



Activity Report 2011

## **Project-Team GEOSTAT**

Geometry and Statistics in acquisition data

RESEARCH CENTER  
**Bordeaux - Sud-Ouest**

THEME  
**Optimization, Learning and Statistical Methods**



## Table of contents

<b>1. Members</b>	<b>1</b>
<b>2. Overall Objectives</b>	<b>1</b>
2.1. Overall objectives	1
2.2. Highlights	3
<b>3. Scientific Foundations</b>	<b>3</b>
<b>4. Application Domains</b>	<b>6</b>
<b>5. Software</b>	<b>6</b>
<b>6. New Results</b>	<b>6</b>
6.1. Multiplicative cascades in real/synthetic oceanographic signals: application to the evaluation of ocean dynamics	6
6.2. Endocardial potential analysis for cardiac arrhythmias	7
6.3. Phonetic segmentation	7
6.4. Optimal wavelets, unpredictable points manifold and the emergence of complexity	8
6.5. Discriminative learning for automatic speaker recognition	8
6.6. A multiscale approach to phase reconstruction for Adaptive Optics	9
6.7. Reconstruction of Speech signal from its Unpredictable Points Manifold	9
6.8. New upwelling indices from complex system methods	9
6.9. A detailed analysis of multisensor fusion of moderate resolution imaging spectroradiometer	10
<b>7. Partnerships and Cooperations</b>	<b>10</b>
7.1. European Initiatives	10
7.2. National Initiatives	10
7.3. International Initiatives	11
7.3.1. Internship	11
7.3.2. Participation In International Programs	11
<b>8. Dissemination</b>	<b>11</b>
8.1. Animation of the scientific community	11
8.2. Teaching	12
<b>9. Bibliography</b>	<b>12</b>



# Project-Team GEOSTAT

**Keywords:** Signal Processing, Multiscale Models, Complexity, Speech, Statistical Methods

## 1. Members

### Research Scientists

Hussein Yahia [Team leader, Research Associate (CR) INRIA, HdR]  
Khalid Daoudi [Research Associate (CR) INRIA]

### PhD Students

Vahid Khanagha [CORDIS Grant, since November 2009]  
Suman-Kumar Maji [CORDIS Grant and Aquitaine Region funding, since November 2010]  
Ayoub Tamim [PhD student in cosupervision, PHC Volubilis France-Morocco]  
Reda Jourani [IMAGEEN Grant]

### Post-Doctoral Fellows

Oriol Pont [LEDUCQ fund post doctoral fellowship, University Bordeaux II, since September 2009]  
Harish Kumar [Ministry of Research post-doctoral grant, since October 2010]

### Administrative Assistant

Josy Baron [Secretary (SAR) INRIA]

### Other

Joshua Winebarger [Internship, from Georgia University, USA]

## 2. Overall Objectives

### 2.1. Overall objectives

Glossary

**singularity exponent** A measure of the unpredictability around a point in a complex signal. Based on local reconstruction around a point, singularity exponents can be evaluated in different ways. GEOSTAT focuses on a microcanonical formulation.

**LPE** Local Predictability Exponent: another name for singularity exponents, that better underlines the relation with predictability.

**Framework of reconstructible systems** Complex systems whose acquisitions can be reconstructed from the knowledge of the geometrical sets that maximize statistical information content. Study of complex signals' compact representations associated to unpredictability.

**MMF** Microcanonical Multiscale Formalism.

**Optimal wavelet (OW)**. Wavelets whose associated multiresolution analysis optimizes inference along the scales in complex systems.

GEOSTAT is a research project in **nonlinear digital signal processing**, with the fundamental distinction that it considers the signals as the realizations of complex dynamic systems. The research in GEOSTAT encompasses nonlinear signal processing and the study of emergence in complex systems, with a strong emphasis on geometric approaches to complexity. Consequently, research in GEOSTAT is oriented towards the determination, in real signals, of quantities or phenomena that are known to play an important role both in the evolution of dynamical systems whose acquisitions are the signals under study, and in the compact representations of the signals themselves. Among these parameters, we can mention:

- Singularity exponents, also called Local Predictability Exponents or LPEs,

- comparison with embedding techniques, such as the one provided by the classical theorem of Takens [41], [35].
- Lyapunov exponents, how they are related to intermittency, large deviations and singularity exponents,
- various forms of entropies,
- the cascading properties of associated random variables,
- persistence along the scales, *optimal wavelets*,
- the determination of subsets where statistical information is maximized, their relation to reconstruction and compact representation,

and, above all, **the ways that lead to effective numerical and high precision determination of these quantities in real signals**. The MMF (Multiscale Microcanonical Formalism) is one of the ways to partly unlock this type of analysis, most notably w.r.t. LPEs and reconstructible systems [5]. We presently concentrate our efforts on it, but GEOSTAT is intended to explore other ways [31]. Presently GEOSTAT explores new methods for analyzing and understanding complex signals in different applicative domains through the theoretical advances of the MMF, and the framework of **reconstructible systems** [42]. Derived from ideas in Statistical Physics, the methods developed in GEOSTAT provide new ways to relate and evaluate quantitatively the *local irregularity* in complex signals and systems, the statistical concepts of *information content* and *most informative subset*. That latter notion is developed through the notion of *transition front* and *Most Singular Manifold*. As a result, GEOSTAT is aimed at providing *radically new approaches* to the study of signals acquired from different complex systems (their analysis, their classification, the study of their dynamical properties etc.). The common characteristic of these signals, as required by *universality classes* [38] [39] [36], being the existence of a *multiscale organization* of the systems. For instance, the classical notion of *edge* or *border*, which is of multiscale nature, and whose importance is well known in Computer Vision and Image Processing, receives profound and rigorous new definitions, in relation with the more physical notion of *transition* and fit adequately to the case of chaotic data. The description is analogous to the modeling of states far from equilibrium, that is to say, there is no stationarity assumption. From this formalism we derive methods able to determine geometrically the most informative part in a signal, which also defines its global properties and allows for *compact representation* in the wake of known problematics addressed, for instance, in *time-frequency analysis*. In this way, the MMF allows the reconstruction, at any prescribed quality threshold, of a signal from its most informative subset, and is able to quantitatively evaluate key features in complex signals (unavailable with classical methods in Image or Signal Processing). It appears that the notion of *transition front* in a signal is much more complex than previously expected and, most importantly, related to multiscale notions encountered in the study of nonlinearity [40]. For instance, we give new insights to the computation of dynamical properties in complex signals, in particular in signals for which the classical tools for analyzing dynamics give poor results (such as, for example, correlation methods or optical flow for determining motion in turbulent datasets). The problematics in GEOSTAT can be summarized in the following items:

- the accurate determination in any n-dimensional complex signal of LPEs **at every point in the signal domain** [43][3].
- The geometrical determination and organization of *singular manifolds* associated to various transition fronts in complex signals, the study of their geometrical arrangement, and the relation of that arrangement with statistical properties or other global quantities associated to the signal, e.g. *cascading properties*[6].
- The study of the relationships between the dynamics in the signal and the distributions of LPEs [44][6].
- The study of the relationships between the distributions of LPEs and other formalisms associated to *predictability* in complex signals and systems, such as cascading variables, large deviations and Lyapunov exponents.
- The ability to compute *optimal wavelets* and relate such wavelets to the geometric arrangement of singular manifolds and cascading properties[7].

- The translation of *recognition, analysis* and *classification problems* in complex signals to simpler and more accurate determinations involving new operators acting on singular manifolds using the framework of reconstructible systems.

In GEOSTAT, the development of nonlinear methods for the study of complex systems and signals is conducted on four broad types of complex signals:

- Ocean dynamics and ocean/atmosphere interactions: generation of high-resolution maps from cascading properties and the determination of optimal wavelets[6], geostrophic or non-geostrophic-complex oceanic dynamics, mixing phenomena.
- Heartbeat signals, in cooperation with IHU LIRYC and Professor M. Haissaguerre (INSERM EA 2668 Electrophysiology and Cardiac Stimulation)[19] [38].
- Speech signal (analysis, recognition, classification)[15].
- Optimal wavelets for phase reconstruction in adaptive optics[17].

## 2.2. Highlights

- In the collaboration with the DynBio team of CNRS LEGOS at Toulouse (UMR 5566), and during the HIRESUBCOLOR contract which is ending in 2011, nonlinear methods for the generation of ocean dynamics from remote sensing have been introduced. They consist in evaluating the cascading properties of physical variables associated to turbulent flows through optimal wavelets and propagate information about the dynamics acquired at low resolution up to higher resolution. H. Yahia has been invited to the AGU (American Geophysical Union) Fall meeting in San Francisco to make a presentation of HIRESUBCOLOR results (December 5-9 2011), and to the EGU meeting to be held in Vienna in 2012. These methods are generic, and can be applied to the determination of high resolution information for ocean/atmosphere interaction. For that matter, the Oceanflux proposal has been submitted and accepted, starting November 1, 2011, and a new proposal called MULTICARO will be submitted to CNES-OSTST in 2012.

## 3. Scientific Foundations

### 3.1. Dynamics of complex systems

GEOSTAT is studying complex signals under the point of view of *nonlinear* methods, in the sense of *nonlinear physics* i.e. the methodologies developed to study complex systems. Linear methods in signal processing refer to the standard point of view under which operators are expressed by simple convolutions with impulse responses. Linear methods in signal processing are widely used, from least-square deconvolution methods in adaptive optics to source-filter models in speech processing. Fundamentally, linear signal processing is deeply rooted in *stationarity hypothesis* of the underlying processes, an hypothesis that is questioned in complex systems whose signals are the acquisitions. Linear methods do not unlock the multiscale structures and cascading variables of primary importance as previewed by the physics of the phenomena. This is the reason why new approaches, such as DFA (Detrended Fluctuation Analysis), Time-frequency analysis etc. have appeared during the last decades. Understanding the dynamics of complex systems is recognized as a new discipline, which makes use of theoretical and methodological foundations coming from nonlinear physics, the study of dynamical systems and many aspects of computer science. One of the challenges is related to the question of *emergence* in complex systems: large-scale effects measurable macroscopically from a system made of huge numbers of interactive agents [33], [30], [45], [37]. Some quantities related to nonlinearity, such as Lyapunov exponents, Kolmogorov-Sinai entropy can be computed at least in phase space [31]. Consequently, knowledge from acquisitions of complex systems (which include *complex signals*) could be obtained from information about the phase space. A result from F. Takens [41] about strange attractors in turbulence has motivated the determination of discrete dynamical systems associated to time

series [35], and consequently the theoretical determination of nonlinear characteristics associated to complex acquisitions. Emergence phenomena can also be traced inside complex signals themselves, by trying to localize information content geometrically. Fundamentally, in the nonlinear analysis of complex signals there are broadly two approaches: characterization by attractors (embedding and bifurcation) and time-frequency, multiscale/multiresolution approaches. Time-frequency analysis [32] and multiscale/multiresolution are the subjects of intense research and are profoundly reshaping the analysis of complex signals by nonlinear approaches [29], [34]. In real situations, the phase space associated to the acquisition of a complex phenomenon is unknown. It is however possible to relate, inside the signal's domain, local predictability to local reconstruction and deduce from that Local Predictability Exponents [5] [20]. The LPEs are defined at any point in the signal's domain, they relate, but are different to other kinds of exponents used in the nonlinear analysis of complex signals. We are working on their relation with:

- properties in universality classes,
- the geometric localization of multiscale properties in complex signals,
- cascading characteristics of physical variables,
- optimal wavelets...

The alternative approach taken in GEOSTAT is microscopical, or geometrical: the multiscale structures which have their "fingerprint" in complex signals are being isolated in a single realization of the complex system, i.e. using the data of the signal itself, as opposed to the consideration of grand ensembles or a wide set of realizations. This is much harder than the ergodic approaches, but it is possible because a reconstruction formula such as the one derived in [42] is local and reconstruction in the signal's domain is related to predictability.

