



---

---

Roma Tre University

Department of Mathematics and Physics  
Doctoral School in Mathematics and Physics XXXII Cycle

Ph.D. Thesis in Physics

---

**Adaptive and Fractal Data Analysis of  
Seismic Monitoring at the Virgo  
Interferometer**

---

Author: Alessandro Longo

Tutor: Prof. Wolfango Plastino  
Coordinator: Prof. Giuseppe Degrassi

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
1.1	Detecting Gravitational waves . . . . .	10
1.2	Virgo Interferometer: Principles of Operation . . . . .	11
1.3	Sensitivity Curve and Fundamental Noises . . . . .	13
1.3.1	Amplitude Noise in the Laser Output Power . . . . .	14
1.3.2	Mechanical Thermal Noise . . . . .	14
1.3.3	Radiation Pressure Noise from the laser Light . . . . .	15
1.4	Seismic Noise . . . . .	15
1.5	The Superattenuator Design and Working Principle . . . . .	16
1.5.1	The Inverted Pendulum . . . . .	17
1.5.2	The Seismic Filters . . . . .	17
1.5.3	The Last Stage . . . . .	18
1.6	Newtonian Noise . . . . .	19
1.6.1	Mitigation of Newtonian Noise . . . . .	20
1.7	Advanced Virgo Interferometer: A second Generation Detector . . . . .	21
<b>2</b>	<b>Methodology: Adaptive and Fractal Algorithms</b>	<b>23</b>
2.1	Outline of adopted Methodology . . . . .	23
2.2	Hilbert Spectral Analysis: The instantaneous frequency, amplitude and bandwidth . . . . .	25
2.2.1	Interpretation of instantaneous frequency and bandwidth . . . . .	27
2.3	Empirical Mode Decomposition . . . . .	27
2.4	Hilbert-Huang Transform: a time-frequency-energy representation of nonlinear nonstationary data . . . . .	29
2.4.1	Marginal Hilbert Spectra, Instantaneous Energy and Degree of Stationarity . . . . .	30
2.5	Mode mixing and noise assisted methods . . . . .	32
2.6	Statistical significance test of IMFs . . . . .	32
2.7	Time varying filter EMD (tvf-EMD) . . . . .	34
2.7.1	B-splines approximation of a signal . . . . .	34
2.7.2	Estimation of local frequency cutoff . . . . .	35
2.7.3	The intermittency problem: frequency realignment to mitigate mode mixing . . . . .	38
2.7.4	Sifting of tvf-EMD: Stopping criterion based on instantaneous bandwidth and local narrow-band signal extraction . . . . .	39
2.8	Fractal Algorithms: Detrended Fluctuation Analysis and Local Hurst Exponent . . . . .	41

2.9	Tool for Adaptive time series analysis based on tvf-EMD algorithm	42
<b>3</b>	<b>Results</b>	<b>45</b>
3.1	Testing: NR Waveform in Purple Noise . . . . .	47
3.2	Adaptive Denoising of Acoustic Noise Injections . . . . .	49
3.2.1	Dataset . . . . .	50
3.2.2	Results of Adaptive Denoising . . . . .	50
3.2.3	Local Hurst Exponent of Seismometer Data . . . . .	53
3.2.4	Completeness of tvf-EMD in the Frequency Domain . . . . .	53
3.2.5	Hilbert Huang Transform and Denoising Threshold Parameter	57
3.3	Local Hurst Exponent for Seismometer Array Monitoring WEB . .	59
3.3.1	Array of seismometers for Newtonian Noise Characterisation	59
3.3.2	Local Hurst Exponent of Seismometer Array . . . . .	60
3.4	Yearly Modulation of surface ${}^7\text{Be}$ activity concentration . . . . .	63
3.4.1	Dataset . . . . .	63
3.4.2	Obtained Results: Trend, Yearly cycles and Outliers . . . . .	63
<b>4</b>	<b>Discussion</b>	<b>69</b>
	<b>Appendices</b>	<b>72</b>
<b>A</b>	<b>Estimate of spectral indexes <math>\beta &gt; 2</math> with the Burg Maximum Entropy Method</b>	<b>73</b>
<b>B</b>	<b>Fractional Gaussian noise (fGn) and fractional Brownian motion (fBm)</b>	<b>77</b>
<b>C</b>	<b>Hurst Exponent of fGn and fBm noises: EMD based estimation</b>	<b>79</b>
<b>D</b>	<b>Bedrosian Product Theorem</b>	<b>82</b>
<b>E</b>	<b>B-splines</b>	<b>83</b>
<b>F</b>	<b>Scattered Light Noise Hunting at Virgo with Empirical Mode Decomposition</b>	<b>85</b>
<b>G</b>	<b><math>1/f^{2.5}</math> Noise Hunting with Multifractal Detrended Fluctuation Analysis: Characterisation of Local Hurst Exponent of C11 data</b>	<b>91</b>
	<b>Bibliography</b>	<b>93</b>

# List of acronyms

Advanced Virgo (AdV)  
Amplitude and frequency modulations (AM-FM)  
Complete Ensemble EMD with adaptive noise (CEEMDAN)  
Comprehensive Nuclear-Test-Ban Treaty Organisation (CTBTO)  
Crosscorrelation (xcorr)  
Detrended Fluctuation Analysis (DFA)  
Differential Arm Motion (DARM)  
Empirical Mode Decomposition (EMD)  
End Mirror (EM)  
Ensemble Empirical Mode Decomposition (EEMD)  
Fractional Gaussian noise (fGn)  
Fractional Brownian motion (fBm)  
Full Width at Half Maximum (FWHM)  
General Relativity (GR)  
Gravitational waves (GW)  
High Purity Germanium Detector (HPGe)  
Hilbert Huang Spectral Analysis (HHSA)  
Hilbert Huang Transform (HHT)  
Hilbert Spectral Analysis (HSA)  
Interferometer (ifo)  
Input Mode Cleaner (IMC)  
Input Mirror (IM)  
Instantaneous amplitude (IA)  
Instantaneous bandwidth (IB)

