

Activity Report 2017

Project-Team GEOSTAT

Geometry and Statistics in acquisition data

RESEARCH CENTER **Bordeaux - Sud-Ouest**

THEME

Optimization, machine learning and statistical methods

Table of contents

1.	Personnel					
2.	Overall Objectives					
3.	Research Program3					
	3.1. General methodology	3				
	3.2. Excitable systems: analysis of physiological time series	5				
	3.2.1. Presentation and objectives	5				
	3.2.2. Results	5				
	3.2.3. Interpretation	5				
	3.3. Multiscale description in terms of multiplicative cascade	6				
	3.4. Data-based identification of characteristic scales and automated modeling	10				
	3.5. Speech analysis	10				
4.	Application Domains	10				
	4.1. Analysis of galactic molecular clouds (GENESIS project)	11				
	4.2. Ocean dynamics, upwelling	11				
	4.3. Non convex optimization for image processing	11				
	4.4. Physiological times series: speech processing	11				
5.	Highlights of the Year	11				
6.	New Software and Platforms	11				
	6.1. Fluex	11				
	6.2. FluidExponents	12				
	6.3. classifemo	12				
	6.4. superres	12				
	6.5. EdgeReconstruct	13				
7.	· ·	13				
	7.1. Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation	13				
	7.2. Super-resolution	13				
	7.3. Surface mixing and biological activity in the Northwest African upwelling	13				
	7.4. Spatio-Temporal Dynamics of Floods	14				
	7.5. Effect of wind stress forcing on ocean dynamics at Air-Sea Interface	14				
	7.6. Ocean dynamics: frontal activity	14				
	7.7. Pathological speech processing	15				
8.		15				
	8.1. Bilateral Contracts with Industry	15				
	8.2. Bilateral Grants with Industry	15				
9.	·	15				
	9.1. Regional Initiatives	15				
	9.2. National Initiatives	16				
	9.3. European Initiatives	16				
	9.4. International Initiatives	17				
	9.4.1. Inria International Partners	17				
	9.4.2. Participation in Other International Programs	17				
	9.5. International Research Visitors	17				
	9.5.1. Visits of International Scientists	17				
	9.5.2. Visits to International Teams	17				
10.		18				
100	10.1. Promoting Scientific Activities	18				
	10.1.1. Scientific Events Organisation	18				
	10.1.2. Journal	18				
	10.1.2.1. Member of the Editorial Boards	18				

10.1.2.2. Reviewer - Reviewing Activities	18
10.1.3. Invited Talks	18
10.1.4. Research Administration	18
10.2. Teaching - Supervision - Juries	18
10.3. Popularization	18
11. Ribliography	19

Project-Team GEOSTAT

Creation of the Team: 2009 November 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

- A3.4.2. Unsupervised learning
- A3.4.7. Kernel methods
- A3.4.8. Deep learning
- A5.3. Image processing and analysis
- A5.3.2. Sparse modeling and image representation
- A5.3.3. Pattern recognition
- A5.3.5. Computational photography
- A5.7. Audio modeling and processing
- A5.7.3. Speech
- A5.7.4. Analysis
- A5.9. Signal processing
- A5.9.2. Estimation, modeling
- A5.9.3. Reconstruction, enhancement
- A5.9.5. Sparsity-aware processing

Other Research Topics and Application Domains:

- B2. Health
- B2.2. Physiology and diseases
- B2.2.1. Cardiovascular and respiratory diseases
- B2.2.6. Neurodegenerative diseases
- B3. Environment and planet
- B3.3. Geosciences
- B3.3.2. Water: sea & ocean, lake & river
- B3.3.4. Atmosphere

1. Personnel

Research Scientists

Hussein Yahia [Team leader, Inria, Researcher, HDR]

Khalid Daoudi [Inria, Researcher]

Nicolas Brodu [Inria, Researcher]

Guillaume Attuel [Inria, Starting Research Position, from Mar 2017]

Post-Doctoral Fellow

Gabriel Augusto Zebadua Garcia [Inria, from Dec 2017]

PhD Students

Anass El Aouni [Inria, Univ Rabat]

Camila Artana [Univ Pierre et Marie Curie]

Gopal Singh [Indian Institute of Technology Roorkee]

Technical staff

Marie Martin [Inria, from Jul 2017]

Interns

Francis Goulamhoussen Satsou [Inria, from Jun 2017 until Aug 2017] Gongfeng Li [Inria, from Apr 2017 until Sep 2017] Quentin Robin [Inria, from May 2017 until Aug 2017]

Administrative Assistants

Sabrina Blondel-Duthil [Inria] Cecile Boutros [Inria]

Visiting Scientists

Khalid Minaoui [Univ Rabat M Agdal, Jul 2017] Dharmendra Singh [IIT Roorkee, from Aug 2017] Ayoub Tamim [Institut Supérieur des Pêches Maritimes, from Jun 2017 until Jul 2017]

2. Overall Objectives

2.1. Overall Objectives

GEOSTAT is a research project which investigates the analysis of some classes of natural complex signals (physiological time series, turbulent universe and earth observation data sets) by determining, in acquired signals, the properties that are predicted by commonly admitted or new physical models best fitting the phenomenon. Consequently, when statistical properties discovered in the signals do not match closely enough those predicted by accepted physical models, we question the validity of existing models or propose, whenever possible, modifications or extensions of existing models. We will give, in the sequel, a detailed example of this approach and methodology (Heart electrical activity signal analysis). An important aspect of this methodological approach is that we don't rely in GEOSTAT, on a predetermined "universal" signal processing model to analyze natural complex signals. Instead, we take into consideration existing approaches in nonlinear signal processing (wavelets, multifractal analysis tools such as log-cumulants or micro-canonical multifractal formalism, time frequency analysis etc.) which are used to determine the micro structures or other micro features inside the acquired signals. Then, statistical analysis of these micro data are determined and compared to expected behaviour from theoretical physical models used to describe the phenomenon from which the data is acquired. From there different possibilities can be contemplated:

- The statistics match behaviour predicted by the model: complexity parameters predicted by the model are extracted from signals to analyze the dynamics of underlying phenomena. Examples: analysis of turbulent data sets in Oceanography and Astronomy.
- The signals displays statistics that cannot be attainable by the common lore of accepted models: how to extend or modify the models according to the behaviour of observed signals? Example: electrical activity of heart signal analysis (see infra).

GEOSTAT is a research project in nonlinear signal processing which develops on these considerations: it considers the signals as the realizations of complex extended dynamical systems. The driving approach is to describe the relations between complexity (or information content) and the geometric organization of information in a signal. For instance, for signals which are acquisitions of turbulent fluids, the organization of information may be related to the effective presence of a multiscale hierarchy of coherent structures, of multifractal nature, which is strongly related to intermittency and multiplicative cascade phenomena; the determination of this geometric organization unlocks key nonlinear parameters and features associated to these signals; it helps understand their dynamical properties and their analysis. We use this approach to derive novel solution methods for super-resolution and data fusion in Universe Sciences acquisitions [10]. Another example can be found in signal analysis of the electrical activity of the heart, where we find the distribution of activation points in a signal during episodes of atrial fibrilation (with strengthening from feature selection and Bayesian learning see below). Specific advances are obtained in GEOSTAT in using this type of

statistical/geometric approach to get validated dynamical information of signals acquired in Universe Sciences, e.g. Oceanography or Astronomy. 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, usually unattainable through linear methods, 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.

Signals studied in GEOSTAT belong to two broad classes:

- Acquisitions in Astronomy and Earth Observation.
- Physiological time series.

3. Research Program

3.1. General methodology

- **Fully Developed Turbulence (FDT)** Turbulence at very high Reynolds numbers; systems in FDT are beyond deterministic chaos, and symmetries are restored in a statistical sense only, and multicale correlated structures are landmarks. Generalizing to more random uncorrelated multi-scale structured turbulent fields.
- **Compact Representation** Reduced representation of a complex signal (dimensionality reduction) from which the whole signal can be reconstructed. The reduced representation can correspond to points randomly chosen, such as in Compressive Sensing, or to geometric localization related to statistical information content (framework of reconstructible systems).
- **Sparse representation** The representation of a signal as a linear combination of elements taken in a dictionary (frame or basis), with the aim of finding as less as possible non-zerio coefficients for a large class of signals.
- Universality class In theoretical physics, the observation of the coincidence of the critical exponents (behaviour near a second order phase transition) in different phenomena and systems is called universality. Universality is explained by the theory of the renormalization group, allowing for the determination of the changes followed by structured fluctuations under rescaling, a physical system is the stage of. The notion is applicable with caution and some differences to generalized out-of-equilibrium or disordered systems. Non-universal exponents (without definite classes) exist in some universal slowing dynamical phenomena like the glass transition and kindred. As a consequence, different macroscopic phenomena displaying multiscale structures (and their acquisition in the form of complex signals) may be grouped into different sets of generalized classes.

