



HAL
open science

Contributions to Nonstationary Spectrum Estimation and Stationarity Tests in the Time-Frequency Plane

Jun Xiao

► **To cite this version:**

Jun Xiao. Contributions to Nonstationary Spectrum Estimation and Stationarity Tests in the Time-Frequency Plane. Signal and Image processing. Ecole normale supérieure de Lyon, 2008. English. NNT: . tel-02167626

HAL Id: tel-02167626

<https://theses.hal.science/tel-02167626>

Submitted on 27 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : 460

N° attribué par la bibliothèque : 07ENSL0 460

THESE

en vue d'obtenir le grade de

Docteur de l'Université de Lyon - Ecole Normale Supérieure de Lyon
en cotutelle avec l'East China Normal University, Shanghai

Spécialité : Physique

Laboratoire de Physique

Ecole doctorale de Physique



présentée et soutenue publiquement le 11/06/2008

par Madame Jun XIAO

Titre :

Contributions to Nonstationary Spectrum Estimation and Stationarity Tests in the Time-Frequency Plane

Directeurs de thèse : Messieurs Patrick FLANDRIN et Zhengqi ZHENG

Après avis de : Monsieur Lotfi SENHADJI, Membre/Rapporteur

Monsieur Ming LI, Membre/Rapporteur

Devant la commission d'examen formée de :

Monsieur Patrick FLANDRIN Membre

Monsieur Pierre BORGNAT Membre

Monsieur Christian JUTTEN Président

Monsieur Lotfi SENHADJI Membre/Rapporteur

Monsieur Ming LI Membre/Rapporteur

Monsieur Zhengqi ZHENG Membre

Abstract

The analysis and processing of nonstationary signals has already led to the development of numerous time-frequency methods and algorithms. The purpose of this Thesis is to contribute new approaches in this area. More precisely, the reported work consists in: (i) improving representation tools for exploratory signal analysis in nonstationary contexts; and (ii) testing stationarity of a signal relatively to an observation scale.

Existing time-frequency tools still call for improvements in terms of nonstationary spectrum estimation, in particular in the case of chirp signals embedded in nonstationary noise. The first part of the Thesis addresses such an issue, with the two-fold objective of a sharp localization for chirp components and a reduced level of statistical fluctuations for the noise background. The technique consists in combining time-frequency reassignment with multitapering, with two variations. The first one, primarily aimed at nonstationary spectrum estimation, is based on sums of reassignment-based estimates with different tapers, whereas the second one makes use of differences between the same estimates for a sake of chirp enhancement. The principle of the technique is outlined, its implementation based on Hermite functions is justified and discussed, and some typical examples (including an application to the problem of Euler's disk) are provided for supporting the efficiency of the approaches, both qualitatively and quantitatively.

Turning to stationarity, the second part of the Thesis proposes to go beyond the classical definition that refers to a strict invariance of statistical properties over time, and to develop an operational framework for testing stationarity, in a wide sense, relatively to an observation scale in both stochastic and deterministic contexts. The statistical test is based on a comparison between local and global time-frequency features and the originality is to make use of a family of stationary surrogates generated from the original signal under test for defining the null hypothesis of stationarity. Within this general framework, two different approaches are proposed. The first one takes advantage of suitably chosen distance measures between local and global spectra and characterizes the null hypothesis of stationarity by constructing a parametric model that is derived from the distribution of surrogates variances. The second approach is implemented by a one-class Support Vector Machine operating on signal features extracted from the time-frequency plane, with those of its stationary surrogates serving as a learning set. The principle of the test and its two variations are presented, some results are shown on typical examples of signals (including speech) which can be considered as stationary or nonstationary, depending on the chosen observation scale.

Key words: time-frequency (TF), nonstationary spectrum, chirp, multitapering, reassignment, stationarity test, Support Vector Machines.

Résumé

L'analyse et le traitement des signaux non-stationnaires a déjà conduit au développement de nombreux algorithmes et méthodes temps-fréquence. Le but de la Thèse est de contribuer à des approches neuves dans ce domaine. Plus précisément, les travaux décrits ici consistent à : (i) améliorer les outils de représentation pour l'analyse des données exploratoires dans le contexte non-stationnaire ; et (ii) tester la stationnarité d'un signal relativement à une échelle d'observation.

L'estimation de spectre non-stationnaire par les outils temps-fréquence existants laisse encore la possibilité d'améliorations, en particulier dans le cas de signaux de type chirps quand s'ajoute un bruit non-stationnaire. La première partie de la Thèse s'intéresse à cette question, avec l'objectif double d'une localisation précise pour les composantes chirp et d'un niveau réduit de fluctuations statistiques pour le bruit. La technique consiste à combiner la réallocation temps-fréquence avec la méthode multi-fenêtrage. Deux variantes sont discutées. La première, visant principalement à l'estimation de spectre non-stationnaire, est basée sur les sommes des estimées réallouées calculées avec des fenêtres différentes, tandis que la deuxième utilise les différences entre ces mêmes estimées avec comme objectif un rehaussement des chirps par rapport au bruit. Le principe de la technique est expliqué, sa mise en œuvre basée sur les fonctions d'Hermite est justifiée et discutée et quelques exemples typiques (y compris une application au problème de disque Euler) sont fournis pour illustrer l'efficacité des approches, à la fois qualitativement et quantitativement.

Abordant la question de stationnarité, la deuxième partie de la Thèse propose d'aller au-delà de la définition classique qui postule une invariance stricte de propriétés statistiques au cours du temps, en développant un cadre opérationnel pour tester la stationnarité, dans un sens moins strict, relativement à une échelle d'observation et dans des contextes à la fois stochastique et déterministe. Un test statistique est construit sur une comparaison entre des attributs temps-fréquence locaux et globaux. L'originalité est d'utiliser une famille de substituts stationnaires générés à partir du signal original à tester, pour définir l'hypothèse nulle de stationnarité. S'appuyant sur ce cadre général, deux approches différentes sont proposées. La première est d'employer des mesures de la distance entre les spectres locaux et le spectre global, puis de caractériser l'hypothèse nulle de stationnarité par un modèle paramétrique dérivé de la distribution des propriétés des substituts. La deuxième approche est mise en œuvre à l'aide de Machines à Vecteur Support Une-Classe, fonctionnant sur les attributs du signal, extraits du plan temps-fréquence. Les attributs des substituts servent alors d'ensemble d'apprentissage de l'hypothèse nulle. Le principe du test et ses deux formulations sont présentés, quelques résultats sont montrés sur des exemples typiques des signaux (en particulier de parole) qui peuvent être considérés comme stationnaires ou non-stationnaires selon l'échelle d'observation choisie.

Mots clefs : temps-fréquence, spectre non-stationnaire, chirp, multi-fenêtrage, réallocation, test de stationnarité, Machines à Vecteur Support.

摘要

非平稳信号的分析和处理已经引起了大量的时频方法和算法的发展，而本论文的目的是希望为该领域的研究提供一些新的方法。具体一点来讲，论文工作主要针对以下两方面的研究内容：（1）改进非平稳信号分析的表象工具；（2）检验信号相对于某一观察范围的平稳性。

就非平稳光谱评估而言，现存的时频工具仍然需要得到进一步的改善，尤其是对于啁啾信号嵌入在噪声中的情况。因此，论文第一部分的工作主要讨论解决的就是该问题，最终的目标是一方面能够加强啁啾成分的局域性，另一方面又能降低噪声背景的统计波动水平。技术的关键在于结合“时频再分配技术”和“多窗技术”。根据这一思想，我们提出了两种不同的方法：第一个方法是基于对使用不同窗口产生的再分配评估求和，主要针对的是非平稳光谱评估；而第二个方法是基于对同样的（不同窗口产生的再分配）评估求差，主要是为了进一步加强啁啾信号。方法的技术原理在文中具体阐述，基于厄密多函数的执行被讨论论证，并且论文也提供了一些典型的例子（包括欧拉盘问题）来支持方法的有效性。

再说平稳性，本论文第二部分的工作在于打破经典的平稳性定义，开发具有实际运作意义的平稳性检验。经典的平稳性是指信号的统计属性随时间严格不变，而本论文中的广义平稳性则是相对于某个观察范围而言的，并且同时适用于统计和确定性的情况。该平稳性检验基于局域时频特性与全局时频特性的比较，原创性地使用了由原始被测信号产生的平稳的“替代品”来定义平稳性的零假设。基于这个框架，两种不同的方法被提出：第一个方法是适当地选择局域谱和全局谱的“距离”度量，构建一个来源于替代品属性分布的参数模型，从而定义平稳性零假设；而第二个方法是从时频平面上提取信号特征量，交由一类支持向量机执行，而把替代品的这一特征量作为训练集。检验的原理和具体方法在文中详细阐述，并且对一些典型的相对于不同观察范围下平稳性情况不同的信号（包括语音信号）进行了测试。

关键字：时频，非平稳光谱，啁啾，多窗，再分配，平稳性检验，支持向量机。

Contents

I. Introduction	11
1 Introduction to Stationarity and Nonstationarity	11
1.1 Classical Definition of Stationarity	11
1.2 Generalized Notion of Stationarity	12
1.3 Issues in Representation of Nonstationary Signals	13
2 Introduction to Time-Frequency Method	14
2.1 Fourier Transform	14
2.2 Characteristics of Time-Frequency Representation	15
2.3 Short-Time Fourier Transform	16
2.4 Wigner-Ville Distribution	16
2.5 Wigner-Ville Spectrum	18
3 Reassignment	19
3.1 Spectrogram	19
3.2 Reassigned Spectrogram	20
3.3 Some Historical Comments	22
4 Multitapering	22
4.1 Power Spectrum Density	22
4.2 Welch Method of Averaged Periodograms	23
4.3 Thomson's Method of Multitapering	23
4.4 Extensions of Multitapering	25
II. Multitaper Time-Frequency Reassignment	27
1 Multitaper Time-Frequency Reassignment for Nonstationary Spectrum Estimation	28
1.1 Principle and Implementation	28
1.1.1 Wedding multitapering with reassignment	28
1.1.2 Choice of tapers	28
1.1.3 Effectiveness measure by Renyi Entropy	29
1.2 Performance Evaluation: Error Measure	31
1.3 Examples	34
2 Variations for Chirp Enhancement	36
2.1 Principle and Implementation	37
2.1.1 Differences between estimates based on different tapers	37

2.1.2	Final estimate combined by "sums" and "differences"	40
2.2	Performance Evaluations: Contrast Measure	40
2.3	Examples	42
3	Application: Euler's Disk	44
3.1	Time-Frequency Estimation	46
3.2	Parameter Estimation of Chirp by Hough Transform	47
3.3	Performance Evaluation on Synthetic Model	49
3.4	Test on Experimental Signal	51
4	Conclusion	54
III. Test of Stationarity with Surrogates		55
1	Time-Frequency Framework: Comparison of Local vs. Global Features	56
2	Surrogates: Stationary Reference	57
2.1	Surrogates in <i>(Non)stationarity</i> Test	57
2.2	Surrogates in <i>(Nonlinear)</i> Physics	59
3	Stationarity Test - Parametric Model Approach	60
3.1	Comparison of Local vs. Global Time-Frequency Spectra	60
3.2	Choice of Distance	60
3.2.1	Test signals	60
3.2.2	Probability laws distances	61
3.2.3	Spectral distances	62
3.2.4	Combined distance	62
3.3	Structure of the Test	62
3.4	Null Hypothesis of Stationarity	64
3.4.1	Distribution based on surrogates	64
3.4.2	Gamma model?	64
3.4.3	How many surrogates needed?	66
3.4.4	Threshold based on a statistical significance	69
3.4.5	Verification of null hypothesis	69
3.5	Index and Scale of Nonstationarity	72
3.6	Examples	72
4	Stationarity Test - One-Class Support Vector Machine Approach	77
4.1	One-Class Support Vector Machines	77
4.1.1	Machine learning	77
4.1.2	Support Vector machines	78
4.1.3	One-class Support Vector machines	78
4.2	Comparison of Local vs. Global Time-Frequency Spectra	80
4.3	Examples	81
5	Application: Speech	84
6	Conclusion	87
IV. Conclusions and Perspectives		89

V. Appendix	91
1 Alternative model	91
2 Choice of Thresholds	94
3 Variations of Test	97
Bibliography	105
Acknowledgment	111

I. Introduction

As an introduction, Chapter I describes the general concepts and definitions, the context of the topics and the history of their developments, as well as the relative approaches and techniques to be used in the work of the Thesis.

1 Introduction to Stationarity and Nonstationarity

Considering *stationarity* is central in many signal processing applications, either because its assumption is a pre-requisite for applying most of standard algorithms devoted to steady-state regimes, or because its breakdown conveys specific information in evolutive contexts. Testing stationarity is therefore an important issue, but addressing it raises some difficulties. The main reason is that the concept itself of “stationarity”, while uniquely defined in theory, is often interpreted in different ways.

1.1 Classical Definition of Stationarity

According to the classical definition, a random signal is called “stationary” if its statistical properties are invariant under a translation in time, i.e. if they are not dependent on an absolute time.

Although such a concept of “stationarity” could be defined for all orders, practically the most important properties of a stationary signal are those of second order, which means only the statistical properties of the first and second orders should be invariant under a translation in time.

In the wide sense, a random signal can be defined as stationary, if (i) it stipulates on the one side

$$\mathbb{E}\{x(t)\} = \mu_x, \quad (\text{I.1})$$

which means that its expectation value is a constant μ_x independent of the time, and on the other side (ii)

$$\mathbb{E}\{x(t)x^*(s)\} = \gamma_x(t-s), \quad (\text{I.2})$$

which means that its autocovariance function depends only on the difference of the two considered instants. (One can assume μ_x to be zero for simplicity without loss of generality and the autocovariance function is the same.)

1.2 Generalized Notion of Stationarity

Indeed, whereas the standard definition of stationarity refers only to *stochastic processes* and concerns the invariance of statistical properties over time, stationarity is also usually invoked for *deterministic signals* whose spectral properties are time-invariant. Moreover, while the underlying invariances (be they stochastic or deterministic) are supposed to hold in theory for *all* times, common practice allows them to be restricted to some *finite* time interval, possibly with abrupt changes in between.

Mandelbrot reconsiders, in the introduction at his first monograph on the “fractals” [54, 55], the notion of dimension by giving it explicitly a *relative* character:

“[...] physical dimension inevitably has a subjective basis. It is a matter of approximation and therefore of degree of resolution. To confirm this last hunch, we will take up an object [...] namely, a ball of 10 cm diameter made of a thick thread of 1 mm diameter. Depending on one’s viewpoint, it possesses (in a latent fashion) several distinct physical dimensions. Indeed, at the resolution possible to an observer placed 10 m away, it appears as a point, that is, as a zero-dimensional figure. At 10 cm it is a ball, that is, a three-dimensional figure. At 10 mm it is a mess of threads, that is, a one-dimensional figure. At 0.1 mm each thread becomes a sort of column and the whole becomes a three-dimensional figure again. At 0.01 mm resolution, each column is dissolved into filiform fibers, and the ball again becomes one-dimensional, and so on, with the dimension jumping repeatedly from one value to another.”

What is talked about here is the issue of dimension, but this idea can be extended to a more general range, in particular to the concept of *stationarity*. By analogy, we can consider now as an example a signal of speech which, seen in long term of a scale of several seconds, is often cited as a prototype of nonstationary signal, with a succession of segments (voiced, unvoiced, silences, etc.) with different characteristics. The same signal however, if one looks at it in short term of a scale of several dozens of milliseconds, will be willingly considered as stationary (in particular in their voiced parts) and amenable to analysis and coding methods (spectral analysis, linear prediction, etc.) resting on such a hypothesis. Clearly, just as the notion of dimension whose values would have had no meaning without the relativity to an observation scale, we can see that the definition of stationarity is also subjected to the relative considerations: a same signal, such as speech, may be described as nonstationary in long term and stationary in short term.

It seems therefore that the standard concept of stationarity is not *operational* in the sense that the practical consideration should include the heuristic adjustment emphasizing the notion of locality which, in the principle, is opposite to the notion of permanence associated with the idea of stationarity. Moreover, the empirical interpretation of stationarity is freed from the theoretical framework of the

stochastic signals by putting the stationary random signals together with the deterministic ones in a same unity. As long as it turns out to be the enough repetition of a certain motif — no matter it is random or not — without presenting too much variability all through the observation, a signal will be probably called stationary. Testing the stationarity of an observed signal is thus a question which deserves to be specified by both a relative sense and a statistical sense.

Several attempts in this direction can be found in the literature, mostly based on concepts such as *local stationarity* [59]. Most of them however share the common philosophy of comparing statistics of adjacent segments, with the objective of detecting changes in the data [27, 44] and/or segmenting it over homogeneous domains [47] rather than addressing the aforementioned issue. Early attempts have nevertheless been made in this direction too by contrasting local properties with global ones [49, 50], but not necessarily properly phrased in terms of hypothesis testing.

This will be the purpose of our work in Chapter III but, in order to properly address this issue, the way nonstationary signals have to be represented deserves some attention.

1.3 Issues in Representation of Nonstationary Signals

Nonstationary spectrum estimation is one of the necessary step in nonstationary signal analysis and processing (e.g. stationarity test) and the choice of an efficient representation is crucial for the ultimate task, because a good representation can offer much desirable information about the structure of signal under analysis without external specifications. As far as only second order evolutions are to be tested, *time-frequency* (TF) distributions and spectra are natural tools [30].

It is worthwhile to mention that Fourier-based methods of TF spectrum estimation are classically faced with intrinsic limitations and two different kinds of trade-offs. On the one hand, from a geometrical perspective, as the methods operate locally by employing windows to guarantee a form of local stationarity, they are necessarily faced with a trade-off between the temporal and the frequential localization in the case of chirp-like signals, called "*time-frequency trade-off*". The interpretation of these two antagonistic resolutions is : the time resolution gets better when the window of analysis becomes shorter, however, the frequency resolution degrades at the same rate, considering the Fourier analysis being confined to the same short time-window. Conversely, analysis by a filter bank with more selective filters has a better frequency-resolution, but leads to a bad time-resolution because the impulse responses of the filters have a longer duration.

On the other hand, from a statistical point of view, the usual "*bias-variance trade-off*" inherent to any estimation procedure is amplified when analyzing nonstationary stochastic processes by the fact that time-averaging, which is aimed at reducing variance, introduces some bias not only in the frequency direction but also in time, i.e. the smaller the variance, the larger the bias.

Such difficulties have been recognized long ago, and numerous studies have tried to address the problems. As far as localization is concerned, Wigner-based approaches have been developed and shown to outperform windowed (Fourier or wavelet-based) methods, at least in the case of noise-free single chirps [30]. In more realistic situations of multi-chirps however, a dramatic improvement over both Fourier and Wigner-based methods has come from the use of the *reassignment* technique [4]. Turning to the estimation issue in a statistical sense, different attempts have been made to take advantage of the idea of *multitapering*, pioneered by D.J. Thomson in a stationary setting [69], and to extend it to nonstationary situations [19, 33], thanks to which an improved statistical stability can be obtained without a time-averaging step.

In the following of this chapter, we will therefore introduce in Section 2 the development of the TF methods for nonstationary signals, and detail particularly the reassignment and multitapering techniques respectively in Section 3 and Section 4, which will pave the way for the presentation of our work in Chapter II — *multitaper time-frequency reassignment* with the two-fold purpose of a sharp localization for chirp components and a reduced level of statistical fluctuations for noise, solving simultaneously the two trade-offs mentioned above.

2 Introduction to Time-Frequency Method

In nonstationary context, TF joint representation, which is a distribution on a two-dimensional (t, f) space, is important and necessary, as it overcomes the drawbacks of conventional time-domain representation and frequency-domain representation by supplying wealthy complementary information of signal, based on which a large number of methods of TF analysis has been continuously developed in the last years (cf. [30]).

2.1 Fourier Transform

For an analyzed signal, two classical representations are the time-domain representation $x(t)$ and the frequency-domain representation $X(f)$ (by *Fourier transform* (FT) of $x(t)$)

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-i2\pi ft} dt. \quad (\text{I.3})$$

FT plays an important role in signal analysis and processing, because it supplies the universal concept of frequency, an essential complement to the exclusively temporal description by $x(t)$, for a better comprehension of the underlying phenomena, and based on which, a lot of algorithms and processors have been developed for frequency analysis of signal.

However, FT does not indicate how the frequency content of a signal changes over time and it is thus only suitable to the study of *stationary* signals where all frequencies have an infinite coherence time. This can be mathematically explained

by (I.3), in which the computation of any value $X(f)$ for one frequency f needs the complete history of the signal ranging from $-\infty$ to ∞ . Conversely, the *inverse Fourier transform* (IFT)

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{i2\pi ft} df \quad (\text{I.4})$$

shows any value $x(t)$ of a signal at one instant t is regarded as an infinite superposition of everlasting and completely nonlocal waves. Obviously, the two variables t and f in both representations are treated as mutually exclusive: to obtain a representation in terms of one variable, the other variable is “infinitely integrated out”. This is too far from the physical reality, so neither of them is sufficient for analyzing real signals, especially in nonstationary context.

Without going to the details, it should be pointed out that there are two mathematical impossibilities resulting from the FT: the first is the *Heisenberg-Gabor uncertainty principle* [38, 13] (expressed by the inequality: $\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$), which shows that a signal cannot be concentrated on arbitrarily small time-frequency regions; the second is the theory of *Slepian-Pollak-Landau* [67, 46] which shows that a signal cannot confine its total energy to finite intervals in the time and frequency domain, no matter how large these intervals might be.

A break out of the strick sense of Fourier analysis would be some local quantities such as *instantaneous frequency* (IF) $f(t)$ [75, 14, 15], which shows the frequential evolution of signal with time. However, the IF is only suitable in quasi-monochromatic situations. In other cases, for example, a signal with two or more FM components, each of which has its own IF, these IF features could not be described by the global IF $f(t)$. Similar situations happen in the inverse function of the IF: *the group delay* (GD) $t(f)$, which shows temporal evolution of signal with frequency.

2.2 Characteristics of Time-Frequency Representation

As the conventional representations in either the time domain or the frequency domain are inadequate in the nonstationary context, a way out is to explore a representation of signal as a function of the two variables or a distribution in the two-dimensional (t, f) space, where the variables t and f are no longer mutually exclusive, but are present together. Such a representation is called *time-frequency (TF) representation*.

With the two variables t and f simultaneously at hand, the two-dimensional space of the TF representation is wealthy in different complementary information, compared to the representation with only one of them, by reading their cross-relations.

First of all, the two-dimensional (t, f) space gives the meaning of a truly joint plane in time and frequency, a concept of no longer “twice in one dimension”, but

instead “once in two dimensions”. The TF duality is important to nonstationary signals, as it underlies all of their relative descriptions.

The second complementary information is its description of the frequency as a function of time, corresponding to the idea of an evolutionary spectral analysis. Superior to the IF mentioned above which has limitations in multicomponent signals, the TF representation supplies a free space, with a whole range frequency values for any instant t , for different IFs to lay out simultaneously.

Dual to the second, the last information is the description of the time as a function of the frequency, connected with the idea of a sequential monitoring of the output of different frequency channels. Consequently, a complete history is offered for each frequency, which permits us to access to events located in time in a frequency-by-frequency manner.

With the many desirable characteristics and advantages, the TF representation is indispensable in the analysis and processing of nonstationary signals.

2.3 Short-Time Fourier Transform

Defining a TF representation for a nonstationary process $\{x(t), t \in \mathbb{R}\}$ is a question that has no unique answer. Among a multitude of solutions, the easiest way may consist in introducing a variable t in the expression of FT, thus leading to *short-time Fourier transform* (STFT) [30]

$$F_x^{(h)}(t, f) = \int_{-\infty}^{+\infty} x(s) h(s-t) e^{-i2\pi fs} ds, \quad (\text{I.5})$$

where the localization of analysis around the instant t is obtained by an arbitrary window $h(t)$.

STFT considers implicitly a non-stationary signal as a succession of quasi-stationary situations on the scale of the short-time window $h(t)$. By such an analysis, the resolution in time is fixed by the window length, while the resolution in frequency is fixed by its FT. Since these two resolutions are antagonist, there exists a necessary trade-off between them: for a highly nonstationary signal, a good resolution in time is necessary, in which case a narrow window of analysis $h(t)$ should be applied, limiting the resolution in frequency; by contrast, if a precise resolution in frequency is required, a wide window should be used, which however results in a bad resolution in time. Such a trade-off in resolution can not be avoided for linear TF approaches.

2.4 Wigner-Ville Distribution

If one emphasizes the concept of adaptable window for STFT, a natural choice would be to take the window as $h(t) = x(-t)$ so as to get the *Wigner-Ville distri-*

tribution (WVD) [76, 75, 16]

$$W_x(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau. \quad (\text{I.6})$$

By construction, the WVD is bilinear in the signal. It is one of the simplest prototype of most other bilinear TF distributions [30].

This definition was proposed by Ville [75] with the motivation of its analogy with a probability density function and it is actually nothing else but the FT in an “acceptable form” of the characteristic function of TF energy distribution. Considering the academic example of a linear chirp

$$x(t) = \exp\left\{i2\pi\left(\frac{at^2}{2} + bt + c\right)\right\}, \quad (\text{I.7})$$

one can show that its WVD ideally localizes to the instantaneous frequency by

$$W_x(t, f) = \delta(f - (at + b)), \quad (\text{I.8})$$

a situation contrary to the STFT.

Despite the desirable property of perfect localization, the WVD suffers some shortcomings. It may attain large negative values, which prevents it from being used as a probability density. Furthermore, due to its bilinearity, it obeys a principle of quadratic (thus nonlinear) superposition

$$W_{x+y}(t, f) = W_x(t, f) + W_y(t, f) + 2 \operatorname{Re}W_{xy}(t, f), \quad (\text{I.9})$$

where

$$W_{xy}(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) y^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau. \quad (\text{I.10})$$

This results in the existence of spurious features called cross-terms, or interferences in WVD, which appear midway between any two signal components in the case of multicomponent signals as well as between any two sub-parts of a monocomponent signal.

On the one side, cross-terms hamper the readability of a TF representation, especially when the components are numerous or close to each other. From this point of view, cross-terms are troublesome and often regarded as the fundamental limitation on the applicability of quadratic TF methods.

On the other side however, the good localization of the WVD for linear chirps can also be viewed as a by-product of interferences. A geometric explanation of the property [9] is that: if a signal has a WVD which is perfectly localized on a given curve of the TF plane, it is necessary that, according to the interference geometry principle, all of the mid-points between any two points on the curve belong to the same curve. It follows that linear chirps are the only admissible solutions in the WVD case, since straight lines are the only curves of the plane that are the geometric locus of all of their mid-points. This advantage may be extended to

other quadratic TF distributions and justifies their localization on nonlinear chirps by suitable modification.

Although being the simplest bilinear TF distribution, WVD plays an important role in the TF analysis because of its good properties and a great deal of work has been devoted later on to the improvement and extensions of the initial definition (I.6).

2.5 Wigner-Ville Spectrum

Let's now introduce the *Wigner-Ville Spectrum* (WVS) [48, 51, 30], whose definition reads:

$$\mathbf{W}_x(t, f) = \int_{-\infty}^{+\infty} \mathbb{E} \left\{ x \left(t + \frac{\tau}{2} \right) x^* \left(t - \frac{\tau}{2} \right) \right\} e^{-i2\pi f\tau} d\tau, \quad (\text{I.11})$$

where $\mathbb{E}\{\cdot\}$ stands for the expectation operator. This definition of a time-varying spectrum, though not unique, presents the advantage of extending the usual concept of *Power Spectrum Density* (PSD) and making it time-dependent in a rather natural way. Without entering into much details, it is worth recalling that the WVS reduces to the PSD at every instant if the analyzed process happens to be stationary. Moreover, it satisfies the important property of preserving time and frequency supports, and its marginal distributions are directly related to meaningful quantities (variance in time, Loève's distribution function in frequency).

Actually, it can be shown that, under mild conditions, the WVS of a process is nothing else but the ensemble average of the WVDs of its possible realizations:

$$\mathbf{W}_x(t, f) = \mathbb{E} \{ W_x(t, f) \}. \quad (\text{I.12})$$

Given one observed realization of a nonstationary process, estimating the WVS amounts to find a substitute for the unattainable ensemble average entering (I.12). One standard way is to assume for $x(t)$ a form of local stationarity in both time and frequency, i.e., some locally slow evolution of the WVS in the two directions. Such an assumption paves the road for a replacement of the ensemble average at a given TF location by a local smoothing over a neighbouring domain. This idea can be formalized by introducing a TF smoothing kernel $\Pi(t, f)$ and defining as a WVS estimator the quantity [30]:

$$\hat{\mathbf{W}}_x(t, f) = \int \int_{-\infty}^{+\infty} W_x(s, \xi) \Pi(s - t, \xi - f) ds d\xi = C_x(t, f; \Pi). \quad (\text{I.13})$$

This expression coincides with a general class of TF estimators called *Cohen's class* $C_x(t, f; \Pi)$ [23, 24, 25, 30] for the observed realization, which indicates that different smoothing kernel results in different TF distribution, in other words, each of the other bilinear TF distributions can be expressed as a smoothed WVD by using a specific smoothing kernel.

Compared with the WVD (I.6), the introduction of the smoothing kernel $\Pi(t, f)$ in WVS (I.13) will on the one hand, help reduce the cross terms but on the other hand, lead to a new trade-off between fluctuations and localization: If we consider for instance the toy example of a linear chirp embedded in broadband noise, the fluctuations of the WVD due to noise and the interferences between two components of signals will be smoothed out provided that $\Pi(t, f)$ is low-pass. However, the WVD of the linear chirp (which has the unique property of being perfectly localized along the instantaneous frequency) will be smoothed out too. A way out is however possible by reconsidering the apparently contradictory issues of fluctuations reduction and localization at the light of the two refinements offered by *reassignment* and *multitapering*.

3 Reassignment

Reassignment is a nonlinear technique which is an efficient means of getting sharply localized TF distribution [4, 31]. The technique consists in moving the value of a TF distribution from the point where it has been computed to a new location which is more representative of the local signal energy distribution.

3.1 Spectrogram

Begin with the *spectrogram* [58] of a signal $x(t)$ with window $h(t)$, which is usually rather expressed as:

$$S_x^{(h)}(t, f) = \left| F_x^{(h)}(t, f) \right|^2 = \left| \int_{-\infty}^{+\infty} x(s) h(s-t) e^{-i2\pi fs} ds \right|^2, \quad (\text{I.14})$$

where $F_x^{(h)}(t, f)$ stands for the STFT.

This definition of spectrogram as a squared STFT is explicitly dependent on the short-time window $h(t)$, aimed at guaranteeing local stationarity by limiting its estimation to some specified neighborhood of the instant t . The window may be thought of as a measurement tool, the resulting spectrogram depending not only on the signal but also on the window. Spectrograms are therefore faced with the same TF resolution trade-off as in the case of STFTs.

An alternative interpretation of the spectrogram could be a smoothed WVD, rather than as a squared STFT:

$$S_x^{(h)}(t, f) = C_x(t, f; W_h) = \iint_{-\infty}^{+\infty} W_x(s, \xi) W_h(s-t, \xi-f) ds d\xi, \quad (\text{I.15})$$

with the smoothing kernel $\Pi(t, f)$ in (I.13) chosen to be the WVD of some signal $h(t)$ supposed to be well localized in both time and frequency, a property that carries over to its WVD.

Such an interpretation gives the clue for improving upon its localization limitations. Indeed, if we recall that the WVD of a linear chirp perfectly localizes on a TF line, the spreading of any corresponding spectrogram just comes from the fact that, when centering the analysis window at some TF point (t, f) that does not belong to this line, a non-zero contribution is nevertheless obtained as long as the line passes through the local TF window (whose joint support cannot be made arbitrarily small). Therefore, an important point of view from the smoothing formula in (I.15) is that the value on a given TF point of spectrogram cannot be considered as pointwise, but instead results from the summation of a continuum of WVD contributions within the TF domain selected by the window.

3.2 Reassigned Spectrogram

Reasoning by a mechanical analogy identifying energy with mass, the situation is as if a whole distribution of mass within a domain (here, the TF window) would be replaced by one single number (the total mass) assigned to the *geometrical* center of the domain. Such an assignment is clearly not well adapted to situations where the distribution is not uniform over the domain, a much more meaningful assignment being the *center of mass* within the domain.

This is precisely the essence of the “reassignment” technique [4, 31], which consists in

(i) evaluating for each TF location, not only the integrated signal WVD within the TF domain of the window (in other words, the spectrogram value at this point), but also the center of mass of the signal WVD;

(ii) reassigning the spectrogram value to this location.

In the idealized case where one single linear chirp intersects the TF window, it is clear that the center of mass necessarily belongs to the line along which the WVD is localized, thus guaranteeing a perfect localization of the spectrogram after its reassignment.

Defining the local centers of mass of the WVD distribution W_x within the domain selected by the TF window W_h centered in each TF point (t, f) as $(\hat{t}_{t,f}, \hat{f}_{t,f})$, where

$$\begin{cases} \hat{t}_{t,f} &= \frac{1}{S_x^{(h)}(t,f)} \iint_{-\infty}^{+\infty} s W_x(s, \xi) W_h(s-t, \xi-f) ds d\xi, \\ \hat{f}_{t,f} &= \frac{1}{S_x^{(h)}(t,f)} \iint_{-\infty}^{+\infty} \xi W_x(s, \xi) W_h(s-t, \xi-f) ds d\xi, \end{cases} \quad (\text{I.16})$$

previous studies of Auger and Flandrin [4, 31] have shown that an efficient evaluation of the local centroids $(\hat{t}_{t,f}, \hat{f}_{t,f})$ can be made implicitly, according to

$$\begin{cases} \hat{t}_{t,f} &= t + \text{Re}\{F_x^{(\mathcal{T}h)}(t, f)/F_x^{(h)}(t, f)\}, \\ \hat{f}_{t,f} &= f - \text{Im}\{F_x^{(\mathcal{D}h)}(t, f)/F_x^{(h)}(t, f)\}. \end{cases} \quad (\text{I.17})$$

where the two additional windows needed in the computation are defined from the mother window $h(t)$ as $(\mathcal{T}h)(t) = th(t)$ and $(\mathcal{D}h)(t) = (dh/dt)(t)$.

