathbb{E}|d_X(j,k)|^q \equiv C2^{qjH}, \forall j, \forall q > 0.$$

It is, however, quite common on empirical data to observe, in a given, finite, range of scales, a behavior of the type:

$$\mathbb{E}|d_X(j,k)|^q \simeq C2^{jH(q)}, j_1 < j < j_2,$$

where the exponents $H(q)$ may differ from the linear qH behavior. This situation, that we proposed to refer to as *multiscaling* is very close to that encountered when analyzing multifractal processes. Multifractal processes present sample paths with a local regularity h that varies very irregularly from point to point and with each realizations. Those fluctuations of the local regularity are described through the so-called multifractal spectrum $D(h)$, (which consists in the Hausdorff dimension of the set of points where the local regularity is h). A major consequence of multifractality is that, for multifractal processes, quantities called partition functions,

$$(1/n)\sum_{k=1}^{n}|d_X(j,k)|^q$$

present in the limit of small scales power law behaviors,

$$(1/n)\sum_{k=1}^{n}|d_X(j,k)|^q \sim C2^{j\zeta(q)}, j \to 0.$$

The measurement of the $\zeta(q)$ exponents offer, through a Legendre transform, a possibility, amongst others, to estimate the multifractal spectrum. Details on multifractal are beyond the scope of this note and the interested reader is referred to e.g., [29, 28]. Celebrated examples of multifractal processes are, for instance, obtained with the so-called binomial cascades.

Reading the partition functions $(1/n)\sum_{k=1}^{n}|d_X(j,k)|^q$ as estimators of the moments $\mathbb{E}|d_X(j,k)|^q$, the scaling relation above is very close to the equation defining multiscaling in the limit of small scales. Therefore multifractal can be seen as the very example for multiscaling. An extension of the estimation procedure described in Section 2 has been proposed to estimate the $H(q)$ exponents (see [2]).

Beyond power laws Multiscaling offers an extension to self-similarity insofar as a collection of exponents is needed rather than a single one. It, yet, maintains a major feature of self-similarity: moments behaves as power laws of the scales. However, when analyzing actual data, it may very well be observed that this is not the case, see e.g., [38]. To account for those situations, the infinitely divisible cascade model, on which we concentrate here, proposes to gain an extra degree of freedom by giving up the requirement that moments behaves as power laws of the scales. It maintains, however, a fundamental feature, in common with self-similarity and multi-scaling, the dependence of the moments in the variables q (order of the moment) and 2^j (scale) is separable. The equations below summarize those relations between self-similarity, multiscaling and infinitely divisible cascade:

$$\begin{array}{lllll} \text{Self-Sim.} & \mathbb{E}|d_X(j,k)|^q = & C_q(2^j)^{qH} & = C_q\exp(qH\ln(2^j)) \\ \text{MultiScaling} & \mathbb{E}|d_X(j,k)|^q = & C_q(2^j)^{H(q)} & = C_q\exp(H(q)\ln(2^j)) \\ \text{Inf. Div. Casc.} & \mathbb{E}|d_X(j,k)|^q = & --- & = C_q\exp(H(q)n(2^j)), \end{array} \tag{21}$$

where the function $n(2^j)$ is not necessarily the ln function, as the function $H(q)$ is not a priori the linear function qH.

The concept of infinitely divisible cascade was first introduced by B. Castaing in [9, 10] and rephrased in the wavelet framework in [5]. We now briefly recall its intuition, definition, consequences and relations to other models. Starting again from the self-similar case, one can write the probability density function (pdf) of the wavelet coefficients at scale $a = 2^j$, as a dilated version of the pdf of the wavelet coefficients at a larger scale a': $p_a(d) = (1/\alpha_0)\,p_{a'}(d/\alpha_0)$ where the dilation factor is unique : $\alpha_0 = (a/a')^H$. In the cascade model, the key ingredient is that there is no longer a unique factor but a collection of dilation factors α ; consequently p_a will result from a weighted sum of dilated incarnations of $p_{a'}$:

$$p_a(d) = \int G_{a,a'}(\ln\alpha)\,\frac{1}{\alpha}\,p_{a'}\left(\frac{d}{\alpha}\right)\,d\ln\alpha.$$