Every signal conveys, as a measure experiment, information on the physical system whose signal is an acquisition of. As a consequence, it seems natural that signal analysis or compression should make use of physical modelling of phenomena: the goal is to find new methodologies in signal processing that goes beyond the simple problem of interpretation. Physics of disordered systems, and specifically physics of (spin) glasses is putting forward new algorithmic resolution methods in various domains such as optimization, compressive sensing etc. with significant success notably for NP hard problem heuristics. Similarly, physics of turbulence introduces phenomenological approaches involving multifractality. Energy cascades are indeed closely related to geometrical manifolds defined through random processes. At these structures' scales, information in the process is lost by dissipation (close to the lower bound of inertial range). However, all the cascade is encoded in the geometric manifolds, through long or short distance correlations depending on cases. How do these geometrical manifold structures organize in space and time, in other words, how does the scale entropy cascades itself? To unify these two notions, a description in term of free energy of a generic physical model is sometimes possible, such as an elastic interface model in a random nonlinear energy landscape: This is

for instance the correspondence between compressible stochastic Burgers equation and directed polymers in a disordered medium. Thus, trying to unlock the fingerprints of cascade-like structures in acquired natural signals becomes a fundamental problem, from both theoretical and applicative viewpoints.

To illustrate the general methodology undertaken, let us focus on an example conducted in the study of physiological time series: the analysis of signals recorded from the electrical activity of the heart in the general setting of Atrial Fibrillation (AF). AF is a cardiac arrhythmia characterized by rapid and irregular atrial electrical activity with a high clinical impact on stroke incidence. Best available therapeutic strategies combine pharmacological and surgical means. But when successful, they do not always prevent long-term relapses. Initial success becomes all the more tricky to achieve as the arrhythmia maintains itself and the pathology evolves into sustained or chronic AF. This raises the open crucial issue of deciphering the mechanisms that govern the onset of AF as well as its perpetuation. We have developed a wavelet-based multi-scale strategy to analyze the electrical activity of human hearts recorded by catheter electrodes, positioned in the coronary sinus (CS), during episodes of chronic AF. We have computed the so-called multifractal spectra using two variants of the wavelet transform modulus maxima method, the moment (partition function) method and the magnitude cumulant method (checking confidence intervals with surrogate data). Application of these methods to long time series recorded in a patient with chronic AF provides quantitative evidence of the multifractal intermittent nature of the electric energy of passing cardiac impulses at low frequencies, i.e. for times ($\gtrsim 0.5$ s) longer than the mean interbeat ($\simeq 10^{-1}$ s). We have also reported the results of a two-point magnitude correlation analysis which infers the absence of a multiplicative time-scale structure underlying multifractal scaling. The electric energy dynamics looks like a "multifractal white noise" with quadratic (log-normal) multifractal spectra. These observations challenge concepts of functional reentrant circuits in mechanistic theories of AF. A transition is observed in the computed multifractal spectra which group according to two distinct areas, consistently with the anatomical substrate binding to the CS, namely the left atrial posterior wall, and the ligament of Marshall which is innervated by the ANS. These negative results challenge also the existing models, which by principle cannot explain such results. As a consequence, we go beyond the existing models and propose a mathematical model of a denervated heart where the kinetics of gap junction conductance alone induces a desynchronization of the myocardial excitable cells, accounting for the multifractal spectra found experimentally in the left atrial posterior wall area (devoid of ANS influence).

GEOSTAT is focusing on the analysis of turbulent datasets in which the multiscale description can be understood in the form of the multiplicative cascade or, in the case of physiological times series, as excitable systems (cardiac electrophysiology: study of intermittency phenomena). The methodological tools used in reaching these objectives place GEOSTAT at the forefront of nonlinear signal processing and analysis of complex systems. We cite: singularity exponents [56], [7] [11], sparse representations with reconstruction formulae [13] [57], [5], super-resolution in Oceanography and Earth Observation [10], [2], comparison with embedding techniques such as the one provided by the classical theorem of Takens [54], [44], the use of Lyapunov exponents [27] [20], how they are related to intermittency, persistence along the scales [6], comparison with other approaches such as sparse representations and compressive sensing [https://hal.inria. fr/tel-01239958], and the ways that lead to effective numerical and high precision determination of nonlinear characteristics in real signals. Derived from ideas in Statistical Physics, complex signals and systems are studied in relation to the statistical concepts of information content and most informative subsets. As a result, GEOSTAT aims to provide radically new approaches to the study of signals acquired from different complex systems (their analysis, their classification, the study of their dynamical properties etc.). A common characteristic of these signals, which is related to universality classes [48] [49] [45], 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 fronts/singularities and fits adequately to the case of chaotic data. The description is analogous to the modeling of states far from equilibrium. From this formalism we derive methods able to determine geometrically the most informative part in a signal scale by scale, which also defines its global properties and allows for compact representation. 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 [52]. 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).

3.2. Excitable systems: analysis of physiological time series

The research described in this section is a collaboration effort of GEOSTAT, CNRS LOMA (Laboratoire Ondes et Matière d'Aquitaine) and Laboratory of Physical Foundation of Strength, Institute of Continuous Media Mechanics (Perm, Russia Federation).

3.2.1. Presentation and objectives

Provide state of the art, cutting-edge tools to intra-cardiac multiscale analysis of the electrical activity of fibrillating hearts; offer physiological hypotheses likely to account for the new quantitative observations together with quantitative simulations.

3.2.2. Results

Wavelet-based methods (WTMM, log-cumulants, two point scale correlations), and confidence statistical methodology, have been applied to catheter recordings in the coronary sinus vein right next to the left atria of a small sample of patients with various conditions, and exhibit clear multifractal scaling without cross-scale correlation, which are coined "multifractal white noise," and that can be grouped according to two anatomical regions. We show that this is incompatible with the common lore for atrial fibrillation based on so-called circuit reentries. In a new description, we propose that circuit reentries may well exist before the onset of fibrillation, favoring onset but not contributing directly to the onset and perpetuation. By contrast, cell-to-cell coupling is considered fundamentally dynamical. The rationale stems from the observation that multifractal scaling necessitates a high number of degrees of freedom (tending to infinity with system size), which can originate in excitable systems in hyperbolic spatial coupling. In other words, common mathematical models for fibrillation which insist on the intrinsic chaotic dynamics of excitable cells coupled by elliptic propagators (like diffusion) are immune to multifractal scaling. Within this framework, we have developed a new hypothesis in physiology, backed by a mathematical model of gap junction conductance kinetics, that is capable of yielding correct spectra, all based on otherwise known physiology.

3.2.3. Interpretation

Atrial Fibrillation (AF) is an arrhythmia originating in the rapid and irregular electrical activity of the atria (the heart's two upper chambers) that causes their pump function to fail, increasing up to fivefold the risk of embolic stroke. The rate of AF recurrences after an initial ablation procedure treating paroxysmal AF increases with time, necessitating multiple redos, and most patients suffering persistent AF are resistant to treatment. The prevailing electrophysiological concepts describing tachy-arrhythmias are more than a century old. They involve abnormal automaticity and conduction. Initiation and maintenance are thought to arise from a vulnerable substrate prone to the emergence of multiple self-perpetuating reentry circuits, also called "multiple wavelets". We have analyzed the complexity of voltage signals recovered with bipolar electrodes in the CS during AF. We used two declinations of a wavelet-based multi-scale method, the moment (partition function) method and the magnitude cumulant method, as originally introduced in the field of fully developed turbulence. In the context of cardiac physiology, this methodology was shown to be valuable in assessing congestive heart failure from the monitoring of sinus heart rate variability (Ivanov, et al. Nature 399, 461–465) [42]. We develop a model such that the substrate function is modulated by the kinetics of conduction. A simple reversible mechanism of short term remodeling under rapid pacing is demonstrated, by which ionic overload acts locally (dynamical feedback) on the kinetics of gap junction conductance. The whole process may propagate and pervade the myocardium via electronic currents, becoming desynchronized. Contrary to existing mathematical models based on circuit reentries, a spatio-temporal multifractal intermittent dynamics emerges similar to the one found in the CS, opening a new avenue towards the understanding of AF mechanisms of perpetuation. We have shown that the wavelet-based multifractal analysis of long time series of the local impulse energy recorded in the CS of a patient with chronic AF was able to reveal and quantify the intermittent nature of these signals at low frequency (f < 2 Hz). To our knowledge, this research is the first to report on the observation and quantification of such multifractal dynamics of the endocavitary electrical activity during AF which is found more complex than previously suspected. Two main observations can be made: (i) the local impulse energy displays different multifractal properties in the left atrial wall area than in the ligament of Marshall area consistently with different anatomical substrate conditions, and (ii) while recorded along the CS vein, the local impulse energy does not exhibit long-range dependence associated with an underlying multiplicative cascade, or in other words the multifractal distribution of the singularities inferred by the two point magnitude analysis does not display any correlation across scales just like a log-normal "multifractal white noise".