Instantaneous frequency (IF)  
International Monitoring System (IMS)  
Intertropical Convergence Zone (ITCZ)  
Intrinsic Mode Functions (IMFs)  
Mean squared error (MSE)  
Mean absolute error (MAE)  
Multistep Time Series Analysis (MTsA)  
North End Building (NEB)  
Newtonian noise (NN)  
Numerical Relativity (NR)  
Peak signal to noise ratio (PSNR)  
Power Recycling cavity length (PRCL)  
Power Recycling Mirror (PRM)  
Signal Recycling (SR)  
Signal Recycling Mirror (SRM)  
Signal to noise ratio (SNR)  
Standard Temperature and Pressure (STP)  
Superattenuator (SA)  
Time varying filter Empirical Mode Decomposition (tvf-EMD)  
Weighted Average Instantaneous Frequency (WAIF)  
West End Building (WEB)

# List of Figures

Figure 1.1: GW strain

Figure 1.2: Optical layout of Michelson interferometer

Figure 1.3: Sensitivity curve and noises

Figure 1.4: Schematic of Virgo Superattenuator

Figure 1.5: Seismic filters

Figure 1.6: Newtonian noise budget

Figure 1.7: Sensitivity of Advanced Virgo

Figure 1.8: Optical configuration of Advanced Virgo

Figure 2.1: EMD algorithm procedure

Figure 2.2: Mode mixing

Figure 2.3: Spread lines for significance against white noise

Figure 2.4: Lowpass behaviour of B-splines

Figure 2.5: Frequency realignment step of tvf-EMD

Figure 2.6: Detrended Fluctuation Analysis

Figure 2.7: Flowchart of adopted methodology

Figure 3.1: Algorithm testing on NR waveform in purple noise

Figure 3.2: Hilbert transform of NR waveform in purple noise

Figure 3.3: Recordings from fourth acoustic noise injection

Figure 3.4: First acoustic noise injection

Figure 3.5: Second acoustic noise injection

Figure 3.6: Third acoustic noise injection

Figure 3.7: Fourth acoustic noise injection

Figure 3.8: Local Hurst exponent of seismometer data

Figure 3.9: FFT of persistent IMFs

Figure 3.10: FFT of antipersistent IMFs

Figure 3.11: Hilbert Huang transform of triaxial seismometer data

Figure 3.12: Coordinates of 38 seismometer array

Figure 3.13: Local Hurst exponent of two sensors of the array

Figure 3.14: Average values of Hurst exponent for the seismometer array

Figure 3.15: Trends of  ${}^7\text{Be}$  time series

Figure 3.16: Annual oscillation of  ${}^7\text{Be}$  time series

Figure 3.17: Occurrence of outliers in  ${}^7\text{Be}$  residual time series

Figure 3.18 Cross-correlations between  ${}^7\text{Be}$  and temperature annual cycle

Figure A.1 Flowchart of MTsA algorithm

Figure A.2 Comparison of spectral index estimation

Figure A.3 Noises of different spectral indexes

Figure B.1 Fractional gaussian noise and fractional Brownian motion dualism

Figure C.1 Fractional noises with different Hurst exponent

Figure C.2 Average PSD of fGn time series and comparison with IMFs spectra

Figure E.1 B-splines of degrees 0 to 3

Figure F.1 Fourier spectra of DARM with and without scattered light

Figure F.2 Hardware injection and spectrogram of PRCL

Figure F.3 Results of scattered light noise hunting with EMD

Figure F.4 Hardware injections

Figure F.5 Results of scattered light noise hunting with EMD

Figure G.1  $1/f^{2.5}$  noise hunt with fractal algorithms

## Outline

Detector characterisation plays a crucial role in the search for Gravitational waves (GW) with a ground based interferometer (ifo) such as Virgo [1]. Both environmental and anthropogenic noise sources can affect the sensitivity in the detection frequency band, a relevant example being Newtonian noise (NN) [2][3]. For this reason, sensors are deployed both indoor and outdoor at the detector site, in order to monitor stable conditions of the detector over time. Data are hence acquired in the form of time series and various algorithms are employed for detector characterisation purposes a useful and widely used tool to investigate transient noise being Omicron [4].

In the framework of this Thesis, an algorithm for time series analysis have been developed and tested on data from different physical systems. It is based on a combination of both fractal [5] and adaptive techniques [6]. The algorithm developed for this Thesis has been applied to characterise data of

- Seismometer array recording at the Virgo West End Building (WEB) for Newtonian Noise (NN) characterisation purposes. In this case is found that fractal analysis allows to characterise and discriminate among sensors placed in different parts of the room.
- Seismometer recording during four different acoustic noise injections performed at the North End Building (NEB) of the Virgo ifo for detector characterisation purposes. In this case, is found that the adopted approach allows to separate the waveform due to the acoustic noise injections from the underlying nonlinear non stationary seismic noise.
- Activity concentration of  ${}^7\text{Be}$ , sampled at ground level by the International Monitoring System (IMS) of the Comprehensive Nuclear-Test-Ban Treaty Organisation (CTBTO). In this case, the annual oscillatory modes are extracted and characterised. It is found that the time of occurrence of the yearly maxima of  ${}^7\text{Be}$  is shifted in latitude, possibly due to the seasonal shift of the Hadley cell and of the Intertropical Convergence Zone (ITCZ), as also reported in [7].

The remaining of this Thesis is organised as follows.

- In Chapter1 the main sources of noise affecting the sensitivity curve of Virgo are briefly summarised and a description of NN and seismic noise is given.
- In chapter 2, the adopted methodology for time series analysis is widely described. In particular, the Empirical Mode Decomposition (EMD) and its recently introduced time varying version (tvf-EMD), are described in 2.3 and 2.7, respectively. The adopted methodology is described in 2.9.
- Chapter 3 presents the results obtained applying such empirical algorithms to seismometer and  ${}^7\text{Be}$  data
- Conclusions and discussion of possible future developments of the present research work is carried out in Chapter 4.