The function $G_{a,a'}$ is called the kernel or the *propagator* of the cascade. Obviously, if $G_{a,a'}$ is a Dirac function, $G_{a,a'}(\ln\alpha) = \delta(\ln\alpha - H\ln(a/a'))$, infinitely divisible cascade reduces to SS, therefore understood as a special case. The definition of the cascade above shows that the pdf's of $\underline{p}_a$ and $\underline{p}_{a'}$ of the wavelet log-coefficients $\ln|d|$ are related by a convolution with the propagator :

$$\begin{array}{rcl} \underline{p}_a(\ln\alpha) & = & \displaystyle\int G_{a,a'}(\ln\alpha)\,\underline{p}_{a'}(\ln|d| - \ln\alpha)\,d\ln\alpha \\ & = & (G_{a,a'} * \underline{p}_{a'})(\ln\alpha). \end{array} \tag{22}$$

If cascades exist between scales a'' and a' and between scales a and a'', then a cascade between scales a and a' exists, and the corresponding propagator results from the convolutions of the two

propagators: $G_{a,a'} = G_{a,a''} * G_{a'',a'}$. Infinite divisibility (also called continuous self similarity) means that no scale between a and a' plays any characteristic role (i.e., a'' in the above statement can be any scale between a and a'). Infinite divisibility therefore implies that the propagator consists of an elementary function G_0 convolved with itself a number of times, where that number depends on a and a':

$$G_{a,a'}(\ln \alpha) = [G_0(\ln \alpha)]^{*\left(n(a)-n(a')\right)}.$$

Using the Laplace transform $\tilde{G}_{a,a'}(q)$ of $G_{a,a'}$, this can be rewritten as $\tilde{G}_{a,a'}(q) = \exp\{H(q)(n(a) - n(a'))\}$, with $H(q) = \ln \tilde{G}_0(q)$ and $a := 2^j$; this implies that $\mathbb{E}|d_X(j,k)|^q = C_q \exp\{H(q)n(a)\}$, thus validating eq. (21). The main consequences of infinitely divisible cascade (in other words, of the separability of the variables q and a), read therefore :

$$\ln \mathbb{E}|d_X(j,k)|^q = H(q)n(a) + K_q \tag{23}$$

$$\ln \mathbb{E}|d_X(j,k)|^q = \frac{H(q)}{H(p)} \ln \mathbb{E}|d_X(j,k)|^p + \kappa_{q,p}. \tag{24}$$

This last equation implies that moments behave as power-laws relative to each other. Such relations are sometimes called "extended self-similarity", in turbulence mainly. Note that, in the relation (23) above, there is an arbitrary element, indeed:

$$\begin{aligned} H(q)n(a) + K_q &= \left(\tfrac{H(q)}{\beta}\right)(\beta n(a) + \gamma) + (K_q - \tfrac{H(q)\gamma}{\beta}) \\ &= H'(q)n'(a) + K'_q \end{aligned}$$

where $\beta \neq 0$ and γ are arbitrary constants. It clearly indicates that the function $H(q)$ is defined up to a multiplicative constant while n is defined up to multiplicative and additive constants.

If it is moreover required that the function $n(a) \equiv \ln a$, the infinitely divisible cascade is called *scale invariant* and this implies that :

$$\tilde{G}_{a,a'}(q) = (a/a')^{\ln \tilde{G}_0(q)} \text{ and } \mathbb{E}|d_X(j,k)|^q = (2^j)^{\ln \tilde{G}_0(q)},$$

proving that scale invariant infinitely divisible cascade reduces to MS. If, moreover, the power-laws are observed in the limit of small scales ($a = 2^j \to 0$), then, multiscaling is equivalent to *multifractal*, and the exponents ζ_q—from which the Legendre multifractal spectrum is obtained through a Legendre transform [28]—are related to the propagator through $\zeta_q = H(q) = \ln \tilde{G}_0(q)$. In this framework, multifractal is therefore understood as a special case of infinitely divisible cascade. The stochastic multiplicative cascades introduced by Mandelbrot [23], constitute the canonical example of such situations. In a scale invariant infinitely divisible cascade, one can also inquire as to whether ζ_q is a linear function of q or not, in which case the cascade reduces to the even more special case of self-similarity. It is, therefore, natural to consider the function $\zeta_q/q = H(q)/q$ and test its constancy.