This analysis definitely challenges current knowledge in physical, physiological and clinical fundamentals of AF arrhythmia.

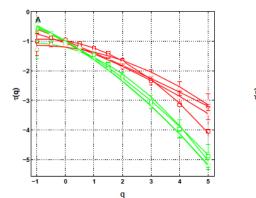
The absence of an underlying cascading process is not such a surprise since underlying the multifractal properties displayed by the local impulse energy at low frequencies (f < 2 Hz), there is no clear 3D "fragmentation" process inducing some cascading of energy from large to small time scales and also no obvious 2D "aggregation, coalescence or growth" process bringing energy from small to large time scales. What are the physical and physiological mechanisms that drive the multifractal nature of local impulse energy and give rise to the observed differences according to area is still an open question. Nonetheless, these results already undermine the commonly accepted concepts revolving around circuit reentries, and a fortiori spiral waves, as being basic mechanisms for the onset and perpetuation of AF. The mechanistic "wavelength" criterion indeed conveys the idea that random spatio-temporal dispersion of refractoriness, or more generally of functional properties, leads to random mixing of circuit reentries. The "wavelength" scale adjusts naturally to the typical scale λ of dispersion when it exists $c \times RP \lesssim \lambda$ as would be the case for Gaussian statistics of dispersion. In that case, the statistics of the local impulse energy remains Gaussian throughout scales. On the contrary, to fit our new observations we see that the statistics is not Gaussian and evolves across scales through a log-normal propagation law, which accounts for the intermittency, over the range of a few beat cycles (~ 0.6 s) to several tens (~ 10 s and possibly more), therefore spanning the whole atria. Although the ligament of Marshall area is highly innervated, it is quite unlikely that modulations by the ANS, that affects primarily heart rate, play a significant role in the intermittent dynamics, since the documented three peak frequencies at 0,4 Hz, 0,15 Hz and 0,04 Hz do not show up in our analysis. Furthermore, we have found at least two areas with different multifractal regimes. See figure 1.

Therefore, our findings raise new challenging questions calling for ongoing efforts to develop physiological heart tissue models that account for the low frequency intermittent nature of local impulse energy. In this spirit, in an ongoing research, we propose a model of gap junction conduction remodeling in a denervated heart that accounts for the observed intermittent dynamics over large time scales, as resulting from incoherent random back scatterings, leading to the desynchronization of the network of cardiac excitable cells.

These results have been accepted in a Frontiers in Physiology paper to be published in 2018 https://hal.inria.fr/hal-01673364.

3.3. Multiscale description in terms of multiplicative cascade

GEOSTAT is studying complex signals under the point of view of methods developed in *statistical physics* to study complex systems, with a strong emphasis on multiresolution analysis. 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. Because of the absence of localization of the Fourier transform, linear methods are not successful to unlock the multiscale structures and cascading properties of variables which are of primary importance as stated by the physics of the phenomena. This is the reason why new approaches, such as DFA (Detrented Fluctuation Analysis), Time-frequency analysis, variations on curvelets [53] etc. have appeared during the last decades. Recent advances in dimensionality reduction, and notably in Compressive Sensing, go beyond the Nyquist rate in sampling theory using nonlinear reconstruction, but data reduction occur at random places, independently of geometric localization



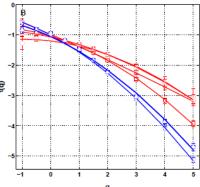


Figure 1. $\tau(q)$ spectra of local impulse energy time-series recorded along the CS vein at the electrodes Pt2 (red), Pt3 (blue) and Pt5 (green). The curves represent quadratic polynomial fit of the data. (A) The symbols correspond to the reference Patient 1 (chronic AF, ∇), and to Patients 2 (chronic AF, \circ) 3 (paroxysmal AF, \square) and 4 (persistent AF, \triangle). (B) The symbols correspond to the reference Patient 1 (∇) and to three different time-series for Patient 4 (\circ , \square , \triangle) recorded at different periods of time preceding ablation procedure.

of information content, which can be very useful for acquisition purposes, but of lower impact in signal analysis. One important result obtained in GEOSTAT is the effective use of multiresolution analysis associated to optimal inference along the scales of a complex system. The multiresolution analysis is performed on dimensionless quantities given by the singularity exponents which encode properly the geometrical structures associated to multiscale organization. This is applied successfully in the derivation of high resolution ocean dynamics, or the high resolution mapping of gaseous exchanges between the ocean and the atmosphere; the latter is of primary importance for a quantitative evaluation of global warming. 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 [38], [26], [58], [47]. Some quantities related to nonlinearity, such as Lyapunov exponents, Kolmogorov-Sinai entropy etc. can be computed at least in the phase space [27]. 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 [54] about strange attractors in transition turbulence has motivated the determination of discrete dynamical systems associated to time series [44], 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 [37] and multiscale/multiresolution are the subjects of intense research and are profoundly reshaping the analysis of complex signals by nonlinear approaches [25], [41]. 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 singularity exponents [11] [7]. We are working on:

- the determination of quantities related to universality classses,
- the geometric localization of multiscale properties in complex signals,
- cascading characteristics of physical variables.

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 [55] is local and reconstruction in the signal's domain is related to predictability. This approach is analogous to the consideration of "microcanonical ensembles" in statistical mechanics.

A multiscale organization is a fundamental feature of a complex system, it can be for example related to the cascading properties in turbulent systems. We make use of this kind of description when analyzing turbulent signals: intermittency is observed within the inertial range and is related to the fact that, in the case of FDT, symmetry is restored only in a statistical sense, a fact that has consequences on the quality of any nonlinear signal representation by frames or dictionaries.

The example of FDT as a standard "template" for developing general methods that apply to a vast class of complex systems and signals is of fundamental interest because, in FDT, the existence of a multiscale hierarchy \mathcal{F}_h which is of multifractal nature and geometrically localized can be derived from physical considerations. This geometric hierarchy of sets is responsible for the shape of the computed singularity spectra, which in turn is related to the statistical organization of information content in a signal. It explains scale invariance, a characteristic feature of complex signals. The analogy from statistical physics comes from the fact that singularity exponents are direct generalizations of critical exponents which explain the macroscopic properties of a system around critical points, and the quantitative characterization of *universality classes*, which allow the definition of methods and algorithms that apply to general complex signals and systems, and not only turbulent signals: signals which belong to a same universality class share common statistical organization. In GEOSTAT, the approach to singularity exponents is done within a microcanonical setting, which can interestingly be compared with other approaches such that wavelet leaders, WTMM or DFA. During the past decades, classical approaches (here called "canonical" because they use the analogy taken from the consideration of "canonical ensembles" in statistical mechanics) permitted the development of a well-established analogy taken from thermodynamics in the analysis of complex signals: if \mathcal{F} is the free energy, \mathcal{T} the temperature measured in energy units, \mathcal{U} the internal energy per volume unit \mathcal{S} the entropy and $\widehat{\beta} = 1/\mathcal{T}$, then the scaling exponents associated to moments of intensive variables $p \to \tau_p$ corresponds to $\widehat{\beta} \mathcal{F}$, $\mathcal{U}(\widehat{\beta})$ corresponds to the singularity exponents values, and $S(\mathcal{U})$ to the singularity spectrum.

The singularity exponents belong to a universality class, independently of microscopic properties in the phase space of various complex systems, and beyond the particular case of turbulent data (where the existence of a multiscale hierarchy, of multifractal nature, can be inferred directly from physical considerations). They describe common multiscale statistical organizations in different complex systems [52], and this is why GEOSTAT is working on nonlinear signal processing tools that are applied to very different types of signals.