Furthermore in appendix are reported results from characterisation of noise properties, such as spectral index and persistency and results from noise hunting performed using data of the Virgo ifo.

- In A are reported estimates of spectral indexes  $\beta > 2$  using the Burg Maximum Entropy Method.
- In B Fractional Gaussian noise (fGn) and fractional Brownian motion (fBm) are described.
- In C the Hurst Exponent of fGn and fBm noises are estimated using EMD.
- In D the Bedrosian Product Theorem is described.
- In E B splines are introduced and some of their properties described
- In F results from scattered light noise hunting at Virgo with EMD are presented.
- Finally, in G some results from the  $1/f^{2.5}$  noise hunt using fractal algorithms are presented. The analysis was carried out using downsampled data from the commissioning run C11 of the Virgo ifo.

# Chapter 1

## Introduction

In this Chapter, a brief overview of how GW are generated and which sources are accessible to ground based detectors is given in section 1.1 based on [8] while operational principles of the Virgo Interferometer to detect them are described in section 1.2, based on [9, 10]. In 1.3 the fundamental noises limiting detector sensitivity are briefly summarised based on [1]. Seismic and Newtonian Noise are described in more detail in 1.4 and 1.6 based on [3] and references therein.

### 1.1 Detecting Gravitational waves

GW are generated each time a massive object is accelerated and are predicted by Einstein General Relativity (GR) theory [11]. GR geometrical interpretation states that gravitation is the result of the curvature of space-time caused by mass or energy density. GW are then deformations of space-time and are predicted to travel at the speed of light. Considering small deviations of the metric tensor  $g_{\mu\nu}$  from Minkowski metric  $\eta_{\mu\nu}$  Einstein's equations can be linearised giving a wave equation for the quantity

$$h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu} \quad (1.1)$$

with propagation speed equal to  $c$ . GW detection involves measuring time variations of  $h$  i.e. variations in the geometry of space-time locally. The order of magnitude of the change of distance  $\delta L$  between two freely-falling test masses located  $L$  meters apart caused by  $h_{\mu\nu}$  is

$$\delta L \propto \frac{h}{2} L \quad (1.2)$$

The effect on test masses due to a passing GW is illustrated in figure 1.1. The value of the gravitational constant being very small, only compact astrophysical objects, subject to high accelerations, produce waves with an amplitude that can be probed with ground based detectors. The amplitude of a GW is inversely proportional to the distance from the source which typically are coalescing binary neutron stars, supernovae and black hole interactions. Nearby events have a higher amplitude but are rare. Instead, events occurring at higher distance have lower amplitude but are more frequent, a larger portion of the Universe being probed [8].

If events in nearby galaxies are considered, i.e. at a distance of around 20 Mpc,

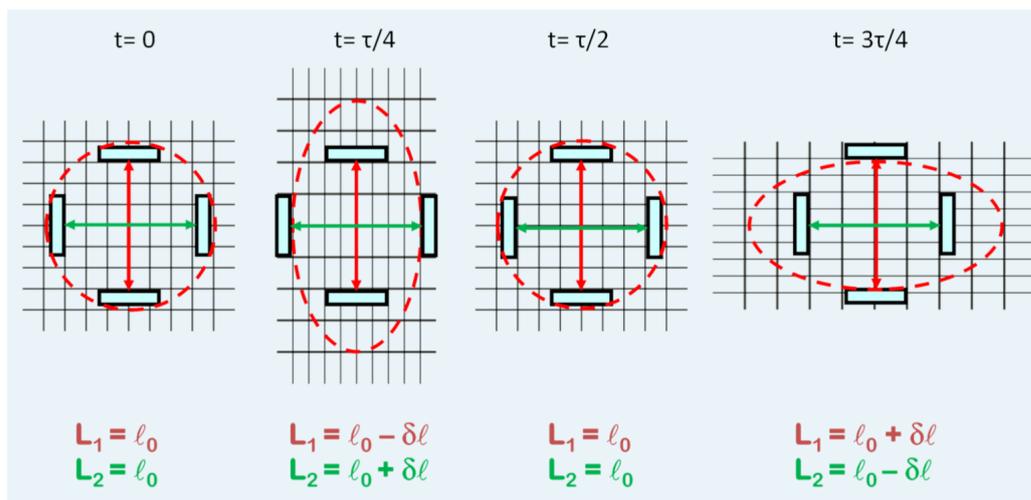


Figure 1.1: GW change distances between objects, while the objects themselves locally remain at rest, by changing the metric of space-time. These changes occur with opposing sign for orthogonal directions. In this figure this is illustrated for one polarisation of a GW incident perpendicular to the paper. Shown is the effect of a sinusoidal gravitational wave with period  $\tau$ , for different times  $t$ . The distances measured between the mirrors change by  $\pm\delta l$  [10]. Figure taken from [10].

an amplitude of the order of  $h = 10^{-21}$  was taken as a reference when Virgo was conceived [8]. Equation 1.2 shows that the measurement of  $\delta L$  has to be precise to the order of  $10^{-18}m$  for an interferometer with arm length  $L = 1km$ .

Historically, GW detection has been attempted using two approaches, i.e. either measuring the amplitude of oscillations of a resonant bar or measuring space-time geometry variations as detected by freely falling masses using Michelson based interferometers such as Virgo and LIGO [8]. The second approach is described in more detail in the next sections.