4 Conclusion

We have tried to show here that scaling or scaling phenomena may cover a wide variety of different situations, but still that it can be gathered under the intuition that all the scales involved in the system or the process play equivalent roles or equivalently, that none plays a specific role. We have introduced self-similarity that can be thought of as the reference model for scaling. Then, we have shown how and why the wavelet transform offers a versatile, powerful and efficient tool to study scaling, there is somehow an intimate adequacy between the analyzed problem, scaling, and the analyzing tool, that study a process simultaneously at all scales. Other scaling models were then introduced, such as long range dependence, fractal sample path, multifractal processes or infinitely divisible cascade, they can be regarded as variations around the model of self-similarity too limited to account for the whole variety of scaling encountered in empirical data.

In this text, the results and techniques were presented without detailed proofs. The emphasis was on providing the reader with both a comprehensive survey of the relations between the various mathematical models involved in the analysis of scaling phenomena and practical simple-to-use wavelet based analysis techniques. So that a reader willing to apply these tools on data of its own can find here *a quick start*. The corresponding MATLAB routines are available on the WEB site `www.ens-lyon.fr/ ~pabry`. Most of the material covered in this note is described and presented in large details in [2].

References

[1] P. Abry, D. Veitch, Wavelet analysis of long-range dependent traffic, *IEEE Trans. on Info. Theory*, 44(1):2–15, 1998.

[2] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch. Wavelets for the analysis, estimation and synthesis of scaling data. A chapter in [27].

[3] P. Abry, B. Pesquet-Popescu, M. S. Taqqu, Wavelet Based Estimators for Self Similar α-Stable Processes, Int. Conf. on Signal Proc., 16th World Computer Congress 2000, Pékin, Chine.

[4] P. Abry, P. Gonçalvès and J. Lévy Véhel, Eds., *Lois d'échelle et Fractales*, Traité Information, Commandes, Contrôle, Hermès, France, to appear, 2002.

[5] A. Arnéodo, J.F. Muzy, S.G. Roux, Experimental analysis of self-similar random cascade processes: application to fully developped turbulence, *J. Phys. II France*, 7:363–370, 1997.

[6] P. Abry, L. Delbeke and P. Flandrin, Wavelet-based estimator for the self-similarity parameter of α-stable processes. *IEEE-ICASSP-99*, Phoenix (AZ), 1999.

[7] V. Bareikis, R. Katilius, Eds., Noise in Physical Sustems and $1/f$ Fluctuations, World Scientific, 1995

[8] J.Beran, *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.

[9] B. Castaing, Y. Gagne, E. Hopfinger, Velocity probability density functions of high Reynolds number turbulence. *Physica D*,46:177, 1990.

[10] B. Castaing. The temperature of turbulent flows. *J. Phys. II France*, 6:105–114, 1996.

[11] P. Chainais, P. Abry, and J.F. Pinton, "Intermittency and coherent structures in a turbulent flow: a wavelet analysis of joint pressure and velocity measurements," *Phys. Fluids*, Vol. 11, no 11, pp. 3524-3539, 1999.

[12] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia (PA), 1992.

[13] L. Delbeke, P. Abry, Stochastic integral representation and properties of the wavelet coefficients of the linear fractional stable motion, Stochastic Processes and their Applications, 86, pp. 177–182, 2000.

[14] P. Frankhauser, L'approche fractale : un nouvel outil dans l'analyse spatiale des agglomerations urbaines, Population, 4:1005–1040, 1997.

[15] Frisch, U., *Turbulence. The legacy of A. Kolmogorov*. Cambridge University Press, UK: Cambridge University Press, 1995.

[16] A.C. Gilbert, W. Willinger, A. Feldmann, Scaling analysis of random cascades, with applications to network traffic, *IEEE Trans. on Info. Theory*, Special Issue on Multiscale Statistical Signal Analysis and its Applications, 45(3):971–991, 1999.

[17] P. Gonçalvès and R. H. Riedi. Wavelet analysis of fractional Brownian motion in multifractal time. Proc. 17ème Colloque GRETSI, Vannes, France, 1999.

[18] M. Keshner, $1/f$ noises, proc. of the IEEE, 70(3):212–218, 1982.