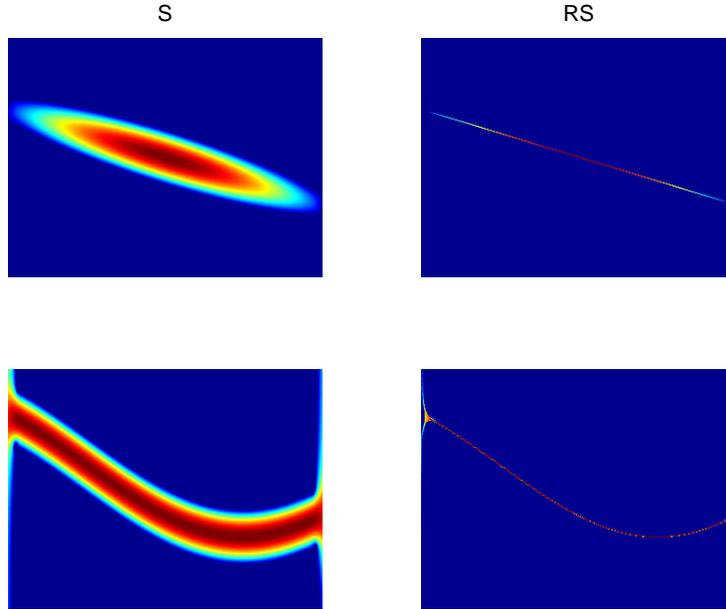


Fig. I.1: **Reassignment.** This figure displays two examples of TF spectrograms before (left column) S and after reassignment (right column) RS . In each diagram, time is horizontal, frequency vertical. In the case of a linear frequency modulation (top row), the reassignment yields a perfect alignment with the instantaneous frequency law. Roughly, the same happens for nonlinear modulations (with an example of a sinusoidal instantaneous frequency displayed on the bottom row), assuming that the chosen short-time window permits a local linear approximation of the modulation.

Given the field of all above centroids, the reassigned spectrogram $RS_x^{(h)}(t, f)$ attached to the conventional spectrogram $S_x^{(h)}(t, f)$ follows as:

$$RS_x^{(h)}(t, f) = \int \int_{-\infty}^{+\infty} S_x^{(h)}(s, \xi) \delta(t - \hat{t}_{s, \xi}) \delta(f - \hat{f}_{s, \xi}) ds d\xi. \quad (\text{I.18})$$

Conceptually, the reassignment can be summed up to a two-step process: first, a smoothing makes the interference terms disappear but in return smears the localized components; second, a squeezing helps concentrate the distributions to one single point so as to offer a perfect localization, even for nonlinear chirps.

To have a demonstration, Fig. I.1 illustrates the effectiveness of reassignment technique for linear and nonlinear chirp signals.

3.3 Some Historical Comments

Indeed, the presentation of the reassignment principle above is a modern way [4] of revisiting an older idea [42, 43] of Kodera et al., whose argument of “modified moving window method” was initially related to phase, instead of TF smoothing. The centroids for reassignment in (I.16) happen to be related to the phase of the STFT, an information ignored by a spectrogram as a squared STFT. To be more precise, the local centroids (within the TF domain selected by the window W_h centered in a given point (t, f)) in (I.16) is equivalent to the *local instantaneous frequency* (IF) and *group delay* (GD) of signal :

$$\begin{cases} \hat{t}_{t,f} &= -\frac{1}{2\pi} \frac{\partial}{\partial f} \varphi(t, f), \\ \hat{f}_{t,f} &= f + \frac{1}{2\pi} \frac{\partial}{\partial t} \varphi(t, f) \end{cases} \quad (\text{I.19})$$

with $\varphi(t, f)$ being the phase of the STFT. From this interpretation, the rationale for reassignment was to favor energy concentrations in the vicinity of local IFs and GDs.

Although this idea was introduced in the mid-1970s and has been applied successfully in some fields like geophysics, it retained little attention on the TF community during the next 20 years. The main reason is that the methodology had been introduced much before the problems it addressed were extensively studied (during the mid-1980s), and that its applicability remained computationally limited for a while.

Not until the mid-1990s did the idea of reassignment revive [4] with an improved algorithm circumventing an explicit computation of the phase of the STFT. A simple manipulation of previous equations yields the result that the local centroids for the reassignment can be equivalently characterized by (I.17), which amounts to employing three window functions instead of one.

Finally, it should be mentioned that based on empirical studies, despite being lack of a result theoretically justified, the reassignment works on the noise as well, thus leading to a higher level of fluctuations of TF spectrum for the noise compared to that without reassignment. Therefore, solving the TF localization trade-off, reassignment leaves the statistical problem in nonstationary spectrum estimation to multitapering technique.

4 Multitapering

4.1 Power Spectrum Density

In the case of stationary processes, the spectral characterization is fully described by means of the PSD, which could be thought of as:

$$\mathbf{S}_x(f) = \lim_{T \rightarrow \infty} \mathbb{E} \left\{ \frac{1}{T} \left| \int_{-T/2}^{+T/2} x(t) e^{-i2\pi ft} dt \right|^2 \right\}. \quad (\text{I.20})$$

In practice, the above quantity is unattainable when only one realization of finite duration is given. The Squared Fourier Transform (SFT) of a single observation is however a crude, non consistent estimator, whose variance is of the order of the squared PSD [57], thus calling for better approximations of (I.20).

4.2 Welch Method of Averaged Periodograms

Since an improvement can only come from averaging (almost) uncorrelated estimations, an ergodic argument suggests to chop the observation into (almost) disjoint blocks and average their SFTs (a procedure sometimes referred to as the *Welch method of averaged periodograms*).

Adopting the notation of spectrograms, it turns out that the corresponding (Welch) estimator can be written as:

$$\hat{\mathbf{S}}_{x,W}(f) = \frac{1}{K} \sum_{k=1}^K S_x^{(h)}(t_k, f), \quad (\text{I.21})$$

where the spacing $\Delta = t_{k+1} - t_k$ between adjacent t_k 's is of the order of the window length.

Assuming that this spacing ensures an approximate decorrelation between blocks, one can expect that the variance is inversely proportional to the number of blocks K (i.e., roughly T/Δ for an observation of duration T). Although variance can be decreased this way by increasing K , the finite duration constraint necessarily leads to shorten Δ , increasing in turn the bias in frequency since a window of duration Δ has a frequency resolution of the order of $1/\Delta$.

4.3 Thomson's Method of Multitapering

In order to circumvent this trade-off, D.J. Thomson suggested [69] a powerful multiple window spectrum estimator, called *multitaper*, which is to still average SFTs stemming from (almost) uncorrelated sequences in order to reduce variance, but to construct such sequences by using for each of them the whole data set so as to not sacrifice bias.

The way to achieve this program consists in projecting the observation on a family of basis functions $\{h_k(t), k \in \mathbb{N}\}$ that are orthonormal over the observation interval $(-T/2, +T/2)$. This results in a (Thomson) estimator that can be written as:

$$\hat{\mathbf{S}}_{x,T}(f) = \frac{1}{K} \sum_{k=1}^K S_x^{(h_k)}(0, f), \quad (\text{I.22})$$

with the number of windows (tapers) K and all windows $\{h_k(t), k = 1, \dots, K\}$ extending over $(-T/2, +T/2)$.

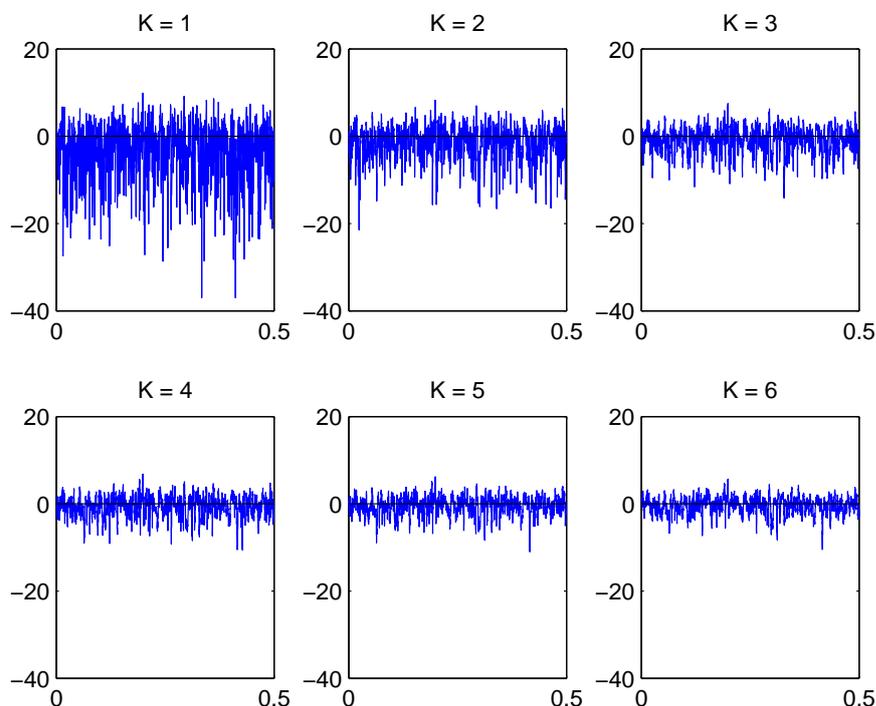


Fig. 1.2: **Multitapering.** Each diagram displays the multitaper spectrum of a same white Gaussian noise with different number of Hermite tapers ($K = 1 : 6$). It shows that the fluctuation of spectrum decreases with the increase of K .

Assuming that the spectrum is flat over a given bandwidth B associated with the expected frequency resolution, the basis can be obtained as the family of orthonormal functions (on the given time interval) that maximize their energy in the given frequency band. The solution to this problem is given by the family of functions known as the *Prolate Spheroidal Wave Functions* or, in a discrete-time setting, as the *Discrete Prolate Spheroidal Sequences* (DPSS) [57]. As windows, they are perfectly suited to stationary spectrum estimation, because they are simultaneously compactly supported in time and optimally concentrated in frequency.

In order to circumvent the statistical trade-off in spectrum without degrading the resolution of line components, Thomson introduced a separate pre-estimate for deterministic sinusoidal components before an average over multiple windows. His original method of multitapering focuses therefore on the estimation of *stationary*

signals which have the form:

$$x(t) = y(t) + \sum_i \mu(f_i) e^{j2\pi f_i t + \rho_i} \quad (\text{I.23})$$

with y a zero-mean, stationary, Gaussian random process having a continuous power spectrum and $\mu(f_i) e^{j2\pi f_i t + \rho_i}$ a sinusoid (stationary deterministic line component) having a line power spectrum. This method consists of three main steps:

1. Detecting and extracting all significant sinusoids from the data x by making use of a statistical significance test to obtain an estimate y

$$y = x - \{\text{sinusoids}\} \quad (\text{I.24})$$

2. Multitapering the spectrum estimate of the sinusoid-free data y by using the family of functions DPSS

$$\hat{\mathbf{S}}_{y,T}(f) = \frac{1}{K} \sum_{k=1}^K S_y^{(h_k)}(0, f), \quad (\text{I.25})$$

(The concentration of the prolate windows in frequency results in a low bias spectrum estimate).

3. Reshaping the spectrum $\hat{\mathbf{S}}_{y,T}$ to account for the excised sinusoids and obtain the final estimate.

4.4 Extensions of Multitapering

Extending Thomson's approach of multitapering to *nonstationary* situations is appealing. The main reason is that the inconsistency (and large variance) of a PSD estimator based on a crude SFT directly carries over to spectrograms considered as WVS estimators. The traditional way out would be to smooth over time and frequency, but at the expense of further increasing bias. In this respect, resorting to multitapers allows for a variance reduction with a bias that only sticks to the common length of the different windows. This is certainly an improvement as compared to (smoothed) spectrograms with respect to statistical efficiency.

In recent years, there have been many attempts at extending Thomson's multitapering to nonstationary signals. One of the natural ways is *multitaper spectrogram*, which computes and averages multiple spectrograms (I.14) by using a family of orthogonal functions, such as DPSS. Essentially, multitaper spectrogram is to make the estimator (I.22) time-dependant [33, 71, 19] by using sliding windows:

$$\hat{\mathbf{S}}_{x,T}(f) \rightarrow S_{x,K}(t, f) = \frac{1}{K} \sum_{k=1}^K S_x^{(h_k)}(t, f), \quad (\text{I.26})$$

where the $\{S_x^{(h_k)}(t, f), k = 1, \dots, K\}$ stand for the K spectrograms computed with the K first orthogonal functions given as a family of short-time windows $\{h_k(t), k \in \mathbb{N}\}$:

$$S_x^{(h_k)}(t, f) = \left| \int x(s) h_k(s - t) e^{-i2\pi f s} ds \right|^2. \quad (\text{I.27})$$

This is a segment estimate indexed by time and the effect is to fit a locally stationary process observed for the duration of the segment.

Further extension [6] has been proposed by Bayram and Baraniuk to make multitapering more suitable for its application in the TF plane. Due to the fact that DPSS have no inherent optimality properties in the joint TF domain, the resulting spectral estimate is not optimum in terms of localization in the two degrees of freedom. The problem was solved however by making use of another family of orthogonal functions: *Hermite functions* (HF) [26], which maximally concentrate in TF domains with elliptic symmetry. The effectiveness of multitapering by means of HF is illustrated in Fig. I.2, which shows the multitaper spectra of a white Gaussian noise, for different number of tapers ($K = 1 : 6$). It is obvious that the statistical fluctuation of noise decreases with the increase of K , which achieves the purpose of multitapering.

Moreover, the chirp rates of the line components pre-estimated in Thomson's method for keeping the resolution must be very low so as to be well approximated as piecewise sinusoidal, so a similar but improved method of identifying and extracting *chirp line components* before applying multitapering was proposed in [6]: first, the chirp line components are extracted from the time-varying data. Then, the spectrogram estimates of the remaining chirp-free data is multitapered by using HF. Finally, the extracted chirp line components are added back to reshape the time-varying spectrum.

In spite of the improvement and the effectiveness, there are limitations to this approach: on the one side, the method assumes that no more than one chirp is present within the analysis region; on the other side, chirp extraction is computationally expensive and fails for crossing chirps. Therefore, we propose in Chapter II a better and simpler way of trading-off TF resolution and statistical stability, that is, wedding multitapering with reassignment. Two variations of the combined technique will be presented in this chapter.

II. Multitaper Time-Frequency Reassignment

As already mentioned, in nonstationary contexts, Fourier-based methods of (time-varying) spectrum estimation [30] are classically faced with a statistical “bias-variance trade-off” and a geometrical “TF localization trade-off”. Solving these two trade-offs is the key for improving the estimation of nonstationary spectrum, since the former prevents the reduction of the level of statistical fluctuations for noise and the latter affects the localization for chirp components. As shown in Chapter I, the modern way of the *reassignment* technique [4] improved dramatically the TF localization with an efficiency that is limited to the cases where the signal-to-noise ratio is high enough; and the appearance of the multitapering technique [69] and its extensions to nonstationary situations [6] improved statistical stability of nonstationary spectrum without a time-averaging step.

When extended in a direct way, the “classical” method of multitapering alone suffers however still from the TF localization trade-off, so some attempts have been made to circumvent this limitation by identifying chirp-like components and excising them prior applying the multitaper machinery [6]. The purpose of this Chapter is to avoid such a complication and to rather combine multitapering (for a sake of fluctuation reduction of noise) with reassignment (for localization of chirp signals) [77, 78].

The general concepts of reassignment and multitapering having been recalled in Chapter I, we will present directly here in this chapter two variations of combination of both ideas.

More precisely, a first combination is discussed in Section 1, whose aim is to improve statistical stability while preserving TF localization [77, 78]. This is achieved by summing estimates based on different tapers, the rationale being that such estimates tend to behave as well-localized, weakly correlated, surrogate data whose sum combines coherently chirp components and incoherently noise contributions. The principle and actual implementation are detailed in Subsection 1.1, while performance evaluations by error measure is provided in Subsection 1.2, together with a typical example in Subsection 1.3, for supporting the efficiency of the approach.

A companion perspective is then envisioned in Section 2 for the purpose of chirp enhancement [78]. The principle in Subsection 2.1 is to exploit differences between estimates, the idea being in this case to get rid of noise by masking those

regions where different tapers lead to significantly different estimates. Quantitative performance evaluations are provided in Subsection 2.2 for illustrating and supporting the novel approach, and possibilities and limitations are illustrated and compared in Subsection 2.3 by a number of numerical experiments.

Finally, the combined technique is applied on the problem of Euler's disk (whose mechanism is to estimate the instantaneous frequency of a spun disk, which is just a chirp signal) in Section 3; conclusions and some of the many possible variations and extensions are briefly outlined in Section 4.

1 Multitaper Time-Frequency Reassignment for Nonstationary Spectrum Estimation

1.1 Principle and Implementation

1.1.1 Wedding multitapering with reassignment

What we propose here is to adopt the same strategy of multitapering on spectrograms as in (I.26), but applied to reassigned spectrograms, i.e., to consider as a WVS estimator the quantity:

$$RS_{x,K}(t, f) = \frac{1}{K} \sum_{k=1}^K RS_x^{(h_k)}(t, f), \quad (\text{II.1})$$

where the $\{RS_x^{(h_k)}(t, f), k = 1, \dots, K\}$ stand for the K reassigned spectrograms computed with the first K short-time windows $h_k(t)$. The rationale for this approach can be justified in a twofold way:

(i) as far as chirp components are concerned, reassignment increases localization in a way that can be made independent of the window, thus permitting (II.1) to act as a *coherent* averaging;

(ii) in noise regions on the contrary, the same windows lead to uncorrelated surrogate data whose TF distributions are different, (II.1) acting in this case as a form of *incoherent* averaging tending to smooth the estimate.

1.1.2 Choice of tapers

In stationary spectrum estimation, multitapers are chosen as DPSSs because the data is of finite duration and estimation concerns frequency only. In the nonstationary case, there is no *a priori* reason to dissymmetrize time and frequency by choosing tapers that would be perfectly localized in the time domain rather than in frequency. Indeed, it makes much more sense to fully exploit the two degrees of freedom offered by the TF plane and, as suggested in [6], to rather pick up those functions that maximally concentrate in TF domains with elliptic symmetry. As shown in [26] in the context of coherent states and in [29] within a Wigner

framework, those functions are the *Hermite* functions (HF), whose definition is given by

$$h_k(t) = \left((t - \mathcal{D})^k g \right) (t) / \sqrt{\pi^{1/2} 2^k k!}, \quad (\text{II.2})$$

with $g(t) = \exp\{-t^2/2\}$. In practice, HFs can be computed recursively, according to

$$h_k(t) = g(t) H_k(t) / \sqrt{\pi^{1/2} 2^k k!}, \quad (\text{II.3})$$

where the $\{H_k(t), k \in \mathbb{N}\}$ stand for the Hermite polynomials that obey the recursion:

$$H_k(t) = 2t H_{k-1}(t) - 2(k-2) H_{k-2}(t), k \geq 2 \quad (\text{II.4})$$

with the initialization $H_0(t) = 1$ and $H_1(t) = 2t$.

Not only the HFs are orthonormal, but they also guarantee a perfect localization of the corresponding reassigned spectrograms in the case of a linear chirp, whatever k . This can be easily understood by noting that the WVD of a HF (which is basically a 2D Laguerre function) has elliptic symmetry [6, 29]. Recalling that the WVD is covariant with respect to dilations and rotations, it is enough to check that reassignment ends up with a perfect localization in the case of a pure tone, which can be done by an elementary calculation.

In the context of reassignment, HFs offer one further advantage, as compared to DPSSs. In the standard implementation of spectrogram reassignment, only the mother window $h(t)$ has to be given and the two additional windows $(\mathcal{T}h)(t)$ and $(\mathcal{D}h)(t)$ that are needed are evaluated numerically [31]. This may cause difficulties when differentiating tapers whose order k is large, since they are highly oscillating. This problem can be circumvented when using HFs since their successive derivatives also obey a recursion that can be explicitly plugged in the algorithm, namely

$$(\mathcal{D}h_k)(t) = (\mathcal{T}h_k)(t) - \sqrt{2(k+1)} h_{k+1}(t). \quad (\text{II.5})$$

1.1.3 Effectiveness measure by Renyi Entropy

The effectiveness of this implementation of multitaper reassignment is illustrated in Fig. II.1 which evaluates the spreading of cumulative Hermite estimates in the case of a linear chirp and of a white Gaussian noise, with the same cases by DPSS estimates illustrated as well for comparison. The measure used for this evaluation is a Rényi entropy of order p , defined as [5]

$$R_p(P) = \frac{1}{1-p} \log_2 \sum_{n=1}^N \sum_{m=1}^N (P[n, m])^p, p > 0, \quad (\text{II.6})$$

for any normalized discrete TF distribution $P[n, m]$ with N points in time and N frequency bins.

In the case where the distribution is perfectly localized on, say, the diagonal of the plane (the situation expected to happen when reassigning the spectrogram

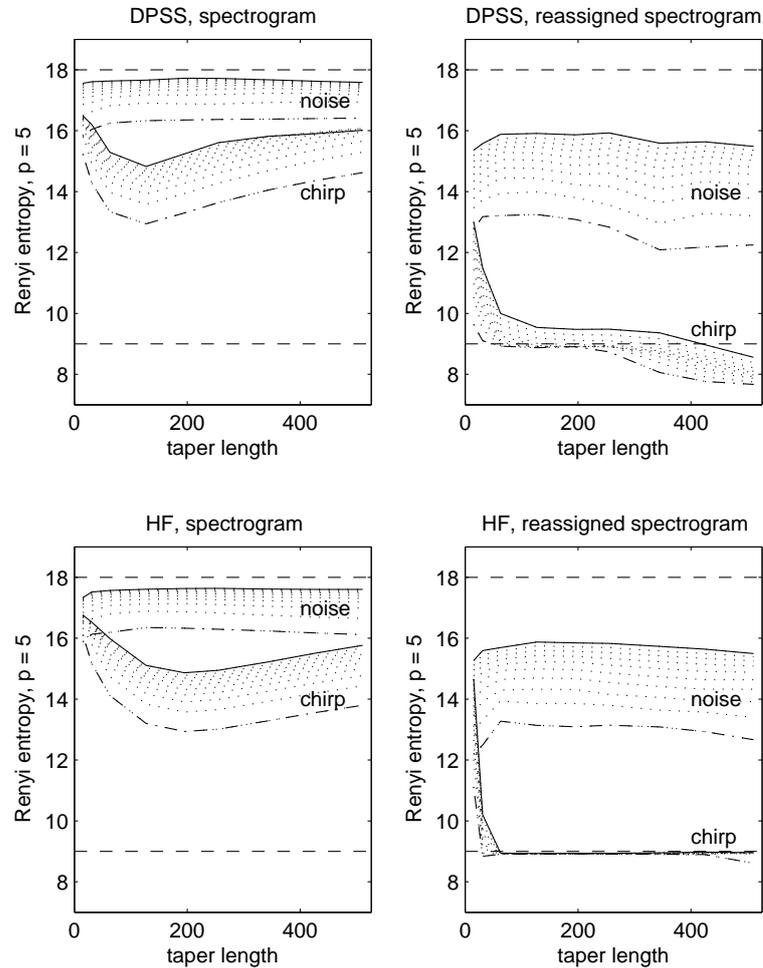


Fig. II.1: **Spreading of Hermite multitaper estimates.** Each diagram (left column: spectrogram; right column: reassigned spectrogram) displays the Rényi entropies R_5 of (top row: DPSS; bottom row: Hermite) multitaper TF estimates, as a function of the taper length, and parameterized by the maximum taper order K , from $K = 1$ (dashed dotted line) to $K = 10$ (full line). A linear chirp signal and a sample of white Gaussian noise (512 points each) are considered, and the theoretical predictions corresponding to a perfect localization and a uniform spreading ($R_5 = 9$ and 18, respectively, see text) are superimposed as dotted lines.

of a linear chirp), we have ideally $P_{chirp}[n, m] = \delta_{n,m}/N$ and $R_p(P_{chirp}) = \log_2 N$. This situation contrasts with that of a white Gaussian noise whose distribution is expected to spread uniformly all over the plane as $P_{noise}[n, m] = 1/N^2$, thus leading to $R_p(P_{noise}) = 2 \log_2 N$.

In the multitaper spectrogram case (without reassignment), Fig. II.1 evidences that the spreading (of either DPSS or Hermite estimates) can approach the theoretical prediction ($R_5 = 18$) in the noise case by increasing the maximum taper order K , but at the expense of increasing at the same time the spreading in the chirp case, with a minimum value (for a taper length matched to the chirp rate) that remains significantly large. This contrasts with multitaper reassigned spectrograms which compare favourably with the theoretical prediction for chirps. Especially, the superiority of Hermite estimates over DPSS ones is obviously shown by their perfect superimposition over the theoretical prediction ($R_5 = 9$) for all Hermite tapers over a wide range of orders and lengths. Comparison is less favourable in the noise case, but it is continuously improved when increasing K , without much degrading by DPSS estimates and with no degrading at all by Hermite estimates the corresponding localization for chirps.

1.2 Performance Evaluation: Error Measure

In order to go beyond the numerical check described above, a simple example concerned with the idealized situation of a bandpass filtered white Gaussian noise within a time-limited support is provided for performance evaluation. Although not strictly attainable (because of the uncertainty relation), the model $M_D(t, f)$ for the WVS of such an observation is the indicator function of a rectangle of area D within the TF plane.

Fig. II.2 illustrates what happens in such a case by comparing the WVD and a sample (reassigned) spectrogram with the corresponding multitaper estimates based on $K = 10$ Hermite functions. Ideally, the estimate should be smooth over the domain of rectangle defined and zero outside. It is clear that by multitapering alone the estimate is smoothed but spreads outside the rectangle, whereas by reassignment alone it is well localized in the domain but suffers large fluctuations. The two effects of reduced fluctuations and support preservation are clearly traded-off in the estimate with the combined techniques, and ensemble averages (based on 10 independent estimates) are also provided for supporting the effectiveness of the approach and its improved convergence rate as compared to an empirical estimate of the WVS.

Fig. II.2 gives a qualitative account of the behaviour of the method, that can be supplemented by the quantitative measure

$$E(K) = \frac{1}{\|M_D\|_1} \int \int_{-\infty}^{+\infty} |\hat{\mathbf{W}}_{x,K}(t, f) - M_D(t, f)| dt df, \quad (\text{II.7})$$

where $\hat{\mathbf{W}}_{x,K}(t, f)$ stands for the WVS estimate ($S_{x,K}(t, f)$ or $RS_{x,K}(t, f)$), the L_1 -

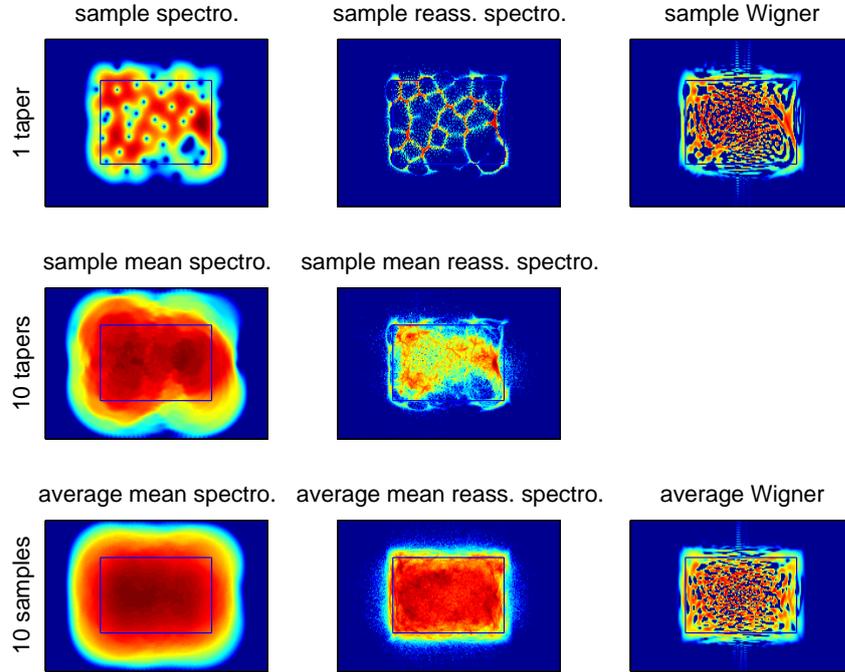


Fig. II.2: **Comparison of noise WVS estimates.** Each diagram represents a WVS estimate in the case of a white Gaussian noise limited in time and frequency within the superimposed rectangular domain. The first row consists of a spectrogram, its reassigned version and the WVD, based on one realization. The corresponding multitaper estimates (10 Hermite functions) are given in the middle row, whereas the bottom row displays ensemble averages of such estimates (10 independent realizations), together with the empirical WVS estimate on the same data set. In each diagram, time is horizontal, frequency vertical and the energy is coded logarithmically with a dynamic range of 30 dB. (Ideally, the estimate should be smooth over the domain of rectangle and zero outside.)

norm being here chosen so as to put emphasis on the localization in the estimates.

Fig. II.3 presents results with different domains, all rectangular and centered in the analyzed TF region, but with different areas D . In the pure white Gaussian noise case where the model support extends over the whole plane (in this case, $D = 256$), we observe for both spectrograms and reassigned spectrograms that the *error measure* behaves asymptotically as $E(K) \propto K^{-1/2}$ when using K tapers.

In the spectrogram case, this can be justified since, for each taper, the values

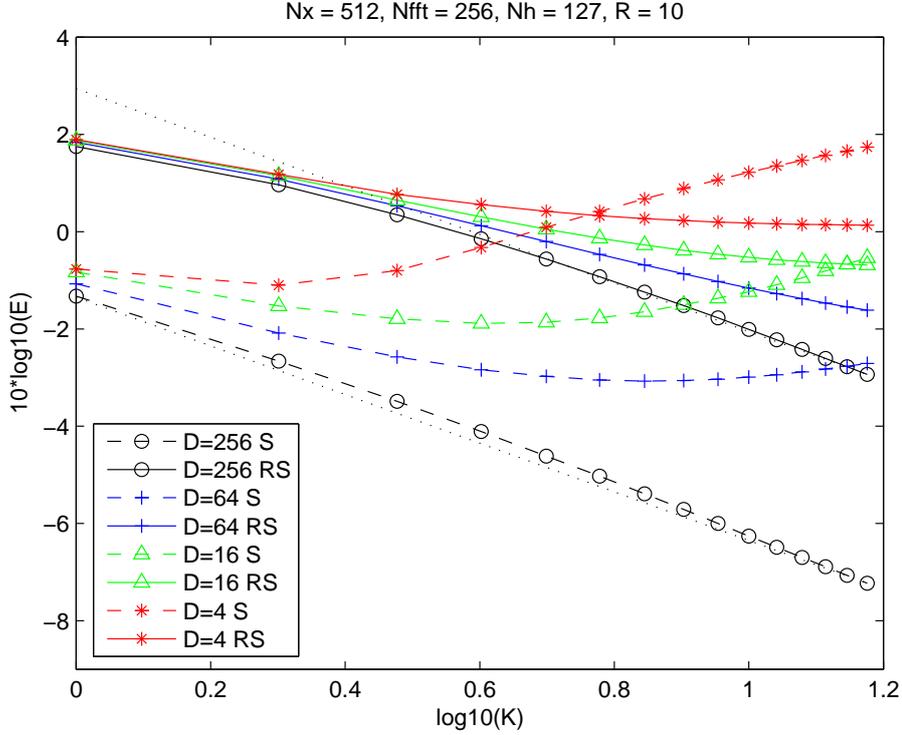


Fig. II.3: **Error measures in WVS multitaper estimates.** The figure plots, as a function of the number of tapers K , the error measure (II.7) attached to multitaper (reassigned) spectrograms when the model is a white Gaussian noise limited in time and frequency over a rectangular domain of area D . The simulations have been conducted (with up to $K = 15$ Hermite tapers, each of length 127) on the basis of 10 independent realizations of 512 data points each, with 256 frequency bins over the whole frequency range $[0, 1/2)$. In the pure white Gaussian noise situation ($D = 256$), asymptotic decays in $K^{-1/2}$ (see text) have been superimposed as dotted lines.

are known to be χ^2 distributed with 2 degrees of freedom [57]. It follows from the orthogonality of the tapers that the sum of the K first Hermite spectrograms is also χ^2 distributed, but with $2K$ degrees of freedom. Given a χ^2 distribution with ν degrees of freedom:

$$p_\nu(y) = \frac{y^{\nu/2-1} e^{-y/2} 2^{-\nu/2}}{\Gamma(\nu/2)}, y \geq 0, \quad (\text{II.8})$$

its absolute mean deviation (the quantity on which the performance measure (II.7) is based) can be evaluated by using properties of the incomplete Gamma function.

A simple calculation ends up with the result:

$$\int_0^\infty |y - \nu| p_\nu(y) dy = \frac{4(\nu/2)^{\nu/2} e^{-\nu/2}}{\Gamma(\nu/2)}, \quad (\text{II.9})$$

and for large ν 's, we can apply Stirling's formula from which we get the asymptotic behavior:

$$\lim_{\nu \rightarrow \infty} \frac{1}{2} \sqrt{\frac{\pi}{\nu}} \int_0^\infty |y - \nu| p_\nu(y) dy = 1. \quad (\text{II.10})$$

Specifying this result to the case where $\nu = 2K$, we therefore justify that the sum of K spectrograms based on orthogonal tapers has an absolute mean deviation which varies as $K^{-1/2}$ for large K 's, leading to the $K^{-1/2}$ behavior for the mean $S_{x,K}(t, f)$. Although no proof is available yet, the experiments reported in Fig. II.3 evidence a similar behaviour for multitaper reassigned spectrograms $RS_{x,K}(t, f)$, but with a higher level of fluctuations.