## 1.2 Virgo Interferometer: Principles of Operation

In this section the principles of operation of the Virgo ifo are briefly described based on [8]. A Michelson ifo is ideally suited for measuring the effect on test masses, described in the previous section, of a travelling GW. The measurement principle is schematically shown in figure 1.2: a laser beam is split into two beams, sent along the orthogonal ifo arms, it reaches the end mirrors where it experiences a phase shift due to the metric change caused by the GW. Then the beam returns to the beam splitter, where it is recombined. The interference condition at the beam splitter, i.e. the phase relation of the two returning beams, determines the intensity on the photo detector [10].

Virgo detector, located in Cascina near Pisa (Italy) is a Michelson interferometer, i.e. two perpendicular arms of equal length designed to cancel wavelength fluctuations of the light source. As described in the previous section, the effect of a travelling GW is to change the geometry of space-time. This results in a different apparent length of each arm. The arm length difference  $\delta L$  gives the gravitational

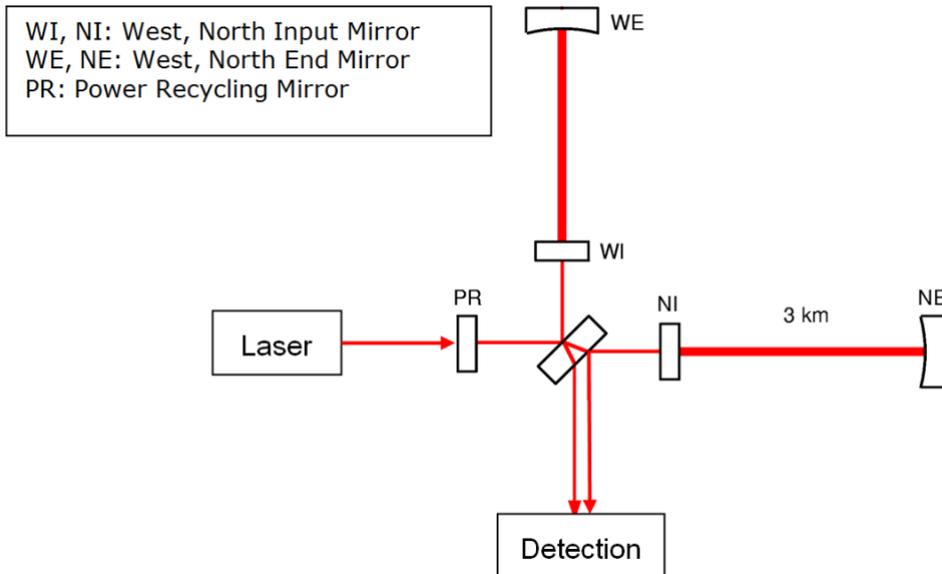


Figure 1.2: Operational principle of a Michelson interferometer with power recycling and Fabry-Perot resonant cavities. Figure from [8]

wave signal in terms of strain  $h$

$$\delta L \propto hL. \quad (1.3)$$

The measurement of  $\delta L$  is obtained from the phase difference of the returning beams, which is deduced from the intensity of the interference figure. The best sensitivity is achieved by setting the interferometer at almost complete destructive interference, i.e. dark fringe.

As described in [8], sensitivity is improved by introducing Fabry-Perot resonant cavities in the arms. The optical path is enhanced, resulting in a larger phase variation for a given arm length change. However, at high frequency, the GW may change sign while the photons are stored in the cavity. This reduce the phase response of the interferometer.

The Virgo arm length is  $L = 3km$ , while the optical path increase is a factor of 30, giving an effective optical length of 90 km [8].

The sensitivity of the phase measurement is limited by photon counting statistics. It can be improved by increasing the intensity of the light circulating in the ifo. This is achieved with a recycling cavity, as can be seen in Figure 1.2. If the ifo is set on dark fringe most of the light power goes back to the light source and this light can be sent back to the ifo by installing an additional semi-reflective mirrors at the entrance of the ifo itself. This increase the total circulating intensity. This consist in forming an overall resonant cavity or recycling cavity [8].

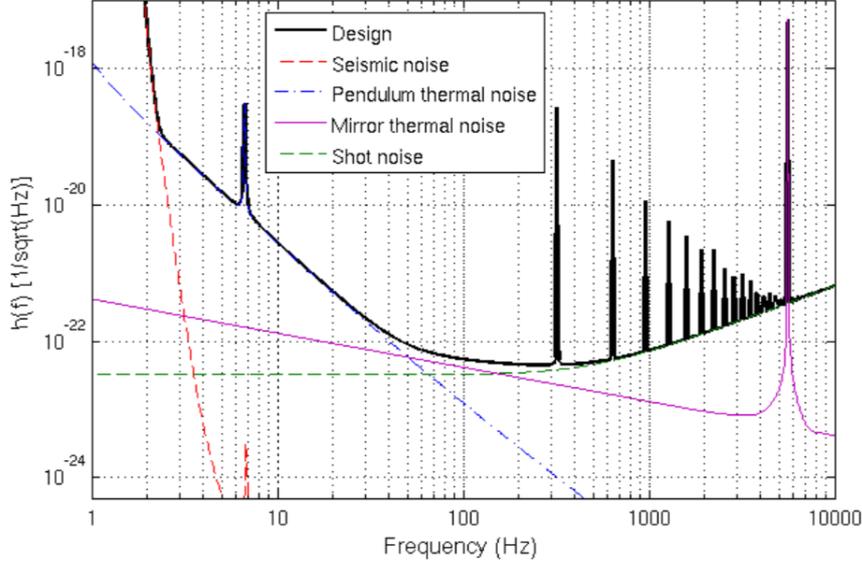


Figure 1.3: Sensitivity curve and main sources of noise affecting GW detection in different frequency bands. Figure from [8]

### 1.3 Sensitivity Curve and Fundamental Noises

In this section, the fundamental noises affecting ground based interferometers are briefly summarised, based on [8] and references therein.