Nonlinear signal processing is making use of quantities related to predictability. For instance the first Lyapunov exponent  $\lambda_1$  is related, from Osedelec's theorem, to the limiting behaviour of the response, after a time  $t$ , to perturbation in the phase space  $\log R_\tau(t)$ :

$$\lambda_1 = \lim_{t \rightarrow \infty} \frac{1}{t} \langle \log R_\tau(t) \rangle \quad (1)$$

with  $\langle \cdot \rangle$  being time average. More refined information is provided by the Kolmogorov-Sinai entropy:

$$h_{KS} = \lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{t} \log N_{\text{eff}}(\varepsilon, t) \quad (2)$$

( $N_{\text{eff}}(\varepsilon, t)$  is related to events which appear with very high probability in long time). In GEOSTAT our aim is to relate these classical quantities (among others) to the behaviour of LPEs, which are defined by a limiting behaviour

$$\mu(\mathcal{B}_r(\mathbf{x})) = \alpha(\mathbf{x}) r^{d+h(\mathbf{x})} + o\left(r^{d+h(\mathbf{x})}\right) \quad (r \rightarrow 0) \quad (3)$$

( $d$ : dimension of the signal's domain,  $\mu$ : multiscale measure, typically whose density is the gradient's norm,  $\mathcal{B}_r(\mathbf{x})$ : ball of radius  $r$  centered at  $\mathbf{x}$ ). For precise computation, LPEs can be smoothly interpolated by projecting wavelets:

$$\mathcal{J}_\Psi \mu(\mathbf{x}, r) = \int_{\mathbb{R}^d} d\mu(\mathbf{x}') \frac{1}{r^d} \Psi\left(\frac{\mathbf{x} - \mathbf{x}'}{r}\right) \quad (4)$$

( $\Psi$ : mother wavelet, admissible or not). LPEs are related to the framework of reconstructible systems, and consequently to predictability. They unlock the geometric localization of multiscale structures in a complex signal:



$$\mathcal{F}_h = \{\mathbf{x} \in \Omega \mid h(\mathbf{x}) = h\}, \quad (5)$$

( $\Omega$ : signal's domain) and are consequently in relation with *optimal wavelets*:

$$\mathcal{T}_\psi[\mathbf{s}](\mathbf{x}, \mathbf{r}_1) = \zeta_{r_1/r_2}(\mathbf{x})\mathcal{T}_\psi[\mathbf{s}](\mathbf{x}, \mathbf{r}_2) \quad (6)$$

( $\mathbf{r}_1 < \mathbf{r}_2$ : two scales of observation,  $\zeta$ : injection variable between the scales,  $\psi$ : optimal wavelet) which are related to persistence along the scales and lead to multiresolution analysis whose coefficients verify

$$\alpha_s = \eta_1\alpha_f + \eta_2 \quad (7)$$

with  $\alpha_s$  and  $\alpha_f$  referring to child and parent coefficients,  $\eta_1$  and  $\eta_2$  are random variables independent of  $\alpha_s$  and  $\alpha_f$  and also independent of each other.

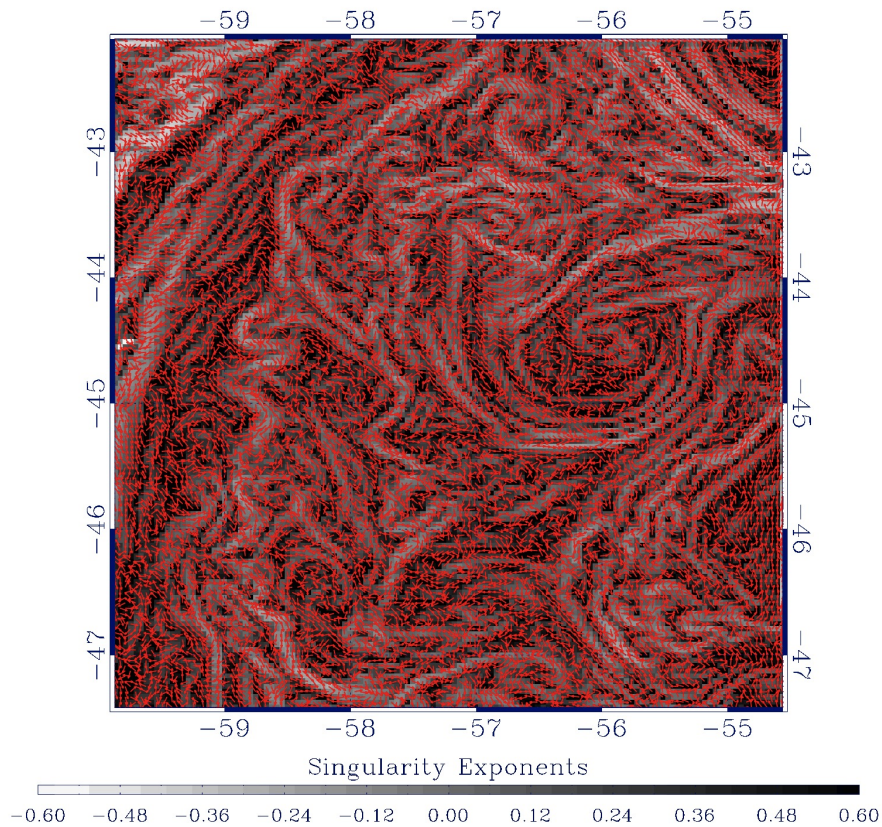


Figure 1. Detail of motion field computed at high spatial resolution (pixel size: 4kms) on Ocean Colour data, in a turbulent area, by propagating along the scales dynamic information obtained from altimetry (spatial resolution of altimetry data: 22 kms) acquired at the same period than the Ocean Colour data.

To take an example, we give some insight about the collaboration with LEGOS Dynbio team about high-resolution ocean dynamics from microcanonical formulations in nonlinear complex signal analysis. LPEs relate to the geometric structures linked with the cascading properties of indefinitely divisible variables in turbulent flows. Cascading properties can be represented by optimal wavelets (OWs); this opens new and fascinating directions of research for the determination of ocean motion field at high spatial resolution. OWs in a microcanonical sense pave the way for the determination of the energy injection mechanisms between the scales. From this results a new method for the complete evaluation of oceanic motion field is introduced; it consists in propagating along the scales the norm and the orientation of ocean dynamics deduced at low spatial resolution (geostrophic from altimetry and a part of ageostrophic from wind stress products). Using this approach, there is no need to use several temporal occurrences as in Optical Flow, Maximum Cross Correlation or FSLE techniques. Instead, the proper determination of the turbulent cascading and energy injection mechanisms in oceanographic signals allows the determination of oceanic motion field at the SST or Ocean colour spatial resolution (pixel size: 4 kms) which often go beyond the results obtained with other models (e.g. SQG models). We use the Regional Ocean Modelling System (ROMS) to validate the results on simulated data and compare the motion fields obtained with other techniques. See figure 1.