However, when the area of the domain D is reduced, the situation evolves quite differently for the two estimates: on the one hand, $S_{x,K}$ is smoother than $RS_{x,K}$; on the other hand, $RS_{x,K}$ is essentially confined to the non-zero support of the model, whereas $S_{x,K}$ spreads outside this domain. The criterion (II.7) can therefore be viewed as a measure of a trade-off between a "bias" term (due mostly to the existence of non-zero values outside the model domain) and a "variance" term (related to non-constant values within the domain). In the case of pure white Gaussian noise, no bias in the aforementioned sense enters the measure and $E(K)$ decreases when K is increased. When D becomes smaller, the variance reduction is balanced by a bias term increasing with K , since the higher the order of an Hermite taper, the larger its TF support. This analysis applies to the spectrogram but not to the reassigned spectrogram, since squeezing smeared spectra concentrates the error on the fluctuations term which globally decreases always the same way. This is illustrated in Fig. II.3, evidencing eventually crossings indicating that multitaper spectrograms may be outperformed by their reassigned counterparts when localized components are to be analyzed.

Actually, the evaluation could be also carried out with an admissible model of a Gaussian modulated Gaussian bandpass filtered white Gaussian noise, which is detailed in Appendix (Chapter V: Section 1).

1.3 Examples

To give an example, we consider now in Fig. II.4 the case already discussed in [6] and [19], with both a (nonlinear) chirp component and a (bandpass) time-varying noise. Several different WVS estimates have been compared: the simplest WVD suffers too much interferences and has a large variance; the spectrogram, thanks to its smoothing kernel, is free from the cross-terms but has a poor resolution; and although its reassigned version outperforms the two formers by providing a perfect localization of the chirp, it suffers a big statistical fluctuation in

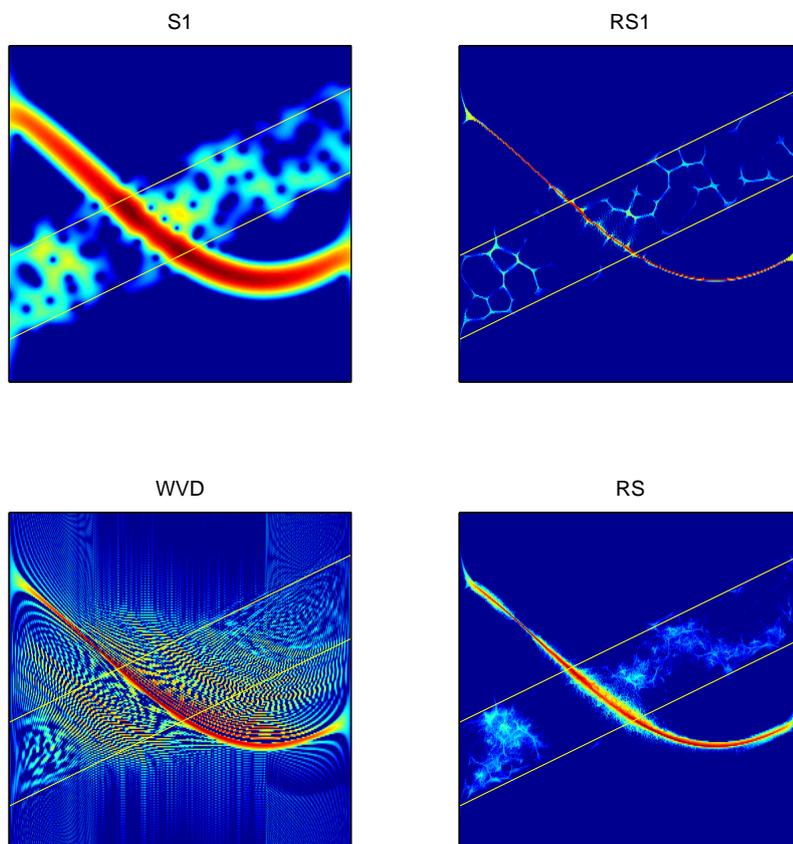


Fig. 11.4: **Comparison of signal + noise WVS estimates.** Each diagram represents a WVS estimate in the case of a nonlinear chirp signal embedded in a bandpass time-varying noise with $\text{SNR} = 10$ dB (Ideally, the estimate should be constant over this band, zero outside and perfectly localized along the chirp instantaneous frequency). The first row consists of a spectrogram and its reassigned version, based on one realization. The corresponding multitaper estimate (8 Hermite functions) is given in the bottom row (right), with the WVD (left) for comparison. In each diagram, time is horizontal, frequency is vertical, and the energy is coded with gray levels on a logarithmic scale with a dynamic range of 30 dB, the limits of the noise band being superimposed as yellow lines.

the noise band. Comparatively, the effectiveness of the combined technique pro-

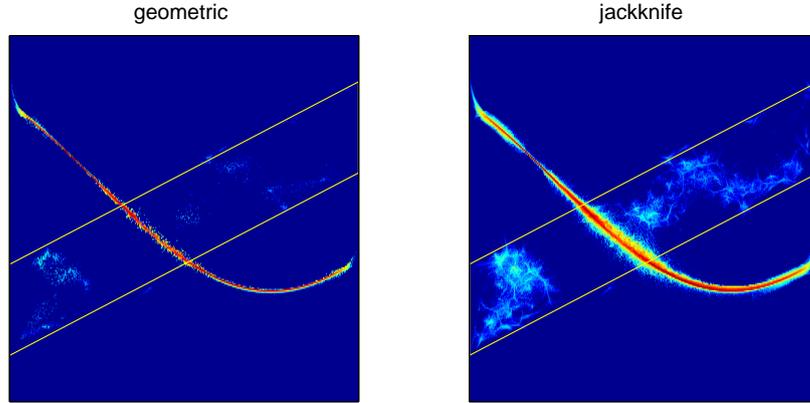


Fig. II.5: **Comparison of signal + noise WVS estimates in different way of means.** Each diagram represents, in the same case of the signal in Fig. II.4, a multitaper reassigned spectrogram in geometrical mean (left) or in jackknifing (right) of the estimates obtained from different tapers.

posed “multitaper reassigned spectrogram” is clearly evidenced by presenting the good trade-off achieved between time-frequency localization along the chirp and smoothness within the (time-varying) frequency band of the noise.

2 Variations for Chirp Enhancement

While preserving the basic idea, different variations upon the method proposed above can be considered. Indeed, the key point in the method is to combine the estimates obtained from different tapers, and the arithmetic means used in the “classical” multitaper approach (I.26) as well as in the definition (II.1) of the method appears as one possibility only amongst others.

Comprehensive approaches have been reported [45] on such generalizations in the general context of optimally combining different TF distributions, given some criterion to minimize. What has been shown is that arithmetic averaging naturally results from the requirement that the combined distribution be at a minimum L_2 -distance of all distributions to be combined. Changing the chosen distance ends up with different ways of averaging and, e.g., replacing the L_2 -distance by a Kullback-Leibler divergence leads to a geometrical mean instead of the arithmetic one.

Given N positive numbers $\{X_n, n = 1, \dots, N\}$, their arithmetic and geometric

means, defined respectively as

$$A(X_1, \dots, X_N) = \frac{1}{N} \sum_{n=1}^N X_n; \quad (\text{II.11})$$

$$G(X_1, \dots, X_N) = \left(\prod_{n=1}^N X_n \right)^{1/N}, \quad (\text{II.12})$$

are clearly such that

$$\log G(X_1, \dots, X_N) = A(\log X_1, \dots, \log X_N), \quad (\text{II.13})$$

evidencing that logarithm of geometrically averaged quantities essentially amounts to arithmetically averaging their logarithms.

In the context of multitaper spectrum estimation, this corresponds to applying the machinery to log-spectra, a possibility that has been explored in [70] in a stationary context. This can be adapted to nonstationary situations, leading to an alternative to the procedure (II.1) previously described

$$RS_{x,K}^{(G)}(t, f) = \exp \left\{ \frac{1}{K} \sum_{k=1}^K \log RS_x^{(h_k)}(t, f) \right\}. \quad (\text{II.14})$$

Given the fact that reassigned spectrograms have a high variability, crude geometric means could tend to favor the small (or even zero) values that might appear in one individual estimate only. A possible improvement is to first jackknifing estimates based on arithmetic means prior their geometric averaging:

$$RS_{x,K}^{(J)}(t, f) = \exp \left\{ \frac{1}{K} \sum_{k=1}^K \log \widetilde{RS}_x^{(h_k)}(t, f) \right\}, \quad (\text{II.15})$$

with

$$\widetilde{RS}_x^{(h_k)}(t, f) = \frac{1}{K-1} \sum_{m=1, m \neq k}^K RS_x^{(h_m)}(t, f). \quad (\text{II.16})$$

These variations have been observed to lead to results (shown in Fig. II.5 for the same example as in Fig. II.4) that were similar, or only incrementally improved. It was however worth mentioning it because the use of *log-distributions* that underlies the idea of geometrical averaging proved useful in the companion approach to be described in the following.

2.1 Principle and Implementation

2.1.1 Differences between estimates based on different tapers

The previous Section 1, aimed at reducing variability in nonstationary spectrum estimation while preserving localization of chirp components, was based on

an idea of averaging, i.e., of an *additive* combination of estimates obtained from different tapers. Making use of similar arguments now based on *differences*, a companion perspective can be envisioned for a sake of chirp enhancement, i.e., of displaying at best localized chirp components in the TF plane while getting rid of noise.

The rationale for this new approach can be justified by the following twofold argument:

(i) as far as signal components are concerned and as mentioned above, reassignment increases localization whatever the window: reassigned spectrograms obtained from different tapers are thus expected to be similar, so that their differences take on small values;

(ii) in noise regions on the contrary, the same set of windows lead to weakly correlated surrogate data whose TF distributions are significantly different, leading in this case to large values for their differences.

What we propose therefore is to consider such differences between the estimates based on the different tapers. As for “*sums*”, “*differences*” can be considered in a generalized sense and, in accordance with the usual way of displaying spectra in dB units, *logarithms* of TF estimates will be used. In fact, evaluating differences between log-distributions just amounts to evaluating ratios between the distributions themselves

$$\exp\{\log RS_x^{(h_{k+1})}(t, f) - \log RS_x^{(h_k)}(t, f)\} = \frac{RS_x^{(h_{k+1})}(t, f)}{RS_x^{(h_k)}(t, f)}, \quad (\text{II.17})$$

leading to defining as a measure of average difference between estimates based on successive tapers the quantity:

$$RSD_{x,K}(t, f) = \frac{1}{K-1} \sum_{k=1}^{K-1} \frac{RS_x^{(h_{k+1})}(t, f)}{RS_x^{(h_k)}(t, f)}. \quad (\text{II.18})$$

It is this quantity that is proposed to be thresholded in order to distinguish between the (“coherent”) signal components and the (“incoherent”) noise regions. More precisely, the idea is

(i) to define a binary mask function M taking on zero values for those TF points where the criterion (II.18) differs significantly from unity and one for the rest, and

(ii) to multiply this mask with the average reassigned spectrogram $RS_{x,K}(t, f)$ (II.1), so as to end up with a masked distribution $RSM_{x,K}(t, f)$ expected to preserve the chirp components while erasing most of the noise.

To demonstrate the function of the mask mentioned above, an example is given in Fig. II.6 which is concerned with a two component chirp signal superimposed to a transient Gaussian noise with SNR = 10 dB. As expected, the average differences $RSD_{x,K}(t, f)$ take small values for the two component chirp signal, whereas big values for the transient Gaussian noise and even favorably for the interference terms

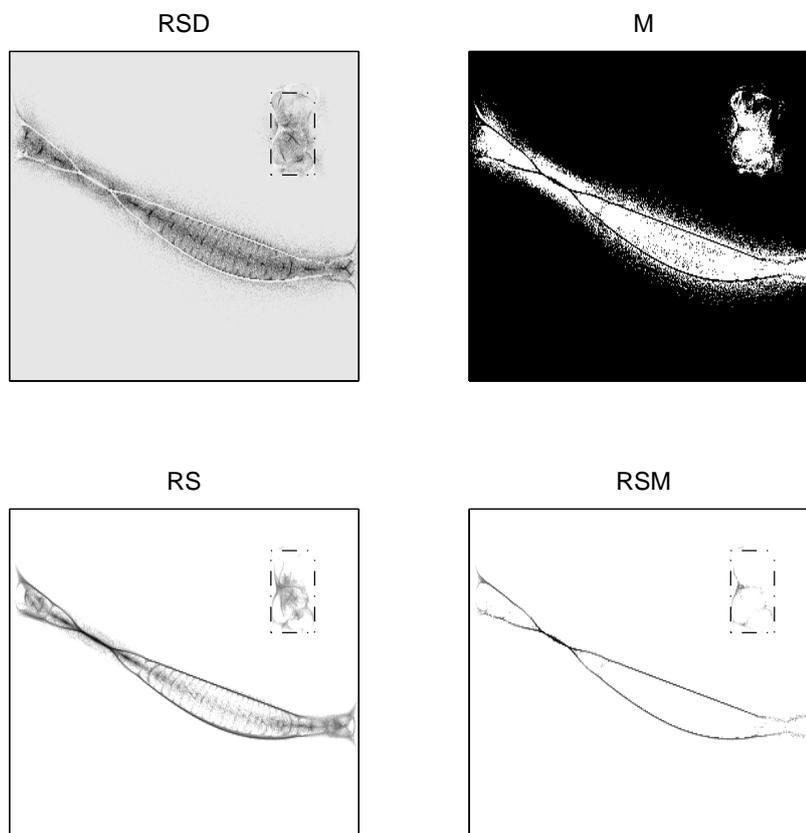


Fig. II.6: Mask function. This figure presents, in the case of one realization of a two component chirp signal superimposed to a transient Gaussian noise with $\text{SNR} = 10$ dB, on the top row the measure of average differences between estimates based on successive tapers (left) and the mask from it (right, “1” – *black*; “0” – *white*), whereas the bottom row displays the average reassigned spectrogram (left) and the corresponding masked estimate (right). In each diagram, time is horizontal, frequency is vertical, the energy is coded logarithmically with a dynamic range of 30 dB, and the limits of the noise TF domain being superimposed as dashed dotted lines.

between two closely spaced chirps, leading to (with a well chosen threshold) a good mask (“0” for big values and “1” for small values) for average reassigned spectrogram $RS_{x,K}(t, f)$. Obviously, the masked distribution $RSM_{x,K}(t, f)$ is a better estimate that not only erases much noise but also preserves chirp components as well.

2.1.2 Final estimate combined by "sums" and "differences"

In practice however, we cannot expect to completely get rid of noise by averaging a limited number of difference distributions. Therefore, a threshold is proposed to be first applied to the average reassigned spectrogram $RS_{x,K}(t, f)$ so as to get a pre-denoised version $RSP_{x,K}(t, f)$.

Proceeding from this point as above, this pre-processing leads, after masking, to a final, combined estimate $RSF_{x,K}(t, f)$. Interestingly, the binary nature of the two masking functions involved in the two thresholding operations detailed above allows us to write

$$RSF_{x,K}(t, f) = \sqrt{RSP_{x,K}(t, f) \cdot RSM_{x,K}(t, f)}, \quad (\text{II.19})$$

thus making of the combined estimate $RSF_{x,K}(t, f)$ the *geometric mean* of the two distributions based on the (thresholded) *sums* and *differences* of the individual taper estimates.

A key question in the approach is of course the choice of the thresholds for both $RSD_{x,K}(t, f)$ and $RS_{x,K}(t, f)$. Since no theory is available yet for the probability distribution function of reassigned spectrograms (and, *a fortiori*, of their ratios), a pragmatic approach is to set the threshold values on the basis of numerical experiments. The nominal values used in Fig. II.6 and in the following examples reported here have been determined this way (for details about this determination, as well as M (V.7), $RSM_{x,K}(t, f)$ (V.8), and $RSP_{x,K}(t, f)$ (V.9), cf. Appendix (Chapter V: Section 2)), and they proved to guarantee fairly robust estimations in a large variety of experimental setups.

2.2 Performance Evaluations: Contrast Measure

In order to test the relevance of the proposed methodology, one of the simplest cases to consider is concerned with a pure tone (signal with constant frequency and amplitude) embedded in a background of zero-mean white Gaussian noise. Fig. II.7 illustrates what happens in such a case with $\text{SNR} = 10$ dB. The two effects of reduced noise and support preservation of the localized signal component thanks to using differences are clearly visible, as well as the impact of pre-denoising for a sake of a further noise reduction.

Fig. II.7 gives a qualitative account of the behaviour of the method, which can be supplemented by the quantitative measure

$$C(\text{SNR}) = \frac{1}{\|\hat{\mathbf{W}}_{x,K}\|_1} \int_{-\infty}^{+\infty} \hat{\mathbf{W}}_{x,K}(t, f_0) dt, \quad (\text{II.20})$$

where $\hat{\mathbf{W}}_{x,K}(t, f)$ stands for the WVS estimate (one chosen among $RS_{x,K}(t, f)$, $RSP_{x,K}(t, f)$, $RSM_{x,K}(t, f)$ or $RSF_{x,K}(t, f)$), and f_0 is the actual frequency of the tone.

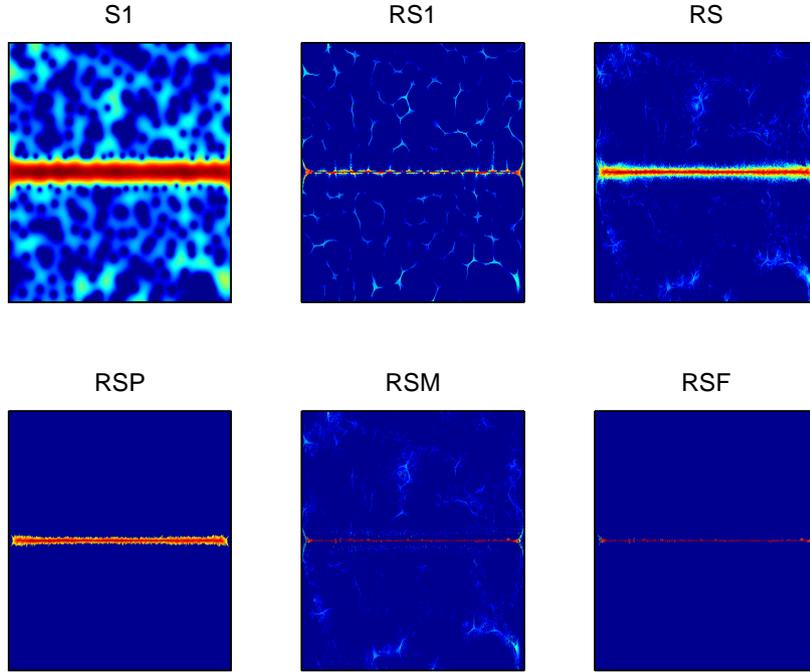


Fig. II.7: **Comparison of WVS estimates: pure tone in stationary noise.** Each diagram represents a WVS estimate in the case of one realization of a pure tone embedded in white Gaussian noise with $\text{SNR} = 10$ dB. The first row consists of a spectrogram (left), its reassigned version (middle) and the corresponding multitaper reassigned estimate (based on 4 Hermite functions, right), whereas the bottom row displays the pre-denoised multitaper reassigned spectrogram (left), the masked multitaper reassigned spectrogram (middle) and the final combined estimate (right). In each diagram, time is horizontal, frequency is vertical and the energy is coded logarithmically with a dynamic range of 30 dB.

The criterion (II.20), which can be interpreted as a *contrast measure*, is illustrated in Fig. II.8 for different SNRs ranging from -40 to $+40$ dB. Comparing the criterion for the 4 different estimates, two regimes are observed: for $\text{SNR} < 0$ dB, the method of pre-denoising ($\hat{\mathbf{W}}_{x,K}(t, f) = RSP_{x,K}(t, f)$) is more efficient than the one based on differences only ($\hat{\mathbf{W}}_{x,K}(t, f) = RSM_{x,K}(t, f)$), whereas the latter outperforms the former when $\text{SNR} > 0$ dB. This situation is uniformly improved when combining the two estimates according to $\hat{\mathbf{W}}_{x,K}(t, f) = RSF_{x,K}(t, f)$.

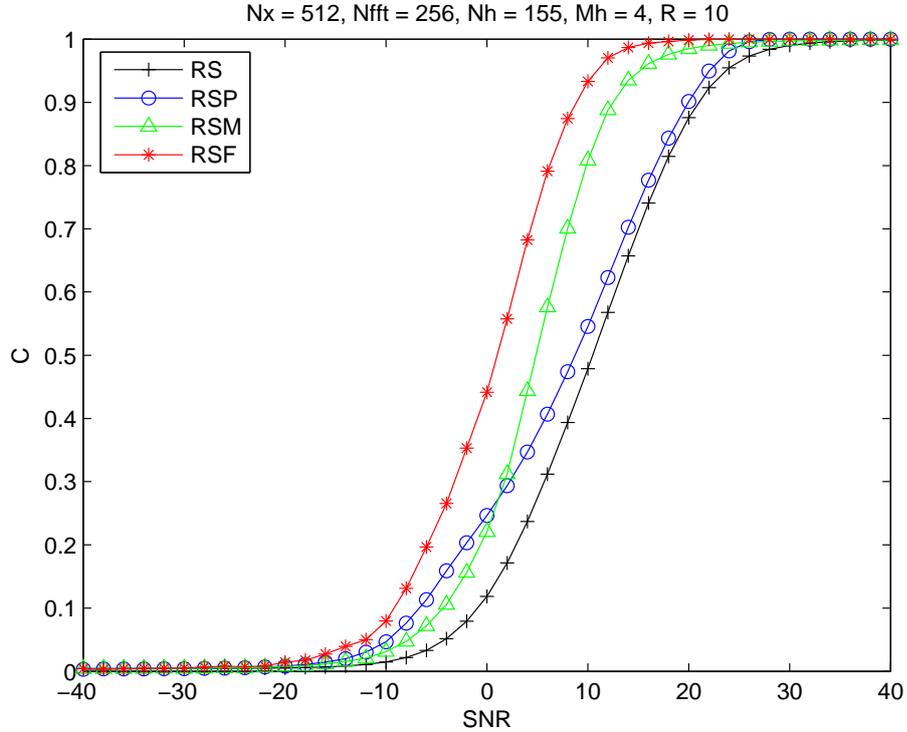


Fig. II.8: **Contrast measure in the case of a pure tone embedded in white Gaussian noise.** The figure plots, as a function of SNR, the measure (II.20) for each WVS estimate (see text). The simulations have been conducted (with 4 Hermite tapers, each of length 155), on the basis of 10 independent realizations of 512 data points each, with 256 frequency bins over the whole frequency range $[0, 1/2)$.

2.3 Examples

Only schematic examples have been considered so far, and this subsection is devoted to slightly more elaborated ones. In this respect and for a sake of comparison, we first consider in Fig. II.9 the same case discussed above in Fig. II.4, with both a (nonlinear) chirp component and a (bandpass) time-varying noise. Concerning spectrum estimation, the effectiveness of the approach is clearly supported by this example which evidences the good trade-off achieved between TF localization along the chirp and smoothness within the (time-varying) frequency band of the noise. As far as chirp enhancement is concerned, the final combined estimate identifies in a very localized way the frequency trajectory while erasing most of the noise.

As a second example, we consider in Fig. II.10 the same case already illustrated

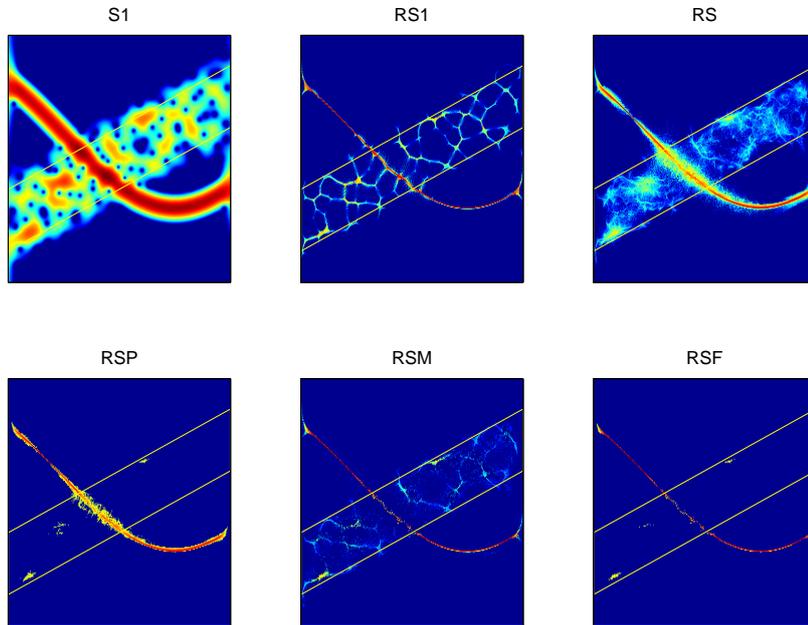


Fig. II.9: **Comparison of WVS estimates: one chirp in time-varying noise.** Each diagram represents a WVS estimate in the case of one realization of a nonlinear chirp signal embedded in time-varying Gaussian noise with $\text{SNR} = 5$ dB. Same display as in Fig. II.7, the limits of the noise TF domain being superimposed as yellow lines.

for the function of mask (cf. Fig. II.6), where two chirps (a linear one added to the nonlinear one of the previous example) are closely superimposed, with crossings of their instantaneous frequencies in the TF plane, some transient noise being added in a disjoint domain. The overall behaviour of the different estimates is similar to what has been observed in Fig. II.9, with the noticeable additional benefit for the combined estimate of improving upon the suppression of the interference terms existing between the two closely spaced chirps.

Also having been discussed in [6], the last example is a bat-echolocation signal [35, 36], which consists of a series of high frequency pulses emitted by bats for obtaining the information about the location and characteristics of target. It is clearly shown in Fig. II.11 that the new approach successfully pulls out all of the four high frequency line components even in the conditions of $\text{SNR} = 10$ dB. In the combined estimate, not only the chirps are enhanced, but also most of noise and the interference terms between the line components are erased as well.

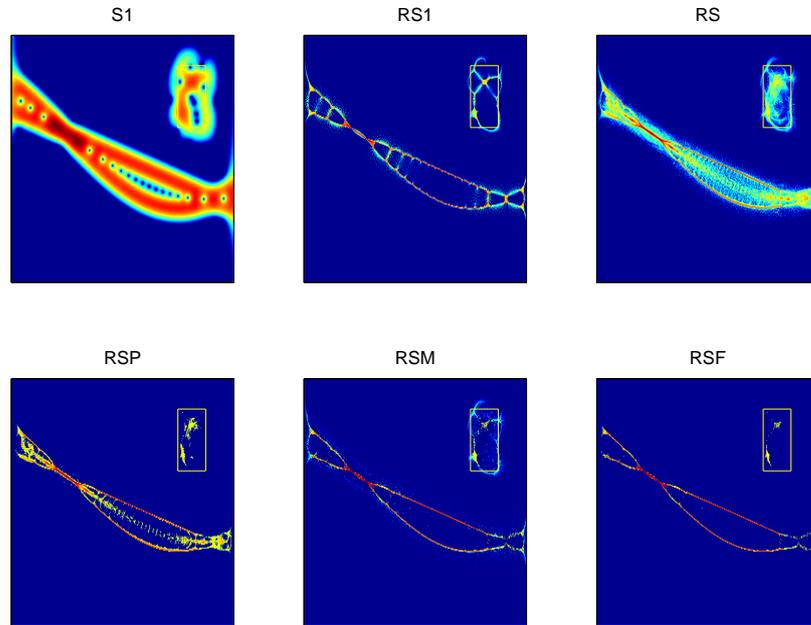


Fig. II.10: **Comparison of WVS estimates: two chirps and transient noise.** Each diagram represents a WVS estimate in the case of one realization of a two component chirp signal superimposed to a transient Gaussian noise with $\text{SNR} = 10$ dB. Same display as in Fig. II.7, the limits of the noise TF domain being superimposed as yellow lines.

3 Application: Euler's Disk

It is a fact of common experience that if a circular disk (for example, a coin) is spun upon a table, then ultimately it comes to rest quite abruptly, the final stage of motion being characterized by a shudder and a whirring sound of rapidly increasing frequency. This complex motion is known as the *Euler's disk* problem [3], which draws a lot of interests in it.

For a disk (see Fig. II.12) of radius R , mass m , and an inertia momentum I , the total energy of the motion is the sum of the kinetic energy and the potential energy

$$E = \frac{1}{2}I\omega^2 + mgR \sin \alpha \quad (\text{II.21})$$

$$\simeq \frac{3}{2}mgR \sin \alpha \quad (\text{II.22})$$

where α is the inclination angle with respect to the horizontal and Ω is the pre-

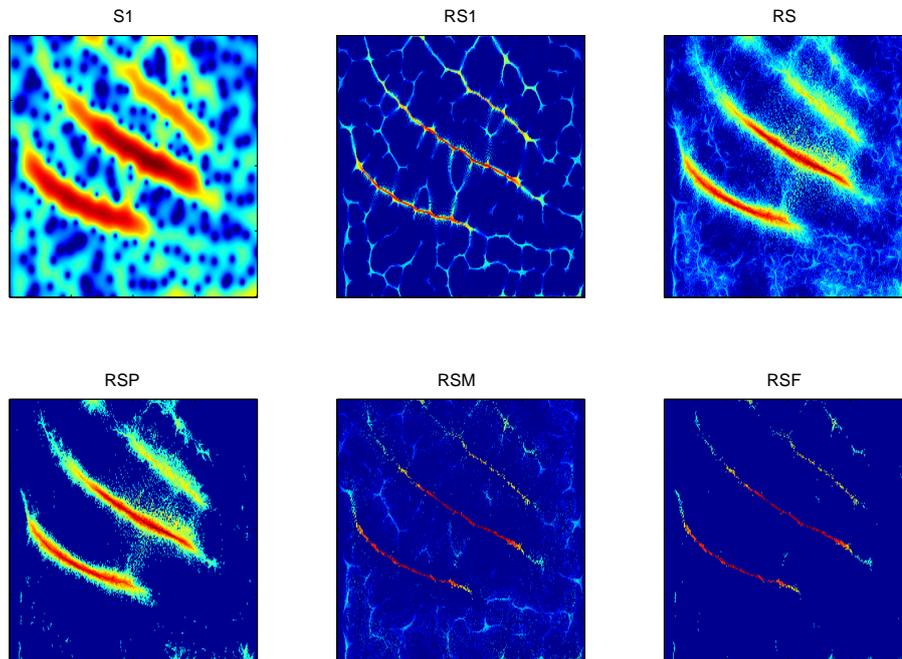


Fig. II.11: **Comparison of WVS estimates: bat-echolocation signal.** Each diagram represents a WVS estimate in the case of one realization of bat-echolocation signal embedded in white Gaussian noise with SNR = 10 dB. Same display as in Fig. II.7.

cession rate.

According to the classical theory of mechanics (conservation of energy), when Euler's Disk is spun, the disk contains both potential and kinetic energy. The

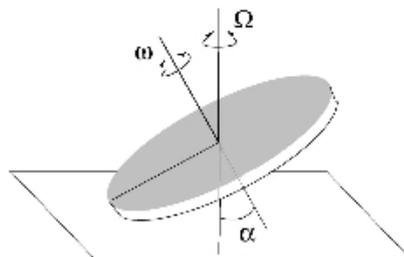


Fig. II.12: **Euler's disk.** Sketch of Euler's disk with its dynamical parameters: inclination angle α , the precession speed Ω , and the angular velocity ω .

potential energy is given to the disk when it is placed upright on its side and the kinetic energy is given to the disk when it is spun on the mirrored base. The disk would "spoll" (i.e. spin and roll) forever, if it were not for friction and vibration.

However, according to observations, it is unrealistic and the energy E is dissipated with a rate ϕ : $\frac{dE}{dt} = -\phi$. To understand why the forever persisting motion of the toy stops, Moffatt used elements of fluid mechanics to predict how Euler's Disk might lose energy during spolling by squeezing out the air between its rolling edge and the mirror base [52]. Many other theoretical and numerical studies have also been devoted to understanding the energy dissipation process leading to the stop of the disk [28, 65, 41, 10, 21].

Here, we are not going to pursue the possible reasons, but would like to enlighten the question by supplying an alternative vision of the motion by means of a direct TF analysis of the accelerations of Euler's disk in 3 dimensions (3D), based on which an parameter estimation of the rapidly increasing frequency chirps is done by means of *Hough Transform*.

3.1 Time-Frequency Estimation

The experimental data are the accelerations of Euler's disk in 3D (X and Y being on the plane of the disk and Z vertical to the plane, with the origin in the center of the disk), which have been measured by Gasteuil [66, 37] from a sensor of acceleration in the center of a spolling disk.

Two realizations of the acceleration signals are studied here in each direction \vec{a}_x , \vec{a}_y , \vec{a}_z , displayed respectively from left to right in Fig. II.13. In each diagram, the oscillations of the signals describe the rotation and precession of the disk and the slow shift (observed most evidently in direction Z) comes from the projection of the gravity on each axis. Moreover, the fluctuation of the amplitude indicates the existence of noise.

As the acceleration signals in each direction are the rapidly increasing frequency chirps, we could take advantage of the good performance of the TF method based on multitaper reassignment, especially the combined estimate RSF in (II.19) to obtain a distribution with a sharp localization for chirp components and a reduced level of statistical fluctuations for noise background.

In Fig. II.14, multitaper spectrogram S , multitaper reassigned spectrogram RS and the combined estimate RSF of the first realization of the acceleration in direction Z (top right diagram of Fig. II.13) of the spolling disk are shown (respectively from left to right) on the top row. Considering the reassignment of chirps being disturbed by zero frequency components which result from the slow shift of signal mentioned above, we apply a high-pass filter before TF estimation in order to eliminate the effect. The corresponding spectrogram estimates for the same signal with the high-pass filter are displayed on the bottom row for comparison, where chirp is shown to be well localized by reassignment and further localized by RSF .