Due to local gravity fluctuations ground-based interferometers are sensitive to gravitational waves with frequencies higher than a few Hz and at lower frequency a GW detector has to operate in space while on the high side of the frequency band the GW signal decreases due to source dynamics. The Virgo interferometer has been designed to cover a frequency band starting from 10 Hz up to a few kHz. Expected sources in this range are coalescing binary systems like neutron star or black hole binaries, stellar collapses, rotating neutron stars, and possibly cosmological background radiation [8].

The Virgo sensitivity is limited by noise sources that can be grouped into different categories

- The position of the test masses can fluctuate due to local perturbations, such as residual seismic noise, local gravity fluctuations and thermal motion
- Other noise sources affect the detected signal without a real mirror displacement, the main source being photon shot noise.

This provide with a sensitivity curve which is illustrated in Figure 1.3, which shows the expected Virgo sensitivity, which is the  $h$  spectral amplitude of a wave that would generate the same signal as the noise present in the detector, as a function of frequency. Typical values are  $\bar{h} = 5 \frac{10^{-23}}{\sqrt{\text{Hz}}}$  at 100 Hz, and  $\bar{h} = 3 \frac{10^{-21}}{\sqrt{\text{Hz}}}$  at 10 Hz [8]. In the following, the main sources of noise affecting the sensitivity curve of ifo's such as Virgo are briefly described, based on [12]. Seismic and Newtonian noise are instead discussed more in detail in section 1.4 and 1.6, respectively.

### 1.3.1 Amplitude Noise in the Laser Output Power

A fundamental limit to the amplitude noise in a laser output is the shot noise in the arrival rate of photons, which follows a Poisson statistic, and the noise generated in the process of detection. The equivalent spectral noise displacement squared per unit frequency is given by [12]

$$\frac{\Delta h^2(f)}{\Delta f} \geq \frac{hc\lambda}{8\pi^2\epsilon P b^2 e^{-b(1-R)}} \quad (1.4)$$

where  $h$  is the Planck constant,  $c$  the velocity of light,  $\lambda$  is the laser wavelength,  $\epsilon$  the quantum efficiency of the photodetector,  $P$  the laser output power,  $b$  the number of passes in each ifo arm and  $R$  the reflectivity of the mirrors.

### 1.3.2 Mechanical Thermal Noise

As described in [12], there is two kinds of thermal noise, one is due to the thermal motion of the center of mass of the masses on the horizontal suspensions and the other is thermal excitation of the internal normal modes of the masses about the center of mass. Thermal noise is modelled assuming the mechanical system is driven by a stochastic driving force with spectral power density given by

$$\frac{\Delta F^2(f)}{\Delta f} = 4k_B T a \quad (1.5)$$

where  $k_B$  is Boltzmann constant,  $T$  the absolute temperature of the damping medium and  $a = m\omega_0/Q$  is the damping coefficient, which can be expressed as a function of the mass  $m$  and resonant frequency  $\omega_0$  of the mechanical system, and  $Q$  is a quality factor. The spectral power density of the displacement, squared, because of the stochastic force on an harmonic oscillator is

$$\frac{\Delta h^2(f)}{\Delta f} = \frac{1}{m^2\omega_0^4} \frac{1}{(1-z^2)^2 + (z/Q)^2} \frac{4k_B T \omega_0 m}{Q} \quad (1.6)$$

where  $z = \omega/\omega_0$ . It is important to notice that the suspension should have a resonant frequency much lower than that of the GW to be detected. In this case  $z \gg 1$  and  $Q \gg 1$  and one obtain

$$\frac{\Delta h^2(f)}{\Delta f} = \frac{4\omega_0 k_B T}{\omega^4 m Q} \quad (1.7)$$

Instead the lowest normal mode frequencies of the internal motions of the masses, i.e. the mirrors and other suspended optical components, should be higher than the GW frequency. The entire suspended optical system should be as rigid as possible. Hence, for the internal motions  $z \ll 1$  while  $Q \gg 1$ , giving

$$\frac{\Delta h^2(f)}{\Delta f} = \frac{4k_B T}{\omega_0^3 m Q} \quad (1.8)$$

Beside reducing the temperature, thermal noise can be minimized using high  $Q$  materials and suspensions with high  $Q$ . The GW frequency should not fall near one of the mechanical resonance.

One problem with suspensions is that they have many degrees of freedom which tend to cross-couple nonlinearly with each other [12].

### 1.3.3 Radiation Pressure Noise from the laser Light

Fluctuation in the output power of the laser can induce mirror motion through the radiation pressure of light. Radiation pressure noise can be treated similarly to thermal noise. The spectral power density of a stochastic radiation pressure force on one mirror is given by [12]

$$\frac{\Delta F_{rad}^2(f)}{\Delta f} = \frac{4b^2 h P}{\lambda c} \quad (1.9)$$

where  $b$  is the number of times the light beam hits the mirror and  $P$  is the average laser power

## 1.4 Seismic Noise

Below a few tens of  $Hz$  the detection is limited by seismic noise consisting of vibrations of the ground causing motions on the optical components. In this way the small mirror displacement caused by a gravitational wave will be masked by seismic noise signals many orders of magnitude larger. Since many astrophysical sources such as pulsars and coalescing binaries are expected to emit mainly low-frequency gravitational waves, from a fraction of  $Hz$  to a few tens of  $Hz$ , the low frequency detection threshold should be lowered as much as possible [13].