## 4. Application Domains

### 4.1. Application Domains

The following application domains are investigated by the GEOSTAT team:

- Complex signal acquired from remote sensing and earth observation, notably in Oceanography, and ocean/climate interaction.
- Speech signal.
- Astronomical imaging and turbulence (in collaboration with LATT and ONERA).
- Analysis of heartbeat signals (in collaboration with M. Haissaguerre, head of team INSERM EA3668 *Electrophysiologie et Stimulation Cardiaque*, Bordeaux University, R. Dubois (ESPCI) and the CARMEN team).

## 5. Software

### 5.1. Software

**Participants:** Hussein Yahia [correspondant], Antonio Turiel, Joel Sudre.

FluidExponents: software implementation of the MMF, written in Java, in a cooperative development mode on the INRIA GForge, deposited at APP in 2010. Contact: hussein.yahia@inria.fr. FluidExponents implements nonlinear signal processing on various types of input data (including NETCDF).

- Version: 0.8

## 6. New Results

### 6.1. Multiplicative cascades in real/synthetic oceanographic signals: application to the evaluation of ocean dynamics

**Participants:** Hussein Yahia, Oriol Pont, Joel Sudre, Véronique Garçon, Claire Pottier [CNES], Antonio Turiel, Christine Provost.

This work is performed during the final year of the HIRESUBCOLOR contract with CNES/NASA. From a fundamental point of view, significant advances have been worked out in the application of complex systems methods for the derivation of new methods for computing ocean dynamics at high spatial resolution (high resolution Sea Surface Temperature, pixel size: 4 x 4 kms). No temporal information is used. Instead, the norms and orientations of low resolution vector fields derived from altimetry and scatterometers are propagated along the scales of turbulent signals. This year, specific study on the propagation of the norm of the vector field has been conducted, resulting in the complete mapping (norm and orientation) of ocean dynamics at the high resolution of Sea Surface Temperature data. Validation is performed by comparison with the output of the ROMS 3D simulation model, with excellent results, and buoy validation is under way. H. Yahia has been invited to the AGU (American Geophysical Union) Fall meeting in San Francisco to make a presentation of HIRESUBCOLOR results (December 5-9 2011) and also to the EGU meeting to be held in Vienna in 2012. In 2011 the complete method for the determination of ocean dynamics has been finalized and it includes:

- the determination of both norm and orientation of the vector fields,
- the propagation along the scale of both geostrophic and ageostrophic dynamics.

These methods are generic, and can be applied to the determination of high resolution information for ocean/atmosphere interaction. For that matter, the Oceanflux proposal has been submitted and accepted, starting November 1, 2011, and a new proposal called MULTICARO will be submitted to CNES-OSTST. Related publications: [28], [21].

## 6.2. Endocardial potential analysis for cardiac arrhythmias

**Participants:** Oriol Pont, Hussein Yahia, Harish Kumar Goddabanahalli, Michel Haissaguerre, Nicolas Derval, Mélèze Hocini.

Cardiac diseases are the main cause of morbidity and mortality in western countries. Both the pathogenic areas and the evolution of the condition are complex to detect and estimate in the case of arrhythmias, specially in atrial and ventricular fibrillation. In these cases, the dynamics of the cardiac potential behaves chaotically in a highly complex way that challenges its description. Under this context, we are working in the direction of the characterization of the heartbeat dynamics in a model-agnostic way. We have performed analysis of heartbeat dynamics through singularity spectrum, empirical analysis of LPEs for the deart-beat Dynamics, and a novel innovative method to distinguish arrhythmic heartbeat with rhythmic heartbeat has been studied.

Intracardial potential is measured by means of electrode catheters used during the radiofrequency ablation procedure for cases of atrial fibrillation. We have found that in the electric potential signal, a simple fast-changing three-state orientation signal can be sifted from its complex but slow-changing modulation. The fast dynamics experimentally fits a Markovian process which can be described in a simple and compact way and its parameters are robustly estimated. This shows a clear change of signature even with small statistics that can be used to detect transitions in the arrhythmia and identify different regimes in them.

Related publications: [18], [23], [19], [24].

## 6.3. Phonetic segmentation

**Participants:** Vahid Khanagha, Joshua Winebarger, Khalid Daoudi, Oriol Pont, Hussein Yahia, Régine André-Obrecht.

Previously we had developed a novel phonetic segmentation method based on Microcanonical Multiscale Formalism (MMF). The algorithm was based on precise computation of Local Predictability Exponents (LPEs) at each point, and then using their integration over time axis (ACC) as a quantitative representative of changes in behavior of distribution of these exponents between neighboring phonemes. The piecewise linear estimation of ACC had provided very good segmentation precision. By performing error analysis of the original algorithm, we proposed a 2-step technique which better exploits LPEs to improve the segmentation accuracy. In the first step, we detect the boundaries of the original signal and of a low-pass filtered version, and we consider the union of all detected boundaries as candidates. In the second step, we use a hypothesis

test over the local LPE distribution of the original signal to select the final boundaries. In summary following steps have been taken:

- Detailed error analysis of the original method, which resulted in the realization of the fact that a high-pass filtering can help to detect some of the missed boundaries.
- Development of the hypothesis test method, using the Log Likelihood Ratio Test for final decision over a list of candidates.
- Evaluation of the overall 2-step algorithm on the whole *train* part of the TIMIT database, to compare with the original method.
- Evaluation on *test* part of TIMIT database to compare with the state of the art methods.

Related publications: [13], [14].

We continued and improved the adaptation of speaker segmentation methods to develop new (nonlinear) techniques for phonetic segmentation. We succeeded in proposing simple and efficient new algorithms that outperform existing ones. Even with new approaches, our nonlinear approach was still competitive.

Related publications: [22], [27].

## 6.4. Optimal wavelets, unpredictable points manifold and the emergence of complexity

**Participants:** Oriol Pont, Hussein Yahia, Suman Maji.

We have found new theoretical developments that link the optimal wavelet description with the information transfer that characterizes the singularity exponents in complex signals. This fact is particularly relevant when there exists a microcanonical cascade as an effective dynamics for the underlying complex system. The implication of this is that under a multiscale hierarchy, the unpredictable set of a signal can be described in terms of its optimal wavelet coefficients and the multiplicative cascade relations between them, easing the information inference or reconstructability between resolution levels.

A new method for the detection of the unpredictable points manifold has been developed, enhancing previous implementations. The algorithm exploits the basic signal symmetries that can be easily verified and has the advantage of not assuming any underlying model. That work is the result of the collaboration between our team and A. Turiel's team at Institute of Marine Sciences of Barcelona.