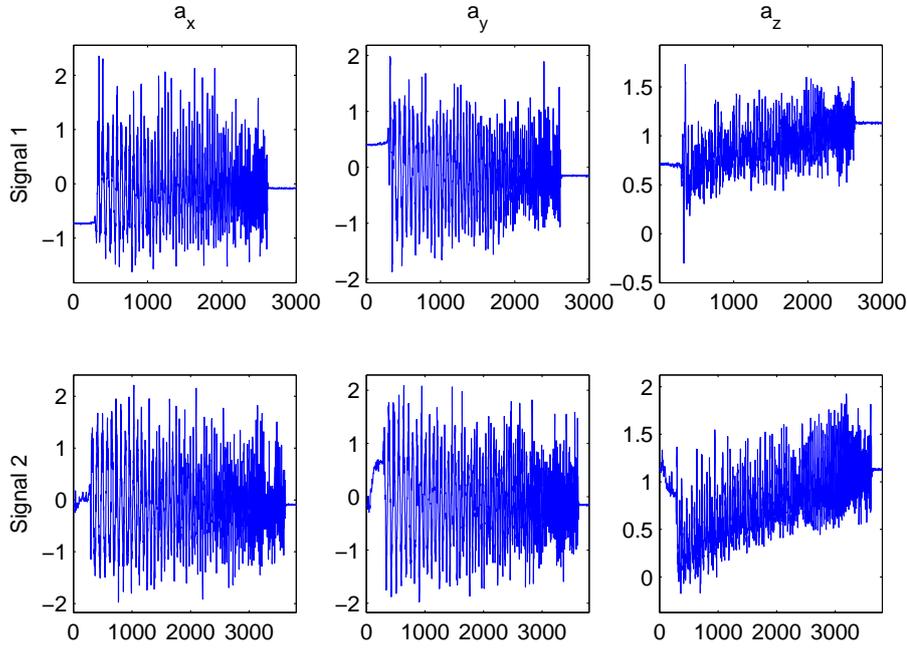


Fig. II.13: **Acceleration signals of Euler's disk.** This figure plots two realizations of the accelerations of Euler's disk from a spolling disk (on the top and bottom row) in each direction of the 3 dimensions X, Y, Z (respectively from left to right).

3.2 Parameter Estimation of Chirp by Hough Transform

With a good TF estimation, it is much easier to detect chirp signals [17] and we could even estimate their parameters as well. The parameter estimation of chirps in the TF plane can be realized by a pattern recognition technique called *Hough transform* (HT) [53, 11, 12] which recognizes specific curves in the TF plane of the signal under analysis, i.e. linear and nonlinear FM trajectories.

Given the TF representation $RSF_x(t, f)$ of a signal $x(t)$, HT is defined as the following mapping rule

$$Q_x(\theta) = \int RSF_x(t, f(t; \theta)) dt, \quad (\text{II.23})$$

where $f(t; \theta)$ is a parametric function (model) with unknown parameters θ . Assuming the instantaneous frequency of signal has the form $f(t; \theta)$, the detection and parameter estimation of signals can then be turned into the search for the peaks of $Q_x(\theta)$, leading to the estimator of parameter of the signal being

$$\hat{\theta} = \arg \max_{\theta} (Q_x(\theta)). \quad (\text{II.24})$$

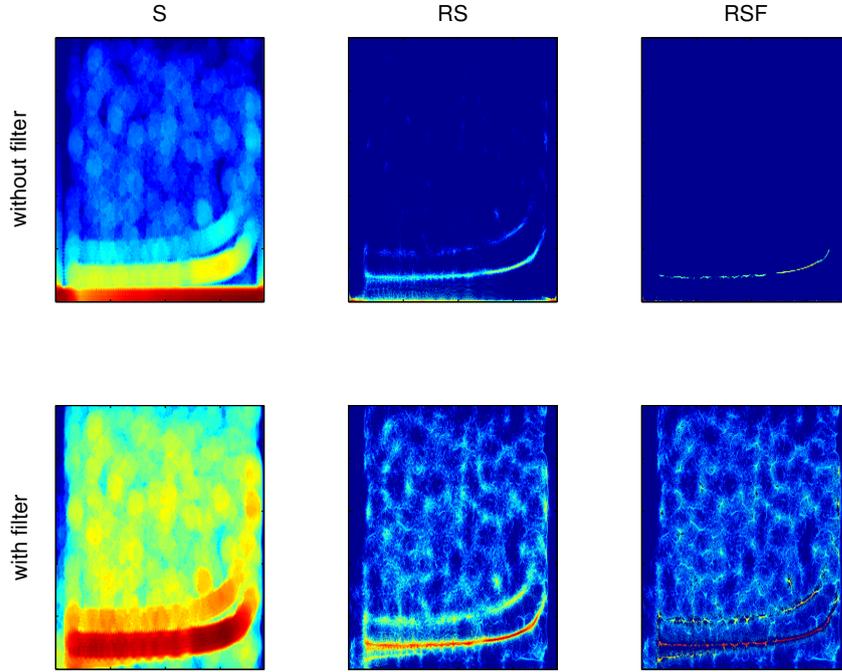


Fig. II.14: **Time-frequency estimation.** This figure displays the multitaper spectrogram S (on the first column), multitaper reassigned spectrogram RS (on the second column), and the combined estimate RSF (on the third column) in the case of one realization of the acceleration in direction Z (top right diagram of Fig. II.13) of a spolling disk (on the top row), and the corresponding estimates of the same signal with a high-pass filter (on the bottom row) added for eliminating zero frequency components. In each diagram, time is horizontal, frequency is vertical and the energy is coded logarithmically with a dynamic range of 40 dB.

In this case of Euler's disk, the instantaneous frequency of the acceleration in each direction is supposed to have the form of a power-law chirp :

$$f(t; t_0, \beta, f_0) = f_0 \cdot \left(\frac{t_0 - t}{T}\right)^\beta. \quad (\text{II.25})$$

As f_0 and T are mutually dependent parameters, we have to fix one of these two. Considering that the power law has no natural characteristic time, we thus fix T as a characteristic time with a reasonable magnitude that is comparable to the duration of the part of signal where the instantaneous frequency increases rapidly.

Here $T = 100$ points is a typical length of the scale, leading to an exploratory space for (t_0, β, f_0) that could be conveniently shaped in all directions.

In this way, t_0, β, f_0 are the only 3 necessary parameters for the construction of such an instantaneous frequency (chirp), so the mapping rule is now

$$Q_x(t_0, \beta, f_0) = \int RSF_x(t, f(t; t_0, \beta, f_0)) dt, \quad (\text{II.26})$$

and hence the estimators of parameters are

$$(\hat{t}_0, \hat{\beta}, \hat{f}_0) = \arg \max_{t_0, \beta, f_0} (Q_x(t_0, \beta, f_0)). \quad (\text{II.27})$$

Since we are only interested in and search among all of the possible chirps whose trajectories fall in the TF plane observed, the scale of the parameters for searching (i.e. the exploratory space of (t_0, β, f_0)) should be well chosen so that the main region of study would be parallelepipedally explored and $\{(t, f(t)), t = 1, \dots, T\}$ falls always in the TF plane, thus being practical for the numerical studies.

3.3 Performance Evaluation on Synthetic Model

To test the effectiveness of the *RSF – Hough* approach, we apply it first on a synthetic power-law chirp signal

$$x(t) = A \cos\left(2\pi \frac{f_0 \cdot \left(\frac{t_0 - t}{T}\right)^{\beta+1}}{\beta + 1} + \phi_0\right), \quad (\text{II.28})$$

where $A = 10$ is the amplitude of the signal, $T = 100$, and the initial phase $\phi_0 = 0$ for simplicity's sake.

If chirp's parameters are set to be: $t_0 = 990$, $\beta = -0.52$, $f_0 = 6.2$ with SNR = 100, the estimates of parameters by HT (II.27) are exactly the same values: $(\hat{t}_0, \hat{\beta}, \hat{f}_0) = (990, -0.52, 6.2)$; but with SNR = 5, the estimates have a little difference: $(\hat{t}_0, \hat{\beta}, \hat{f}_0) = (990, -0.53, 6.2)$.

To have a performance evaluation of the method, Fig. II.15 illustrates means and standard deviations of the 3 estimates $\hat{t}_0, \hat{\beta}, \hat{f}_0$ of the synthetic chirp as a function of SNR. The computation is based on 10 realizations, and from the results we would say that the minimum SNR at which the method provides reliable estimates is around 5.

Besides, it should be pointed out that a *weighting* has been added in the HT estimation to compensate the low power at the end of chirp, which results from the decreasing amplitude of signal with time and sparser points of estimation for the area of rapidly increasing frequency, therefore the mapping rule becomes:

$$Q_x(t_0, \beta, f_0) = \int RSF_x(t, f(t; t_0, \beta, f_0)) \cdot P(t) dt, \quad (\text{II.29})$$

where the weighting $P(t) = t^n$, n being chosen generally in our studies from 2 to 7, depending on the signal.

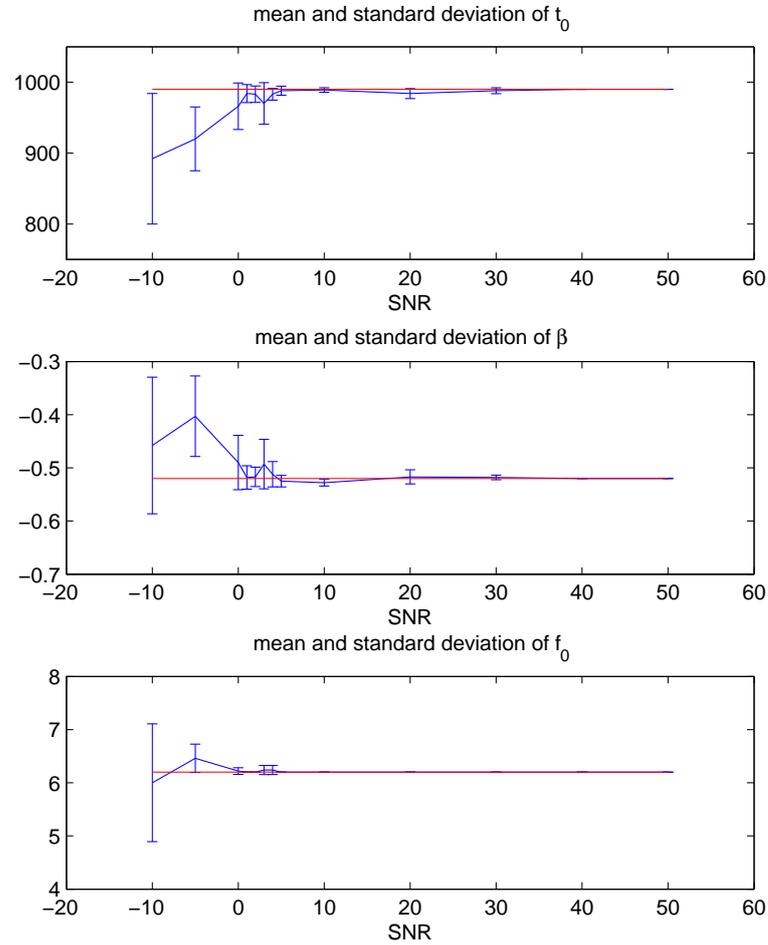


Fig. II.15: **Performance evaluation in the case of a synthetic chirp.**

Each diagram displays, as a function of SNR, the mean and standard deviation of the parameter estimate \hat{t}_0 , $\hat{\beta}$, or \hat{f}_0 (respectively from top to bottom) by Hough transform (with a weighting $P(t) = t^2$) in the TF representation RSF of a synthetic chirp signal (II.28), based on 10 realizations. The true parameter values of the chirp are plotted by red lines.

Empirically speaking, the weighting helps counterpoise the estimation of chirp over the TF plane, without which we would see that the resulted chirp, constructed by the estimates from HT ($\hat{t}_0, \hat{\beta}, \hat{f}_0$), matches very well in the TF representation at the beginning and in the middle of time with the trajectory of the real chirp

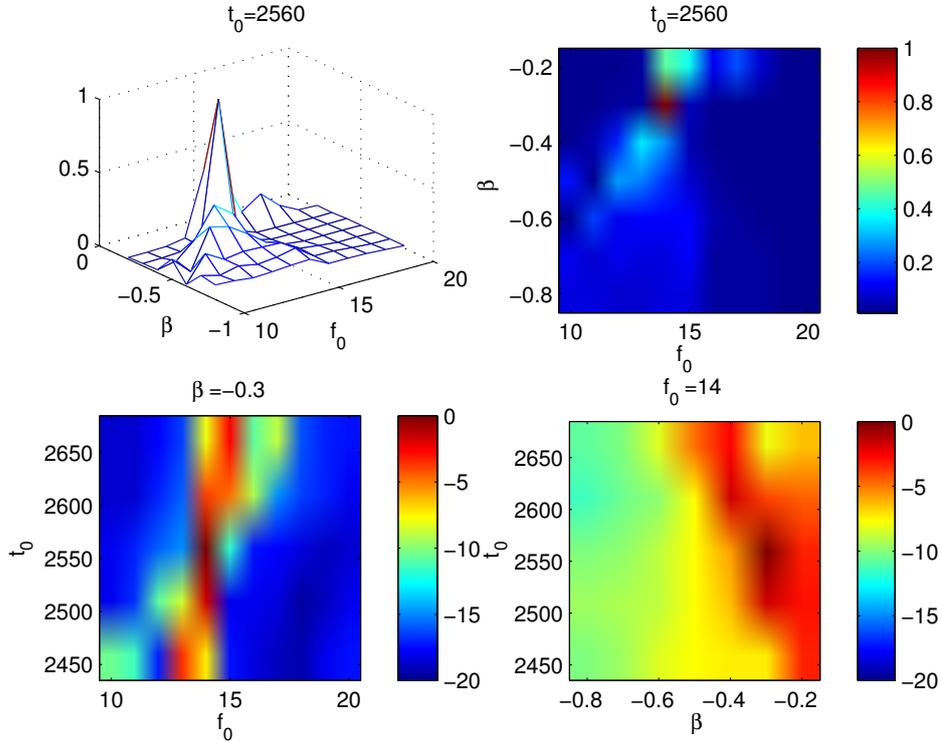


Fig. II.16: **Hough transform in the 1st step.** This figure displays, in the case of the first acceleration signal in direction Z (right top diagram in Fig. II.13), 2D images of Hough transform as a function of any two parameters among t_0, β, f_0 with the third parameter fixed to the resulted estimate ($\hat{t}_0, \hat{\beta}$, or \hat{f}_0 , shown on the top of each diagram) in the 1st wide but coarse step of calculation. A 3D mesh surface of Hough transform as a function of f_0 and β is also illustrated (on the left top) with t fixed to the resulted estimate \hat{t}_0 .

signal, but is far from it at the end due to the low power.

3.4 Test on Experimental Signal

The *RSF – Hough* approach is now applied to the experimental signals of Euler's disk, i.e. the two realizations of the 3D accelerations of spolling disk shown in Fig. II.13.

In order to accelerate the numerical calculation but keep enough precision of parameters as well, we propose a *2-step calculation* of HT: (*i*) first, a wide but coarse estimation; (*ii*) and then a refined estimation around the estimates of

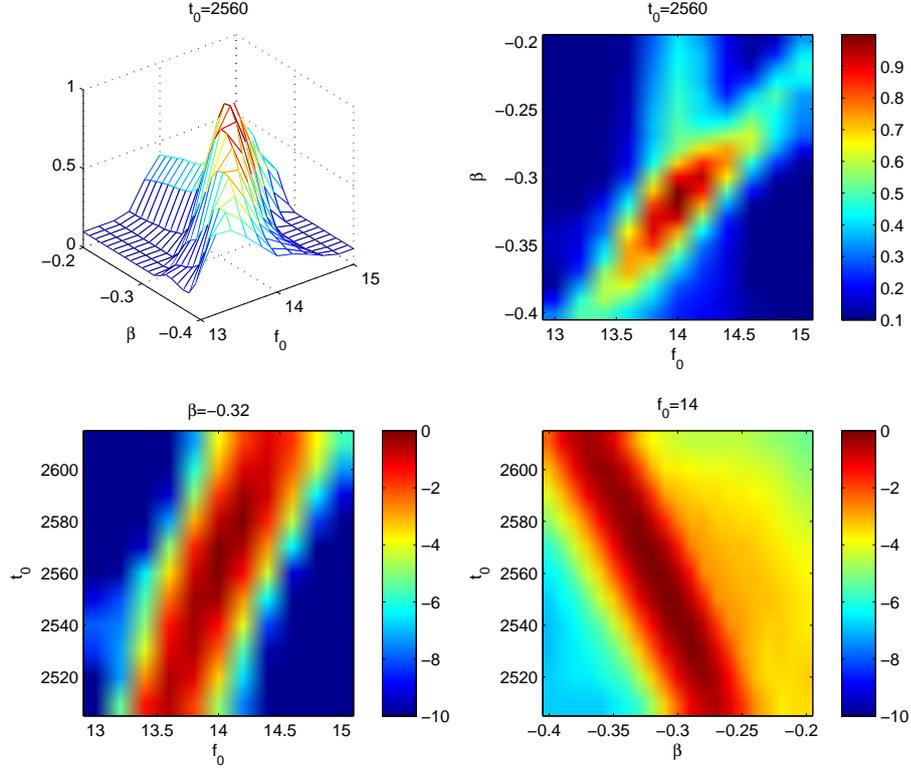


Fig. II.17: **Hough transform in the 2nd step.** This figure presents Hough transform in the 2nd refined step of calculation. Same display as in Fig. II.16.

parameters got from the first step.

Fig. II.16 displays as an example, in the case of the first acceleration signal in direction Z in Fig. II.13, the HT as a function of any two parameters among t_0, β, f_0 with the third parameter fixed to the resulted estimate ($\hat{t}_0, \hat{\beta}$, or \hat{f}_0) in

$P(t) = t^n$	\vec{a}_x $(\hat{t}_0, \hat{\beta}, \hat{f}_0)_n$	\vec{a}_y $(\hat{t}_0, \hat{\beta}, \hat{f}_0)_n$	\vec{a}_z $(\hat{t}_0, \hat{\beta}, \hat{f}_0)_n$
# 1	$(2640, -0.42, 14.2)_5$	$(2640, -0.42, 14.2)_7$	$(2560, -0.32, 14)_3$
# 2	$(3650, -0.46, 14)_5$	$(3620, -0.41, 14)_7$	$(3500, -0.27, 13.8)_3$

Tab. II.1: **Estimates of chirp parameters.** This table shows the RSF-Hough estimates of parameters $(\hat{t}_0, \hat{\beta}, \hat{f}_0)$ (II.27) in the case of the two realizations of the 3D accelerations shown in Fig. II.13.

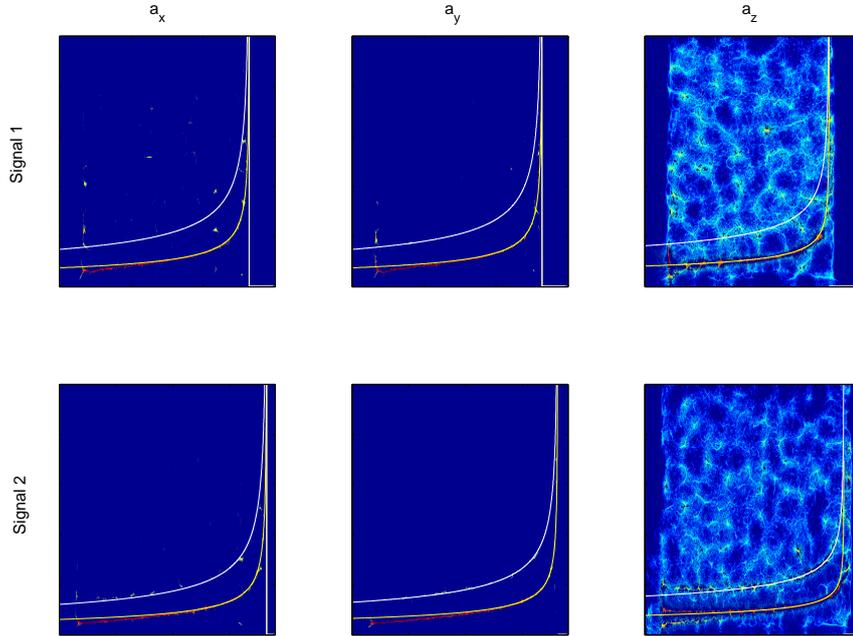


Fig. II.18: **Time-frequency estimation of chirp vs. constructed instantaneous frequency.** Each diagram displays the corresponding combined estimate RSF of the 3D acceleration signals in Fig. II.13 with the constructed instantaneous frequency by HT estimates of parameters $f(t; \hat{t}_0, \hat{\beta}, \hat{f}_0)$ (yellow curve) overlapped on the chirp and the 2-times value $2f(t; \hat{t}_0, \hat{\beta}, \hat{f}_0)$ (white curve) on its 2nd harmonic. In each diagram, time is horizontal, frequency is vertical and the energy is coded logarithmically with a dynamic range of 40 dB.

the 1st wide but coarse step of calculation; the same display in Fig. II.17 for the 2nd narrow, precise step of HT.

The final results of the 3 parameter estimation $(\hat{t}_0, \hat{\beta}, \hat{f}_0)$ from the 2-step HT in the TF plane RSF are shown in Table II.1. We find that the values of the estimate β is compatible with the results of a measure in [21] on a metal surface.

Finally, in order to show the performance of $RSF - Hough$ estimation in the acceleration signals of Euler's disk, each diagram in Fig. II.18 presents the good overlapping of the instantaneous frequency constructed by the 3 estimates of parameter $f(t; \hat{t}_0, \hat{\beta}, \hat{f}_0)$ (in yellow) on the real experimental acceleration signals in the TF representation. Besides, 2-times value of the constructed instantaneous frequency $2f(t; \hat{t}_0, \hat{\beta}, \hat{f}_0)$ (in white) overlaps favorably on the second curve above

the former, verifying that it is the 2nd harmonic of the acceleration.

4 Conclusion

A novel approach has been proposed in this Chapter for better estimating time-frequency spectra of chirp signals embedded in nonstationary noise. The successful wedding between time-frequency reassignment and multitapering results in a sharp localization for the chirp components and a reduced level of statistical fluctuations for the noise. Two variations of the technique are based on different ways of combining the estimates obtained from different Hermite tapers: “sums” mainly for nonstationary spectrum estimation, whereas “differences” for a sake of chirp enhancement. The principles and the implementations of the technique and both variations have been outlined and justified, both qualitatively and quantitatively, with several examples provided for supporting the efficiency of the approach. Finally, an application is made on some experimental chirp data of Euler’s disk.

III. Test of Stationarity with Surrogates

Testing stationarity is an important issue in several respects. As far as methods are concerned, signal processing is equipped with many powerful algorithms devoted to stationary processes, whose applicability should therefore be first assessed prior using them. Turning to interpretation, rejecting stationarity (and measuring some degree of nonstationarity) is of primary importance in numerous applications ranging from exploratory data analysis to diagnosis or surveillance.

However, whereas the concept of (weak, or second-order) stationarity is well-defined in theory, it turns out that it generally cannot be used as such in practice. Indeed, stationarity refers classically to a strict invariance of statistical properties over time, but common practice generally considers this invariance in a looser sense, relatively to some (explicit or implicit) observation scale. This certainly agrees with the physical intuition according to which a signal might appear, e.g., as both “short-term stationary” and “long-term nonstationary”. As an example, we can think of speech that is routinely “segmented into stationary frames”, the “stationarity” of voiced segments relying in fact on periodicity structures within restricted time intervals.

Those remarks call for a better framework aimed at dealing with “stationarity” in an operational sense, with a definition that would both encompass *stochastic* and *deterministic* variants, and include the possibility of its test *relatively to a given observation scale* and/or even to given frequency subbands as well. Two different approaches are therefore proposed for such a stationarity test, both of which are based on a comparison between local and global TF features, and originally make use of *surrogates* for defining *null hypothesis of stationarity*. One method is based on a parametric model derived from the distribution of surrogates variances [79, 80, 81], while the other is a statistical test implemented by *one-class Support Vector machine* [82], with surrogates serving as a learning set.

More precisely, this Chapter is organized as follows. The time-frequency framework as a plane for comparison of local and global features is outlined in Section 1, whereas Section 2 is concerned with the introduction of surrogate data for defining the null hypothesis of stationarity, being the kernel of the test.

The two approaches of the test themselves are discussed and detailed respectively in Section 3 and Section 4: for the first one which is based on a parametric model, local and global TF spectra being compared in Subsection 3.1, by a suitable *distance* measure chosen in Subsection 3.2 for typical nonstationary signals. The structure of the test in Subsection 3.3 can thus be carried out, based on a null

hypothesis of stationarity characterized in Subsection 3.4, for deciding whether an observed signal is stationary, and if not, an index and a scale of nonstationarity are defined in Subsection 3.5. Some illustrations are given in Subsection 3.6 for supporting the efficiency of the method.

On the other hand, the second method by means of one-class Support Vector Machine is presented by first introducing one of the powerful tools of kernel-based machine learning in Subsection 4.1, and then detailing in Subsection 4.2 the extraction of time-frequency features by its local vs. global comparison. Typical examples are provided for illustrating and supporting the approach in Subsection 4.3.

Finally, an actual application of these two approaches is made on a real speech signal in Section 5 and some of the many possible variations and extensions are briefly outlined in Section 6.

1 Time-Frequency Framework: Comparison of Local vs. Global Features

The first necessary ingredient of the stationarity test is a suitable representation for later processing. As far as only second order evolutions are to be tested, TF distributions and spectra are natural tools [30].

Well-established theories (cf. Chapter I) exist for justifying the choice of a given TF representation. In the case of stationary processes, the WVS is not only constant as a function of time but also equal to the PSD at each instant. From a practical point of view, the WVS is a quantity that has to be estimated. Given a signal $x(t)$, we choose to make use of a multitaper spectrogram estimate (I.26) of its WVS, computed with the K first Hermite functions (II.2) (typically chosen to be 5 to 10) as short-time windows $h_k(t)$. As already mentioned, the multitaper approach is preferred to a classical spectrogram in order to reduce the level of statistical fluctuations of the nonstationary spectrum without recouring to a time-averaging step.

The TF interpretation suggesting that suitable representations should undergo no evolution in stationary situations, stationarity tests can be thus envisioned on the basis of some comparison between local and global features. Relaxing the assumption that stationarity would be some *absolute* property, the basic idea underlying the approach proposed here is that, when considered over a given duration, a process will be referred to as stationary relatively to this observation scale if its time-varying spectrum undergoes no evolution or, in other words, if the local spectra (within the window of analysis) at all different time instants are statistically similar to the global spectrum obtained by marginalization.

Moreover, the definition of stationarity based on a comparison of statistics in the TF plane offers a unified framework of stationarity, no matter whether the stationarity is invoked in a stochastic sense or a deterministic sense. The latter could be in fact connected with the repetition of a certain motif over time,

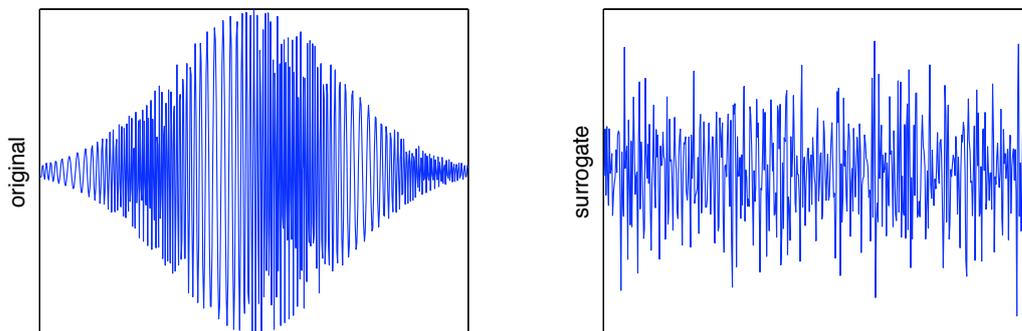


Fig. III.1: **Signal and surrogates.** This figure displays a Frequency Modulated signal with a Gaussian amplitude, which is a nonstationary signal (left), and one of its surrogates obtained by replacing the phase of the Fourier transform by an i.i.d. uniform phase (right).

i.e. a periodicity. The proprieties of the TF representation will thus, seeing that the window of analysis is not too short compared with this motif, represent this periodicity as an constancy of the TF distribution of deterministic signal; The TF marks of the periodicity and stationarity of random signals turns out to be identical by representing its constancy of the local spectrum over time.

2 Surrogates: Stationary Reference

2.1 Surrogates in (Non)stationarity Test

Revisiting stationarity within the TF perspective has already been pushed forward [50], but the idea of the test is to identify the notion of stationarity with the equivalence of local and global spectral properties. In practice however, even if a signal under test is stationary, the global characteristics and the local ones will never be completely identical, so the question is to know whether the observed differences between them are significant. A null hypothesis of stationarity is thus necessary to be specified.

The TF interpretation of stationarity amounts to say that, for an identical marginal spectrum over the same observation interval, nonstationary processes differ from stationary ones by some structured organization in time. Indeed, distinguishing between stationarity and nonstationarity would therefore be made easier if, besides the analyzed signal itself, we also have at our disposal some reference having the same marginal spectrum while being stationary. Since such a reference is generally not available, one possibility is to *create it from the data*: this is the rationale behind the idea of “*surrogate data*” [68, 60].

Up to some related proposal reported in [40], the technique of surrogate data,

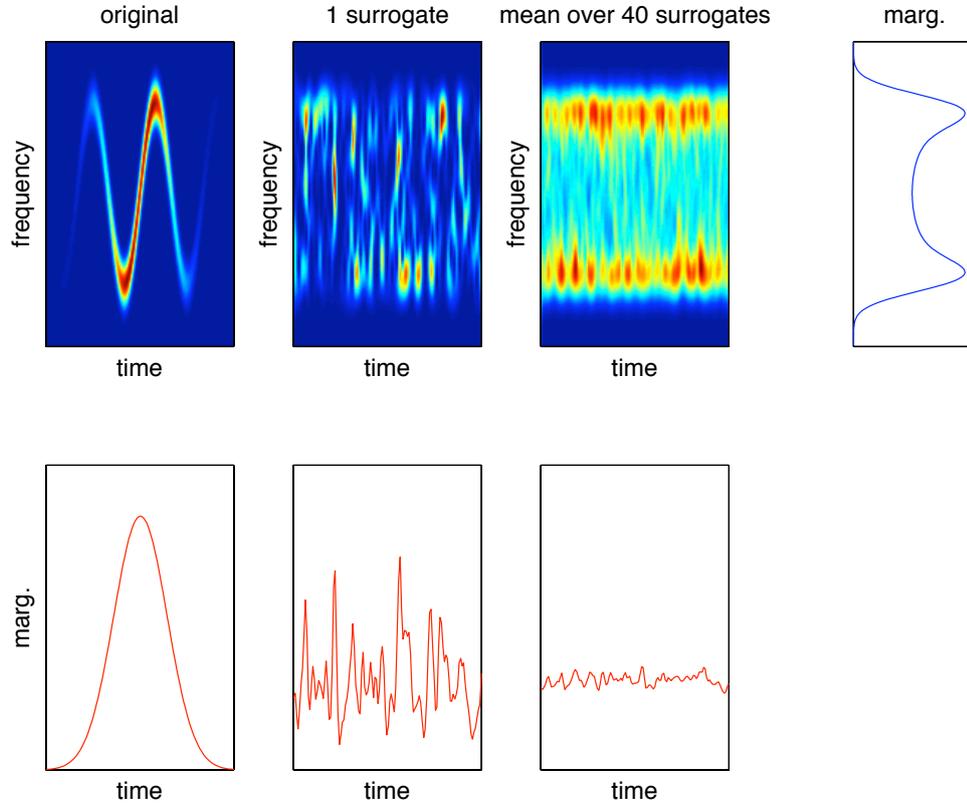


Fig. III.2: **Surrogates.** This figure compares the TF structure of the non-stationary FM signal in Fig. III.1 (1st column), with that of one of its surrogates (2nd column) and that of the mean over $J = 40$ surrogates (3rd column). The spectrogram is presented (time being horizontal, frequency vertical) in each case on the 1st row, with the corresponding marginal in time on the 2nd row. The marginal in frequency, which is for all of the three spectrograms, is displayed on the far right of the 1st row.

widely used in nonlinear physics (cf. Subsection 2.2), seems to have never been used for testing nonstationarity. Only a recent approach [40] took an interest in the surrogates in connection with the properties of nonstationarity, which was devoted to the preservation of the mean and variance of the temporal evolution in the construction of the surrogates. However, this approach is orthogonal to what we propose here in the following because our objective is just to, with the help of surrogates, make this temporal evolution disappear.

According to the interpretation that having the same averaged spectrum, a nonstationary signal is distinguished from a stationary counterpart by a temporal *structuration* whose mark is found in the spectral phase, it is possible to produce from a given single signal $x(t)$ a battery of *surrogates* $\{s_j(t); j = 1, \dots, J\}$, each of which has the same PSD as the original signal but has a *stationarized* temporal content.

This can be achieved by destroying the organized phase structure controlling the nonstationarity of $x(t)$, if any. In practice, $x(t)$ is first Fourier transformed to $X(f)$, and the modulus of $X(f)$ is then kept unchanged while its phase is replaced by a random one, uniformly distributed over $[-\pi, \pi]$. This modified spectrum is finally (inverse) Fourier transformed, leading to as many (stationary) surrogate signals as phase randomizations are operated.

Fig. III.1 shows a nonstationary signal and one of the surrogates resulting from this operation. The effect of the surrogate procedure is further illustrated in Fig. III.2, displaying both signal and surrogate spectrograms, together with their marginals in time and frequency. It clearly appears from this figure that, while the original signal undergoes a structured evolution in time, the recourse to phase randomization in the Fourier domain ends up with stationarized (i.e., time unstructured) surrogate data with identical spectrum. Therefore, the surrogates generated in this way can be natural candidates for quantifying statistically what should be the behaviour of a stationary signal.

2.2 Surrogates in (Nonlinear) Physics

The technique of surrogates was originally proposed as a method for testing experimentally the nonlinear proprieties of systems (via the signals from the systems) in the domain of nonlinear physics [68].

In this context, the surrogates are generated from the original data in a way that the spectrum is not changed, neither are the correlations, thus having the same linear characteristics as the original data; yet they destroy all statistical correlations of higher order, thus no signature of nonlinearity, considering that the nonlinearity are generated from the statistical properties of the orders larger than two.

The method first specifies some linear process as a null hypothesis, then generates surrogate data sets which are consistent with this null hypothesis, and finally computes a discriminating statistic for the original data and the surrogate data sets. If the value computed for the original data is significantly different from the ensemble of values computed for the surrogate data, the null hypothesis is rejected and nonlinearity is thus detected.