If the test masses were firmly attached to ground, the seismic noise would be larger than any other source of noise previously described. The power spectrum due to seismic noise can be approximated by [12]

$$\frac{\Delta h^2(f)}{\Delta f} \propto f^{-4} \quad (1.10)$$

For this reason, the test masses are suspended and seismically isolated to minimize coupling with ground motion. The isolation provided by a single degree of freedom suspension is given by

$$|\Delta x_m(f)/\Delta x_l(f)|^2 = \frac{[(1-z^2) + (2/Q)^2]^2 + (z^3/Q)^2}{[(1-z^2)^2 + (z/Q)^2]^2} \quad (1.11)$$

where  $z = f/f_0$  with  $f_0$  being the resonant frequency of the suspension. In Equation 1.11,  $\Delta x_m(f)$  is the displacement of a test mass at frequency  $f$  relative to an inertial frame while  $\Delta x_l(f)$  is the motion of the Earth measured in that reference frame. At frequencies for which it holds  $z \gg 1$ , the isolation ratio is

$$|\Delta x_m(f)/\Delta x_l(f)|^2 \sim \frac{f_0^4}{f} + \frac{f_0^2}{f} \frac{1}{Q^2} \quad (1.12)$$

To increase the rejection of seismic noise the suspension can be made longer or alternatively several shorter periods suspensions can be used in series, since their isolation factors multiply. The disadvantage is that in this case cross coupling among moving objects is enhanced [12]. Regarding Virgo, an elaborate suspension system has been developed, the superattenuator (SA), to support each optical component of the ifo. The SA, described in 1.5, has been designed to suppress the seismic noise transmission to the last stage of the mirror suspension system by more than ten orders of magnitude starting from about 4  $Hz$  [13].

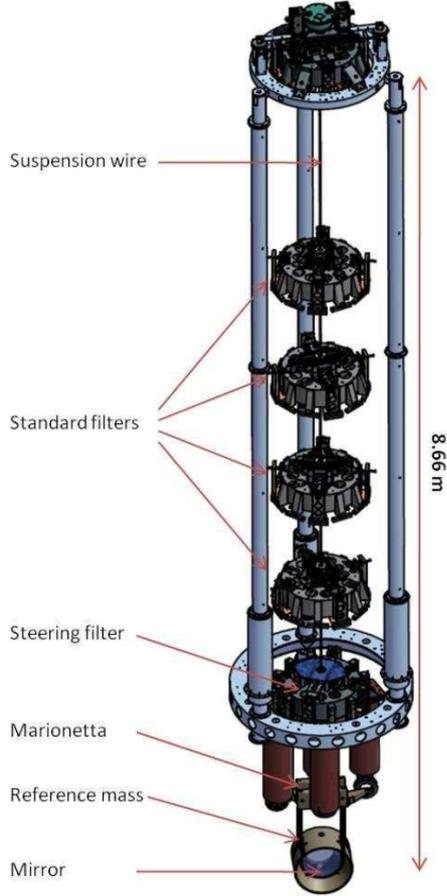


Figure 1.4: Schematic of the Virgo SA multi stage pendulum configuration. Figure taken from [14].

## 1.5 The Superattenuator Design and Working Principle

The SA design is based on a multi stage pendulum configuration. Regarding an  $n$ -stages pendulum, it can be shown that horizontal motion of its suspension point, at a frequency  $f$  higher than the frequencies of the chain normal modes ( $f > f_0 > f_1 \dots f_n$ ), is transmitted to the suspended mass with an attenuation proportional to  $f^{-2n}$ . Furthermore, the ratio between the linear spectral density of the last mass horizontal displacement, i.e. the optical component, and the linear spectral density of the suspension point horizontal displacement, i.e. where the excitation is applied, decreases as  $A/f^{2n}$  where  $A = f_0^2 \cdot f_1^2 \dots f_n^2$ . This way, a very large attenuation of the seismic noise, horizontal component, can be obtained at a frequency above the highest pendulum resonance [13].

Unfortunately the end mirrors, suspended 3 km away in the ifo, are misaligned with respect to one another by about  $3 \cdot 10^{-4}$  rad because of the different directions of the plumb line on the curved Earth surface, meaning that they are not perfectly perpendicular to the laser beam. Since any vertical vibration will be partially

transmitted to the interferometer horizontal axis, i.e. along the laser beam direction, due to coupling among different degrees of freedom, vertical motion cause a phase change of the laser beam [13].

Hence a vertical attenuation of seismic noise comparable with the horizontal one is needed to reduce the frequency detection threshold. With a multistage pendulum this is obtained by replacing each suspension wire with a spring, acting as vertical oscillators. The spring should support a heavy load and, at the same time, it should be soft enough to exhibit a low resonant frequency [13]. The SA of the Virgo ifo consists of three fundamental elements

- The inverted pendulum
- The seismic filters, connected to each other by metallic suspension wires
- The last stage, i.e. the pay-load of the chain.

These are hereafter described based on [13].

### 1.5.1 The Inverted Pendulum

As described in [13], an ideal inverted pendulum can be conceived as a massless vertical bar of length  $l$  connected to a ground by means of an elastic joint with stiffness  $k$  and supporting a mass  $M$  on its top. In such a pendulum the gravity acts as an anti-spring, which resonant frequency, given by

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{k}{M} - \frac{g}{l}} \quad (1.13)$$

can be lowered by increasing the mass  $M$  until mechanical instability is reached. Such a mechanical structure can operate at very low frequency, down to a few tens of mHz. Attaching the suspension point of the multistage pendulum chain on the top of an inverted pendulum, this acts as an ultralow frequency pre-isolator stage, suppressing a large part of the input horizontal seismic noise. The inverted pendulum of Virgo SA can be seen in Figure 1.4.

### 1.5.2 The Seismic Filters

In the multi stage pendulum configuration of the SA, each pendulum mass has been replaced by a rigid drum-shaped metallic structure acting as an oscillator in the vertical direction, referred to as the seismic filters. Due to the working principle of a multistage pendulum, a sequence of five mechanical filters provides with a good seismic noise isolation of the optical components. In this Section, a brief description of the seismic filters is given based on [13] while a detailed description can be found in [15].