For example we give some insight about the collaboration with LEGOS Dynbio team 1 about high-resolution ocean dynamics from microcanonical formulations in nonlinear complex signal analysis. Indeed, synoptic determination of ocean circulation using data acquired from space, with a coherent depiction of its turbulent characteristics remains a fundamental challenge in oceanography. This determination has the potential of revealing all aspects of the ocean dynamic variability on a wide range of spatio-temporal scales and will enhance our understanding of ocean-atmosphere exchanges at super resolution, as required in the present context of climate change. We show that the determination of a multiresolution analysis associated to the multiplicative cascade of a typical physical variable like the Sea Surface Temperature permits an *optimal inference* of oceanic motion field across the scales, resulting in a new method for deriving super resolution oceanic motion from lower resolution altimetry data; the resulting oceanic motion field is validated at super resolution with the use of Lagrangian buoy data available from the Global Drifter Program 2 . In FDT, singularity exponents range in a bounded interval: $]h_{\infty}, h_{\max}[$ with $h_{\infty} < 0$ being the most singular exponent. Points \mathbf{r} for which $h(\mathbf{r}) < 0$ localize the strongest transition fronts in the turbulent fluid, where an intensive

¹ http://www.legos.obs-mip.fr/recherches/equipes/dynbio.
² http://www.aoml.noaa.gov/phod/dac/index.php.

physical variable like sea surface temperature behaves like $1/\mathbf{r}^{|h(\mathbf{r})|}$. The links between the geometricaly localized singularity exponents, the scaling exponents of structure functions, the multiplicative cascade and the multiscale hierarchy \mathcal{F}_h is the following:

$$\begin{cases}
\mathcal{F}_{h} = \{\mathbf{r} \mid h(\mathbf{r}) = h\} \\
D(h) = \dim \mathcal{F}_{h} \\
\tau_{p} = \inf_{h} \{ph + 3 - D(h)\} \\
D(h) = \inf_{p} \{ph + 3 - \tau_{p}\}
\end{cases} \tag{1}$$

Let $\mathfrak{S}(\mathbf{x})$ be the bidimensionnal signal recording, for each sample point \mathbf{x} representing a pixel on the surface of the ocean of given resolution, the sea surface temperature (sst). To this signal we associate a measure μ whose density w.r.t Lebesgue measure is the signal's gradient norm, and from which the singularity exponents are computed [6]. It is fundamental to notice here that, contrary to other types of exponents computed in Oceanography, such as Finite Size Lyapunov exponents, singularity exponents are computed at instantaneous time, and do not need time series.

Having computed the singularity exponents at each point of a SST signal, a microcanonical version of the multiplicative cascade associated to the scaling properties of the sst becomes available. The idea of the existence of a geometrically localized multiplicative cascade goes back to [51]. The multiplicative cascade, written pointwise, introduces random variables $\eta_{l'/l}(\mathbf{x})$ for 0 < l' < l such that

$$\mathfrak{I}_{\psi}\mu(\mathbf{x},l') = \eta_{l'/l}(\mathbf{x})\mathfrak{I}_{\psi}\mu(\mathbf{x},l) \tag{2}$$

in which the equality is valid pointwise and not only in distribution. Inference of physical variables across the scales is optimized and consequently we describe the multiplicative cascade at each point \mathbf{x} in the signal domain. The injection variables $\eta_{l'/l}(\mathbf{x})$ are indefinitely divisible: $\eta_k(\mathbf{x})\eta_{k'}(\mathbf{x}) \doteq \eta_{kk'}(\mathbf{x})$. It is possible to optimize cross-scale inference of physical variables by considering a multiresolution analysis associated to a discrete covering of the "space-frequency" domain. Denoting as usual $(V_j)_{j\in\mathbb{Z}}$ and $(W_j)_{j\in\mathbb{Z}}$ the discrete sequence of approximation and detail spaces associated to a given scaling function, and denoting by $\psi \in L^2(\mathbb{R}^2)$ a wavelet which generates an Hilbertian basis on each detail space W_j , it is known that the detail spaces encode borders and transition information, which is ideally described in the case of turbulent signals by the singularity exponents $\mathbf{h}(\mathbf{x})$. Consequently, a novel idea for super-resolution consists in computing a multiresolution analysis on the signal of singularity exponents $\mathbf{h}(\mathbf{x})$, and to consider that the detail information coming from spaces W_j is given the signal $\mathbf{h}(\mathbf{x})$. The associated orthogonal projection $\pi_j: L^2(\mathbb{R}^2) \to W_j$ defined by $\pi_j(\mathbf{h}) = \sum_{n \in \mathbb{Z}} \langle \mathbf{h} \mid \psi_{j,n} \rangle \psi_{j,n}$ is then used in the reconstruction formula for retrieving a physical

variable at higher resolution from its low resolution counterpart. If $\mathfrak{S}(\mathbf{x})$ is such a variable, we use a reconstruction formula: $A_{j-1}\mathfrak{S} = A_j\mathfrak{S} + \pi_j(\mathbf{h})$ with $A_j: L^2(\mathbb{R}^2) \to V_j$ is the orthogonal projection on the space V_j (approximation operator) and π_j is the orthogonal projection on the detail spaces W_j associated to the signal of singularity exponents $\mathbf{h}(\mathbf{x})$. Validation is performed using Lagrangian buoy data with very good results [10]. We have realized a demonstration movie showing the turbulent ocean dynamics at an SST resolution of 4 km computed from the SST microcanonical cascade and the low-resolution GEKCO product for the year 2006 over the southwestern part of the Indian Ocean. We replace the missing data in the SST MODIS product (clouds and satellite swath) by the corresponding data available from the Operational SST and Sea Ice Analysis (OSTIA) provided by the Group for High-Resolution SST Project [11], which, however, is of lower quality. Two images per day are generated for the whole year of 2006. The resulting images show the norm of the vector field in the background rendered using the line integral convolution algorithm. In the foreground, we show the resulting vector field in a linear gray-scale color map. See link to movie (size: 800 Mo).

3.4. Data-based identification of characteristic scales and automated modeling

Data are often acquired at the highest possible resolution, but that scale is not necessarily the best for modeling and understanding the system from which data was measured. The intrinsic properties of natural processes do not depend on the arbitrary scale at which data is acquired; yet, usual analysis techniques operate at the acquisition resolution. When several processes interact at different scales, the identification of their characteristic scales from empirical data becomes a necessary condition for properly modeling the system. A classical method for identifying characteristic scales is to look at the work done by the physical processes, the energy they dissipate over time. The assumption is that this work matches the most important action of each process on the studied natural system, which is usually a reasonable assumption. In the framework of time-frequency analysis [36], the power of the signal can be easily computed in each frequency band, itself matching a temporal scale.

However, in open and dissipative systems, energy dissipation is a prerequisite and thus not necessarily the most useful metric to investigate. In fact, most natural, physical and industrial systems we deal with fall in this category, while balanced quasi-static assumptions are practical approximation only for scales well below the characteristic scale of the involved processes. Open and dissipative systems are not locally constrained by the inevitable rise in entropy, thus allowing the maintaining through time of mesoscopic ordered structures. And, according to information theory [40], more order and less entropy means that these structures have a higher information content than the rest of the system, which usually gives them a high functional role.

We propose to identify characteristic scales not only with energy dissipation, as usual in signal processing analysis, but most importantly with information content. Information theory can be extended to look at which scales are most informative (e.g. multi-scale entropy [31], ε -entropy [30]). Complexity measures quantify the presence of structures in the signal (e.g. statistical complexity [33], MPR [46] and others [35]). With these notions, it is already possible to discriminate between random fluctuations and hidden order, such as in chaotic systems [32], [46]. The theory of how information and structures can be defined through scales is not complete yet, but the state of art is promising [34]. Current research in the team focuses on how informative scales can be found using collections of random paths, assumed to capture local structures as they reach out [29].

Building on these notions, it should also possible to fully automate the modeling of a natural system. Once characteristic scales are found, causal relationships can be established empirically. They are then clustered together in internal states of a special kind of Markov models called ε -machines [33]. These are known to be the optimal predictors of a system, with the drawback that it is currently quite complicated to build them properly, except for small system [50]. Recent extensions with advanced clustering techniques [28], [39], coupled with the physics of the studied system (e.g. fluid dynamics), have proved that ε -machines are applicable to large systems, such as global wind patterns in the atmosphere [43]. Current research in the team focuses on the use of reproducing kernels, coupled possibly with sparse operators, in order to design better algorithms for ε -machines reconstruction. In order to help with this long-term project, a collaboration with J. Crutchfield lab at UC Davis was initiated in 2017.

3.5. Speech analysis

Our research in speech processing focus on the development of novel nonlinear analysis methods for the characterization and classification of pathological and affective speech. For the latter, classical linear methods do not generally capture the nonlinearity, aperiodicity, turbulence and noise that can be present in pathological voices. We thus aim to design and extract new features that allow better characterization/classification of such voices, while being easy to interpret by clinicians. For the former, recent research have shown that the voice source signal information allow significant improvement of speech emotion detection systems. Our goal is to develop novel nonlinear techniques to extract relevant voice source features and to design efficient machine learning algorithms for robust emotion classification.

4. Application Domains

4.1. Analysis of galactic molecular clouds (GENESIS project)

GENESIS (GENeration and Evolution of Structures in the ISm) is a German/French collaboration project, supported for three years (start 1.5. 2017) by the Deutsche Forschungsgemeinde (DFG) and the Agence national de recherche (ANR). The objective of this research project is to better understand the structure and evolution of molecular clouds in the interstellar medium (ISM) and to link cloud structure with star-formation. For that, far-infrared observations of dust (Herschel) and cooling lines (SOFIA) are combined with ground-based submillimetre observations of molecular lines. Dedicated analysis tools will be used and developed to analyse the maps and compared to simulations in order to disentangle the underlying physical processes such as gravity, turbulence, magnetic fields, and radiation.