[19] A.N. Kolmogorov, *a) Dissipation of energy in the locally isotropic turbulence. b) The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. c) On degeneration of isotropic turbulence in an incompressible viscous liquid.* In S.K. Friedlander and L. Topper, editors, *Turbulence, Classic papers on statistical theory.* 151-161. Interscience publishers (1941).

[20] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, On the self-similar nature of Ethernet traffic, Extended-Version, *IEEE/ACM Trans. on Networking*, 2:1–15, 1994.

[21] L. Liebovitch, A. Todorov, Invited editorial on fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations, The American Physiological Society, 1446:1447, 1996.

[22] B.B. Mandelbrot, J.W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev, 10:422-437, 1968.

[23] B.B. Mandelbrot, Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, J. of Fluid Mech., 62(2):331–358, 1974.

[24] B.B. Mandelbrot, *Fractales, Hasard et Finance*, Flammarion, 1997.

[25] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, Boston, 1997.

[26] B. Pesquet-Popescu, Statistical properties of the wavelet decomposition of some non Gaussian self-similar processes, *Signal Processing* 75(3): 1999.

[27] K. Park, W. Willinger, eds, *Self-Similar Network Traffic and Performance Evaluation*, Wiley (Interscience Division), 2000.

[28] R. Riedi. Multifractal processes. 1999. preprint.

[29] R. Riedi, M.S. Crouse, V.J. Ribeiro, R.G. Baraniuk, A Multifractal Wavelet Model with Application to Network Traffic, *IEEE Trans. on Info. Theory*, Special Issue on Multiscale Statistical Signal Analysis and its Applications, 45(3):992–1018, April, 1999.

[30] M. Roughan, D. Veitch, and P. Abry, Real-time estimation of parameters of long-range dependence,
IEEE Trans. on Networking, 8(4), pp.467–478, 2000.

[31] M. Roughan, D. Veitch, J. Yates, M. Ashberg, H. Elgelid, M. Castro, M. Dwyer, P. Abry, Real-Time Measurement of Long-Range Dependence in ATM Networks, Passive and Active Measurement Workshop, Hamilton, New-Zealand, 2000.

[32] G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, New York, London, 1994.

[33] A. Saucier, Méthodes multifractales pour l'analyse d'images et de signaux, in [4].

[34] D. Sornette, *Critical Phenomena in Natural Sciences*. Springer, 2000.

[35] A. Torcini, M. Antoni, Equilibrium and dynamical properties of N-body systems with long-range interactions, Physical Review E, 59(3):2746–2763, 1999.

[36] M. Teich, S. Lowen, B. Jost and K. Vibe-Rheymer, Heart rate Variability: Measures end Models, preprint, 2000.

[37] D. Veitch, P. Abry, A wavelet-based joint estimator of the parameters of long-range dependence, *IEEE Trans. on Info. Theory*, 45(3):878–897, Special Issue on Multiscale Statistical Signal Analysis and its Applications, April 1999.

[38] D. Veitch, P. Abry, P. Flandrin and P. Chainais, Infinitely Divisible Cascade Analysis of Network Traffic Data. Proceedings of *ICASSP 2000*, Istanbul, June 2000.

[39] D. Veitch, P. Abry, A statistical test for the constancy of scaling exponents, To appear in *IEEE Trans. on Sig. proc.*, 2001.

[40] C. Walter, Lois d'échelle en finance, in [4].

[41] B. West, A. Goldberger, Physiology in fractal dimensions, *American Scientist.*, 75:354–365, 1987.

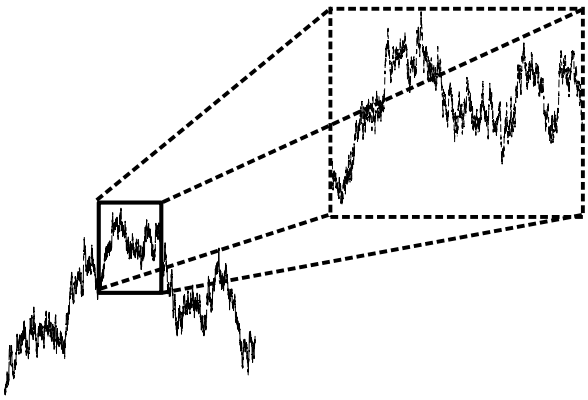


Figure 1: **Sample path of a self-similar process.** Starting with the sample path of a self-similar process, if one performs a dilation of the time axis of factor $1/c$ and a dilation on the amplitude axis of factor c^H, one obtains a new sample path that is (statistically) indistinguishable from the original one.