The mechanical filter has a 70 cm diameter, is 18.5 cm high, and is suspended as close as possible to its center of mass. On the outer circumference of the filter body, bottom part, a set of triangular cantilever spring blades is clamped, as can be seen in Figure 1.5. Each blade, 3.5 mm thick and 385.5, mm long is bent at a constant curvature radius and with a different base width according to the load to

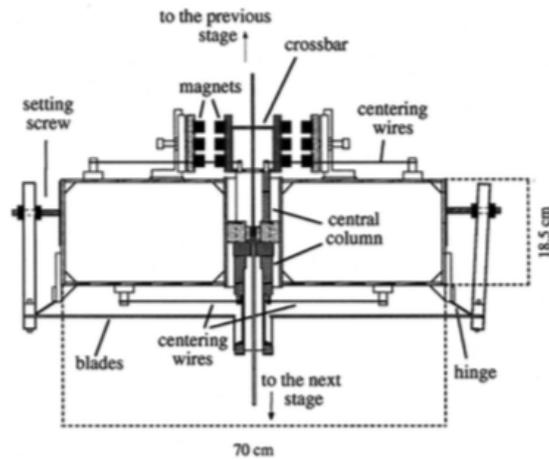


Figure 1.5: Schematic of the Virgo SA seismic filters. Figure taken from [13].

be supported. A nominal load, ranging between 48 and 96 kg hung on the blade tip, forces it in a flat and horizontal position. The blade tip is connected by a 1 mm diameter wire to a central column, inserted through a hole in the center of the filter body. Any movement of the central column, apart from the vertical direction, is prevented by two sets of four centering wires accommodated on the top and on the bottom of the filter body. A crossbar, bolted on the upper part of the central column, is used as a mechanical support for the magnetic antispring system, developed in order to reduce the vertical stiffness of the blades and to confine the main vertical resonant frequency of each filter below the pendulum one. The central column and the crossbar represent the moving part of the mechanical filter from which the load of the lower stages is suspended by a steel suspension wire. Connecting each filter to the next, a chain of mechanical oscillators in the vertical direction is obtained [13].

### 1.5.3 The Last Stage

The last stage of the SA is suspended from the last mechanical filter of the chain, which is historically called "Filter 7". It consists of a special anvil-shaped element with four wings named "Marionette", a reference mass, and a mirror (see Figure 1.4). The marionette has been designed with four wings on which four small permanent magnets are attached. In front of these magnets four coils attached at the end of four aluminum pipes screwed on the bottom part of filter 7 are placed. Four thin wires start from the marionette. The first pair supports the mirror in a cradle and the second one, with the same technique, supports the reference mass forming the last stage with a pendulum length of 0.7 m. The magnet-coil system drives the marionette allowing control of the interferometer optical component in three relevant degrees of freedom: the displacement along the beam direction, the rotation around the vertical axis, and the rotation around the horizontal axis perpendicular to the beam. The final control of the mirror displacement in the beam direction is made by four coils acting on the magnets glued on the back side of the mirror [13].

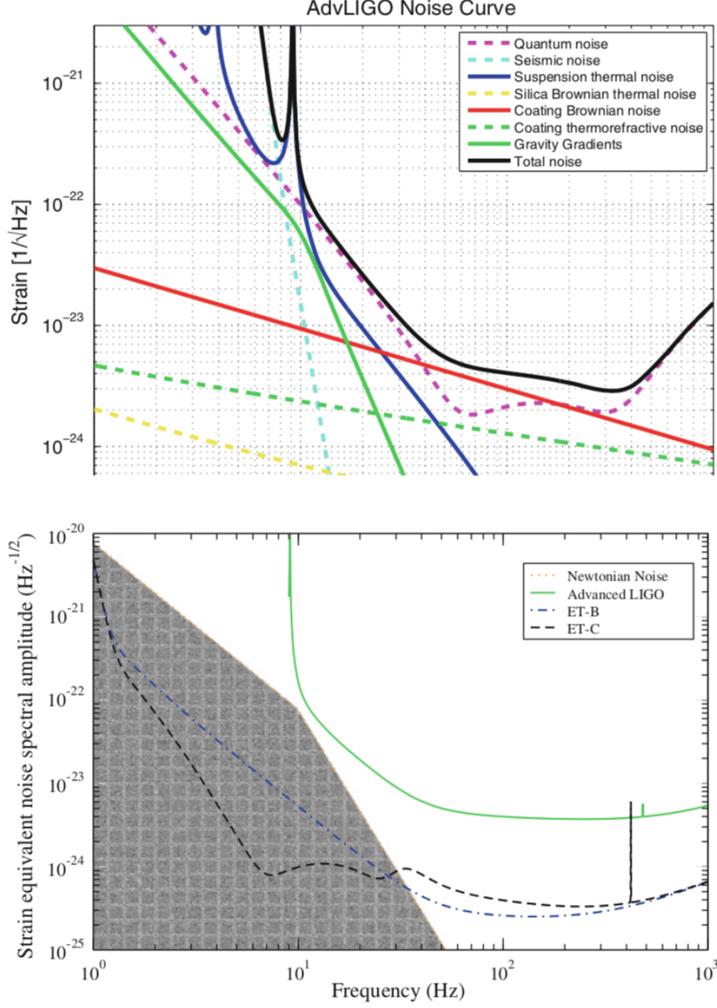


Figure 1.6: NN estimate in 2 and 3 generation detectors. Figure taken from [16].