4.2. Ocean dynamics, upwelling

Ocean dynamics from remote sensing data is studied in collaboration with LEGOS CNRS Laboratory in Toulouse (SYSC O_2) team, with CRTS and Rabat University in Morocco. The following thematics are studied:

- Large time series analysis on global synoptic ocean data to provide fine characteristics of Ocean turbulence, analysis of altimetry data and ocean/atmosphere exchanges (supported by ICARODE project and submitted LEFE-EC2CO IMECO project).
- Upwelling, Lyapunov exponents and Lagrangian Coherent Structures (supported by TOUBKAL project).

4.3. Non convex optimization for image processing

Three years contract with I2S company on sparse representations and non-convex optimizations methods for image processings problems such as: deconvolution, stiching, noise reduction, HDR mapping etc. This contract is taking place after award-winning results of former GEOSTAT PhD student H. Badri.

4.4. Physiological times series: speech processing

Differential diagnosis between Parkinson's disease and Multiple System Atrophy using complex physiological tile series analysis. Supported by ANR Voice4PD-MSA project.

5. Highlights of the Year

5.1. Highlights of the Year

Innovation LAB GEOSTAT-I2S based on 3 year contract with I2S company on non convex optimization methods for image processing.

5.1.1. Awards

A. Tamim, PhD Student in Geostat, wins the gold medal of Hubert Curien PhD prize 2017. A. Tamim's PhD title: "Segmentation et classification des images satellitaires : application à la détection des zones d'upwelling côtier marocain et mise en place d'un logiciel de suivi spatiotemporel". See https://www.inria.fr/centre/bordeaux/actualites/prix-de-these-pour-ayoub-tamim.

6. New Software and Platforms

6.1. Fluex

KEYWORDS: Signal - Signal processing

SCIENTIFIC DESCRIPTION: Fluex is a package consisting of the Microcanonical Multiscale Formalism for 1D, 2D 3D and 3D+t general signals.

FUNCTIONAL DESCRIPTION: Fluex is a C++ library developed under Gforge. Fluex is a library in nonlinear signal processing. Fluex is able to analyze turbulent and natural complex signals, Fluex is able to determine low level features in these signals that cannot be determined using standard linear techniques.

Participants: Hussein Yahia and Rémi Paties

Contact: Hussein Yahia

• URL: http://fluex.gforge.inria.fr/

6.2. FluidExponents

KEYWORDS: Signal processing - Wavelets - Fractal - Spectral method - Complexity

FUNCTIONAL DESCRIPTION: FluidExponents is a signal processing software dedicated to the analysis of complex signals displaying multiscale properties. It analyzes complex natural signals by use of nonlinear methods. It implements the multifractal formalism and allows various kinds of signal decomposition and reconstruction. One key aspect of the software lies in its ability to evaluate key concepts such as the degree of impredictability around a point in a signal, and provides different kinds of applications. The software can be used for times series or multidimensional signals.

• Participants: Antonio Turiel and Hussein Yahia

• Contact: Hussein Yahia

• URL: https://fluidexponents@scm.gforge.inria.fr/svn/fluidexponents/FluidExponents

6.3. classifemo

KEYWORDS: Classification - Audio

FUNCTIONAL DESCRIPTION: Classifies vocal audio signals. Classifemo extracts characteristics from vocal audio signals. These characteristics are extracted from signals of different type: initially these were emotion databases, but it can also process signals recorded from patients with motor speach disorders. The software can train usual classifiers (SVM, random forests, etc) on these databases as well as classify new signals.

Participants: Khalid Daoudi and Nicolas Brodu

Contact: Khalid Daoudi

• URL: https://allgo.inria.fr/app/emotionclassifierprototype

6.4. superres

Super-Resolution of multi-spectral and multi-resolution images

KEYWORD: Multiscale

SCIENTIFIC DESCRIPTION: This resolution enhancement method is designed for multispectral and multiresolution images, such as these provided by the Sentinel-2 satellites (but not only). Starting from the highest resolution bands, band-dependent information (reflectance) is separated from information that is common to all bands (geometry of scene elements). This model is then applied to unmix low-resolution bands, preserving their reflectance, while propagating band-independent information to preserve the sub-pixel details.

FUNCTIONAL DESCRIPTION: This super-resolution software for multi-spectral images consists of: - A core C++ library, which can be used directly - A Python module interface to this library - A Java JNI interface to the library - An end-user Python script for super-resolving Sentinel-2 images - An end-user plugin for the widely used SNAP software of the ESA.

• Participant: Nicolas Brodu

• Contact: Nicolas Brodu

• URL: http://nicolas.brodu.net/recherche/superres/index.html

6.5. EdgeReconstruct

Edge Reconstruction With UPM Manifold

KEYWORDS: 2D - Fractal - Signal processing

FUNCTIONAL DESCRIPTION: EdgeReconstruct is a software that reconstructs a complex signal from the computation of most unpredictible points in the framework of the Microcanonical Multifractal Formalism. The quality of the reconstruction is also evaluated. The software is a companion of a paper published in 2013: https://hal.inria.fr/hal-00924137.

• Contact: Suman Kumar Maji

• URL: https://geostat.bordeaux.inria.fr/index.php/downloads.html

7. New Results

7.1. Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation

Participants: G. Attuel, H. Yahia.

We compute the so-called multifractal spectra using two variants of the wavelet transform modulus maxima method, the moment (partition function) method and the magnitude cumulant method. Application of these methods to long time series recorded in a patient with chronic AF provides quantitative evidence of the multifractal intermittent nature of the electric energy of passing cardiac impulses at low frequencies, i.e. for times (> 0.5s) longer than the mean interbeat. We also report the results of a two-point magnitude correlation analysis which infers the absence of a multiplicative time-scale structure underlying multifractal scaling. The electric energy dynamics looks like a "multifractal white noise" with quadratic (log-normal) multifractal spectra. These observations challenge concepts of functional reentrant circuits in mechanistic theories of AF, still leaving open the role of the autonomic nervous system (ANS). A transition is indeed observed in the computed multifractal spectra which group according to two distinct areas, consistently with the anatomical substrate binding to the CS, namely the left atrial posterior wall, and the ligament of Marshall which is innervated by the ANS. In a companion paper (II. Modeling), we propose a mathematical model of a denervated heart where the kinetics of gap junction conductance alone induces a desynchronization of the myocardial excitable cells, accounting for the multifractal spectra found experimentally in the left atrial posterior wall area.

Publication: G. Attuel *et al.*, Multifractal desynchronization of the cardiac excitable cell network during atrial fibrillation. I. Multifractal analysis of clinical data. Accepted in **Frontiers in Physiology**, publication beginning 2018.

7.2. Super-resolution

Participant: N. Brodu.

N. Brodu, "Super-Resolving Multiresolution Images With Band-Independent Geometry of Multispectral Pixels". IEEE Transactions on Geoscience and Remote Sensing 55(8), p4610-4617, 2017. In this work we introduce a new method to enhance Sentinel-2 images. Sentinel is a leading ESA program in satellite-based environmental monitoring. N. Brodu has included a super-resolution algorithm in the form of a plug-in in the ESA SNAP software. SNAP is a major reference in satellite image processing. Free software (Cecill-C and GPL), APP deposit under way, https://bil.inria.fr/fr/software/view/2984/tab.

Publication: [15].

7.3. Surface mixing and biological activity in the Northwest African upwelling

Participants: A. El Aouni, K. Daoudi, H. Yahia.

The aim of this work is to study the horizontal stirring and mixing in different upwelling areas of the Northwest African margin using attracting/repelling Lagrangian coherent structures (LCS) obtained as subsets of hyperstreamline of the Cauchy-Green strain tensor, whose normal repulsion rate is larger than tangential stretch over backward/forward time interval, and their link to the chlorophyll fronts concentrations, based on 10 years satellite data. The temporal variability of surface stirring is compared to the fronts chlorophyll concentration. Two of the four studied areas exhibit negative correlation between mixing and the chlorophyll concentration. The other two regions show similar seasonal variations, nearly coincident maxima and minima, leading to a global positive correlation. These results are compared to other works that make use of Finite Size Lyapunov Exponent (FSLE) whose output is a plot of scalar distributions. Furthermore, we compare the chlorophyll concentrations with both compressing and stretching lines. Results show different regions with different properties. The surface mixing and chlorophyll concentrations correlation is governed by stretching lines in two regions, by compressing lines in one region with, while no different is shown between attracting and repelling LCS in the northenrn region of the studied area.

Publication: [20].

7.4. Spatio-Temporal Dynamics of Floods

Participant: N. Brodu.