Additionally, we have developed a new algorithm that allows for the first time a very robust detection of the optimal wavelet in 2D signals. The main advantage of this new algorithm is that it optimizes the wavelet shape in a totally unconstrained way, therefore not restricting to specific wavelet families.

Related Publications: [7], [20].

## 6.5. Discriminative learning for automatic speaker recognition

**Participants:** Khalid Daoudi, Reda Jourani, Régine André-Obrecht.

Most of the speaker recognition systems rely on generative learning of Gaussian Mixture Models (GMM). During the last decade, discriminative approaches have been an interesting and valuable alternative to address directly the classification problem. For instance, Support Vector Machines (SVM) combined with GMM supervectors are among state-of-the-art approaches in speaker recognition. Recently a new discriminative approach for multiway classification has been proposed, the Large Margin Gaussian mixture models (LM-GMM). These latter methods have the same advantage as SVM in term of the convexity of the optimization problem to solve. However they differ from SVM because they draw nonlinear class boundaries directly in the input space, and thus no kernel trick is required. We continued our work on investigating simplified versions of LM-GMM for speaker recognition that can handle large scale databases. We developed a new and efficient learning algorithm and evaluated it on NIST-SRE data. The results show that this new algorithm not only outperforms both the original LM-GMM and the traditional GMM, but also outperforms state-of-the-art discriminative methods such as GMM-supervectors SVM.

Related Publications: [9], [10], [11], [12], [8].

## 6.6. A multiscale approach to phase reconstruction for Adaptive Optics

**Participants:** Suman Kumar Maji, Hussein Yahia, Oriol Pont, Thierry Fusco, Vincent Michau, Joel Sudre.

Atmospheric turbulence in Earth's atmosphere upper layers plays a fundamental role in limiting the resolution of ground based instruments. These turbulent layers perturbate to a great extent incoming light from outer space. One of the best known solutions to overcome this hurdle is Adaptive Optics (AO). It provides real-time compensation by deforming a mirror through a servo-loop, according to phase measurements provided by a wavefront sensor (WFS). We propose and experiment with a new model for phase reconstruction from an acquired subimage of the perturbed phase: instead of reconstructing the phase gradient using conventional methods of AO, we propagate along the scales phase information, from the low resolution of the WFS to higher resolution, using specific wavelet projections that mimic inference along the scales associated to cascading properties of fully developed turbulence.

Related Publication: [17].

## 6.7. Reconstruction of Speech signal from its Unpredictable Points Manifold

**Participants:** Vahid Khanagha, Khalid Daoudi, Oriol Pont, Hussein Yahia.

Local Predictability Exponents (LPEs) can be used to classify a given signal's samples according to their predictability. In particular, the Unpredictable Points Manifold, the subset of less predictable points can be formed as the ensemble of points having the least value of singularity exponents. We call these exponents the Local Predictability Exponents since they are computed according to a procedure based on the evaluation of the degree of reconstruction at a given point. We demonstrate in the case of Speech signal that LPEs are key quantities related to predictability in the framework of reconstructible systems: it is possible to reconstruct the whole Speech signal by applying a reconstruction kernel to the UPM. This provides a strong indication of the importance of the UPM, already demonstrated for other types of complex signals. Experiments show that a UPM containing a small number of the points provides very good perceptual reconstruction quality. In summary following steps have been taken:

- Using the LPEs to form the UPM for Speech signal and coping with the implementation issues in particular case of Speech signal.
- Successful reconstruction of Speech signal from the UPM. The performance was measured using objective measures of reconstruction quality.
- Detailed study of geometrical implications of the points in UPM, and proposition of a new multiscale measure, to be used in estimation procedure of exponents, which is more appropriate for speech analysis. In fact, the same quality of reconstruction is achieved with a quite smaller UPM.
- Development of a very simple compression algorithm (8-bit differential nonuniform quantizer) which overperforms the traditional DPCM coding method.

Related Publication: [15].

## 6.8. New upwelling indices from complex system methods

**Participants:** Hussein Yahia, Ayoub Tamim, Khalid Minaoui, Driss Aboutajdine, Véronique Garçon, Joel Sudre.

We started the Volubilis project on the study of upwelling in the Moroccan coast by satellite imaging. A Ph.D. student, Ayoub Tamim, was recruited in January 2011. We organized the kick-off meeting of the project in Rabat from June 6 to June 10 with the presence of all partners (INRIA, LEGOS, FSR and CRTS). A. Tamim implemented a new version of the software CRTS is using to compute an upwelling index. He then joined GEOSTAT (from September 5 to December 15, 2011) where he first constituted a small validation database of SST and Chlorophyll images. He then implemented some algorithms such as Morand index and multiscale entropy. The goal being to define in a first step a coarse segmentation of upwelling regions.



## 6.9. A detailed analysis of multisensor fusion of moderate resolution imaging spectroradiometer

**Participants:** Harish Kumar Goddabanahalli, Dharmendra Singh, Hussein Yahia.

Related publication: [16].

# 7. Partnerships and Cooperations

## 7.1. European Initiatives

### 7.1.1. Collaborations in European Programs, except FP7

**Program: Support to Science Element, ESRIN/AO/1-6668/11/I-AM, Fund: E/0029-01-L**

Project acronym: Oceanflux

Project title: Oceanflux

Duration: November 2011 - May 2013

Coordinator: Dr. Christoph S. Garbe, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg.

Other partners: INRIA, GEOSTAT (FRANCE); KIT, IKM-ASF (Allemagne); CNRS, LEGOS (France); IRD (France); Université Paul Sabatier, (France).

Abstract: Mapping at high spatial resolution of GHGs exchange flux between ocean and atmosphere using model outputs and nonlinear techniques in signal processing.

**Program: PHC**

Project acronym: Volubilis

Project title: Study of upwelling in the Moroccan coast by satellite imaging

Duration: November 2010 - October 2013

Coordinator: K. Daoudi, INRIA, GEOSTAT (France)

Other partners: Rabat University, CRTS.

Abstract: Multiscale methods for the characterization of coastal upwelling from remote sensing data.

**Program: OSTST CNES-NASA**

Project acronym: Hiresubcolor

Project title: Multiscale methods for the evaluation of high resolution ocean surface velocities and subsurface dynamics from ocean color, sst and altimetry

Duration: November 2008 - December 2011

Coordinator: H. Yahia, INRIA, GEOSTAT (France)

Other partners: CNRS LEGOS (France), ICM-CSIC (Spain), LOCEAN (France)

Abstract: nonlinear signal processing methods for high resolution mapping of ocean dynamics.