To be short, the surrogate data, generated to represent the *null hypothesis of linearity*, are compared to the original data under a discriminating statistic so as to approve or reject the null hypothesis. The technique of surrogates for testing nonlinearity has become popular in the last years and various null hypotheses and

discriminating statistics have been discussed [68, 60, 40]. Its application ranges from the research of chaotic physical systems to many other experimental time series which arise in the measurement of superfluids, brain waves, and sunspots, etc..

3 Stationarity Test - Parametric Model Approach

3.1 Comparison of Local vs. Global Time-Frequency Spectra

As mentioned in Section 1, stationarity test is based on some comparison between local and global spectra. The multitaper spectrogram (I.26) is in fact evaluated only at N time positions $\{t_n, n = 1, \dots, N\}$, with a spacing $t_{n+1} - t_n$ which is an adjustable fraction of the temporal width N_h of the K windows $h_k(t)$. Given therefore the set of spectral “slices” $\{S_{x,K}(t_n, f), n = 1, \dots, N\}$, we can compute an average spectrum via the marginalization in time:

$$\langle S_{x,K}(t_n, f) \rangle_{n=1, \dots, N} = \frac{1}{N} \sum_{n=1}^N S_{x,K}(t_n, f), \quad (\text{III.1})$$

and compare it to each spectral slice according to some dissimilarity measure $d(\cdot, \cdot)$ (possible choices for this measure will be discussed in Section 3.2), thus leading to the series of “distances” over the time interval from t_1 to t_N :

$$\{c_n^{(x)} := d(S_{x,K}(t_n, \cdot), \langle S_{x,K}(t_n, \cdot) \rangle_{n=1, \dots, N}), n = 1, \dots, N\}. \quad (\text{III.2})$$

In the idealized case where $x(t)$ would be stationary and the estimation perfect, all those coefficients would be zero. In practice however, they of course fluctuate and the issue is to determine whether the observed fluctuation is significant or not: this is where *surrogates*, the stationarized references, enter into the play.

Let us label $\{s_j(t), j = 1, \dots, J\}$ the J surrogates obtained as just described. When they are analyzed as explained above for the original signal $x(t)$, we finally end up with a new collection of distances

$$\{c_n^{(s_j)} := d(S_{s_j,K}(t_n, \cdot), \langle S_{s_j,K}(t_n, \cdot) \rangle_n), n = 1, \dots, N, j = 1, \dots, J\}, \quad (\text{III.3})$$

depending on both time indexes and randomizations.

3.2 Choice of Distance

3.2.1 Test signals

Setting specific parameters in the implementation is likely to end up with performance depending on the type of nonstationarity of the signal under test.

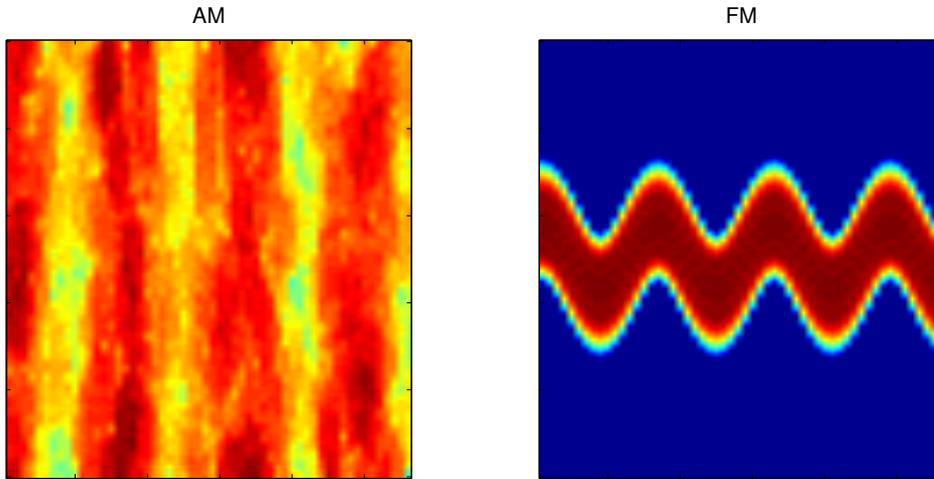


Fig. III.3: **AM and FM.** This figure displays TF spectrograms of a signal AM (left) and a signal FM (right). In each diagram, time is horizontal, frequency vertical.

Whereas no general framework can be given for encompassing all possible situations, two main classes of nonstationarities can be distinguished, which both give rise to a clear picture in the time-frequency plane (shown in Fig. III.3): *amplitude modulation* (AM) and *frequency modulation* (FM). We will base the following discussions on such classes.

In the first case (AM), a basic, stochastic representative of the class can be modelled as:

$$x(t) = (1 + \alpha \sin \frac{2\pi t}{T_0}) e(t), t \in T, \quad (\text{III.4})$$

with $\alpha \leq 1$ and where $e(t)$ is white Gaussian noise, T_0 is the period of the AM, and T the observation duration.

In the second case (FM), a deterministic model can be defined as:

$$x(t) = \sin 2\pi(f_0 t + \alpha \sin \frac{2\pi t}{T_0}), t \in T, \quad (\text{III.5})$$

with $\alpha \leq 1$ and where f_0 is the central frequency of the FM, T_0 its period, and T the observation duration.

3.2.2 Probability laws distances

Within the chosen time-frequency perspective, the stationarity test amounts to compare local spectra with their average over time thanks to some “*distance*”

(see Equations III.2 and III.3), the choice of a distance (or dissimilarity) measure is therefore instrumental for contrasting local vs. global features.

Many approaches are available in the literature [7] that, without entering into too much details, can be broadly classified in two groups. In the first one, the underlying interpretation is that of a *probability density function*, such as distance of Kolmogorov, distance of Kullback and distance of Jensen. One of the most efficient candidate being the well-known *Kullback-Leibler* (KL) divergence defined as

$$d_{\text{KL}}(G, H) := \int_{\Omega} (G(f) - H(f)) \log \frac{G(f)}{H(f)} df, \quad (\text{III.6})$$

where, by assumption, the two distributions $G(\cdot)$ and $H(\cdot)$ to be compared are positive and normalized to unity over the domain Ω . In our context, such a measure can be envisioned for (always positive) spectrograms thanks to the probabilistic interpretation that can be attached to distributions of time and frequency [30].

3.2.3 Spectral distances

A second group of approaches, which is more of a *spectral* nature, is aimed at comparing distributions in both shape and level, such as log spectral deviation, Itakura-Saito distance and Itakura distance. The simplest examples in this respect is the *log-spectral deviation* (LSD) defined as

$$d_{\text{LSD}}(G, H) := \int_{\Omega} \left| \log \frac{G(f)}{H(f)} \right| df. \quad (\text{III.7})$$

3.2.4 Combined distance

Intuitively, the KL measure (III.6) should perform poorer than the LSD one (III.7) in the AM case (III.4), because of normalization. It should however behave better in the FM case (III.5), because of its recognized ability at discriminating distribution shapes. In order to take advantage of both measures, it is therefore proposed to combine them in some ad hoc way as

$$d(G, H) := d_{\text{KL}}(\tilde{G}, \tilde{H}) \cdot (1 + \lambda d_{\text{LSD}}(G, H)), \quad (\text{III.8})$$

with \tilde{G} and \tilde{H} the normalized versions of G and H , and where λ is a trade-off parameter to be adjusted. In practice, the choice $\lambda = 1$ ends up with a good performance, as justified in Fig. III.4 (the performance measure used in this Figure is the inverse of the maximum value (over N_h) of the index of nonstationarity INS (III.17) (i.e., an inverse measure of contrast) to be defined and explained later in Section 3.5.

3.3 Structure of the Test

Ideally, stationarity corresponds to an equivalence of local and global spectra. In practice however, a close *similarity* is expected, instead of an absolute equiva-

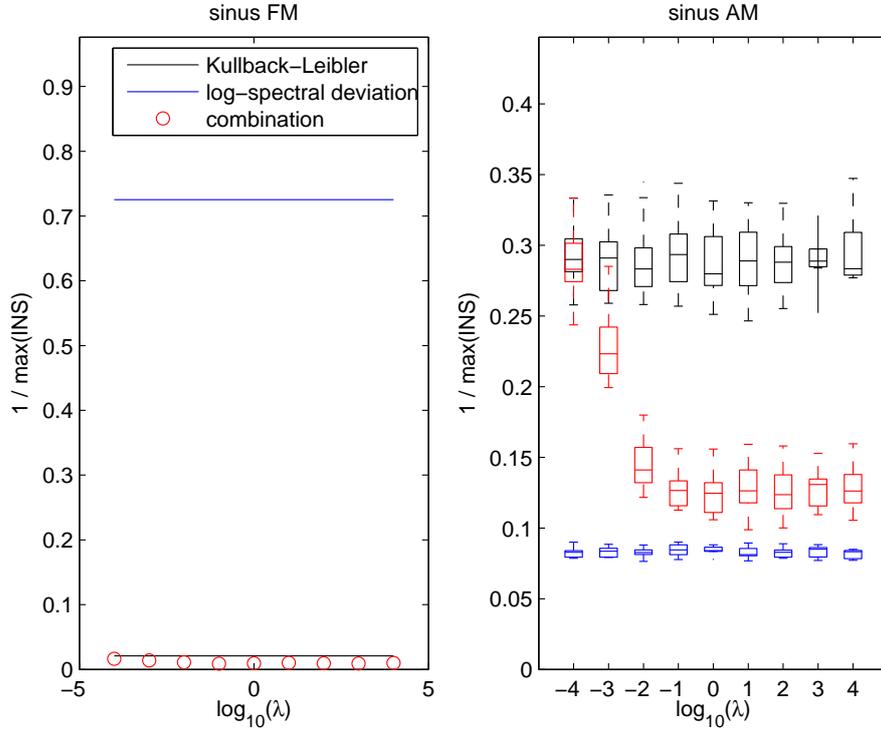


Fig. III.4: **Choosing a distance.** The inverse of the maximum value (over N_h) of the index of nonstationarity INS defined in (III.17) is used as a performance measure. Comparing the Kullback-Leibler (KL) divergence with the log-spectral deviation (LSD), a better result (i.e., a lower value) is obtained with KL (black) in the FM case (left, with $\alpha = 0.03$), and with LSD (blue) in the AM case (right, with $\alpha = 0.5$). A better balanced performance is obtained when using the combined distance (red) defined in (III.8): in the FM case, this measure performs best, and in the AM case it achieves a good contrast when $\lambda \geq 1$. In the AM case, the boxplots resulting from 10 realizations of the process are displayed.

lence, which can be represented by the fluctuations in time of the divergences $c_n^{(\cdot)}$ (III.2 and III.3) between local spectra and global one.

In order to measure the fluctuations, we make use of l_2 -distance defined by

$$L(g, h) := \frac{1}{N} \sum_{n=1}^N (g_n - h_n)^2 \quad (\text{III.9})$$

for any pair of sequences $\{(g_n, h_n), n = 1, \dots, N\}$.

As far as the intrinsic variability of surrogate data is concerned, the dispersion of divergences under the null hypothesis of stationarity can be measured by the distribution of the J empirical variances

$$\left\{ \Theta_0(j) = L \left(c^{(s_j)}, \langle c^{(s_j)} \rangle_{n=1, \dots, N} \right), j = 1, \dots, J \right\}. \quad (\text{III.10})$$

This distribution allows for the determination of a threshold γ above which the null hypothesis is rejected.

The effective test is therefore based on the statistics

$$\Theta_1 = L \left(c^{(x)}, \langle c^{(x)} \rangle_{n=1, \dots, N} \right) \quad (\text{III.11})$$

and takes on the form of the one-sided test:

$$\begin{cases} \Theta_1 > \gamma & : \text{“nonstationarity”}; \\ \Theta_1 \leq \gamma & : \text{“stationarity”}. \end{cases} \quad (\text{III.12})$$

3.4 Null Hypothesis of Stationarity

3.4.1 Distribution based on surrogates

The basic ingredient (and originality) of the approach is the use of surrogate data for creating signals whose spectrum is identical to that of the original one while being stationarized by getting rid of a well-defined structuration in time. Since those surrogates can be viewed as distinct, independent realizations of the stationary counterpart of the analyzed signal, the central part of the stationary test is based on the statistical distribution of the J surrogates variances (i.e. histogram of $\Theta_0(j)$ (III.10)), according to which the null hypothesis of stationarity and threshold γ will be set.

A histogram of Θ_0 is therefore illustrated in Fig. III.5, in the case of a white Gaussian noise without or with sinus FM, and compared with test Θ_1 (III.11). It is obvious that for a stationary signal, Θ_1 is close to the center of the histogram of Θ_0 ; on the contrary, Θ_1 is far away from the center for nonstationary case. Therefore, a null hypothesis based on the distribution is feasible.

However, calculating an actual histogram of Θ_0 in the asymptotic regime ($J = 5000$) is definitely time-consuming, so some fast way should be found out for practical use by studying the characteristics of the distribution. Moreover, it should be pointed out that for the same signal under test, the distance of Θ_1 from the center of histogram of Θ_0 , which means the degree of non-stationarity, depends not only on the observation interval but also (as shown in Fig. III.5) on the window length N_h as well (which will be detailed soon in Subsection 3.5).

3.4.2 Gamma model?

When using the combined distance (III.8) suggested above in Section 3.2, an empirical study (on both AM and FM signals) shows that such a distribution

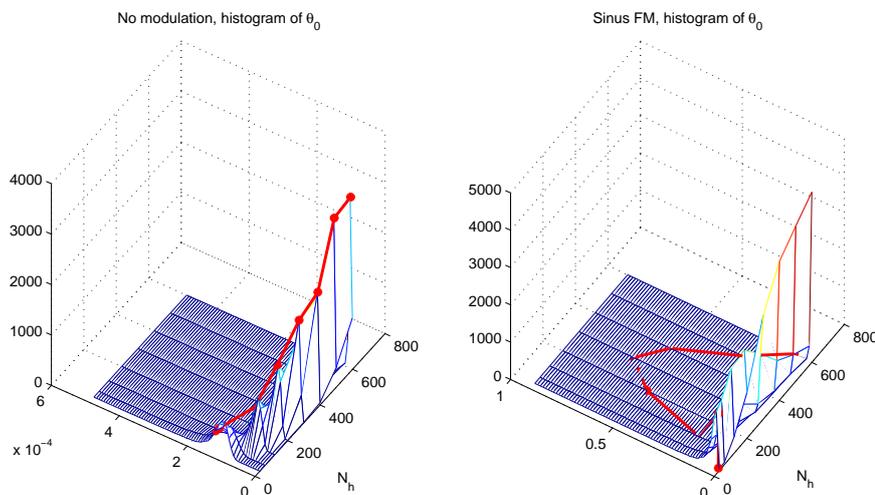


Fig. III.5: Statistical characterization of null hypothesis of stationarity 1. This figure makes a comparison of histogram of Θ_0 ($J = 5000$) with test Θ_1 (the red curve) as a function of window length N_h . The simulations have been conducted on a white Gaussian noise of 2048 data point without (left) and with (right) sinus FM with a period of $T_0 = 512$, for different window lengths N_h (of 5 tapers). As expected, Θ_1 is close to the center of the histogram for stationary case (left) and far from the center for nonstationary case (right).

can be fairly well approximated by a *Gamma* distribution. This makes sense since, according to (III.9), the test statistics basically sums up squared, possibly dependent quantities which themselves result from a strong mixing likely to act as a Gaussianizer. An illustration of the relevance of this modeling is given in Fig. III.6, where Gamma fits are superimposed to an actual histogram in the asymptotic regime ($J = 5000$ surrogates).

Assuming the Gamma model to hold, it is now possible to estimate its shape and scale parameters α and β directly from the J surrogates, e.g., in a maximum likelihood sense. As illustrated in Fig. III.7, the study on the estimators' behaviour of parameters $\hat{\alpha}$ and $\hat{\beta}$ as a function of the number of surrogates J for different taper lengths N_h shows a rapid convergence towards the asymptotic regime. The quick convergence with the increase of J is also in favor of the assumption that the statistical distribution of Θ_0 can be fairly modeled by a Gamma distribution with its two parameters.

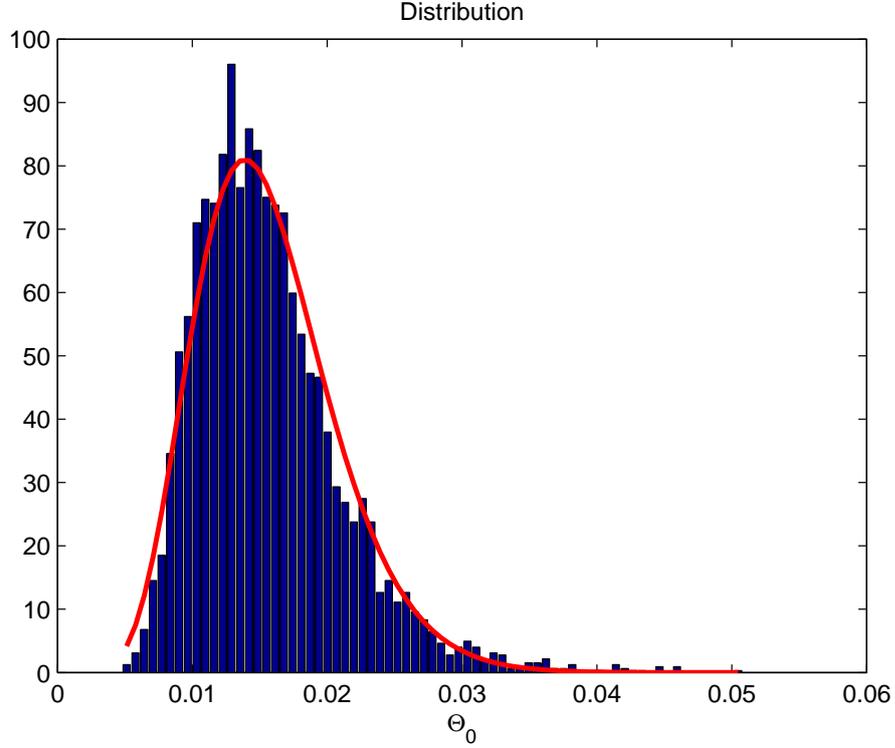


Fig. III.6: **Statistical characterization of null hypothesis of stationarity 2.** This figure presents a typical asymptotic histogram ($J = 5000$) of Θ_0 and its adjustment by a Gamma law (red curve). Experimental conditions: the signal $x(t)$ is a white Gaussian noise of 1024 data points, modulated in amplitude with a period of $T = 256$; the simulation has been conducted with $K = 5$ tapers, each of length $N_h = 61$, on the basis of 5000 realizations.

3.4.3 How many surrogates needed?

Small number of surrogates is required for accelerating further the numerical calculation, so the next question is how many surrogates is a good trade-off between complexity and precision.

Assuming that the mean of infinite realizations approximate the real value and that 1000 realizations could be regarded as almost infinite, the normalized bias and variance of α and β as a function of window length N_h and of the number of surrogates J could be defined as follows:

$$\widetilde{Bias}(\hat{\alpha}(N_h, J)) = \frac{\mathbb{E}\{\hat{\alpha}(N_h, J)\} - \mathbb{E}\{\hat{\alpha}(N_h, 1000)\}}{\mathbb{E}\{\hat{\alpha}(N_h, 1000)\}}; \quad (\text{III.13})$$

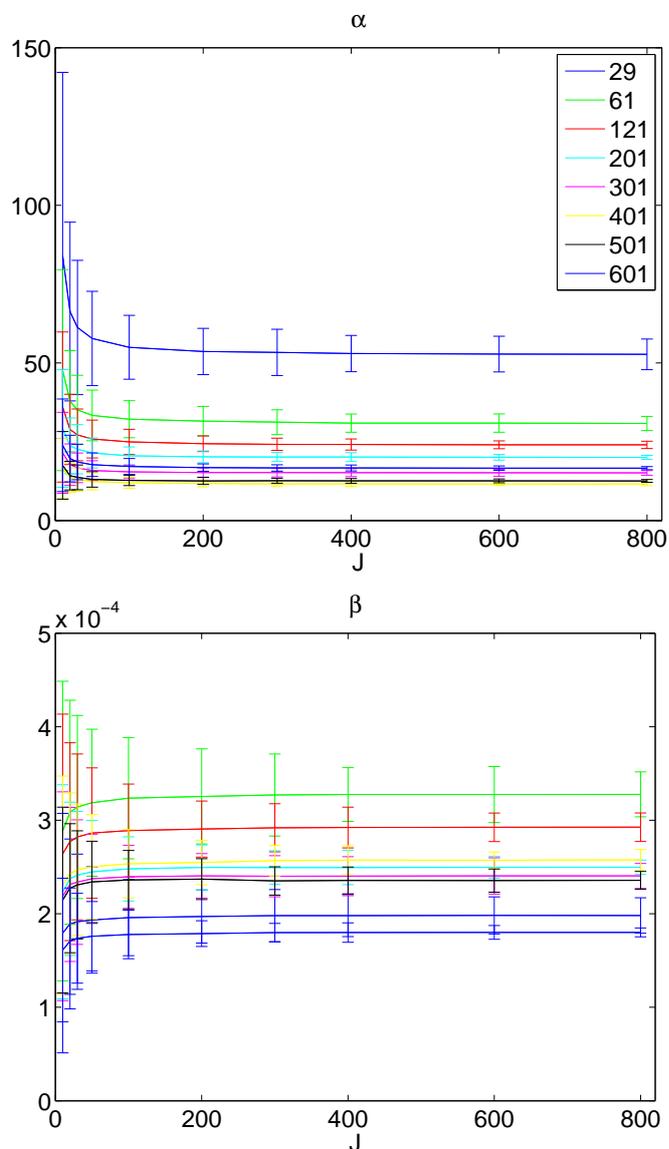


Fig. III.7: **Statistical characterization of null hypothesis of stationarity 3.** This figure presents the mean and standard deviation of the parameter estimators α (top) and β (bottom) of the Gamma distribution of Θ_0 as a function of the number of surrogates J for different taper length N_h (indicated in the category) of 5 tapers. The simulations have been conducted on the same signal as in Fig. III.6, and a rapid convergence towards the asymptotic regime is observed.

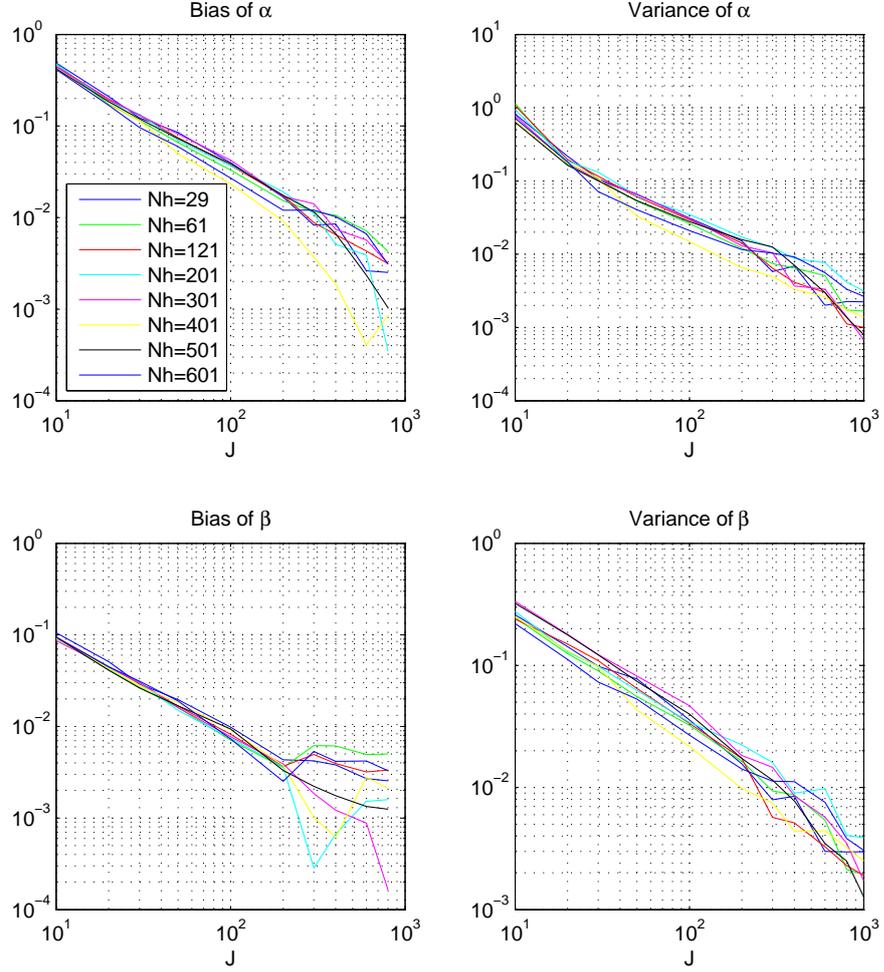


Fig. III.8: **Statistical characterization of null hypothesis of stationarity 4.** The normalized bias (left) and variance (right) relative to estimators of parameters α (on top row) and β (on bottom row) of the supposed Gamma law of Θ_0 are drawn as a function of the number of surrogates J and of taper length N_h (same experimental conditions as in Fig. III.7).

$$\widetilde{Var}(\hat{\alpha}(N_h, J)) = \frac{\mathbb{E}\{(\hat{\alpha}(N_h, J) - \mathbb{E}\{\hat{\alpha}(N_h, J)\})^2\}}{(\mathbb{E}\{\hat{\alpha}(N_h, 1000)\})^2}; \quad (\text{III.14})$$

$$\widetilde{Bias}(\hat{\beta}(N_h, J)) = \frac{\mathbb{E} \left\{ \hat{\beta}(N_h, J) \right\} - \mathbb{E} \left\{ \hat{\beta}(N_h, 1000) \right\}}{\mathbb{E} \left\{ \hat{\beta}(N_h, 1000) \right\}}; \quad (\text{III.15})$$

$$\widetilde{Var}(\hat{\beta}(N_h, J)) = \frac{\mathbb{E} \left\{ (\hat{\beta}(N_h, J) - \mathbb{E} \left\{ \hat{\beta}(N_h, J) \right\})^2 \right\}}{(\mathbb{E} \left\{ \hat{\beta}(N_h, 1000) \right\})^2}. \quad (\text{III.16})$$

It appears (in Fig. III.8) that the (normalized) bias and variance are very weakly dependant on N_h and decrease, as a function of J , with a slope of -1 in logarithmic unit of base 10. Moreover, the error becomes inferior to 5% for J larger than 50: we may thus keep this quantity as a good trade-off between complexity and precision.

3.4.4 Threshold based on a statistical significance

To sum up, Fig. III.9 supports again the claim that the distribution based on surrogates can be well approximated by a Gamma model (top row) in both stationary (left) and non-stationary (right) case, and also the “theoretical” probability density function (more precisely, its estimate in the asymptotic regime $J \approx 5000$) can be reasonably well approached by a reduced number of surrogates (typically, $J \approx 50$) (bottom row).

Finally, the value of the test Θ_1 (III.11), computed on the actual signals under study, is also plotted in the same figure and (as expected,) shown to stand in the middle of the distribution in the stationary case (bottom left) while clearly appearing as an outlier in the considered nonstationary situation (bottom right). Therefore, given the parametric model of Gamma law for the distribution based on a reduced number of surrogates, it becomes straightforward to derive rapidly a threshold γ for (III.12) above which the null hypothesis of stationarity is rejected with a *given statistical significance*.

3.4.5 Verification of null hypothesis

Furthermore, we would like to verify that the approach of the test described above reproduces well the null hypothesis of stationarity. Two elements are worth attention in this respect:

- (i) the fact that any stationary signal will be regarded as nonstationary in a probability that is directly imposed by the threshold with a given significance;
- (ii) and the fact that the surrogates follows themselves well the null hypothesis of stationarity.

The latter point is not difficult to verify, as we have already studied before that their distribution can be well described by the parametrical model of Gamma law. Turning back to the former, if one starts with a stationary signal, it is expected that the decision made from the test itself will be that: the signal is not stationary

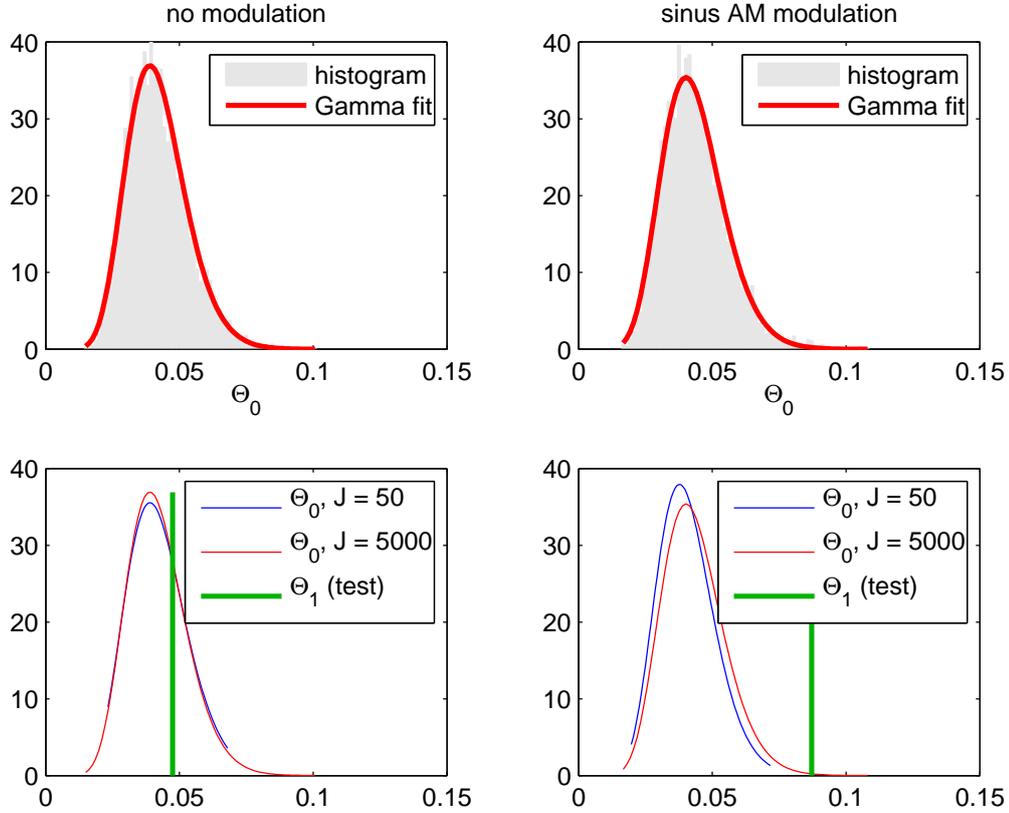


Fig. III.9: Statistical characterization of null hypothesis of stationarity 5. The top row superimposes empirical histograms of the variances Θ_0 (III.10) based on $J = 5000$ surrogates (grey) and their Gamma fits (red), in the case of a white Gaussian noise without (left) and with (right) a sinusoidal AM (with $\alpha = 0.5$). The bottom row compares the corresponding probability density functions, as parameterized by using $J = 50$ (blue) and 5000 (red) surrogates. The values of the test statistics Θ_1 (III.11) computed on the analyzed signal are pictured in both cases as green lines.

in a certain probability of the studied cases, which should be *a priori* the false alarm of the threshold that one settles.

For that, given a random stationary process (here, type AR as a good example strictly stationary), it is possible to take from it an ensemble of different realizations and to put the test (III.12) proposed in action by constructing as many collections of surrogates. We thus verify that, with the Monte-Carlo simulations, the result

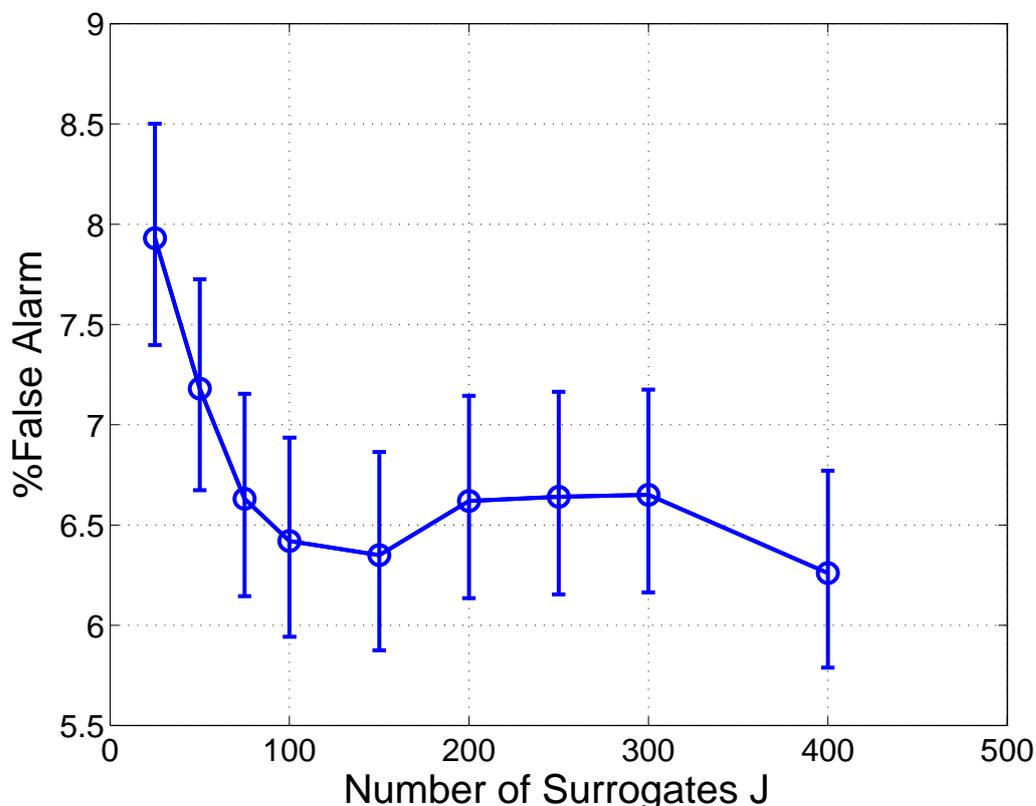


Fig. III.10: **Verification of the null hypothesis of stationarity.** A stationary signal of type AR(2) (pole at the reduced frequency 0.3 and with modulus 0.8) with 1024 data points, is submitted to the test with $K = 5$ tapers. Monte-Carlo estimate (based on a threshold with a given significance level of 95%, and 10^4 independent realizations) of the number of bad decision (“nonstationary”) is drawn as a function of the number J of surrogates used. The real level of confidence, when $J \geq 75$, remains around 6.5% for the threshold with a level of false alarm fixed *a priori* to 5 %.

of the test (illustrated in Figure III.10) leads to a *false alarm* of nonstationarity in around 6.5 % of the cases, for a threshold with a significance level of 95 %, (i.e. a level of false alarm fixed *a priori* to 5 %), given that one uses 75 or more surrogates.