The floods are an annual phenomenon on the Pacific Coast of Ecuador and can become devastating during El Niño years, especially in the Guayas watershed (32,300 km²), the largest drainage basin of the South American western side of the Andes. As limited information on flood extent in this basin is available, this study presents a monitoring of the spatio-temporal dynamics of floods in the Guayas Basin, between 2005 and 2008, using a change detection method applied to ENVISAT ASAR Global Monitoring SAR images acquired at a spatial resolution of 1 km.

Publication: [16].

7.5. Effect of wind stress forcing on ocean dynamics at Air-Sea Interface

Participant: H. Yahia.

At first order, oceanic currents are generated by the balance of Coriolis and pressure gradient (geostrophic current) and the balance of Coriolis and the frictional force dominated by wind stress in the surface ocean (Ekman current). We aim at studying the difference in term of turbulent hydrodynamics carried by the wind forcing at the air-sea interface. We explore the statistical properties of singularity spectra computed from velocity norms and vorticity data, notably in relation with kurtosis information to underline differences in the turbulent regimes associated with both kinds of velocity fields. This study is conducted over 1 year of daily data and demonstrates the differences in terms of turbulent property of wind forcing.

Publication: [24].

7.6. Ocean dynamics: frontal activity

Participant: H. Yahia.

A highresolution (1km spatial and daily temporal resolutions) dataset of 11 years (2003 to 2013) remotely sensed SST by MODIS sensor onboard Aqua and Terra satellites has been investigated and compared with coastal numerical model experiments. The detection and characterization fronts with fluctuating amplitudes is achieved through the Singularity Analysis (i.e. the process of calculating the degree of regularity or irregularity of a function at each point in a domain).

Publication: [18].

7.7. Pathological speech processing

Participants: K. Daoudi, G. Li, Q. Robin, F. G. Satsou.

- Small amount of training data in learning robust classifiers for differential diagnosis between progressive supranuclear palsy (PSP) and multiple system atrophy (MSA). We showed that factorial discriminant analysis and logistic regression can lead to such robust classifiers. Moreover, we showed that these models provide good insights on the multivariate variability and (un)correlation of acoustic features, which can facilitate clinical interpretation.
- We investigated the problem of extracting ground thruth of glottal closure instants (GCI) from electroglottographic (EGG) signals of healthy and pathological speakers. We carried out a large experimental study which showed that existing methods are not robust to recording settings and material. We then proposed a method to overcome this problem. On the other hand, this problem highlighted the non robustness of state of the art methods in automatic detection of GCI from speech.
- We made an experimental evaluation of state of the art methods in automatic extraction of the
 excitation source from voiced speech. To carry out this evaluation, we used a very recent sourcefilter model of sustained phonations. The results showed that these methods are reliable only in very
 particular cases and fail in most.
- Matching pursuit (MP), particularly using the Gammatones dictionary, has become a popular tool in sparse representations of speech/audio signals. The classical MP algorithm does not however take into account psychoacoustical aspects of the auditory system. Recently two algorithms, called PAMP and PMP have been introduced in order to select only perceptually relevant atoms during MP decomposition. We compared the performance these two algorithms on few speech sentences. The results showed that PMP, which also has the strong advantage of including an implicit stop criterion, always outperforms PAMP as well as classical MP. We then raised the question of whether the Gammatones dictionary is the best choice when using PMP. We thus compared it to the popular Gabor and damped-Sinusoids dictionaries. The results showed that Gammatones always outperform damped-Sinusoids, and that Gabor yield better reconstruction quality but with higher atoms rate.

Publications: [22], [23], [21], [19].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

• Three year contract with I2S company on the transfert of award winning H. Badri PhD results (AFRIF PhD price in 2016). The contract is being transformed in 2018 in the form of an Inria Innovation Lab. The Innovation Lab is focused on non convex optmization methods in image processing and digital acquisition devices. People involved in GEOSTAT: H. Yahia, N. Brodu, K.Daoudi, M. Martin, A. Zebadua.

8.2. Bilateral Grants with Industry

- Transfert in the analysis of hearbeat data. Discussion and collaboration with Cardiologs company https://cardiologs.com/.
- Contacts for a partnership strategy on heartbeat database utilization with Parly II Hospital (F. Halimi).
- Patent 185 "Dispositif analyseur de rythme cardiaque" extended for France in 2018.

9. Partnerships and Cooperations

9.1. Regional Initiatives

GEOSTAT is working with the following regional partners:

- CNRS LOMA (Laboratoire Ondes et Matière d'Aquitaine) and RAS Institute (Russia): collaboration on the analysis/modeling of heartbeat physiological time series (A. Arneodo, E. Gerasimova, F. Argoul).
- GEOSTAT has a decade-long close scientific collaboration with team SYSCO2 (LEGOS LAboratoryUMR 5566): V. Garçon, B. Dewitte, J. Sudre.
- Laboratoire d'Astrophysique de Bordeaux (S. Bontemps, N. Schneider, GENESIS project).
- Collaboration with L. Bourrel (GET Toulouse / IRD) and F. Frappart (GET/UMR EPOC)Flood monitoring in Equator.
- With Bruno Castelle (EPOC).
- With D. Gibert (OSUR) on signal and image processing.
- CHU Bordeaux : Prof. Wassilios Meissner (IMN), Dr. Solange Milhé de Saint Victor (service ORL).
- CHU Toulouse : Dr. Anne Pavy Le traon (service Neurologie), Prof. Virginie Woisard (service ORL)
- IRIT : Prof. Régine André-Obrecht, Dr. Julie Mauclair.
- IMT (Institut de Mathématique de Toulouse) : Dr. Sébastien Déjean, Dr. Laurent Risser.
- Mercator Océan: Dr. A. El Moussaoui. UMR EPOC).

9.2. National Initiatives

- ANR project *Voice4PD-MSA*, led by K. Daoudi, which targets the differential diagnosis between Parkinson's disease and Multiple System Atrophy. The total amount of the grant is 468555 euros, from which GeoStat has 203078 euros. The duration of the project is 42 months. Partners: CHU Bordeaux (Bordeaux), CHU Toulouse, IRIT, IMT (Toulouse).
- PhD grant for C. Artnana from UPMC University, under co-supervision with H. Yahia and C. Provost (LOCEAN, Paris).
- PhD grant for G. Singh from IIT Roorkee, under co-supervision with D. Singh (IIT Roorkee).
- The PHC-Toubkal project "Caractérisation multi-capteurs et suivi spatio-temporel de l'Upwelling sur la côte atlantique marocaine par imagerie satellitaire", led by K. Daoudi, is in its second year. The partners in this project are: Faculté des sciences de Rabat, Centre Royal de Télédetection Spatiale, Mercator-Ocean and GEOSTAT.
- GEOSTAT is a member of ISIS (Information, Image & Vision) and AMF (Multifractal Analysis) GDRs.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: supported by Deutsche Forschungsgemeinde (DFG) and the Agence national de recherche (ANR).

Project acronym: GENESIS.

Project title: GENeration and Evolution of Structures in the ISm.

Duration: start 1.5. 2017, 3 years.

Coordinator: N. Schneider (I. Physik, Cologne).

Other partners: Cologne (R. Simon, N. Schneider, V. Ossenkopf, M. Roellig), LAB (S. Bontemps, A. Roy, L. Bonne, F. Herpin, J. Braine, N. Brouillet, T. Jacq), ATN Canberra (Australia), LERMA Paris (France), MPIfR Bonn (Germany), CEA Saclay (France), ITA/ZAH Heidelberg (Germany), Institute of Astronomy, Cardiff (UK), ESO (Germany, Chile), CfA Harvard (USA), IPAG Grenoble (France), Argelander Institut Bonn (Germany), CASS San Diego (USA), University of Sofia (Bulgaria).

Abstract: The formation of stars is intimately linked to the structure and evolution of molecular clouds in the interstellar medium (ISM). We propose to explore this link with a new approach by combining far infrared maps of dust (Herschel) and cooling lines(C+ with SOFIA) with molecular line maps. Dedicated analysis tools will be used and developed to analyze the maps and compare them to simulations in order to identify for the underlying physical processes. This joint project relies on the complementary expertise of the members of the Cologne KOSMA group (structure identification methods and SOFIA), the Bordeaux LAB star formation group (Herschel and spectroimaging maps), and the Bordeaux GEOSTAT team of Inria. To understand the genesis of stars, it is necessary to disentangle the relative importance of gravity, turbulence, magnetic fields, and radiation from diffuse gas, to molecular clouds and collapsing cores, and to study the role of filaments. Using innovative new analying tools developed by the GeoStat team, we will analyze the Herschel images as well as new spectro-imaging surveys from ground-based telescopes, and THz spectroscopy using SOFIA. The comparison with similar analysis on simulated clouds will allow us to derive the underlying physical process which explains cloud evolution and the formation of dense structures. The project does not aim at a full understanding of star formation within 3 years, but it constitutes an important step forward as it will make systematic use of a wealth of existing, yet not fully exploited archival data, carefully chosen new observations, and sophisticated tools to analyze and interpret the data. As such, it will shed new light on how molecular clouds and stars form and may well be the starting point for many studies to follow.