## 7.2. National Initiatives

- Region Aquitaine research call. Funding of the OPTAD project on adaptive optics.
- Region Aquitaine research call. Funding for equipment in Speech databases and software.
- GEOSTAT is funded by *Conseil Région Aquitaine* for acquiring Speech databases and software.

- Marie-Curie post doctoral grant with *Conseil Région Aquitaine* on heartbeat dynamics.
- ARC FIBAUR with INRIA, ESPCI and (INSERM EA 2668 Electrophysiology and Cardiac Stimulation).

## 7.3. International Initiatives

### 7.3.1. Internship

Joshua Winebarger (from Apr 2011 until Sep 2011)

Subject: Fusion of different speech segmentation algorithms

Institution: Georgia Tech. (United States)

### 7.3.2. Participation In International Programs

Program: Canadian CRSNG

Project acronym: Profilage à partir des données hétérogènes du Web pour la cybersécurité

Project title: Profilage à partir des données hétérogènes du Web pour la cybersécurité

Duration: March 2011 - February 2014

Coordinator: Concordia University

Other partners: University of Sherbrooke, E-Profile Company, S. d. Quebec, GEOSTAT (INRIA)

Abstract: Use of various complex signals for cybersecurity.

## 8. Dissemination

### 8.1. Animation of the scientific community

- H. Yahia is now a member of Elsevier's Digital Signal Processing journal editorial board.
- H. Yahia has been an invited speaker at AGU (American Geophysical Union) Fall Meeting (San Francisco, December 5-9 2011). Section: Nonlinear Geophysics. Title of presentation: *High-resolution ocean dynamics from microcanonical formulations in nonlinear complex signal analysis*.
- H. Yahia has been an invited speaker for the Assyst *Workshop on Mathematics in Network Science: Implications to Socially Coupled Systems*, organized by ISI foundation at Torino (Italy) on November 23-21 2011. Title: *Local Predictability Exponents and universality classes in the framework of reconstructible complex systems*<sup>1</sup>.
- H. Yahia has joined an official visit of INRIA CEO Michel Cosnard in India in September 2011, organized by INRIA Direction of International Relations. The mission in India consisted first in accompanying the Direction of International Relations for official visits to: Chennai Mathematical Institute, Institute of Mathematical Sciences (Chennai), International Institute of Information Technology (Hyderabad), University of Hyderabad (where a plaque has been inaugurated in the presence of M. Cosnard in the Institute of Sanskrit Studies about the collaboration with G. Huet), IIT Mumbai, Tata Institute of Fundamental Research (Mumbai), Indian Institute of Science (Bangaluru), and second, to participate to the Indo-French seminar *Modelling and analysis of complex systems for communication, health and sustainable development* organized by INRIA and CEFIPRA in Bangaluru on September 26-28, 2011. Titles of presentations: *Multiscale methods in the analysis of heartbeat data. Application to Atrial Fibrillation* and *nonlinear signal processing for Earth Observation. A case study for a physical approach to complex signals*.

<sup>1</sup>See <http://www.isi.it/events/assyst-workshop>.

- H. Yahia has given a presentation at CRTS and FSR, Rabat, during the first PHC Volubilis meeting, June 7, 2011. Title: : *Analysis of complex signals: new trends in nonlinear approaches to complexity, and their application to specific classes of signals.*
- H. Yahia has given a presentation during a workshop organized by MeteoFrance on October 18, at the MeteoFrance centre in Toulouse. Title: *Evaluation de la dynamique océanique à haute résolution satellitaire SST par fusion de données à différentes échelles spatiales* (joint work with LEGOS partners J. Sudre, V. Garçon, and CNES partner C. Pottier).
- H. Yahia has been selected to give a presentation of the GEOSTAT research thematic during info days at the European Commission in Brussels, October 12, 2011.
- O. Pont gave a lecture at a meeting on CARDIOSENSE action, with ANUBIS team, GEOSTAT team, partners involved in CARDIOSENSE, and Prof. M Haissaguerre team. Title: *Analyse non-linéaire multi-échelle des signaux cardiaques.* January 7, 2011, at Bordeaux University.
- H. Yahia has given a presentation at CEA on November 15, 2011. Title: *Méthodes non-linéaires en analyse des signaux complexes: évaluation de la dynamique océanique à haute résolution spatiale.*
- O. Pont has given a presentation at CEA on November 15, 2011. Title: *Nonlinear analysis of heartbeat dynamics.*
- H. Yahia has presented the results of HIRESUBCOLOR contract and the presentation for a MULTICARO proposal at CNES DSP/LEGOS 2011 meeting on November 18 in Toulouse. Title: *Méthodes non-linéaires en analyse de la dynamique océanique et interaction océan/atmosphère: résultats de Hiresubcolor, perspectives de Multicar et Oceanflux*
- O. Pont gave a lecture at INRIA Bordeaux Sud-Ouest *Unithé ou Café* on January 21, 2011. Title: *L'ordre complexe qui émerge du chaos.*

## 8.2. Teaching

M2 : K. Daoudi, 10 hours, Speech processing, Master2 InfoTelecom, University of Rabat, Morocco.

PhD in progress : V. Khanagha, Nonlinear methods for speech processing, beginning November 2009, supervisors: K. Daoudi, O. Pont, H. Yahia,

PhD in progress : S. Kumar Maji, Optimal wavelets for adaptive optics, beginning November 2010, supervisors: O. Pont, H. Yahia,

PhD in progress : A. Tamim, Coastal Upwelling, beginning October 2010, supervisors: H. Yahia, K. Daoudi, K. Minaoui, D. Aboutajdine.

## 9. Bibliography

### Major publications by the team in recent years

- [1] K. DAOUDI. *Iterated function systems and some generalizations: local regularity analysis and multifractal modeling of signals*, in "Chapter in the book: Scaling, Fractals and Wavelets", 2009.
- [2] V. KHANAGHA, H. YAHIA, K. DAOUDI, O. PONT, A. TURIEL. *Reconstruction of Speech Signals from their Unpredictable Points Manifold*, in "NOLISP 2011, published in Advanced in Nonlinear Speech Processing, LNAI 7015, Springer Verlag", Las Palmas de Gran Canaria, November 2011, <http://hal.inria.fr/hal-00647197/fr>.
- [3] O. PONT, A. TURIEL, H. YAHIA. *An optimized algorithm for the evaluation of local singularity exponents in digital signals*, in "Combinatorial Image Analysis, Lectures Notes in Computer Science 6636", 2010, vol. 43, p. 346-357, <http://dx.doi.org/10.1007/978-3-642-21073-0>, <http://hal.inria.fr/inria-00581057/fr>.