We have therefore an acceptable level of false alarm, and its difference from the prescribed level comes from the fact that modelling the distribution of surrogates in Gamma law is a convenient approximation – but not a result theoretically justified. Therefore, the cases where the law fails to perfectly represent the statistics of surrogates will increase the false alarm of nonstationarity by the test. In a way, we

can say that this approach is pessimistic, regarding the attribution of stationarity, because it rejects this hypothesis a little more often than necessary.

3.5 Index and Scale of Nonstationarity

Assuming that the null hypothesis of stationarity is rejected, an *index of nonstationarity* can be introduced as a function of the ratio between the test statistics (III.11) and the mean value (or the median) of its stationarized counterparts (III.10):

$$\text{INS} := \sqrt{\frac{\Theta_1}{\frac{1}{J} \sum_{j=1}^J \Theta_0(j)}}. \quad (\text{III.17})$$

If the signal happens to be stationary, INS is expected to take a value close to unity whereas, the more nonstationary the signal, the larger the index. The corresponding threshold for INS would be $\gamma' = \sqrt{\gamma / (\sum_j \Theta_0(j) / J)}$.

Finally, it has to be remarked that, whereas the tested stationarity is globally relative to the time interval T over which the signal is chosen to be observed, the analysis still depends on the window length N_h of the spectrogram. Given T , the index INS will therefore be a function of N_h and, repeating the test with different window lengths, we can end up with a typical *scale of nonstationarity* SNS, depending on the relative value of N_h/T , defined as:

$$\text{SNS} := \arg \max_{\frac{N_h}{T}} \left\{ \text{INS}\left(\frac{N_h}{T}\right) \right\}, \quad (\text{III.18})$$

with N_h in the whole range of window lengths such that the prescribed threshold is exceeded in (III.12).

3.6 Examples

In order to illustrate the proposed approach and to support its effectiveness, a simple example is given in Fig. III.11. The analyzed signal consists of one realization of an AM process of the form (III.4). Depending on the relative values of the period T_0 and the time interval T , three regimes can be intuitively distinguished:

1. if $T \gg T_0$ (macroscale), many oscillations are present in the observation, creating a sustained, well-established quasi-periodicity that corresponds to a form of stationarity;
2. if $T \approx T_0$ (mesoscale), emphasis is put on the local evolutions due to the AM, suggesting to rather consider the signal as nonstationary, with some typical scale;
3. if $T \ll T_0$ (microscale), no significant difference in amplitude is perceived, turning back to stationarity.

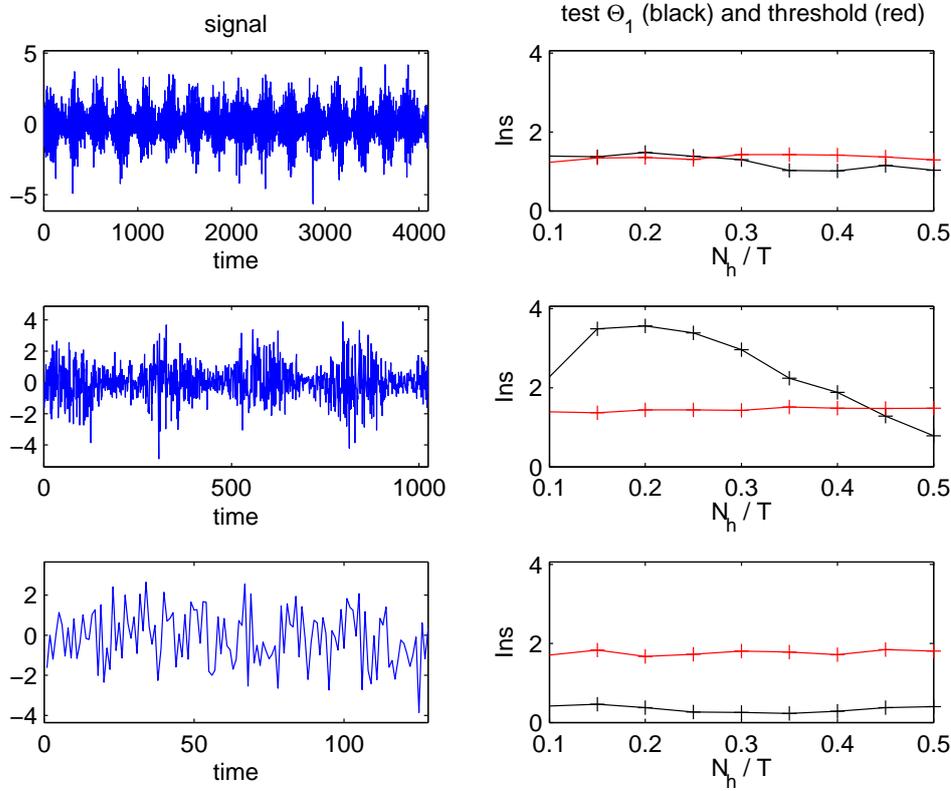


Fig. III.11: **AM example** ($\alpha = 0.5$). In the case of the same signal (III.4) observed over different time intervals (left column), the indice of nonstationarity INS (right column, in black) are consistent with the physical interpretation according to which the observation can be considered as stationary at macroscale (top row), nonstationary at mesoscale (middle row) and stationary again at microscale (bottom row). The threshold (in red) of the stationarity test (III.12) is calculated with a confidence level of 95% and represented in term of INS as γ' , with $J = 50$ surrogates. In the nonstationary case, the position of the maximum of INS at the relative analysis window N_h/T also gives an indication of a typical scale of nonstationarity.

What is shown in Fig. III.11 is that such interpretations of *relative stationarity* are precisely evidenced by the proposed test. They are moreover quantified in the sense that, when the null hypothesis of stationarity is rejected (middle diagram), both an *index* and a *scale* of nonstationarity can be defined according to (III.17) and (III.18). In the present case, the maximum value of INS is obtained for SNS

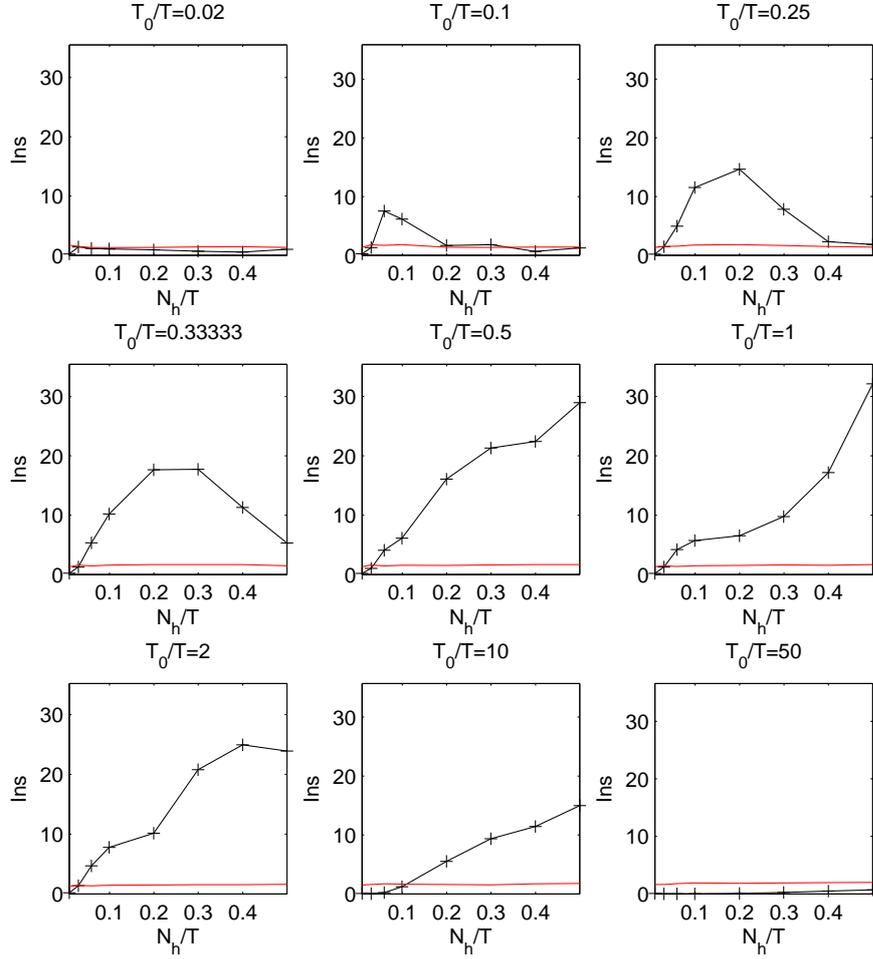


Fig. III.12: **Periodic correlation of SNS.** Each diagram displays the index of nonstationarity INS (in black) as a function of relative window length N_h/T (with 5 Hermite tapers, T being the observation length), in the case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of relative period (T_0/T from 0.02 to 50), based on 2048 data points. The threshold (in red) is calculated as γ' with a confidence level of 95% and $J = 50$. In the nonstationary cases, the typical scale of nonstationarity SNS indicates the relative period of the signal with respect to the observation interval.

≈ 0.2 , in qualitative accordance with the 4 AM periods entering the observation window.

In this specific example, the data could be referred to as *cyclostationary* and analyzed by tools dedicated to such processes [61]. However, it has to be stressed that no such a priori modeling is assumed in the proposed methodology, and that the existence of a typical scale of nonstationarity (related to the periodic correlation) naturally emerges from the analysis.

The interesting correlation between the period of signal T_0 and the typical scale of nonstationarity SNS is shown in Fig. III.12, in the case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of period. First similarly, when $T \ll T_0$ (left top) or $T \gg T_0$ (right bottom), the results of test are always stationary. In other cases (all the rest diagrams) however that $T \approx T_0$, local evolutions due to the FM is emphasized, leading to nonstationarity. Then in nonstationary cases, the typical scale of nonstationarity SNS (i.e. according to (III.18), the maximum value of the relative window length to observation interval N_h/T in the whole range such that the prescribed threshold is exceeded) indicates the relative period of signal to observation interval T_0/T . For example (left middle diagram), SNS is around 0.3 when $T_0/T = 1/3$. However, in the case when the value of the relative period T_0/T goes beyond the range of the relative window length N_h/T calculated, SNS can no longer be displayed in the diagram and the index of nonstationarity INS increases continuously with N_h/T (e.g. right middle diagram). To sum up, SNS (if shows) indicates the relative period of signal under test, in other words, the most nonstationary degree happens when window length is close to *one* period of signal (if any).

In order to illustrate that the proposed approach could also be applied to a stationarity test relative to frequency subbands of interest, Fig. III.13 gives another particular example of a signal with two frequency components: one of them (pure frequency) could be considered as stationary; while the stationarity of the other one (sinusoïdal FM III.5 with a given period of modulation) depends on the observation interval. The application of the test and the corresponding evaluation of index (III.17)-(III.18) illustrate a good agreement with what the intuition would expect from such a situation: The component of high frequency bands is always detected as stationary; by contrast, the null hypothesis of stationarity is rejected for the component of low frequency bands within the domain where the observation interval is neither long nor short compared to the period of modulation, thus similar to the situation in the AM example above (cf. Fig. III.11) in this aspect.

In the latter case, the value of index (which is drawn here in logarithm for a better readability) offers a measure of nonstationarity relative to an interval of analysis (i.e. relative window length N_h/T). Given fixed duration of observation T , the position of maximum gives furthermore an indication of the typical scale of nonstationarity SNS.

Finally, some variations of this approach of stationarity test are proposed and compared in Appendix (Chapter V: Section 3)

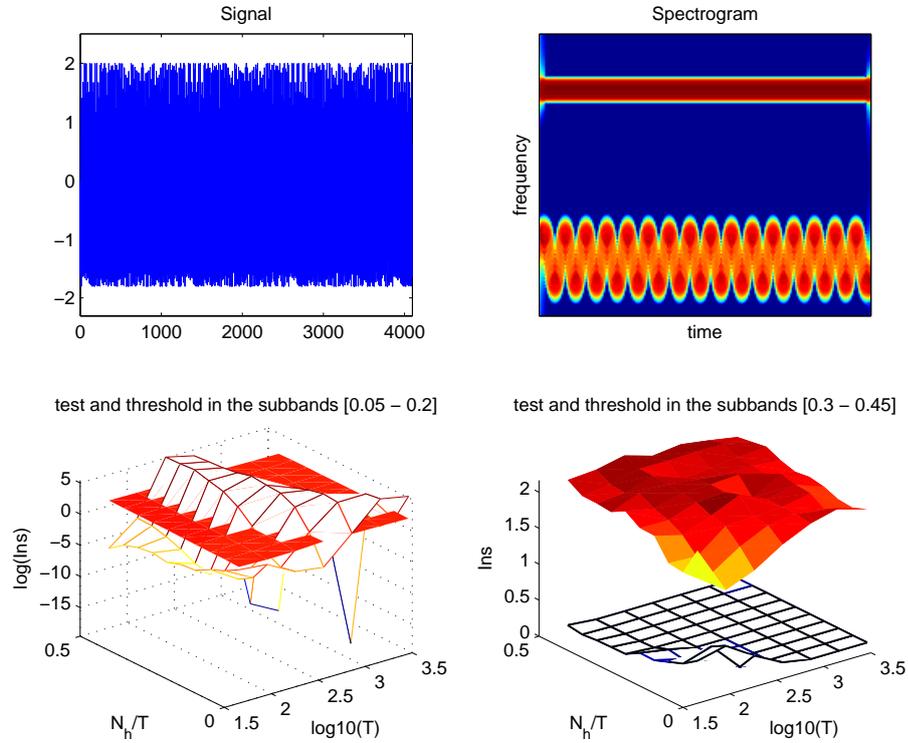


Fig. III.13: **FM example** ($\alpha = 0.03$). The top diagrams present a signal (left) and its multitaper spectrogram (right) with 5 tapers, the analysis based on the reduced frequency band $[0, 0.5]$. The bottom diagrams present the corresponding index of nonstationarity (white surfaces) within the subbands $[0.05, 0.2]$ (left) and $[0.3, 0.45]$ (right). In these two cases, the results are drawn as a function of the observation interval T of signal and of the relative interval of analysis N_h/T , N_h being the taper length of spectrogram. The thresholds γ' deduced from the $J = 50$ surrogates (for a confidence of 95 %) are drawn by the surface of colored facet, in agreement with the interpretation according to which the signal is stationary in the high band of frequency but nonstationary in low band when the observation interval is neither long nor short compared to the period of modulation.

4 Stationarity Test - One-Class Support Vector Machine Approach

A new approach has been proposed in Section 3 for testing stationarity from a TF viewpoint, relatively to a given observation scale, in both stochastic and deterministic contexts. A key point is that the null hypothesis of stationarity (which corresponds to time-invariance in the TF spectrum) is statistically characterized on the basis of a set of surrogates which all share the same average spectrum as the analyzed signal while being stationarized.

However, the proposed framework still leaves room for more thorough investigations as well as some variations. As to the test itself, the way that the estimated time-frequency spectrum fluctuates in time has been considered here by comparing local features (frequency “slices”) to a global one (the average spectrum resulting from marginalization) thanks to some distance measure. This is a classical approach, but it has the drawback of calling for the choice of a (more or less arbitrary) distance and the evaluation of associated distributions. Moreover, the null hypothesis of stationarity is characterized by a ad-hoc parametric model, empirically but not theoretically verified.

A different possibility would be to look at the statistical decision problem from a learning perspective [18, 39, 62]. Another operational framework of stationarity test relative to an observation scale is therefore developed in this Section by means of a powerful machinery of kernel methods, namely *one-class support vector machines*, where the stationarized surrogates are originally considered as a learning set and suitable descriptors is extracted from the TF plane by a comparison of local TF features to global one.

4.1 One-Class Support Vector Machines

4.1.1 Machine learning

Usually, when we ask computers to help us to solve some practical problem, we provide them with a sequence of instructions so that they can follow these instructions to accomplish the given task. In this case, the method and steps of implementation should be described explicitly for the computers.

In real world however, we are faced with a large variety of problems which can not be easily expressed by a traditional programming approach, such as classification of protein types based on the DNA sequence, or the classification of credit applications into those who will default and those who will repay the loan, or modeling a complex chemical reaction, where the precise interactions of the different reactants are not known, etc..

A way out of such problems is to let the computer learn the input/output functionality from examples, in the same way that babies and children learn many new things. This way of learning from experiences for machines is called *machine learning* [18], a concept that is first defined by Mitchell [56].

The availability of reliable learning systems is of strategic importance. On the one side, a wide range of applications that can not be reached by classical programming techniques depending on mathematical model can potentially be solved by such an approach. On the other side, we may also avoid much laborious design and programming inherent in the traditional solution methodology, at the expense of collecting some labelled data and running an off-the-shelf algorithm for learning the *input/out mapping*.

4.1.2 Support Vector machines

Among various machine learnings stands first *Support Vector Machine* (SVM) [22, 64]. First introduced by Vapnik [73, 74], SVM are learning systems that use a hypothesis space of *linear* functions in a *high dimensional (kernel-induced) feature space*, trained with a learning algorithm from *optimization* theory. This statistical learning strategy is such a powerful method that it outperforms most of other systems and has been widely applied to a variety of fields.

Generally, there are three main classes of SVM: *pattern recognition* (or *multi-class classification*) was historically the first that has been developed. It distinguishes between (or among) at least two classes of patterns; in the case of only one class, by contrast, one calls for *one-class classification*, which belongs to the realm of unsupervised learning and which is for the problem of novelty detection. Apart from classifications, another class of SVM is *regression* estimation which retains most of the properties of other SVMs but is for the problems of real-valued outputs.

4.1.3 One-class Support Vector machines

Turning back now to the topic of stationarity test itself, once a collection of stationarized surrogate data has been synthesized (as in Section 2), different possibilities are offered. Apart from the first parametric approach proposed in Section 3, we will here rather focus on an alternative viewpoint rooted in statistical learning theory: the collection of surrogates will be considered as a *learning set* and used to detect departure from stationarity.

In this context, the classification task is fundamentally a one-class classification problem and differs from conventional two-class pattern recognition problems in the way how the classifier is trained. The latter uses only target data to estimate a boundary which encloses most of them. The machinery of *one-class Support Vector Machines* (1-class SVM), which was introduced for outlier detection [63], can be used.¹ This technique has been successfully applied to a number of problems, including audio and biomedical signal segmentation [27, 34].

Let $\mathcal{Z} = \{z_1, \dots, z_J\}$ be a set of J surrogates (or a collection of features derived from it), and let $\kappa : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{R}$ be a *kernel function* that satisfies Mercer

¹The part of work in the use of 1-class SVM is cooperated with Cédric Richard from l'Université de Technologie de Troyes.

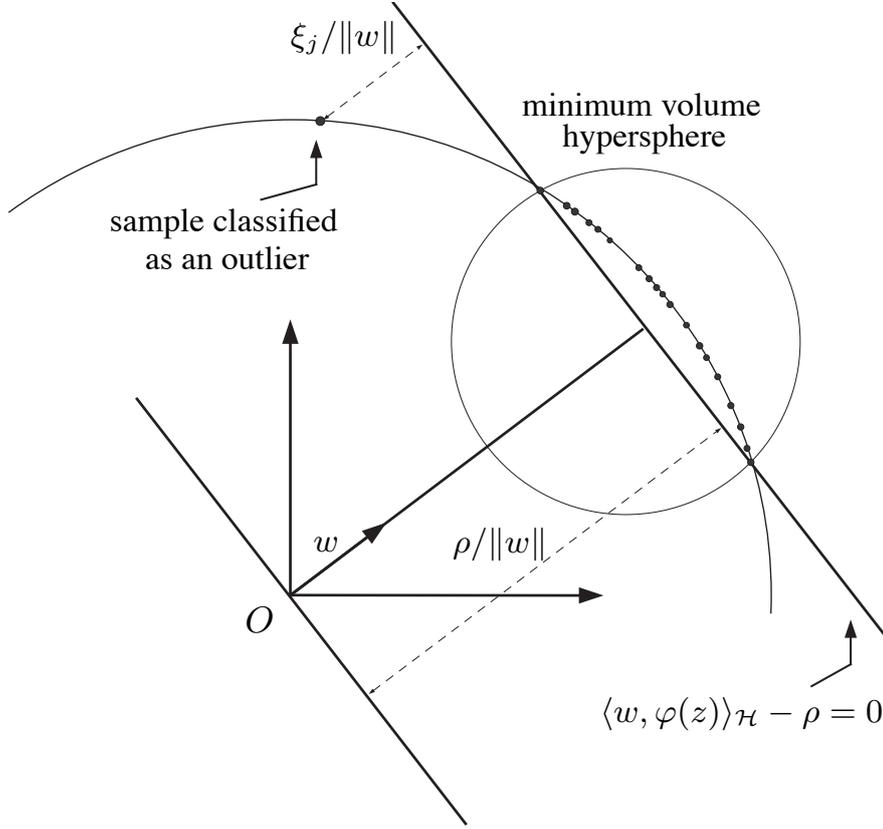


Fig. III.14: **1-class SVM**. This figure displays 1-class SVM with kernel $\kappa(z_i, z_j)$ depending only on $z_i - z_j$.

conditions. The latter can be used to map the z_j 's into a *feature space* denoted by \mathcal{H} via $\varphi : \mathcal{Z} \rightarrow \mathcal{H}$, with the definition $\varphi(z) = \kappa(z, \cdot)$.

The space \mathcal{H} is shown to be a reproducing kernel Hilbert space of functions with dot product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. The reproducing kernel property states that

$$\langle \kappa(z_i, \cdot), \kappa(z_j, \cdot) \rangle_{\mathcal{H}} = \kappa(z_i, z_j), \quad (\text{III.19})$$

which means that $\kappa(z_i, z_j)$ can be interpreted as the dot product between z_i and z_j mapped to \mathcal{H} by $\varphi(\cdot)$.

A classical example of Mercer kernel is the Gaussian kernel defined as

$$\kappa(z_i, z_j) = \exp(-\|z_i - z_j\|^2 / 2\sigma_0^2), \quad (\text{III.20})$$

where σ_0^2 is a bandwidth parameter. Note that it maps any data point onto a hypersphere of radius 1 since $\kappa(z_j, z_j) = 1$ for all z_j .

The learning strategy adopted by 1-class SVM is to map the data into the feature space corresponding to the kernel function, and determine the *hyperplane*

$$\langle w, \varphi(z) \rangle_{\mathcal{H}} - \rho = 0, \quad (\text{III.21})$$

which separates them from the origin with maximum margin. The *decision function*

$$d(z) = \text{sgn}(\langle w, \varphi(z) \rangle_{\mathcal{H}} - \rho) \quad (\text{III.22})$$

then gives on which side of the hyperplane any new point z falls in feature space, and determine if it may be considered as an outlier.

For kernels $\kappa(z_i, z_j)$ depending only on $z_i - z_j$ such as the Gaussian kernel, which map data onto a hypersphere, this strategy is equivalent to finding the minimum volume hypersphere enclosing the data [72] (See Fig. III.14).

Now, let us focus on the *optimization* problem solved to get the hyperplane parameters w and ρ . On the one hand, the distance $\rho/\|w\|$ that separates the hyperplane from the origin must be maximized. On the other hand however, the number of target samples wrongly classified as outliers must be minimized. Such samples z_j satisfy inequalities of the form $\langle w, \varphi(z_j) \rangle_{\mathcal{H}} \geq \rho - \xi_j$ with $\xi_j > 0$.

Based on these results, the decision function is found by minimizing the weighted sum of a regularization term $\|w\|^2$, and an empirical *error* term depending on the margin variable ρ and individual errors ξ_j

$$\begin{aligned} \min_{w, \rho, \xi} \quad & \frac{1}{2} \|w\|^2 + \frac{1}{\nu J} \sum_{j=1}^J \xi_j - \rho \\ \text{subject to} \quad & \langle w, \varphi(z_j) \rangle_{\mathcal{H}} \geq \rho - \xi_j, \quad \xi_j \geq 0, \end{aligned} \quad (\text{III.23})$$

with $\nu \in [0, 1]$. Basic properties of 1-class SVM are reported in [63]. An important result is that the ν parameter may be used to incorporate prior information about the frequency of novelty occurrences.

We shall now use 1-class SVM for stationarity test with (stationarized) surrogates from original signal as a learning set. The resulting decision rules will help us to distinguish between stationary and nonstationary processes.

4.2 Comparison of Local vs. Global Time-Frequency Spectra

Given a signal $x(t)$, the TF features can be extracted from its multitaper spectrogram (I.26). Interpreting the normalized TF distribution (for $f > 0$ only) at time t_n ($n = 1, \dots, N$):

$$\tilde{S}_{x,K}(t_n, f) := \frac{S_{x,K}(t_n, f)}{\int_0^\infty S_{x,K}(t_n, f) df} \quad (\text{III.24})$$

as a probability distribution, temporal features such as the time evolution of the local power P_n of the signal and that of the local frequency content F_n can be

described as:

$$\begin{cases} P_n &= \int \tilde{S}_{x,K}(t_n, f) df \\ F_n &= \int f \tilde{S}_{x,K}(t_n, f) df, \quad n = 1, \dots, N \\ F_n^2 &= \int f^2 \tilde{S}_{x,K}(t_n, f) df. \end{cases} \quad (\text{III.25})$$

By introducing the simplified notation of local average $\langle \cdot \rangle_{\tilde{S}_n}$ of some function $\phi(f)$ with respect to $\tilde{S}_{x,K}(t_n, f)$:

$$\langle \phi(f) \rangle_{\tilde{S}_n} := \int \phi(f) \tilde{S}_{x,K}(t_n, f) df, \quad (\text{III.26})$$

the local features in (III.25) could be expressed compactly by

$$P_n = \langle 1 \rangle_{\tilde{S}_n}; \quad F_n = \langle f \rangle_{\tilde{S}_n}; \quad F_n^2 = \langle f^2 \rangle_{\tilde{S}_n}. \quad (\text{III.27})$$

In order to keep a small number of features for a sake of clarity, we retain from this the following two characteristics for comparing local TF features to global ones:

$$\begin{cases} P = \text{std}(\{P_n\}_{n=1..N}) / \text{mean}(\{P_n\}_n) \\ F = \text{std}(\{F_n\}_{n=1..N}) / \text{mean}(\{\sqrt{\{F_n^2 - (F_n)^2\}}_n\}). \end{cases} \quad (\text{III.28})$$

The first one (P) is a measure of the fluctuations in time of the local power of the signal, whereas the second one (F) operates the same way with respect to the local mean frequency. These two characteristics of the signal are used as features: $z = (P, F)$, for the 1-class SVM with those of its stationary surrogates serving as a learning set.

4.3 Examples

Two test signals (cf. Section 3.2.1) are used as simple illustrations of possible nonstationary evolutions: AM of a random noise and deterministic FM. The models are expressed, for $t \in [0, T]$, as

$$\begin{aligned} (\text{AM}) \quad x(t) &= (1 + \alpha \sin 2\pi t/T_0) e(t); \\ (\text{FM}) \quad x(t) &= \sin 2\pi(f_0 t + \alpha \sin 2\pi t/T_0) + e(t), \end{aligned}$$

where $e(t)$ is white Gaussian noise and (FM case) f_0 is the central frequency. For both models, T_0 is the period of the modulation and $0 \leq \alpha \leq 1$ is the modulation factor. All nonstationarities cannot be subsumed under these two categories, but they are believed to give meaningful examples.

By using the SVM toolbox proposed in [20], the outputs of the test for AM and FM signals are displayed respectively in Fig. III.15 and Fig. III.16, for $T_0 = T/20$, T and $20T$, allowing to consider stationarity relatively to the ratio between the modulation period T_0 and the observation time T . They can be summarized as follows:

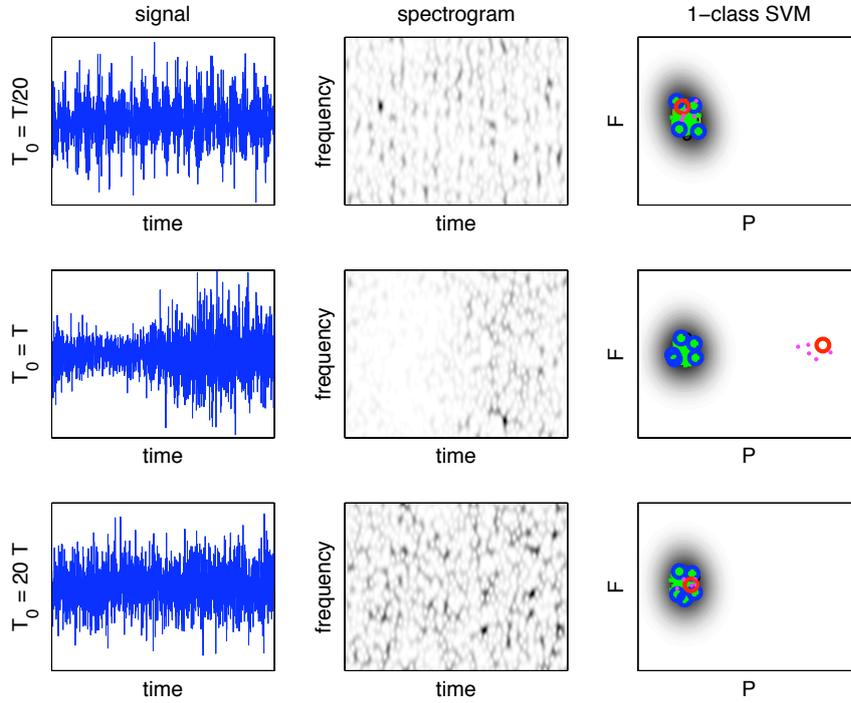


Fig. III.15: **AM example** ($\alpha = 0.5$). Signal (left), spectrogram (middle) and space (P, F) of the TF features (right) in AM situation. From top to bottom, the ratio between the modulation period and the observation time is respectively $T_0 = T/20$, T and $20T$, with $T = 1600$. In each case, the red circle corresponds to the (P, F) pair of one test signal used to produce the surrogates. Those surrogates ($J = 40$ in the experiments reported here) are plotted as green dots which, with 1-class SVM, define the domain of stationarity represented here as the gray shaded region, the blue circles corresponding to the support vectors. Magenta dots are independent realizations of the same test model. Other parameters are as follows — number of tapers: $K = 5$, length of tapers: $N_h = 387$, signal-to-noise ratio: $\text{SNR} = 10$ dB, SVM kernel κ is Gaussian, with $\sigma = 0.07$, $\nu = 0.05$.

Macroscale — For a large observation time (or a small modulation period, i.e., when $T_0 \ll T$), the situation can be considered as stationary, due to the observation of many similar oscillations over the observed time scale. This is reflected by the (P, F) feature of the tested signal (red circle, see caption) which lies inside the region defined by the 1-class SVM for the stationary

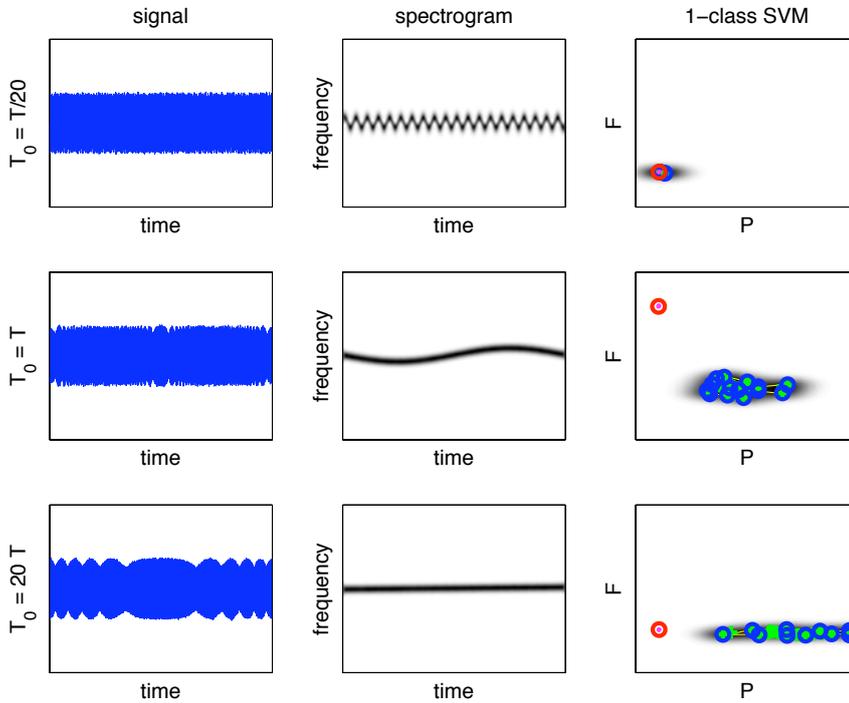


Fig. III.16: **FM example** ($\alpha = 0.02$). Signal (left), spectrogram (middle) and space (P, F) of the TF features (right) in FM situation. From top to bottom, $T_0 = T/20$, T and $20T$, with $T = 1600$. Same marks and same parameters as in Fig. III.15.

surrogates.

Mesoscale — For a medium observation time ($T \approx T_0$), the local evolution due to the modulation is prominent and thus the red circle for the modulated signal is well outside the stationary region, in accordance with a situation that can be referred to as nonstationary.

Microscale — Finally, if $T_0 \gg T$, the result turns back to stationarity because no significant change in the amplitude or the frequency is observed over the considered time scale.

Two remarks about the type of signals can be made with respect to the results shown in Fig. III.15 and and Fig. III.16.

First, although both AM and FM signals are seen as nonstationary in the mesoscale regime, in the AM case the nonstationarity manifests through a deviation of the local power P from the stationary domain (in gray) defined by

surrogates, whereas in the FM case, the deviation is observed mainly in the local frequency F . This is quite reasonable because the former has a modulation in amplitude, leading naturally to the fluctuations in power and thus the difference in power between nonstationary and stationary situations, and the latter has a modulation in frequency, resulting directly in the difference in frequency itself.