9.4. International Initiatives

9.4.1. Inria International Partners

Funding from French-Indian IFCAM program (Visit of Prof. D. Singh in GEOSTAT, 2017).

9.4.1.1. Informal International Partners

- Visit of N. Brodu to Univ. UC Davis in the team of Prof. J. Crutchfield. Setting up of a collaboration
 on a formalism of statistical reconstruction from dynamic empirical data; the formalism involves
 markovian automata called Epsilon machines. The internal states of these machine correspond to
 equivalence classes of a physical system having similar causal relations.
- Laboratory LRIT from Rabat University (K. Minaoui, D. Aboutajdine).

9.4.2. Participation in Other International Programs

Participation in the IFCAM project with India (funding of the visit of Prof. D. Singh in 2017).

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- N. Schneider (Cologne University, GENESIS project).
- Prof. D. Singh (IIT roorkee, on CEFIPRA-CNRS funding). Duration: 8 weeks, August and December 2017. Co-supervision of G. Singh PhD stdudent, scientific collaboration with N. Brodu and K. Daoudi.

9.5.1.1. Internships

- G. Li. Master2, University Paris-Saclay.
- Q. Robin. Engineer, INP-Grenble.
- F. G. Satsou. Master1, University Bordeaux 1.

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

June-July 2017: PhD student A. El Aouni was invited in the MERCATOR project by A. Moussaoui in the framework of the Toubkal project on ocean modeling.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

• H. Yahia with S. Bontemps (Bordeaux Astrophysics Laboratory) and N. Schneider (Cologne University): GENESIS/MOBS kick-off meeting 28. and 29.9.2017 in Bordeaux, https://hera.ph1.uni-koeln.de/~nschneid/genesis-meeting-sep2017.html.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

H. Yahia is a review editor of Frontiers in Fractal Physiology.

10.1.2.2. Reviewer - Reviewing Activities

H. Yahia: Frontiers in Physiology.

10.1.3. Invited Talks

- N. Brodu: Presentation given to the 8th École Interdisciplinaire sur les Systèmes Complexes : « Extraction de données et information: applications à l'imagerie, à l'environnement et à la physique ». 11-13 october 2016, Rennes.
- K. Daoudi has been invited to the Wolrd Voice Consortium 2017 to present the project on differential diagnosis in Parkinsonism.
- H. Yahia: presentation given to the Laboratoire d'Astrophysoique de Bordeaux on March 8, 2017.
- H. Yahia is invited to the TOSCA (http://www.tosca2017.fr/) group in march 2017.
- A. Tamim is invited in June 2017 for the reception of his PHC Hubert Curien PhD gold medal.
- K. Minaoui is invited in GEOSTAT in July 2017.
- Visit of D. Singh in GEOSTAT in August 2017 on the co-supervision of G. Singh PhD thesis.
- Invitation of H. Badri in the framework of I2S kickoff meeting, on Inria funding.

10.1.4. Research Administration

- H. Yahia participated in the first scientific meeting of "Centre d'Excellence sur les Maladies Neurodégénératives", projet « Bordeaux Initiative for Neurodegenerative Disorders », May 11 2017, Bordeaux (NEUROCAMPUS auditorium building).
- Participation of H. Yahia, H. Badri, K. Daoudi and N. Brodu to the I2S-GEOSTAT kickoff meeting in September 2017.
- Participation of K. Daoudi to the SABOR project organization.

10.2. Teaching - Supervision - Juries

10.2.1. Supervision

PhD in progress: B. Das, supervised of H. Yahia in the framework of the Toubkal project (starts 01/01/18).

PhD in progress: A. El Aouni, co-supervised by K. Daoudi, H. Yahia and K. Minaoui in the framework of the Toubkal project.

PhD in progress : G. Singh, co-supervised by N. Brodu in the framework of OPTIC assoicated team anf IFCAM collaboration.

PhD in progress: C. Artana, co-supervised by H. Yahia in a collaboration with LOCEAN team (Univ. Paris 6).

10.3. Popularization

Diffusion of the GENESIS project in the magazine Inria PLUGIN (published beginnig 2018) and in the Inria website (national and INRIS BSO), see https://www.inria.fr/centre/bordeaux/actualites/lancement-du-projet-genesis.

11. Bibliography

Major publications by the team in recent years

- [1] H. BADRI, H. YAHIA, K. DAOUDI. Fast and Accurate Texture Recognition with Multilayer Convolution and Multifractal Analysis, in "European Conference on Computer Vision", Zürich, Switzerland, ECCV 2014, September 2014, https://hal.inria.fr/hal-01064793
- [2] I. HERNÁNDEZ-CARRASCO, J. SUDRE, V. GARÇON, H. YAHIA, C. GARBE, A. PAULMIER, B. DEWITTE, S. ILLIG, I. DADOU, M. GONZÁLEZ-DÁVILA, J. SANTANA CASIANO. Reconstruction of super-resolution ocean pCO 2 and air-sea fluxes of CO 2 from satellite imagery in the Southeastern Atlantic, in "Biogeosciences", September 2015, 20 p., https://hal.inria.fr/hal-01193339
- [3] V. KHANAGHA, K. DAOUDI, H. YAHIA. Detection of Glottal Closure Instants based on the Microcanonical Multiscale Formalism, in "IEEE Transactions on Audio, Speech and Language Processing", December 2014, https://hal.inria.fr/hal-01059345
- [4] V. KHANAGHA, K. DAOUDI, H. YAHIA. Efficient and robust detection of Glottal Closure Instants using Most Singular Manifold of speech signals, in "IEEE Transactions on Acoustics Speech and Signal Processing", 2014, forthcoming, https://hal.inria.fr/hal-00802014
- [5] S. K. Maji, H. Yahia. Edges, Transitions and Criticality, in "Pattern Recognition", January 2014, http://hal. inria.fr/hal-00924137
- [6] O. PONT, A. TURIEL, 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, pp. 35-61
- [7] O. PONT, A. TURIEL, H. YAHIA. Singularity analysis in digital signals through the evaluation of their Unpredictable Point Manifold, in "International Journal of Computer Mathematics", 2012, http://hal.inria.fr/hal-00688715
- [8] O. PONT, H. YAHIA, R. DUBOIS, M. HAÏSSAGUERRE. A Singularity-analysis Approach to characterize Epicardial Electric Potential, in "Computing in Cardiology", 2012, http://hal.inria.fr/hal-00750003
- [9] O. PONT, H. YAHIA, B. XU. Arrhythmic dynamics from singularity analysis of electrocardiographic maps, in "EBMC'13-The 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society", Osaka, Japan, IEEE, July 2013 [DOI: 10.1109/EMBC.2013.6610955], http://hal.inria.fr/hal-00823856
- [10] J. SUDRE, H. YAHIA, O. PONT, V. GARÇON. Ocean Turbulent Dynamics at Superresolution From Optimal Multiresolution Analysis and Multiplicative Cascade, in "IEEE Transactions on Geoscience and Remote Sensing", June 2015, vol. 53, n^o 11, 12 p. [DOI: 10.1109/TGRS.2015.2436431], https://hal.inria.fr/hal-01166170
- [11] A. TURIEL, H. YAHIA, 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

- [12] B. Xu, S. Jacquir, G. Laurent, S. Binczak, O. Pont, H. Yahia. *In Vitro Arrhythmia Generation by Mild Hypothermia a Pitchfork Bifurcation Type Process*, in "Physiological Measurement", January 2015, 15 p., https://hal.inria.fr/hal-01076401
- [13] 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, pp. 3591-3604, http://dx.doi.org/10.1016/j.patcog.2010.04.011