- [4] E. SANCHEZ-SOTO, A. POTAMIANOS, K. DAOUDI. *Unsupervised Stream-Weights Computation in Classification and Recognition Tasks*, in "IEEE Trans. on Audio, Speech and Language Processing", 2009, vol. 17, n<sup>o</sup> 3, p. 436–445.
- [5] A. TURIEL, H. YAHIA, J. C. PEREZ-VICENTE. *Microcanonical multifractal formalism: a geometrical approach to multifractal systems. Part I: singularity analysis*, in "Journal of Physics A: Math. Theor", 2008, vol. 41, <http://dx.doi.org/10.1088/1751-8113/41/1/015501>.
- [6] H. YAHIA, J. SUDRE, C. POTTIER, V. GARÇON. *Motion analysis in oceanographic satellite images using multiscale methods and the energy cascade*, in "Pattern Recognition", 2010, vol. 43, p. 3591-3604, <http://dx.doi.org/10.1016/j.patcog.2010.04.011>, <http://hal.inria.fr/inria-00423748>.

## Publications of the year

### Articles in International Peer-Reviewed Journal

- [7] O. PONT, A. TURIEL, J. C. PEREZ-VICENTE. *On optimal wavelet bases for the realization of microcanonical cascade processes*, in "International Journal of Wavelets Multiresolution and Information Processing", 2011, vol. 9, n<sup>o</sup> 1, p. 35-61, International Journal of Wavelets, Multiresolution and Information Processing [DOI : 10.1142/S0219691311003943], <http://hal.inria.fr/inria-00582372/en>.

### International Conferences with Proceedings

- [8] M. AL MASHRGY, N. BOUGILA, K. DAOUDI. *A Robust Approach for Multivariate Binary Vectors Clustering and Feature Selection*, in "ICONIP 2011", Shangai, China, November 2011, volume 7063 of Lecture Notes in Computer Science, <http://hal.inria.fr/hal-00647990/fr>.
- [9] K. DAOUDI, R. JOURANI, R. ANDRÉ-OBRECHT, D. ABOUTAJDINE. *Speaker Identification Using Discriminative Learning of Large Margin GMM*, in "ICONIP 2011", Shangai, China, November 2011, volume 7063 of Lecture Notes in Computer Science, <http://hal.inria.fr/hal-00647990/fr>.
- [10] R. JOURANI, K. DAOUDI, R. ANDRÉ-OBRECHT, D. ABOUTAJDINE. *Apprentissage discriminant des GMM à grande marge pour la vérification automatique du locuteur*, in "XXIIIe colloque GRETSI", Bordeaux, France, September 2011, <http://hal.inria.fr/hal-00647978/fr>.
- [11] R. JOURANI, K. DAOUDI, R. ANDRÉ-OBRECHT, D. ABOUTAJDINE. *Fast training of Large Margin diagonal Gaussian mixture models for speaker identification*, in "6th Conference on Speech Technology and Human-Computer Dialogue (SpeD)", 2011, <http://hal.inria.fr/hal-00647213/fr>.
- [12] R. JOURANI, K. DAOUDI, R. ANDRÉ-OBRECHT, D. ABOUTAJDINE. *Speaker verification using Large Margin GMM discriminative training*, in "International Conference on Multimedia Computing and Systems (ICMCS)", 2011, <http://hal.inria.fr/hal-00647232/fr>.
- [13] V. KHANAGHA, K. DAOUDI, O. PONT, H. YAHIA. *Improving text-independent phonetic segmentation based on the Microcanonical Multiscale Formalism*, in "ICASSP 2011", Prague, Czech Republic, May 2011, <http://hal.inria.fr/inria-00557661/en>.
- [14] V. KHANAGHA, O. PONT, K. DAOUDI, H. YAHIA. *Une nouvelle approche non-linéaire pour la segmentation phonétique*, in "XXIIIe colloque GRETSI", Bordeaux, France, September 2011, <http://hal.inria.fr/hal-00647585/fr>.

- [15] V. KHANAGHA, H. YAHIA, K. DAOUDI, O. PONT, A. TURIEL. *Reconstruction of Speech Signals from their Unpredictable Points Manifold*, in "NOLISP 2011, published in Advanced in Nonlinear Speech Processing, LNAI 7015, Springer Verlag", Las Palmas de Gran Canaria, November 2011, <http://hal.inria.fr/hal-00647197/fr>.
- [16] H. KUMAR GODDABANAHALLI, D. SINGH, H. YAHIA. *A detailed analysis of multi-sensor fusion of moderate resolution imaging spectroradiometer*, in "19th European Signal Processing Conference (EUSIPCO 2011)", 2011, <http://hal.inria.fr/hal-00649466/fr>.
- [17] H. MAJI, O. PONT, T. FUSCO, V. MICHAU, J. SUDRE. *A multiscale approach to phase reconstruction for Adaptive Optics*, in "IEEE Workshop on Electronics, Control, Measurement and Signals ECMS", Liberec, Czech Republic, Technical university of Liberec, 2011, <http://hal.inria.fr/inria-00591097/en>.
- [18] O. PONT, M. HAÏSSAGUERRE, H. YAHIA, N. DERVAL, M. HOCINI. *A novel analysis method to characterize heartbeat dynamics through the Microcanonical Multiscale Formalism*, in "6th International Conference on Mass Data Analysis of Image and Signals in Medicine, Biotechnology, Chemistry and Food Industry MDA", New York, United States, Lecture Notes in Computer Science (LNCS), Springer, 2011, Full paper and oral talk Accepted, <http://hal.inria.fr/inria-00590696/en>.
- [19] O. PONT, M. HAÏSSAGUERRE, H. YAHIA, N. DERVAL, M. HOCINI. *Heartbeat Dynamics from a Microcanonical Multifractal Approach*, in "Computing in Cardiology", Hangzhou, China, September 2011, <http://hal.inria.fr/hal-00648301/en>.
- [20] O. PONT, A. TURIEL, H. YAHIA. *An optimized algorithm for the evaluation of local singularity exponents in digital signals*, in "14th International Workshop, IWCI", Madrid, Spain, J. AGGARWAL, R. BARNEVA, V. BRIMKOV, K. KOROUTCHEV, E. KORUTCHEVA (editors), Lecture Notes in Computer Science (LNCS), Springer, 2011, vol. 6636, p. 346-357, <http://hal.inria.fr/inria-00581057/en>.
- [21] J. SUDRE, H. YAHIA, O. PONT, C. POTTIER, V. GARÇON. *Inferring fluid motion at submesoscale from high resolution remotely sensed data*, in "OSTST 2011", San Diego, 2011, <http://depts.washington.edu/uwconf/ostst2011/>.
- [22] J. WINEBARGER, K. DAOUDI, H. YAHIA. *Improving SVF with DISTBIC for Phoneme Segmentation*, in "International Conference Speech and Computer (SPECOM) 2011", Kazan, Russian Federation, September 2011, <http://hal.inria.fr/hal-00647985/fr>.