Second, it turns out that the random stationarity (in the AM case) presents a circular domain of stationarity in the space (P, F) , while the deterministic stationarity (in the FM case) naturally ends up with a much larger dispersion in P than in F , especially in the microscale regime. It is because that, for the latter, the spectrum is narrowband by construction. Furthermore, the randomization which underlies the construction of surrogates necessarily ends up with more power fluctuations in the stationarized data than in the original signal under test, and hence with the (P, F) pair of the signal which, at best, lies on the border (in the direction of P) of the support of the stationary class.

All above suggest that, in some sense, the position of the (P, F) pair of the signal under test with respect to the stationary region does not only supply an information about a possible nonstationarity but also, together with the shape of the stationary domain defined by surrogates, gives an indication about its type.

Finally, compared with the test of stationarity based on a parametric model in Section 3, we can see that the nonparametric approach of the test here by means of 1-class SVM gives a more vivid description by offering a 2D image where the domain of stationarity defined by the surrogates is outlined and the distance of the signal feature point from the stationary domain, if any, indicates to some extent the degree of nonstationarity. Another advantage is that, as just mentioned above, it offers the information about the possible type of the signal under analysis. By contrast, despite being faced with a choice of distance and based on a parametric model, the approach of the test in Section 3 can offer us directly the index of nonstationarity INS and additionally the scale of nonstationarity SNS.

5 Application: Speech

The test of stationarity relatively to an observation scale is well adapted to audio signals, such as human speech as well as musical records. An application of the test is therefore made on a simple experimental data of human speech: a continuing pronunciation of vowels “*a, e, i, o, u*”. Both the results from the approach based on the parametric model and those from the approach of 1-class SVM are shown in Fig. III.17 and Fig. III.18 and can be concluded as follows:

1. *Long term* — In a scale of several seconds, a succession of segments (voiced, unvoiced, silences, etc.) in different characteristics are present in the observation, and significant difference in amplitude is perceived, rather considering the signal as nonstationary, with some typical scale;
2. *Short term* — In a scale of several dozens of milliseconds, only one segment

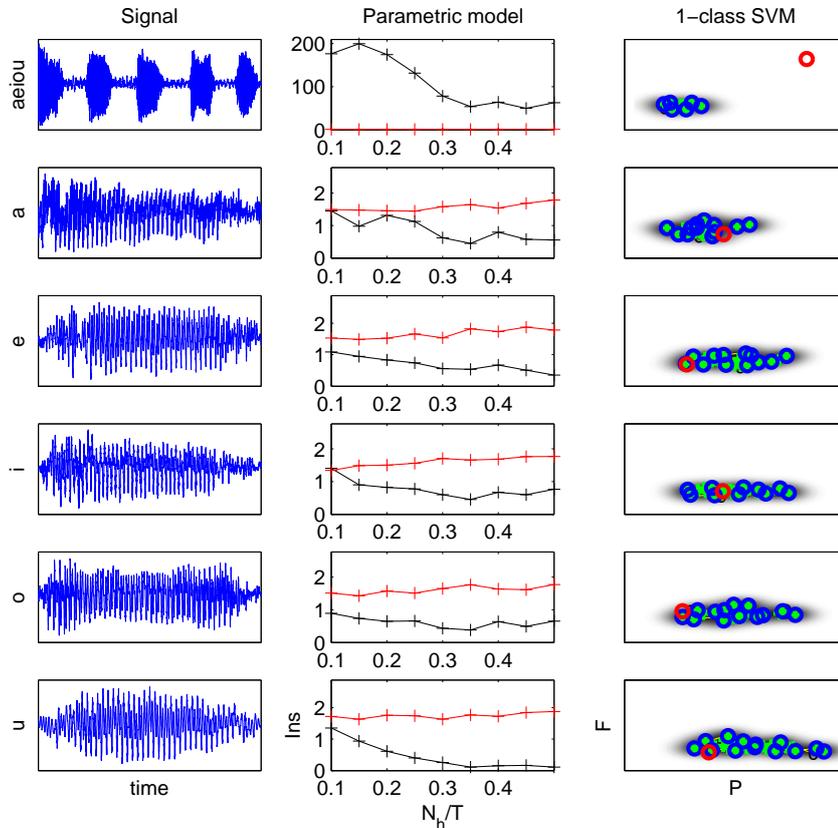


Fig. III.17: **Stationarity test on a signal of speech 1.** Signal (left), test by parametric model (middle) and test by 1-class SVM (right) in the case of a continuing pronunciation of vowels. From top to bottom, the observation intervals are corresponding respectively to the whole sequences of the vowels “a, e, i, o, u”, and each of them “a”, “e”, “i”, “o”, “u”. For the test by parametric model, each diagram displays the index of nonstationarity INS (III.17) (black) as a function of relative window length N_h/T (with 5 Hermite tapers, T being the observation length), and the threshold (red) γ' calculated with a confidence level of 95%, with $J = 40$ surrogates. In the nonstationary case, the position of the maximum of INS gives an indication of a typical scale of nonstationarity SNS (III.18). For the test by 1-class SVM, each diagram displays the space (P, F) of the TF features, where the red circle corresponds to the (P, F) pair of the tested signal. Its surrogates ($J = 40$) are plotted as green dots which, with 1-class SVM, define the domain of stationarity represented here as the gray shaded region, the blue circles corresponding to the support vectors. The simulations have been conducted with 5 tapers, each of length $N_h = T/4$. SVM kernel κ is Gaussian, with $\sigma = 0.07$, $\nu = 0.05$. The results from both of the approaches are consistent with the physical intuition that a speech might appear as “long-term nonstationary” and “short-term stationary”.

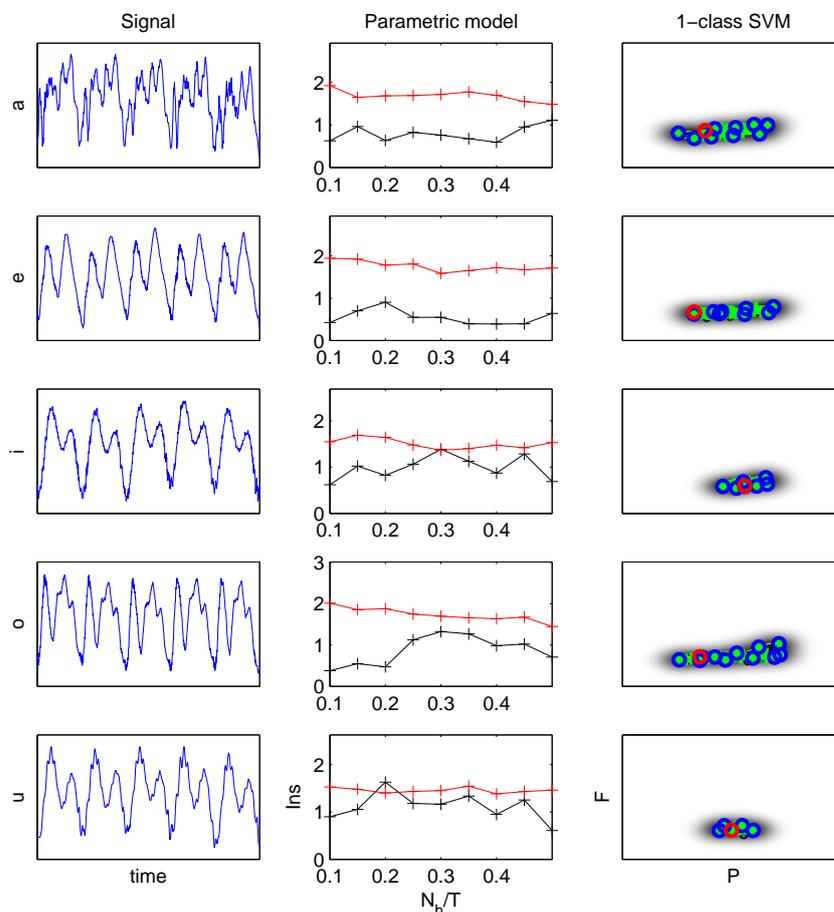


Fig. III.18: **Stationarity test on a signal of speech 2.** Signal (left), test by parametric model (middle) and test by 1-class SVM (right) in the case of the same signal of speech as in Fig. III.17. From top to bottom, the observation intervals are however much shorter, corresponding to only several oscillations within each of the vowels “a”, “e”, “i”, “o”, “u”. For the test by parametric model, each diagram displays the index of nonstationarity INS as a function of relative window length N_h/T ; for the test by 1-class SVM, each diagram displays the space (P, F) of the TF features. Same marks and same parameters as in Fig. III.17. The results show that a speech might appear as “mini-term stationary”.

appears in the observation which consists in fact of large numbers of similar oscillations, creating a sustained, well-established quasi-periodicity that is

willingly considered as a form of stationarity;

3. *Mini term* — In a scale of the millisecond, several oscillations within one segment are present in the observation, and the periodicity that assures in short term a form of temporal permanence is represented. As no significant AM or FM is observed, the signal keeps stationary in this scale.

Clearly, it agrees well with the physical intuition that a same audio signal may be described as nonstationary in long-term but stationary in short-term, and the stationarity is also observed in mini-term. It verifies that the notion of stationarity would have had no meaning without the relativity to an observation scale.

6 Conclusion

Testing for stationarity in signal processing and data analysis has already received some attention, but maybe not as much as might be expected from its ubiquitous nature. In this chapter, a novel test of stationarity has been proposed from a TF viewpoint, relatively to a given observation scale. A key point of the technique is the use of stationarized realizations from the analyzed signal, called surrogates, serving as the null hypothesis of stationarity. Two different methods based on the idea have been proposed for such a test, whose basic principles have been outlined, with a number of considerations related to its implementation and typical illustrations for supporting the effectiveness.

The framework of the first approach is to make use of distance measures between local and global spectra and characterize parametrically the null hypothesis of stationarity by constructing a model of distribution with surrogates. An index of nonstationarity can be provided, as well as a characteristic scale of the detected nonstationarity, when necessary.

By contrast, the second approach is a nonparametric one and comes from a learning perspective, where surrogates are considered as a learning set. Suitable descriptors (i.e. signal features) is extracted from the TF plane by comparing the local TF features with global one. The advantage of a statistical test implemented by 1-class SVM is to put emphasis on a general space of trait, so in this way there is no need to characterize the null hypothesis in a parametric way.

IV. Conclusions and Perspectives

It is known that a great deal of work has already been devoted to the stationary signal processing, which is theoretically well-analyzed and equipped with a multitude of powerful tools. However, the signals in real world are in majority non-stationary, which occur in a wide range of fields such as telecommunications, radar, sonar, speech processing, medical diagnosis and biomedical engineering, etc.. This has therefore led to the development of the tools designed specifically for nonstationary contexts. We have been devoted in this Thesis to the exploration of new approaches in this field.

Apart from an introduction of the basic concepts in Chapter I, we have been dedicated in Chapter II to a better representation tool, i.e., a better estimation of nonstationary spectrum in the case of chirp signals embedded in nonstationary noise. The wedding of reassignment with multitapering succeeds in a sharp localization for the chirp components and a reduced level of statistical fluctuations for the noise.

Two variations have been proposed based on different ways of combinations between the estimates obtained from different Hermite tapers. The sums of the estimates are designed mainly for nonstationary spectrum estimation, while the differences between them for a further chirp enhancement. The principles of the technique with both variations have been outlined, the implementations have been justified, and some examples have been offered for supporting the efficiency of the method.

Certainly, the method still leaves room for many possible extensions. Some possible variations (related to the choice of the averaging type) have been mentioned but, in the specific context of reassignment, other possibilities are offered which are worth to be explored further such as, e.g., combining reassignment vector fields and not only reassigned distributions. More fundamentally, both the performance evaluation in the spectrum estimation context and the tuning of parameters (such as thresholds) for chirp enhancement call for a theoretical analysis of the statistics of reassigned spectrograms.

In Chapter III, we have proposed a novel test of stationarity in a wide sense that is relative to a given observation scale and suitable for both stochastic and deterministic signals. A key point of the test is that the null hypothesis of stationarity (which corresponds to time invariance in the TF spectrum) is statistically characterized on the basis of a set of surrogates which all share the same average spectrum as the analyzed signal while being stationarized.

Two different approaches based on the idea have been proposed. One is to measure the distance between local and global spectra and construct a parametric model by surrogates for defining the null hypothesis of stationarity. An index as well as a characteristic scale of the detected nonstationarity can be provided. The other approach coming from a learning perspective makes use of 1-class SVM, considering surrogates as a learning set. The operation is made on some suitable descriptors which are extracted from the TF plane in a way of a comparison of the local TF features with global one.

Basic principles of the approaches and their implementations have been explained, with some typical illustrations for supporting the effectiveness. It is clear however that besides the variations mentioned in Appendix, a number of extensions from the present work are possible. As far as the test itself, the way that the estimated TF spectrum fluctuates in time has been considered here by comparing local features to a global one thanks to some distance measure, or from a learning perspective. In the latter case, only simple (specific bidimensional) features have been used for a sake of proof of concept, but the extraction of more suitable descriptors from the learning set still remains an open question. For example, using general TF representations emerging from the use of SVM machinery for TF, such as in [39]. Another possibility is to generalize the present approach to other forms of stationarity, which requires to define new specific stationarizing tools and signal representations, in the spirit of [32, 8]. Besides, in terms of TF distributions, one could imagine to go beyond multitaper spectrograms and take advantage of more recent advances (e.g. the combined technique: multitaper reassigned spectrogram in Chapter II) [77, 78].

V. Appendix

1 Alternative model

In practice, the “box” model used in Figures II.2 and II.3 for performance evaluation is unattainable because of uncertainty relations. As a substitute, we can instead consider an admissible model, the WVS of a suitably modulated and filtered white Gaussian noise.

Let us then define the process $x(t)$ to be analyzed as

$$x(t) = m(t) \int_{-\infty}^{+\infty} h(t-s) e(s) ds, \quad (\text{V.1})$$

where $m(t)$ is some amplitude modulation, $h(t)$ the impulse response of some linear filter and $e(t)$ white Gaussian noise such that $\mathbb{E}\{e(t)\} = 0$ and $\mathbb{E}\{e(t)e(s)\} = \sigma^2 \delta(t-s)$.

Writing $x(t) = m(t)y(t)$ with $y(t) := (h * e)(t)$ and using basic covariance properties of the WVD, we have

$$W_x(t, f) = \int_{-\infty}^{+\infty} W_m(t, f-f') W_y(t, f') df' \quad (\text{V.2})$$

and

$$W_y(t, f) = \int_{-\infty}^{+\infty} W_h(t-t', f) W_e(t', f) dt'. \quad (\text{V.3})$$

Since $e(t)$ is white Gaussian noise, we have $\mathbf{W}_e(t, f) := \mathbb{E}\{W_e(t, f)\} = \sigma^2$. This leads to $\mathbf{W}_y(t, f) = \sigma^2 |H(f)|^2$ (with $H(f)$ the Fourier transform of $h(t)$) and it readily follows that the WVS of the output process $x(t)$ reads

$$\mathbf{W}_x(t, f) = \sigma^2 \int_{-\infty}^{+\infty} W_m(t, f-f') |H(f')|^2 df'. \quad (\text{V.4})$$

If we now specify $m(t)$ and $H(f)$ by choosing a Gaussian model for both according to $m(t) = \exp\{-\alpha t^2\}$ and $H(f) = \exp\{-\beta(f-f_0)^2\}$ (where f_0 stands for some central frequency), we get

$$\mathbf{W}_x(t, f) = \sqrt{\frac{2\pi}{\alpha}} \sigma^2 e^{-2\alpha t^2} \int_{-\infty}^{+\infty} e^{-(2\pi^2/\alpha)(f-f')^2} e^{-2\beta(f'-f_0)^2} df', \quad (\text{V.5})$$

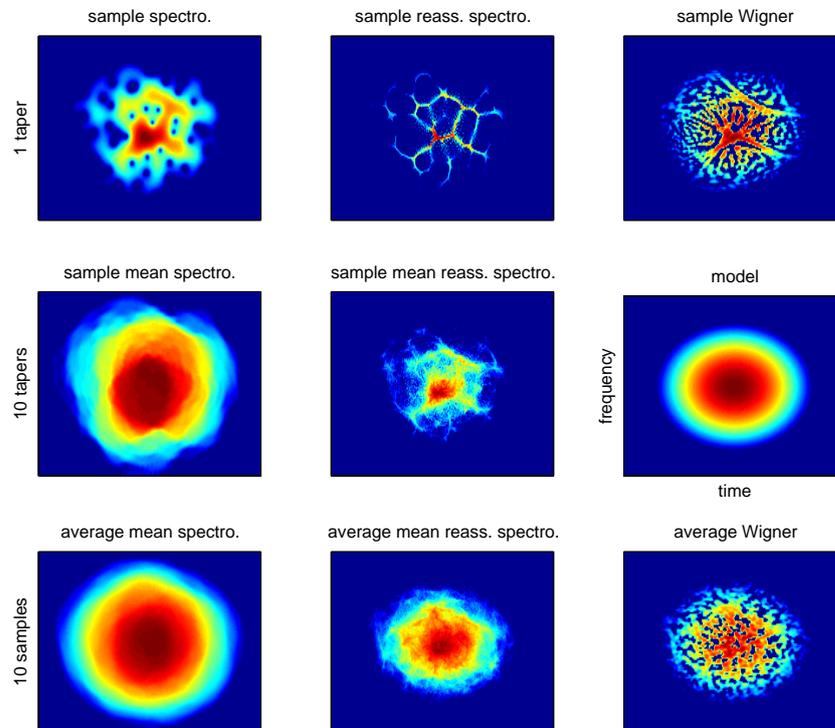


Fig. V.1: **Comparison of noise WVS estimates.** Each diagram represents a WVS estimate in the case of a Gaussian filtered white Gaussian noise modulated in time by a Gaussian window (the corresponding WVS model is given in the middle row of the right column). The first row consists of a spectrogram, its reassigned version and the WVD, based on one realization. The corresponding multitaper estimates (10 Hermite functions) are given in the middle row, whereas the bottom row displays ensemble averages of such estimates (10 independent realizations), together with the empirical WVS estimate on the same data set. In each diagram, time is horizontal, frequency vertical and the energy is coded logarithmically with a dynamic range of 30 dB.

and evaluating the above integral leads to the final result:

$$\mathbf{W}_x(t, f) = \frac{\sigma^2}{\sqrt{1 + \alpha\beta/\pi^2}} e^{-2\alpha t^2} e^{-\frac{2\beta}{1 + \alpha\beta/\pi^2}(f - f_0)^2}. \quad (\text{V.6})$$

It has to be remarked that this (valid) model corresponds to a non-negative

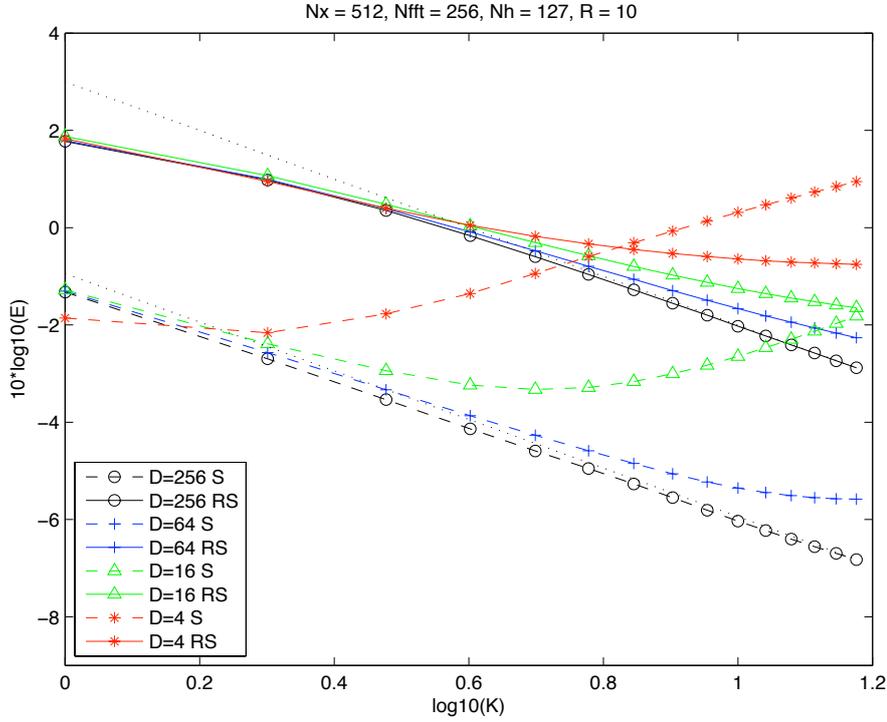


Fig. V.2: **Error measures in WVS multitaper estimates.** The figure plots, as a function of the number of tapers K , the error measure attached to multitaper (reassigned) spectrograms when the model is the WVS of a Gaussian filtered white Gaussian noise modulated in time by a Gaussian window, extending over an equivalent elliptic domain of area D . The simulations have been conducted (with up to $K = 15$ Hermite tapers, each of length 127) on the basis of 10 independent realizations of 512 data points each, with 256 frequency bins over the whole frequency range $[0, 1/2)$. In the pure white Gaussian noise situation ($D = 256$), asymptotic decays in $K^{-1/2}$ (see text) have been superimposed as dotted lines.

WVS and that its equivalent time-frequency support (ellipse area) is controlled by the bandwidth-duration product $1/\alpha\beta$.

Using this model, we can get as a substitute to Figures II.2 and II.3 of the “box” model the following Figures V.1 and V.2, respectively.

2 Choice of Thresholds

In SubSection 2.1, both $RSD_{x,K}(t, f)$ and $RS_{x,K}(t, f)$ are proposed to be thresholded in the approach for the mask of differences and pre-denoising, re-

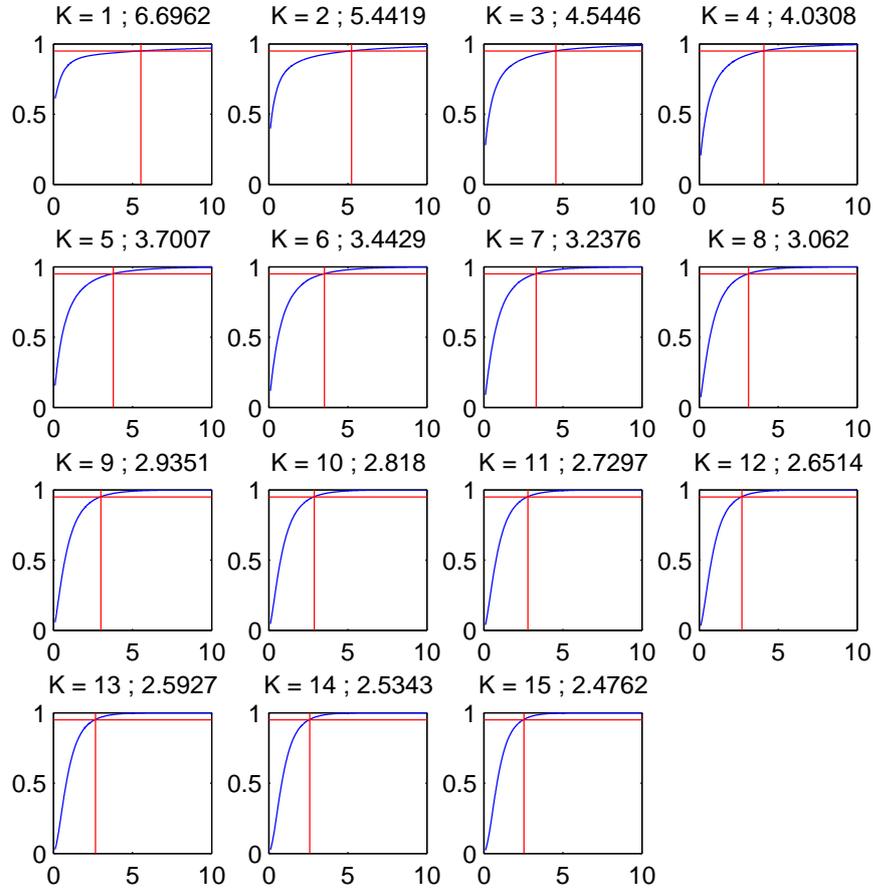


Fig. V.3: **Cumulative distribution function of $RS_{x,K}$.** Each diagram displays a cumulative distribution function of $RS_{x,K}$ based on a different number of tapers ($K = 1 : 15$). In the case of a “detection” probability $P_p = 0.95$ (red horizontal line), the corresponding threshold T_p^* (red vertical line) is given, with the corresponding value of T_p shown above each diagram.

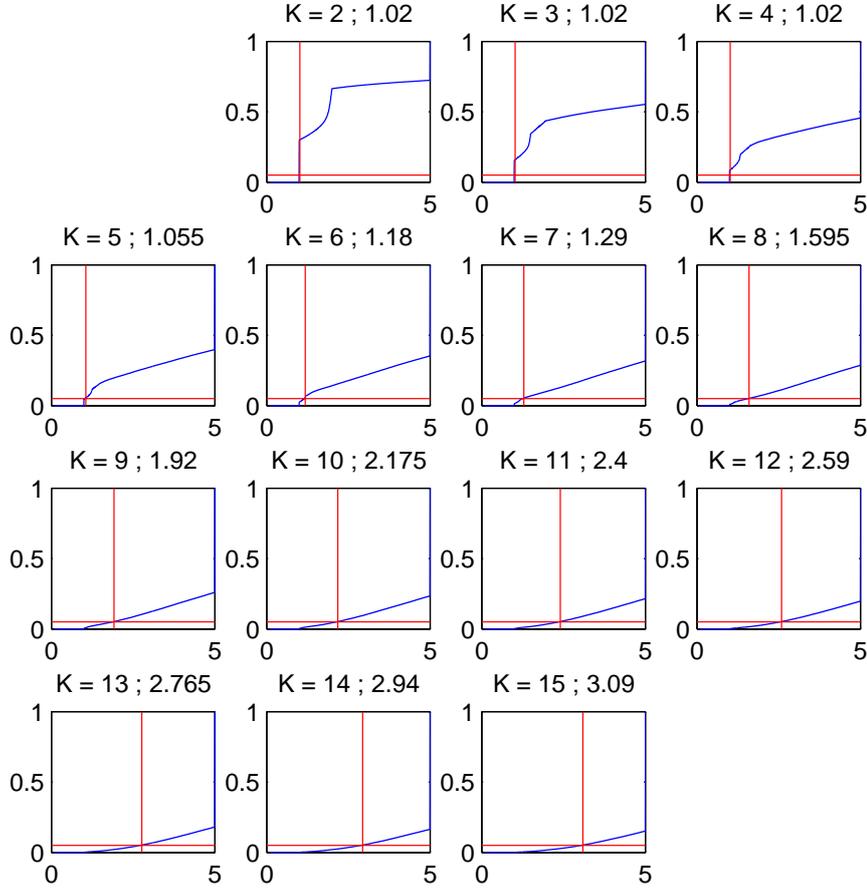


Fig. V.4: **Cumulative distribution function of $RSD_{x,K}$.** Each diagram displays a cumulative distribution function of $RSD_{x,K}$ based on a different number of tapers ($K = 2 : 15$). In the case of a “false alarm” probability $P_d = 0.05$ (red horizontal line), the corresponding threshold T_d (red vertical line) is given, with the exact value shown above each diagram.

spectively. The thresholds T_d and T_p^* work as follows:

$$M(t, f) = \begin{cases} 0, & |RSD_{x,K}(t, f) - 1| + 1 \geq T_d \\ 1, & |RSD_{x,K}(t, f) - 1| + 1 < T_d, \end{cases} \quad (\text{V.7})$$

$$RSM_{x,K}(t, f) = RS_{x,K}(t, f) \cdot M(t, f); \quad (\text{V.8})$$

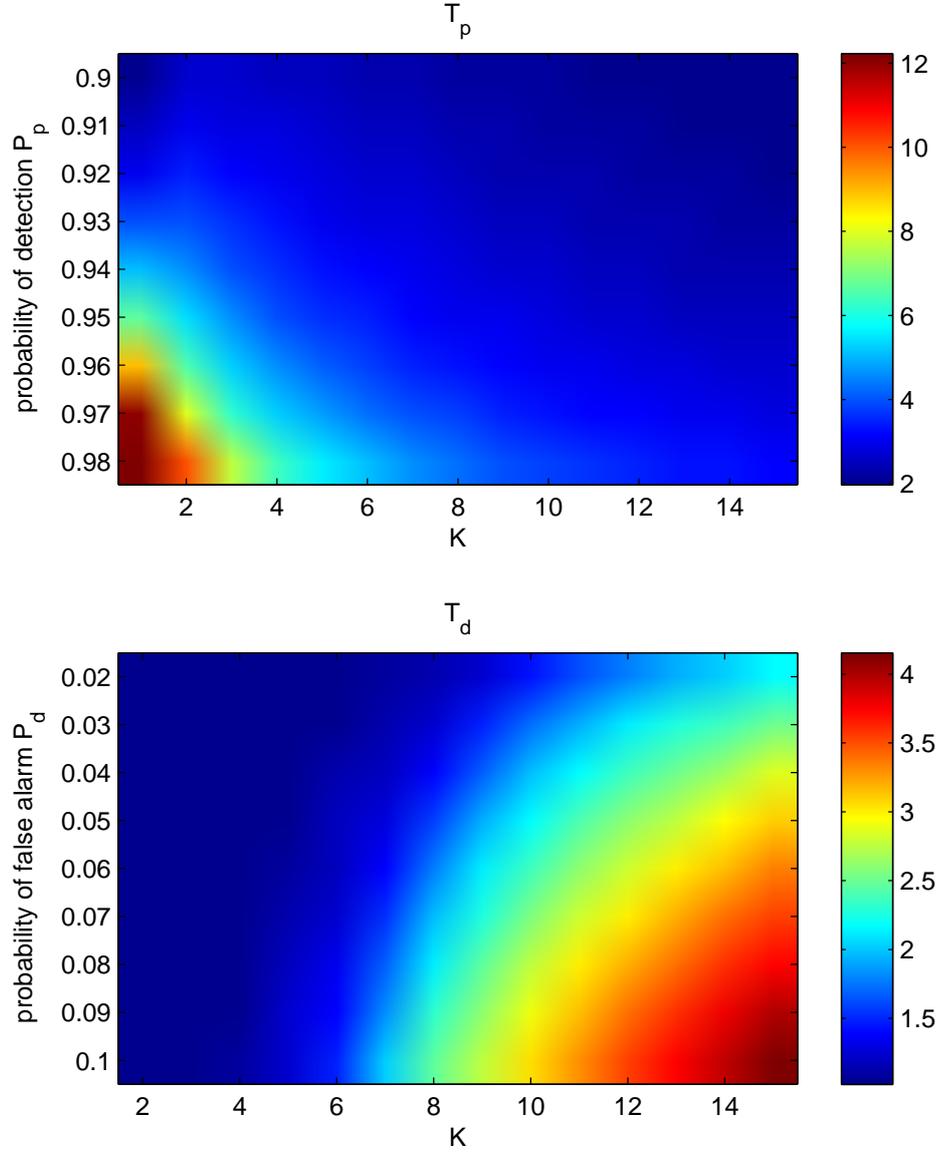


Fig. V.5: **Choosing thresholds T_p and T_d .** The figure displays respectively T_p (on the first row) and T_d (on the bottom row) as a function of “detection” probability P_p (resp. “false alarm” probability P_d) and of the number of tapers K .

$$RSP_{x,K}(t, f) = \begin{cases} RS_{x,K}(t, f), & RS_{x,K}(t, f) \geq T_p^*, \\ 0, & RS_{x,K}(t, f) < T_p^*. \end{cases} \quad (\text{V.9})$$

where $T_p^* := T_p \cdot RS_{x,K}$, $RS_{x,K}$ being the empirical mean of $RS_{x,K}(t, f)$ over the TF plane. We prefer to use in the programs the parameter T_p less fluctuated to different cases as it is independent to $RS_{x,K}$.

The pre-denoising threshold T_p —applied to $RS_{x,K}(t, f)$ in the form of T_p^* —was chosen according to a given (“detection”) probability P_p for rejecting the null hypothesis of white Gaussian noise. Its cumulative distribution function is illustrated in Fig. V.3 based on different numbers of tapers K . For the case of $P_p = 0.95$, the values of the corresponding threshold T_p are shown above each diagram.

Similarly, the masking threshold T_d —applied to $RSD_{x,K}(t, f)$ —was chosen so as to guarantee some prescribed (“false alarm”) probability P_d for the rectified quantity $|RSD_{x,K}(t, f) - 1| + 1$ in the white Gaussian noise case. Fig. V.4 gives its cumulative distribution function based on different numbers of tapers K . The values shown above each diagram is the T_d corresponds to the case of $P_d = 0.05$.

Finally, Fig. V.5 displays these two thresholds T_p (resp. T_d) as a function of different probability P_p (resp. P_d) and of the number of tapers K . It shows that in practice T_p (resp., T_d) turned out to essentially depend on $(1 - P_p)K$ (resp., $P_d K$), with typical values $T_p \approx 7$ times the empirical mean $RS_{x,K}$ (i.e. $T_p^* \approx 7RS_{x,K}$) for $P_p \approx 0.97$ and $T_d \approx 1.5$ for $P_d \approx 0.05$ when K ranges in between 4 and 8.

3 Variations of Test

Apart from (III.12), an alternative decision function of the measure of stationarity, based on the Gamma distribution of surrogate variances (cf. Chapter III: Section 3), would be a P-value of Θ_1 (III.11) in the (probability density) distribution of Θ_0 (III.10).

$$\begin{cases} P(\Theta_0(j) \geq \Theta_1) > \mu, (j = 1, \dots, J) & : \text{“nonstationarity”}; \\ P(\Theta_0(j) \geq \Theta_1) \leq \mu, & : \text{“stationarity”}. \end{cases} \quad (\text{V.10})$$

where P means P-value and μ is the threshold.

If *index of nonstationarity* is taken as

$$\text{INS}_p = -\log_{10}(1 - P), \quad (\text{V.11})$$

the corresponding threshold should be $\mu' = -\log_{10}(1 - \mu)$.