Publications of the year

Articles in International Peer-Reviewed Journals

- [14] R. Balti, M. B. Mansour, N. Sayari, L. Yacoubi, L. Rabaoui, N. Brodu, A. Masse. *Development and characterization of bioactive edible films from spider crab (Maja crispata) chitosan incorporated with Spirulina extract*, in "International Journal of Biological Macromolecules", July 2017 [DOI: 10.1016/J.IJBIOMAC.2017.07.046], https://hal.archives-ouvertes.fr/hal-01571302
- [15] N. BRODU. Super-resolving multiresolution images with band-independant geometry of multispectral pixels, in "IEEE Transactions on Geoscience and Remote Sensing", May 2017, pp. 1 8, https://arxiv.org/abs/1609. 07986 [DOI: 10.1109/TGRS.2017.2694881], https://hal.archives-ouvertes.fr/hal-01555993
- [16] F. FRAPPART, L. BOURREL, N. BRODU, X. R. SALAZAR, F. BAUP, J. DARROZES, R. POMBOSA. Monitoring of the Spatio-Temporal Dynamics of the Floods in the Guayas Watershed (Ecuadorian Pacific Coast) Using Global Monitoring ENVISAT ASAR Images and Rainfall Data, in "Water", 2017, vol. 9, pp. 12 31 [DOI: 10.3390/w9010012], https://hal.archives-ouvertes.fr/hal-01425021
- [17] V. J.-L. RALAIARISOA, A. VALANCE, N. BRODU, R. DELANNAY. *High speed confined granular flows down inclined: numerical simulations*, in "EPJ Web of Conferences", 2017, vol. 140, 03081 p. [DOI: 10.1051/EPJCONF/201714003081], https://hal-univ-rennes1.archives-ouvertes.fr/hal-01580442
- [18] Ö. YELEKÇI, G. CHARRIA, X. CAPET, G. REVERDIN, J. SUDRE, H. YAHIA. Spatial and Seasonal Distributions of Frontal Activity over the French Continental Shelf in the Bay of Biscay, in "Continental Shelf Research", June 2017, vol. 144, pp. 65–79 [DOI: 10.1016/J.CSR.2017.06.015], https://hal.inria.fr/hal-01548361

International Conferences with Proceedings

- [19] K. DAOUDI, N. VINUESA. An analysis of psychoacoustically-inspired matching pursuit decompositions of speech signals, in "International Conference on Natural Language, Signal and Speech Processing", Casablanca, Morocco, December 2017, https://hal.inria.fr/hal-01627106
- [20] A. EL AOUNI, K. DAOUDI, H. YAHIA, K. MINAOUI. *Surface Mixing and Biological Activity in The North African Upwelling*, in "AGU Ocean Sciences Meeting 2018", Portland, Oregon, United States, February 2018, https://hal.inria.fr/hal-01627958

Other Publications

[21] F. GOULAMHOUSSEN SATSOU. *Déconvolution de signaux vocaux*, Université de bordeaux, September 2017, https://hal.inria.fr/hal-01627891

- [22] G. LI. Speech analysis for the differential diagnosis between Parkinson's disease, progressive supranuclear palsy and multiple system atrophy, Université Paris Saclay, September 2017, https://hal.inria.fr/hal-01627868
- [23] Q. ROBIN. Détection automatique des instants de fermeture glottale dans les voixpathologiques, INP-Grenoble/Phema, September 2017, https://hal.inria.fr/hal-01627875
- [24] J. SUDRE, H. M. YAHIA, C. MAES, V. GARÇON. *Effect of wind stress forcing on ocean dynamics at Air-Sea Interface*, May 2017, Frontiers in ocean-atmosphere exchange: Air sea interface and fluxes of mass and energy, Poster, https://hal.inria.fr/hal-01522024

References in notes

- [25] A. ARNEODO, F. ARGOUL, E. BACRY, J. ELEZGARAY, J. F. MUZY. *Ondelettes, multifractales et turbulence*, Diderot Editeur, Paris, France, 1995
- [26] N. BOCCARA. Modeling Complex Systems, Springer, New-York Dordrecht Heidelberg London, 2010
- [27] G. BOFFETTA, M. CENCINI, M. FALCIONI, A. VULPIANI. Predictability: a way to characterize complexity, in "Physics Report", 2002, vol. 356, pp. 367–474, arXiv:nlin/0101029v1, http://dx.doi.org/10.1016/S0370-1573(01)00025-4
- [28] N. Brodu. *Reconstruction of epsilon-machines in predictive frameworks and decisional states*, in "Advances in Complex Systems", 2011, vol. 14, no 05, pp. 761–794
- [29] N. Brodu, H. Yahia. Stochastic Texture Difference for Scale-Dependent Data Analysis, in "arXiv preprint arXiv:1503.03278", 2015
- [30] P. CASTIGLIONE, M. FALCIONI, A. LESNE, A. VULPIANI. *Chaos and coarse graining in statistical mechanics*, Cambridge University Press Cambridge, 2008
- [31] M. COSTA, A. L. GOLDBERGER, C.-K. PENG. *Multiscale entropy analysis of complex physiologic time series*, in "Physical review letters", 2002, vol. 89, n^o 6, 068102 p.
- [32] J. P. CRUTCHFIELD. Between order and chaos, in "Nature Physics", 2012, vol. 8, n^o 1, pp. 17–24
- [33] J. P. CRUTCHFIELD, K. YOUNG. *Inferring statistical complexity*, in "Physical Review Letters", 1989, vol. 63, n^o 2, 105 p.
- [34] G. FALASCO, G. SAGGIORATO, A. VULPIANI. About the role of chaos and coarse graining in statistical mechanics, in "Physica A: Statistical Mechanics and its Applications", 2015, vol. 418, pp. 94–104
- [35] D. P. FELDMAN, J. P. CRUTCHFIELD. *Measures of statistical complexity: Why?*, in "Physics Letters A", 1998, vol. 238, n^o 4-5, pp. 244–252
- [36] P. FLANDRIN. Time-frequency/time-scale analysis, Academic press, 1998, vol. 10
- [37] P. FLANDRIN. Time-frequency/Time-scale analysis, Academic Press, San Diego, USA, 1999

- [38] M. GELL-MANN. *What is Complexity?*, John Wiley and Sons, Inc., 1995, vol. 1, 1 p., http://complexity.martinsewell.com/Gell95.pdf
- [39] G. GOERG, C. SHALIZI. *Mixed LICORS: A nonparametric algorithm for predictive state reconstruction*, in "Artificial Intelligence and Statistics", 2013, pp. 289–297
- [40] R. M. GRAY. Entropy and information theory, Springer Science & Business Media, 2011
- [41] 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^o 5, pp. 6103-6110, http://link.aps.org/doi/10.1103/PhysRevE.62.6103
- [42] P. IVANOV, L. AMARAL, A. GOLDBERGER, S. HAVLIN, M. ROSENBLUM, Z. STRUZIK, H. STANLEY. *Multifractality in human heartbeat dynamics*, in "Nature", 1999, vol. 399, pp. 461–465
- [43] H. JANICKE, A. WIEBEL, G. SCHEUERMANN, W. KOLLMANN. *Multifield visualization using local statistical complexity*, in "IEEE Transactions on Visualization and Computer Graphics", 2007, vol. 13, n^o 6, pp. 1384–1391
- [44] K. Judd, A. Mees. *Embedding as a modelling problem*, in "Non Linear Optimization", 1998, vol. 20, n^o 3-4, pp. 273-286, http://www.sciencedirect.com/science/article/pii/S016727899800089X
- [45] A. LESNE. Méthodes de renormalisation. Phénomènes critiques, chaos, structures fractales, Eyrolles, Paris, 1996, ISBN 2-212-05830-6
- [46] M. MARTIN, A. PLASTINO, O. ROSSO. Generalized statistical complexity measures: Geometrical and analytical properties, in "Physica A: Statistical Mechanics and its Applications", 2006, vol. 369, n^o 2, pp. 439–462
- [47] 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
- [48] 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
- [49] O. PONT, A. TURIEL, C. PEREZ-VICENTE. Empirical evidences of a common multifractal signature in economic, biological and physical systems, in "Physica A, Elsevier", 2009, n^o 388, pp. 2025–2035, http://hal.inria.fr/inria-00438521
- [50] C. R. SHALIZI, K. L. SHALIZI. Blind construction of optimal nonlinear recursive predictors for discrete sequences, in "Proceedings of the 20th conference on Uncertainty in artificial intelligence", AUAI Press, 2004, pp. 504–511
- [51] Z. S. She, E. Leveque. *Universal scaling laws in fully developed turbulence*, in "Physical Review Letters", 1994, vol. 72, pp. 336–339

- [52] H. STANLEY. Introduction to phase transitions and critical phenomena, Oxford Science publications, Oxford, UK, 1987
- [53] J. L. STARCK, F. MURTHAG, J. FADILI. Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity, Cambridge University Press, 2010, ISBN:9780521119139
- [54] F. TAKENS. *Detecting Strange Attractors in Turbulence*, in "Non Linear Optimization", 1981, vol. 898, pp. 366–381, http://www.springerlink.com/content/b254x77553874745/
- [55] A. TURIEL, A. DEL POZO. *Reconstructing images from their most singular fractal manifold*, in "IEEE Trans. on Im. Proc.", 2002, vol. 11, pp. 345–350
- [56] A. TURIEL, 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, no 1, pp. 362–390
- [57] A. TURIEL, C. PEREZ-VICENTE. *Role of multifractal sources in the analysis of stock market time series*, in "Physica A", September 2005, vol. 355, pp. 475–496
- [58] G. WEISBUCH. Complex Systems Dynamics, Addison Wesley Longman, 1991