### National Conferences with Proceeding

- [23] O. PONT, M. HAÏSSAGUERRE, H. YAHIA, N. DERVAL, M. HOCINI. *A microcanonical multifractal approach to the characterization of heartbeat dynamics*, in "FisEs 2011", Barcelona, Spain, Royal Spanish Society of Physics RSEF, 2011, <http://hal.inria.fr/inria-00590976/en>.

### Other Publications

- [24] O. PONT, M. HAÏSSAGUERRE, H. YAHIA, N. DERVAL, M. HOCINI. *Microcanonical processing of ECG and intracardial potential in atrial fibrillation*, in "Transactions on Mass-Data Analysis of Images and Signals", 2012, Invited paper, Special Edition, To appear.
- [25] O. PONT, V. NIEVES, A. TURIEL. *Description of ocean dynamics with optimal wavelet cascade processes*, 2011, Journal of Geophysical Research (JGR). To be submitted, <http://hal.inria.fr/inria-00438524/en>.

- [26] O. PONT, A. TURIEL, J. C. PEREZ-VICENTE. *Volatility forecasting methods with optimal wavelets*, 2011, Physica A. In preparation, <http://hal.inria.fr/inria-00438525/en>.
- [27] J. WINEBARGER. *Phoneme segmentation and Voice activity detection*, in "Internship report", 2011, <http://hal.inria.fr/hal-00647986/fr>.
- [28] H. YAHIA, J. SUDRE, V. GARÇON, C. POTTIER. *High-resolution ocean dynamics from microcanonical formulations in non linear complex signal analysis*, in "AGU Fall Meeting", San Francisco, USA, 2011, Invited session, <http://hal.inria.fr/hal-00650224/en>.

## References in notes

- [29] A. ARNEODO, F. ARGOUL, E. BACRY, J. ELEZGARAY, J. F. MUZY. *Ondelettes, multifractales et turbulence*, Diderot Editeur, Paris, France, 1995.
- [30] N. BOCCARA. *Modeling Complex Systems*, Springer, New-York Dordrecht Heidelberg London, 2010.
- [31] G. BOFFETTA, M. CENCINI, M. FALCIONI, A. VULPIANI. *Predictability: a way to characterize complexity*, in "Physics Report", 2002, vol. 356, p. 367–474, arXiv:nlin/0101029v1, [http://dx.doi.org/10.1016/S0370-1573\(01\)00025-4](http://dx.doi.org/10.1016/S0370-1573(01)00025-4).
- [32] P. FLANDRIN. *Time-frequency/Time-scale analysis*, Academic Press, San Diego, USA, 1999.
- [33] M. GELL-MANN. *What is Complexity ?*, John Wiley and Sons, Inc., 1995, vol. 1, 1, <http://complexity.martinsewell.com/Gell95.pdf>.
- [34] C. HENEGHAN, G. MCDARBY. *Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes*, in "Phys. Rev. E", 2000, vol. 62, n<sup>o</sup> 5, p. 6103-6110, <http://link.aps.org/doi/10.1103/PhysRevE.62.6103>.
- [35] K. JUDD, A. MEES. *Embedding as a modelling problem*, in "Non Linear Optimization", 1998, vol. 20, n<sup>o</sup> 3-4, p. 273-286, <http://www.sciencedirect.com/science/article/pii/S016727899800089X>.
- [36] A. LESNE. *Méthodes de renormalisation. Phénomènes critiques, chaos, structures fractales*, Eyrolles, Paris, 1996, ISBN 2-212-05830-6.
- [37] J. NADAL, P. E. GRASSBERGER. *From Statistical Physics to Statistical Inference and Back*, Springer, New York Heidelberg Berlin, 1994, <http://www.springer.com/physics/complexity/book/978-0-7923-2775-2>.
- [38] O. PONT. *Universality in multiscale, turbulent-like self-organization: from algal blooms to heartbeat to stock market dynamics*, in "IFISC, Emergence and Design of Robustness - General principles and applications to biological, social and industrial networks", Palma de Mallorca, September 2010, <http://hal.inria.fr/inria-00533395/en>.
- [39] O. PONT, A. TURIEL, J. C. PEREZ-VICENTE. *Empirical evidences of a common multifractal signature in economic, biological and physical systems*, in "Physica A, Elsevier", 2009, n<sup>o</sup> 388, p. 2025–2035, <http://hal.inria.fr/inria-00438521>.

- [40] H. STANLEY. *Introduction to phase transitions and critical phenomena*, Oxford Science publications, Oxford, UK, 1987.
- [41] F. TAKENS. *Detecting Strange Attractors in Turbulence*, in "Non Linear Optimization", 1981, vol. 898, p. 366–381, <http://www.springerlink.com/content/b254x77553874745/>.
- [42] A. TURIEL, A. DEL POZO. *Reconstructing images from their most singular fractal manifold*, in "IEEE Trans. on Im. Proc.", 2002, vol. 11, p. 345–350.
- [43] A. TURIEL, J. C. PEREZ-VICENTE, J. GRAZZINI. *Numerical methods for the estimation of multifractal singularity spectra on sampled data: a comparative study*, in "Journal of Computational Physics", 2006, vol. 216, n<sup>o</sup> 1, p. 362–390.
- [44] A. TURIEL, J. C. PEREZ-VICENTE. *Role of multifractal sources in the analysis of stock market time series*, in "Physica A", September 2005, vol. 355, p. 475–496.
- [45] G. WEISBUCH. *Complex Systems Dynamics*, Addison Wesley Longman, 1991.