With the same distribution of Gamma law for Θ_0 , a similar example by P-value measure is given in Fig. V.6, which is also consistent with the physical interpretation according to which the observation can be considered as stationary at macroscale (top row), nonstationary at mesoscale (middle row) and stationary again at microscale (bottom row).

However, it should be pointed out that this measure of stationarity by making use of P-value works not as well as the approach proposed before in (III.12), because, according to an empirical study and as shown in Fig. V.7, the results

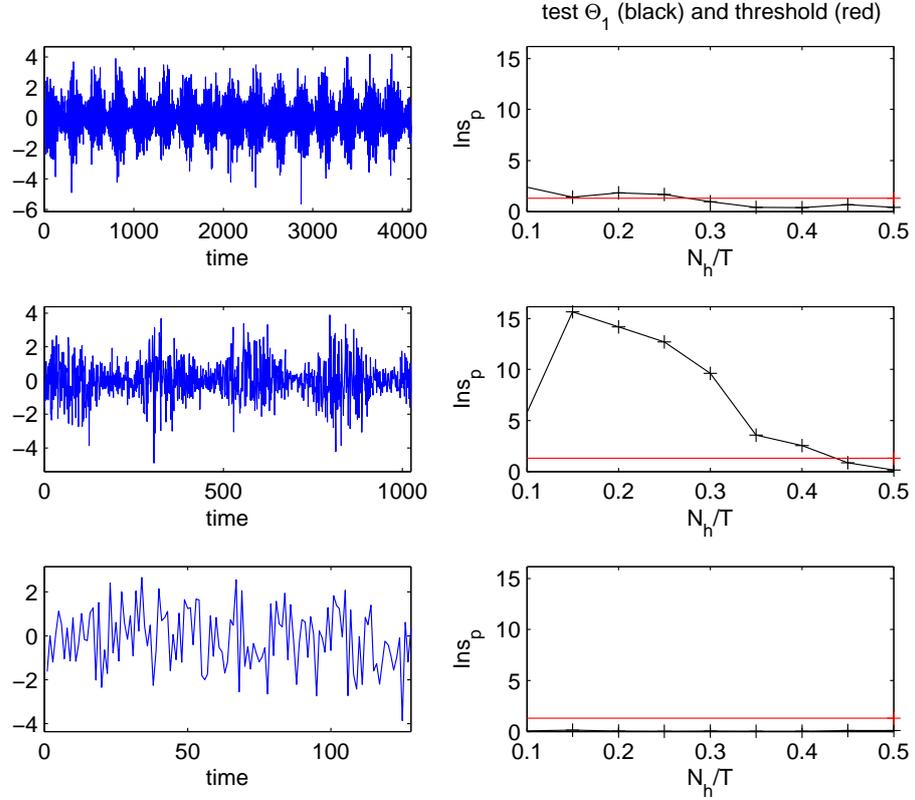


Fig. V.6: **P-value measure.** In the case of the same signal (III.4) observed over different time intervals (left column), the indice of nonstationarity INS_p (V.10) from P-value measure (right column, in black) are consistent with the physical interpretation according to which the observation can be considered as stationary at macroscale (top row), nonstationary at mesoscale (middle row) and stationary again at microscale (bottom row). The threshold (in red) of the stationarity test is calculated with a confidence level of 95% and represented in term of INS_p as μ' , with $J = 50$ surrogates.

from the P-value measure varies too sharply and the problem of finite numerical precision provokes a saturation of INS_p towards 16, leading to the difficulty of finding the typical scale of nonstationarity SNS_p .

On the other hand, the test Θ_1 (III.11) uses the distance between the signal divergences and their mean value in time, but some variations are possible:

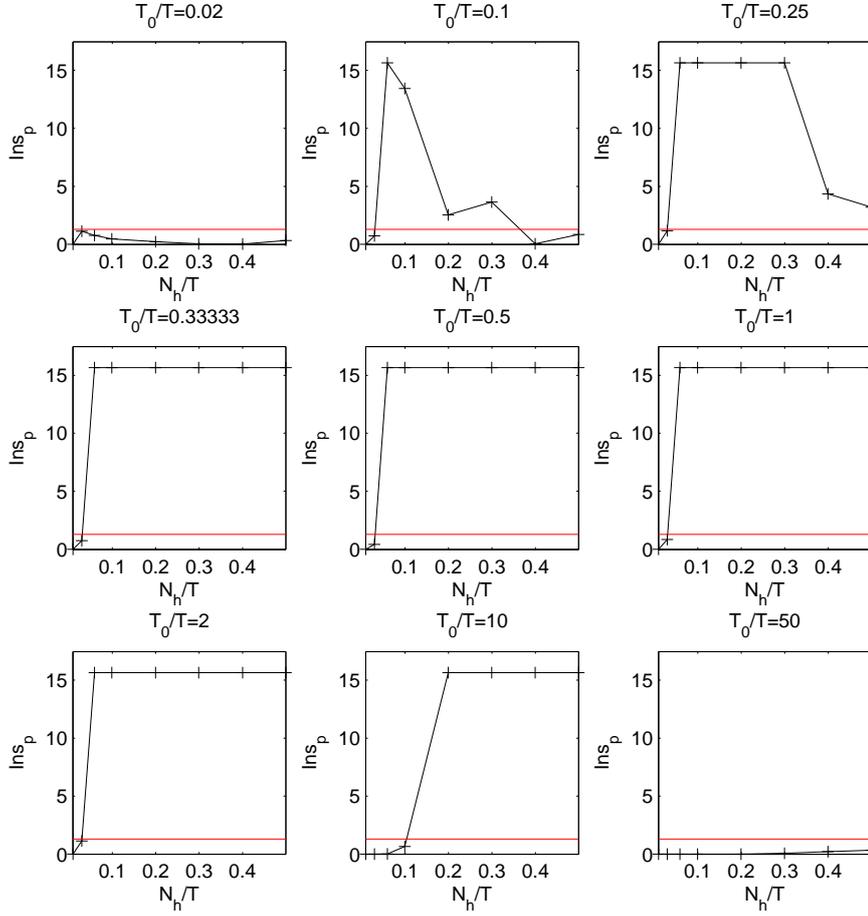


Fig. V.7: **Periodic correlation of SNS_p .** Each diagram displays the index of nonstationarity INS_p from P-value measure (in black) as a function of relative window length N_h/T (with 5 Hermite tapers, T being the observation length), in the same case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of relative period (T_0/T from 0.02 to 50), based on 2048 data points. The threshold (in red) is calculated as μ' with a confidence level of 95% and $J = 50$ surrogates.

Test 2. Use the distribution of the distances between the signal divergences and those of each surrogate data:

$$\left\{ \Theta_2(j) = L \left(c^{(x)}, c^{(s_j)} \right), j = 1, \dots, J \right\},$$

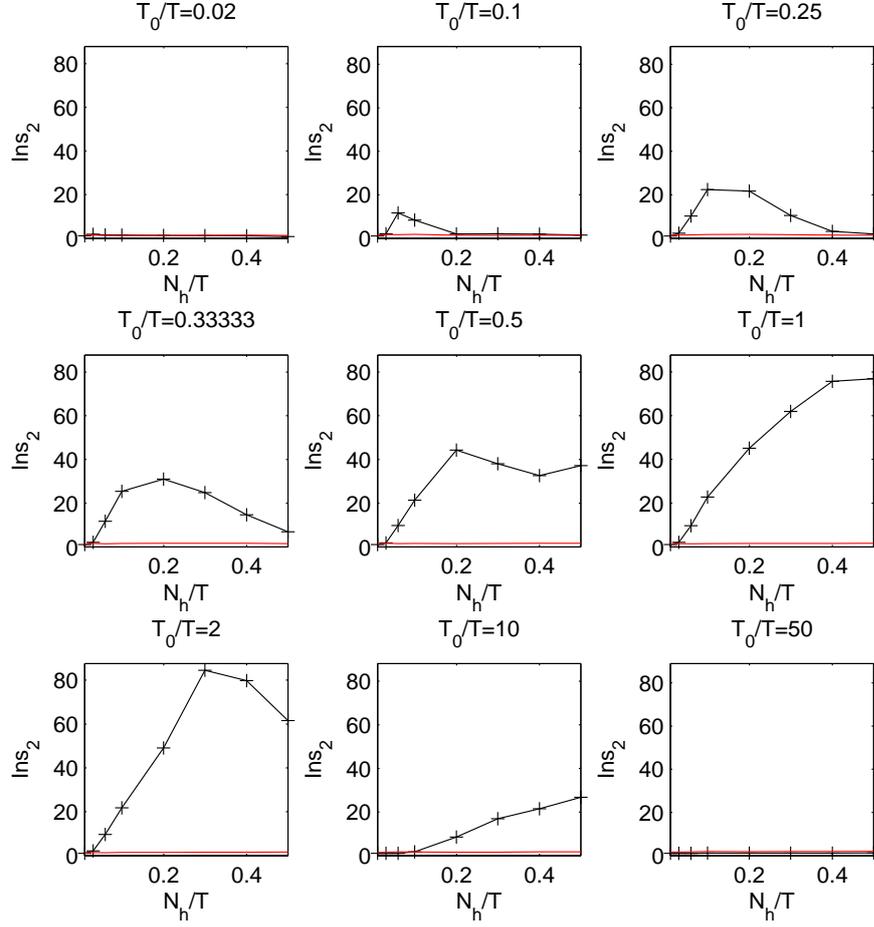


Fig. V.8: **Periodic correlation of SNS for Test Θ_2 .** Each diagram displays the index of nonstationarity INS_2 (V.15) by Test Θ_2 (V.12) (in black) as a function of relative window length N_h/T (with 5 tapers, T being the observation length), in the same case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of relative period (T_0/T from 0.02 to 50), based on 2048 data points. The threshold (in red) is calculated as γ' with a confidence level of 95% and $J = 50$ surrogates.

$$\Theta_2 = \frac{1}{J} \sum_{j=1}^J \Theta_2(j), j = 1, \dots, J. \quad (\text{V.12})$$

Test 3. Use the distribution of the distances between the signal divergences and

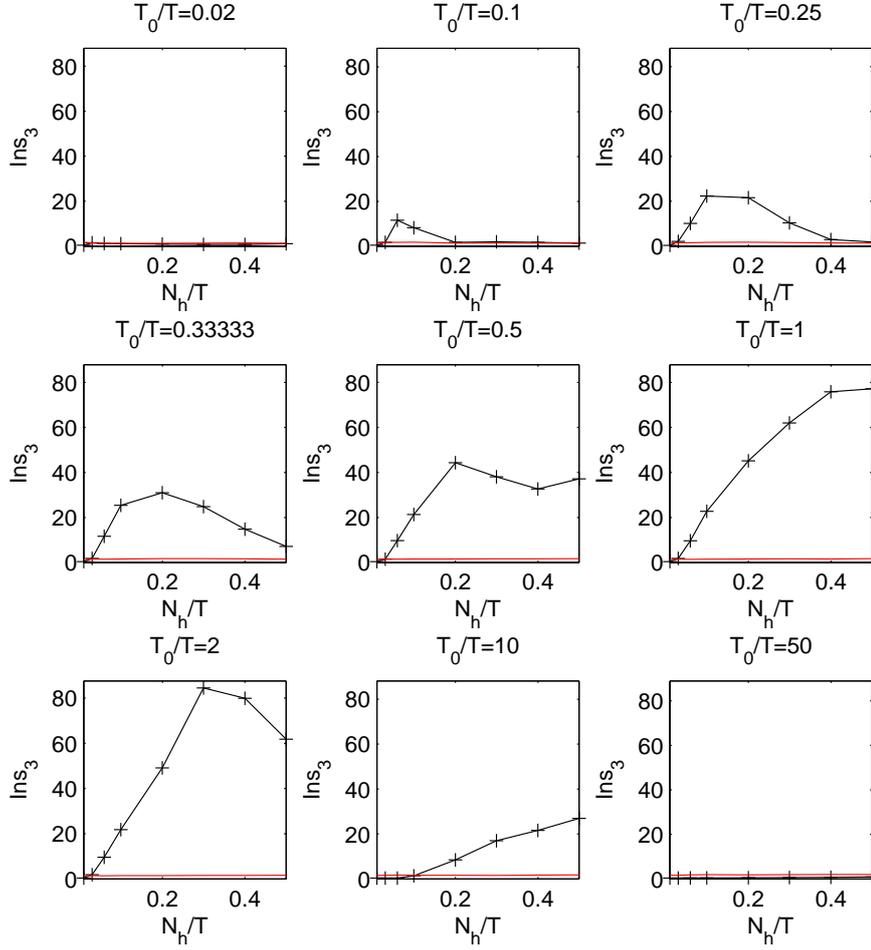


Fig. V.9: **Periodic correlation of SNS for Test Θ_3 .** Each diagram displays the index of nonstationarity INS_3 (V.15) by Test Θ_3 (V.13) (in black) as a function of relative window length N_h/T (with 5 tapers, T being the observation length), in the same case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of relative period (T_0/T from 0.02 to 50), based on 2048 data points. The threshold (in red) is calculated as γ' with a confidence level of 95% and $J = 50$ surrogates.

the mean value in time of those of each surrogate data:

$$\left\{ \Theta_3(j) = L \left(c^{(x)}, \langle c^{(s_j)} \rangle_{n=1, \dots, N} \right), j = 1, \dots, J \right\},$$

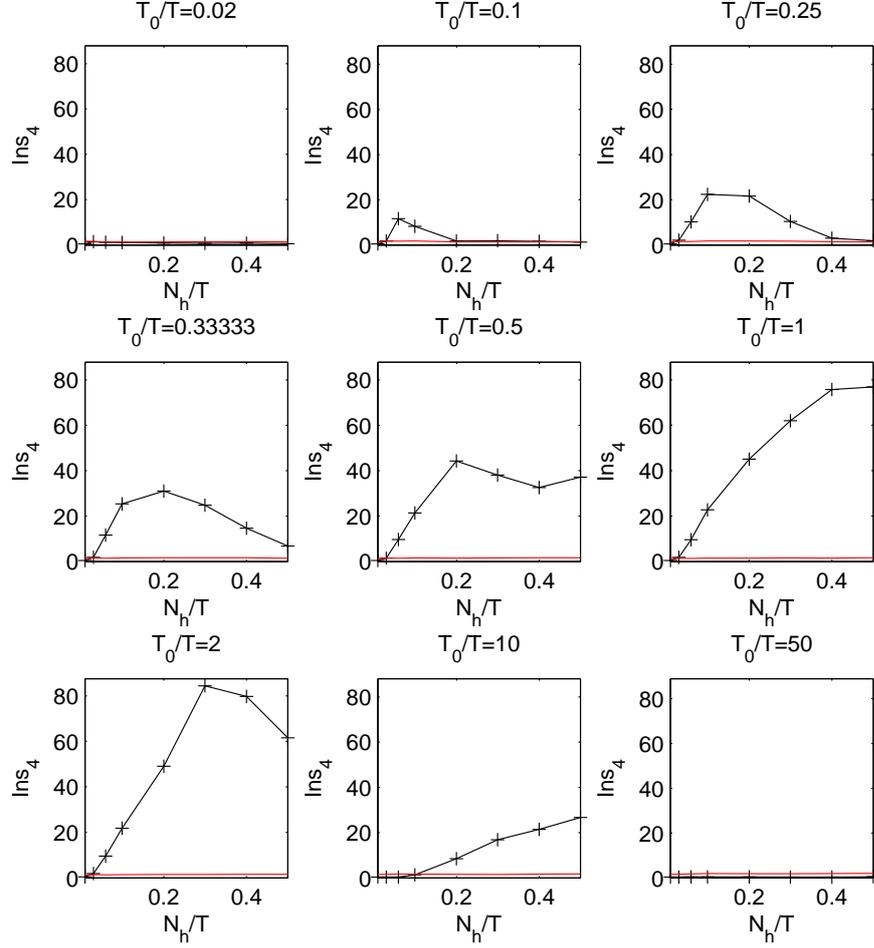


Fig. V.10: **Periodic correlation of SNS for Test Θ_4 .** Each diagram displays the index of nonstationarity INS_4 (V.15) by test Θ_4 (V.14) (in black) as a function of relative window length N_h/T (with 5 tapers, T being the observation length), in the same case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of relative period (T_0/T from 0.02 to 50), based on 2048 data points. The threshold (in red) is calculated as γ' with a confidence level of 95% and $J = 50$ surrogates.

$$\Theta_3 = \frac{1}{J} \sum_{j=1}^J \Theta_3(j), j = 1, \dots, J. \quad (\text{V.13})$$

Test 4. Use the distance between the signal divergences and the mean value over

all randomizations of those of each surrogate data:

$$\Theta_4 = L \left(c^{(x)}, \langle c^{(s_j)} \rangle_{j=1, \dots, J} \right). \quad (\text{V.14})$$

The index of nonstationarity can be generalized for all of the four Test by the definition:

$$\text{INS}_n := \sqrt{\frac{\Theta_n}{\frac{1}{J} \sum_{j=1}^J \Theta_0(j)}}. \quad (\text{V.15})$$

Similarly to test Θ_1 , results from each of the tests Θ_2 , Θ_3 and Θ_4 (V.12, V.13, and V.14), in the same case of FM signal (III.5) ($\alpha = 0.03$) with different modulation of period, also show respectively in Fig. V.8, Fig. V.9 and Fig. V.10: on the one side, stationarity for $T \ll T_0$ or $T \gg T_0$ and nonstationarity for $T \approx T_0$; on the other side, in nonstationarity context, the interesting correlation between the period of signal T_0 and typical scale of nonstationarity SNS.

Now that the results from the four tests are similar, we have the reason, based on economic considerations in the implementation of computers, to choose the simplest one test Θ_1 , which disposes only of the tested signal itself and doesn't request any information of surrogates.

Bibliography

- [1] <http://perso.ens-lyon.fr/patrick.flandrin/multifr.html>
- [2] <http://www-stat.stanford.edu/~donoho/Reports/1995/wavelab.pdf>
- [3] <http://www.eulersdisk.com>
- [4] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Proc.*, vol. SP-43(5), pp. 1068–1089, 1995.
- [5] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, and O. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. on Info. Th.*, vol. 47(4), pp. 1391–1409, 2001.
- [6] M. Bayram and R.G. Baraniuk, "Multiple window time-varying spectrum estimation," in *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerald et al., eds.), pp. 292–316, Cambridge Univ. Press, 2000.
- [7] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Proc.*, vol. 18, no. 4, pp. 349–369, 1989.
- [8] P. Borgnat, P.-O. Amblard, and P. Flandrin, "Stochastic invariances and Lamperti transformations for stochastic processes," *J. Phys. A: Math. Gen.*, vol. 38, no. 10, pp. 2081–2101, Feb. 2005.
- [9] B. Boashash, "Time-frequency signal analysis and processing: a comprehensive reference", *Elsevier*, (London), 2003.
- [10] L. Bildsten, "Viscous dissipation for Euler's disk," *Phys. Rev. E*, 66, 056309, 2002.
- [11] S. Barbarossa, "Analysis of multicomponent LFM signals by a combined Wigner-Hough transform," *IEEE Trans. Signal Processing*, vol. 43, no. 6, June 1995.
- [12] S. Barbarossa and O. Lemoine, "Analysis of nonlinear FM signals by pattern recognition of their time-frequency representation," *IEEE Signal Processing Letters*, vol. 3, no.4, April 1996.

-
- [13] L. Brillouin, “La science et la théorie de l’information,” Paris, Gauthiers-Villars, 1959.
- [14] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal,” Part I: Fundamentals, *Proc. IEEE*, 80, pp. 520-538, 1992.
- [15] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal,” Part II: Algorithms and applications, *Proc. IEEE*, 80, pp. 540-568, 1992.
- [16] J. Bass, “Les fonctions aléatoires et leur interprétation mécanique,” *Revue Scientifique*, 83ème Année, 3, 20, 1945.
- [17] E. Chassande-Mottin and P. Flandrin, “On the time-frequency detection of chirps,” *Appl. Comp. Harm. Anal.*, vol. 6, pp 252-281, 1999.
- [18] N. Cristianini and J. Shawe-Taylor, “An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods”, Cambridge Univ. Press, 2000.
- [19] F. Çakrak and P.J. Loughlin, “Multiple window time-varying spectral analysis,” *IEEE Trans. Signal Proc.*, vol. SP-49(2), pp. 448–453, 2001.
- [20] S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy, “SVM and Kernel Methods Matlab Toolbox”, Perception Systèmes et Information, INSA de Rouen, <http://asi.insa-rouen.fr/~arakotom/toolbox/index.html>, 2005.
- [21] H. Caps, S. Dorbolo, S. Ponte, H. Croisier, and N. Vandewalle, “Rolling and slipping motion of Euler’s disk,” *Phys. Rev. E*, 69, 056610, 2004.
- [22] N. Cristianini and J. Shawe-Taylor, “Support Vector Machines and other kernel-based learning methods,” Cambridge University Press, 2000.
- [23] L. Cohen, “Generalized phase-space distribution functions,” *J. Math. Phys.*, 7, pp. 781-786, 1966.
- [24] L. Cohen, “Time-frequency distributions – a review,” *Proc. IEEE*, 77, pp. 941-981, 1989.
- [25] L. Cohen, “Time-frequency analysis,” Englewood Cliffs: Prentice-Hall, 1995.
- [26] I. Daubechies, “Time-frequency localization operators: A geometric phase space approach,” *IEEE Trans. Info. Th.*, vol. 34(4), pp. 605-612, 1988.
- [27] M. Davy and S. Godsill, “Detection of abrupt signal changes using Support Vector Machines: An application to audio signal segmentation,” in *Proc. IEEE ICASSP-02*, Orlando (FL), 2002.

- [28] G. van der Engh, P. Neslon, and J. Roach, “Numismatic Gyration”, *Nature*, vol. 408, pp. 540, 2000.
- [29] P. Flandrin, “Maximum signal energy concentration in a time-frequency domain,” in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc. — ICASSP-88*, (New York, NY), pp. 2176–2179, 1988.
- [30] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, San Diego: Academic Press, 1999.
- [31] P. Flandrin, F. Auger, and E. Chassande-Mottin, “Time-Frequency Reassignment — From Principles to Algorithms,” in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, ed.), Chap. 5, pp. 179-203, CRC Press, 2003.
- [32] P. Flandrin, P. Borgnat, and P.-O. Amblard, “From stationarity to self-similarity, and back : Variations on the Lamperti transformation,” in *Processes with Long-Range Correlations: Theory and Applications*, G. Raganjara and M. Ding, Eds. June 2003, vol. 621 of *Lectures Notes in Physics*, pp. 88–117, Springer-Verlag.
- [33] G. Frazer and B. Boashash, “Multiple window spectrogram and time-frequency distributions,” in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc. — ICASSP’94*, (Adelaide, AU), vol. IV, pp. 293–296, 1994.
- [34] A.B. Gardner, A.M. Krieger, G. Vachtsevanos, and B. Litt, “One-class novelty detection for seizure analysis from intracranial eeg,” *Journal of Machine Learning Research*, vol. 7, pp. 1025–1044, 2006.
- [35] D.R. Griffin, “Listening in the Dark,” Yale Univ. Press, New Haven, 1958.
- [36] D.R. Griffin, F.A. Webster, and C.R. Michael, “The echolocation of flying insects by bats,” *Animal Behaviour*, 8, pp. 141-154, 1960.
- [37] Y. Gasteuil, Thèse en cours au laboratoire de physique de l’École Normale Supérieure de Lyon (director: J.-F. Pinton).
- [38] D. Gabor, “Theory of communication,” *J. IEE*, 93, pp. 429-457, 1946.
- [39] P. Honeine, C. Richard, and P. Flandrin, “Time-frequency learning machines,” *IEEE Trans. on Signal Proc.*, vol. 55, no. 7 (Part 2), pp. 3930–3936, 2007.
- [40] C.J. Keylock, “Constrained surrogate time series with preservation of the mean and variance structure,” *Phys. Rev. E*, vol. 73, pp. 030767.1–030767.4, 2006.
- [41] P. Kessler and O.M. O’Reilly, “The Ringing of Euler’s Disk,” *Regular Chaotic Dyn.*, 7(1), pp. 49-60, 2002.

-
- [42] K. Kodera, C. de Villedary, and R. Gendrin, "A new method for the numerical analysis of non-stationary signals," *Phys. Earth Planet. Inter.*, vol. 12, pp. 142–150, 1976.
- [43] K. Kodera, R. Gendrin, and C. de Villedary, "Analysis of time-varying signals with small BT values," *IEEE Trans. Acoust., Speech and Signal Proc.*, 26, pp. 64-76, 1978.
- [44] H. Laurent and C. Doncarli, "Stationarity index for abrupt changes detection in the time-frequency plane," *IEEE Signal Proc. Lett.*, vol. 5, no. 2, pp. 43–45, 1998.
- [45] P.J. Loughlin, J. Pitton, and B. Hannaford "Approximating time-frequency density functions via optimal combinations of spectrograms," *IEEE Signal Proc. Lett.*, vol. 1(12), pp. 199–202, 1994.
- [46] H.J. Landau and H.O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-II," *Bell Syst. Tech. J.*, 40, pp. 65-84, 1961.
- [47] S. Mallat, G. Papanicolaou, and Z. Zhang, "Adaptive covariance estimation of locally stationary processes," *Ann. of Stat.*, vol. 24, no. 1, pp. 1–47, 1998.
- [48] W. Martin, "Time-frequency analysis of random signals," in *IEEE Int. Conf. on Acoust., Speech and Signal Proc - ICASSP-82*, Paris, pp. 1325-1328, 1982.
- [49] W. Martin, "Measuring the degree of non-stationarity by using the Wigner-Ville spectrum," in *Proc. IEEE ICASSP-84*, San Diego (CA), pp. 41B.3.1–41B.3.4, 1984.
- [50] W. Martin and P. Flandrin, "Detection of changes of signal structure by using the Wigner-Ville spectrum," *Signal Proc.*, vol. 8, pp. 215–233, 1985.
- [51] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech and Signal Proc.*, 33, pp. 1461-1470, 1985.
- [52] H.K. Moffatt, *Nature* (London), vol. 404, pp. 833, 2000.
- [53] H. Maitre, "Un panorama de la transformation de Hough," *Traitement du Signal*, vol .2, no. 4, pp. 306-317, 1985.
- [54] B. Mandelbrot, "Les Objets Fractals — Forme, Hasard et Dimension," Flammarion, 1975.
- [55] B. Mandelbrot, "Fractals: form, chance, and dimension," San Francisco : W.H. Freeman, 1977.
- [56] T. Mitchell, "Machine Learning," Mc Graw-Hill, New York, 1997.

- [57] D.B. Percival and A.T. Walden, *Spectral Analysis for the Physical Sciences*, Cambridge Univ. Press, 1993.
- [58] L. Pimonow, “Vibrations en régime transitoire,” Paris: Dumod, 1962.
- [59] R.A. Silverman, “Locally stationary random processes,” *IRE Trans. on Info. Theory*, vol. 3, pp. 182–187, 1957.
- [60] T. Schreiber and A. Schmitz, “Improved surrogate data for nonlinearity tests,” *Phys. Rev. Lett.*, vol. 77, no. 4, pp. 635–638, 1996.
- [61] E. Serpedin, F. Panduru, I. Sari, and G.B. Giannakis, “Bibliography on cyclostationarity,” *Signal Proc.*, vol. 85, no. 12, pp. 2233–2303, 2005.
- [62] B. Schölkopf and A.J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press, 2001.
- [63] B. Schölkopf, J.C. Platt, J.S. Shawe-Taylor, A.J. Smola, and R.C. Williamson, “Estimating the support of a high-dimensional distribution,” *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [64] B. Schölkopf and A.J. Smola, “Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond,” MIT Press, Cambridge, MA, 2002.
- [65] A. Stanislavsky and K. Weron, “Nonlinear oscillations in the rolling motion of Euler’s disk,” *Physica D: Nonlinear Phenomena*, vol. 156, Issue 3-4, pp. 247–259, 08/2001.
- [66] W.L. Shew, Y. Gasteuil, M. Gibert, and J.-F. Pinton, “An instrumented tracer for Lagrangian measurements in Rayleigh-Benard convection,” *Rev. Sci. Instr.* 78, 065105, 2007.
- [67] D. Slepian and H.O. Pollak, “Prolate spheroidal wave functions Fourier analysis and uncertainty-I” *Bell Syst. Tech. J.*, 40, pp. 43–64, 1961.
- [68] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and J.D. Farmer, “Testing for nonlinearity in time series: The method of surrogate data,” *Physica D*, vol. 58, no. 1, pp. 77–94, 1992.
- [69] D.J. Thomson, “Spectrum estimation and harmonic analysis,” in *Proc. IEEE*, vol. 70, pp. 1055–1096, 1982.
- [70] D.J. Thomson, “Jackknifing multiple-window spectra,” in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc. — ICASSP-94*, (Adelaide, AU), vol. 6, pp. VI.73–VI.76, 1994.

- [71] D.J. Thomson, "Multitaper analysis of nonstationary and nonlinear time series data," in *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerald et al., eds.), pp. 317–394, Cambridge Univ. Press, 2000.
- [72] D.M.J. Tax and R.P.W. Duin, "Support vector data description," *Machine Learning*, vol. 54, pp. 45–66, 2004.
- [73] V. Vapnik, "The Nature of Statistical Learning Theory," Springer, 1995.
- [74] V. Vapnik, "Statistical Learning Theory," Wiley-Interscience, New York, 1998.
- [75] J.A. Ville, "Théorie et applications de la notion de signal analytique," *Câles et Transm.*, 2A, pp.61-74, 1948.
- [76] E.P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, 40, pp. 749-759, 1932.
- [77] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment," in *Proc. 14th European Signal Proc. Conf. — EUSIPCO-06*, (Florence, I), 2006.
- [78] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement," *IEEE Trans. on Signal Proc.*, vol. 55, no. 6 (Part 2), pp. 2851–2860, 2007.
- [79] J. Xiao, P. Borgnat, and P. Flandrin, "Testing stationarity with time-frequency surrogates," in *Proc. EUSIPCO-07*, Poznan (PL), 2007.
- [80] J. Xiao, P. Borgnat, and P. Flandrin, "Sur un Test Temps-Fréquence de Stationnarité," in *Proc. GRETSI-07*, Troyes, 2007.
- [81] J. Xiao, P. Borgnat, and P. Flandrin, "Sur un Test Temps-Fréquence de Stationnarité," in *Proc. GRETSI-07*, sessions du colloque GretsI 2007, Traitement du Signal, submitted 2008.
- [82] J. Xiao, P. Borgnat, P. Flandrin, and C. Richard, "Testing stationarity with surrogates – A one-class SVM approach," in *Proc. IEEE Stat. Sig. Proc. Workshop SSP-07*, Madison (WI), 2007.

Acknowledgment

Before the end of my thesis, I would like to express my gratitude to all those people who have helped me during all these years of studies.

First of all, thanks to the cooperation program between East China Normal University (ECNU) and Écoles Normales Supérieures Group (ENS-Ulm, ENS-Lyon, ENS-Cachan, and ENS-LSH), I have got the chance to take the French language course for two years and followed some academic courses given by French professors during my master studies, and after that to go to France and begin my brand-new PhD studies at École Normale Supérieure de Lyon (ENS-Lyon).

The deepest gratitude (only if I could find out some other word more than 'gratitude') belongs no doubt to three people: my French supervisor Patrick FLANDRIN, without whom I would even not have entered in this fantastic research area, and without whom you would not have had the chance to read this Thesis; my Chinese supervisor Zhengqi ZHENG, for all of his help and support during my last year of studies in Shanghai; my vice French supervisor Pierre BORGNAT, who never hesitates to lend me a hand and take me out of puzzles. I am deeply touched by their patience in guidance and enthusiasm in research, and much impressed by their intelligence and personality.

I wish to thank Prof. Christian JUTTEN from Université Joseph Fourier who kindly accepted to be the president of my defense, as well as Prof. Lotfi Senhadji from Université de Rennes 1 who took time off to take part in the committee member and act as spokesman.

Sincere gratitude is due to the "Laboratoire de Physique, UMR CNRS-ENSL" and to each member of the group 'Signaux, Systèmes et Physique' (SISYPHE): the director of our group Patrice ABRY, Professor Stéphane ROUX, and my dear colleagues Herwig WENDT, Gabriel RILLING and Nicolas MALLICK. The gratitude to them is necessary, not only because that I have been benefited a lot from the discussion with them, but also because that the group is a big family to me in France, thanks to which I never felt lonely in Lyon.

I appreciate those who are busy with the Franco-Chinese cooperation program for the support at the back from both ECNU and ENS-Lyon: Mrs. Yunhua QIAN, from the Office of International Affaires of ECNU; Mr. Jean-Louis DUCLOS, Mr. Jan MATAS and Ms. Nadine DONGUY, from the Service des Relations Internationales de l'ENS-Lyon; Mr. Wenhui YOU and Mr. Haisheng LI, from the Graduate Institute of ECNU. I should also thank ENS-Lyon for the scholarship of the first year of my studies and to EGIDE from the French government who supported me financially during the second year in France.

Special gratitude to friendly professors Thierry DAUXOIS and Françoise ARGOUL, who also care about my studies and life in France, especially to Thierry DAUXOIS, who introduced me to my excellent supervisor Patrick FLANDRIN and to the nice group SISYPHE, or else the story would have been different.

I am also grateful to many of my friends Yanyi JIANG, Chunyan XU, E WU, Sanjun ZHANG and my roommate Basia PIASECKA for their friendship or help

during my studies in France.

If it is possible, I would also like to thank my French language teachers: Grégoire, Elsa CHALAUX, Hui'er CHEN, who introduced the beautiful language and the culture of the country to me. Actually, I was surprised to meet teacher CHEN at her first class at ECNU, as she was exactly the same French teacher for me ten years before at Shanghai International Children's Cultural and Arts Festival.

Last but not least, I am deeply indebted to my family: my dear parents, who supply the strongest support and encouragement from Shanghai, China; and my husband Yong YANG who is doing his PdD studies at University of Florida, USA. It is their love, in spite of the geographical distance, that smoothes the way of the accomplishment of the thesis. I wish them healthy and happy every day and my husband could successfully accomplish his thesis in one year.

My life of the two years' studies in France is such a colorful experience that I will never forget. Best wishes to all those who have been with me during that special period.