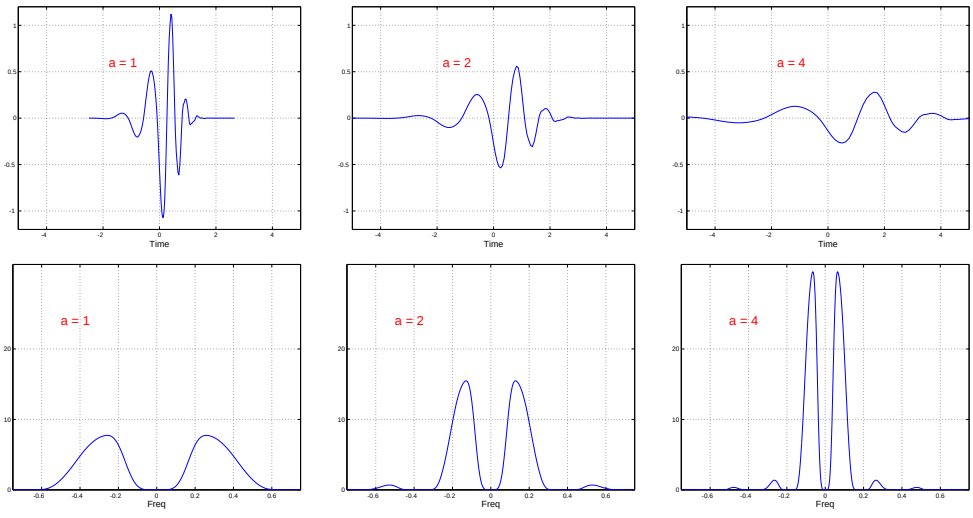


Figure 2: **Translated and dilated wavelets and their corresponding Fourier transforms.** Top, dilated versions of the same mother-wavelet (Daubechies6) with dilation factors 1, 2 and 4, and, bottom, the corresponding Fourier transforms.

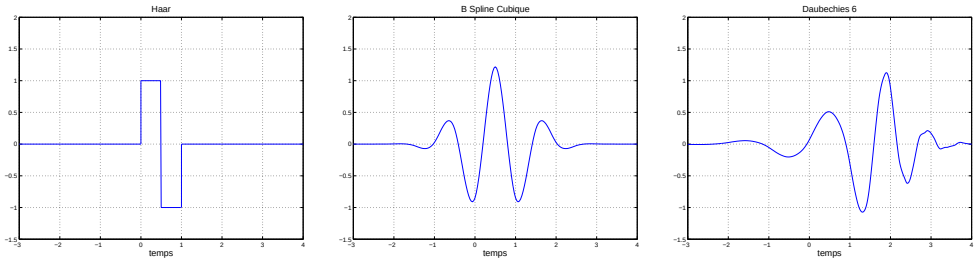


Figure 3: **Vanishing moments.** Examples of mother wavelet with respectively, 1 (Haar, or poor man's, wavelet), 3 (B-spline wavelet) and 6 (Daubechies6) vanishing moments.

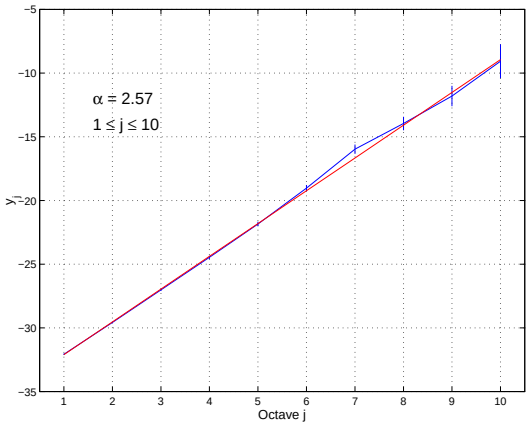


Figure 4: **Logscale Diagram for a fractional Brownian motion.** The fractional Brownian motion is an exact self-similar process, this can be seen on the Logscale diagram through the fact that the linear behaviour of the log of the variance of the wavelet coefficients against the log of the scale holds for all scales. The measurement of the slope enables moreover to precisely estimate the self-similarity parameter.