## 1.6 Newtonian Noise

In this section a brief description of NN main characteristics and mitigation strategies is given based on [10, 3]. The ifo is sensitive to gravitational field gradients or in other word to differential gravitational forces exerted on the end masses placed at the ends of the ifo's arms. Two physical effects are relevant

- Time dependent density variations in both atmosphere and ground
- Motions of inhomogenities in the mass distribution around the detector

A simplified estimate of NN is now given, based on [10]. The acceleration that a test mass located at  $y$  experience, given a distribution of masses  $\rho(x, t)$  is

$$a^{NN}(y, t) = G \int_V \rho(x, t) \frac{x - y}{|x - y|^3} dV \quad (1.14)$$

where the integration is extended over the volume  $V$  and  $G$  is the gravitational constant. For NN estimates is interesting the time variation of such quantity and

the medium considered is an elastic solid. Hence, inserting in Equation 1.14 the equation for the conservation of mass

$$\dot{\rho} + \nabla \cdot J_m = 0 \quad (1.15)$$

where  $J_m = \rho_0(x)\dot{\xi}(x, t)$  is the mass density current,  $\rho_0$  is the density of the medium in static configuration and  $\xi$  is a small displacement, gives

$$a^{NN}(y, \omega) = G \int_V \nabla \cdot [\rho(x)\xi(x, \omega)] \frac{x - y}{|x - y|^3} dV \quad (1.16)$$

Two different effects are contained in this expression, as can be seen expanding the derivative. The term proportional to  $\rho_0 \nabla \cdot \xi$  describes the fluctuation of the local density due to the compression of the medium. Instead  $\xi \cdot \nabla \rho_0$  accounts for movements of density inhomogeneities, e.g. at the surface boundary [10]. A theory describing the connection in between seismic measurements and NN can be found in [16]. As explained also in [10], the general idea is to decompose seismic motion in normal modes, which are expected to behave as oscillators coupled to unknown stochastic forces. By measuring quantities connected to seismic fluctuations such as the power spectrum, information about the excitation of the oscillators can be obtained and using Equation 1.16 it can be converted into an estimate of NN.

Third generation GW interferometric detectors will be underground detectors to extend the GW detection frequency band to below the NN limit. Figure 1.6 (top) shows the expected noise budget for a second generation GW ifo detector while Figure 1.6 (bottom) shows a comparison between NN noise and the planned sensitivity of a third generation detector, i.e. the Einstein Telescope (see <http://www.et-gw.eu/>). It can be seen that NN becomes the most critical sensitivity limit in the low frequency region [16].

### 1.6.1 Mitigation of Newtonian Noise

A brief description of NN mitigation techniques is now given based on [3] and references therein.

As explained in [3], mitigation of NN without employing environmental data are considered to be passive Newtonian-noise mitigation techniques. Among these, site selection is one of them. In this case the quietest detector site should be identified, in terms of seismic noise and atmospheric noise.

Opposite to that are active mitigation strategies. In this case, while active seismic isolation cancels seismic noise before it reaches the suspension stages of a test mass, NN have to be cancelled in the data of the GW detector. Coherent noise cancellation, also known as active noise cancellation, is based on the idea that the information required to model noise in data can be obtained from auxiliary sensors that monitor the sources of the noise. The noise model can then be subtracted from the data in real time or during post processing with the goal to minimize the noise. A well known technique is based on Wiener filters, as described in more details in [3]. In general for active noise mitigation techniques an array of sensors, i.e. seismometers, is needed.

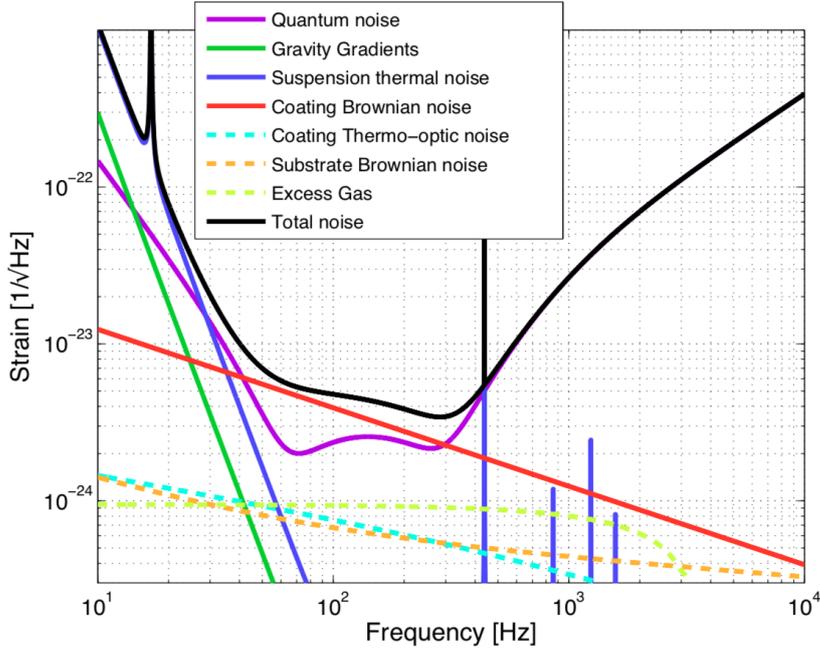


Figure 1.7: Sensitivity curve of Virgo in its advanced configuration and expected noise contributions. Figure taken from [14].

## 1.7 Advanced Virgo Interferometer: A second Generation Detector

Advanced Virgo (AdV) is the project to upgrade the Virgo detector to a second generation instrument. AdV is designed to achieve a sensitivity of about one order of magnitude better than that of Virgo (see Figure 1.7). This corresponds to an increase in the detection rate of three orders of magnitude [1]. Major subsystems upgrades to reach AdV configuration are foreseen. A detailed description can be found in [1, 14]. Some of the upgrades foreseen are now described, based on [1]

- **Interferometer optical configuration.** AdV will be a dual-recycled ifo. Besides the standard power recycling, a signal-recycling (SR) cavity will also be present, as can be seen in Figure 1.8. The tuning of the signal-recycling parameter allows for the changing of the shape of the sensitivity curve and the optimizing of the detector for different astrophysical sources. To reduce the impact of the thermal noise of the mirror coatings in the mid-frequency range, the beam spot size on the test masses has been enlarged.
- **Increased laser power.** Improving the sensitivity at high frequency requires high laser power. The AdV reference sensitivity is computed assuming 125 W entering the interferometer, after the Input Mode Cleaner (IMC).
- **Thermal compensation system.** A sophisticated thermal compensation system has been designed to cope with thermally-induced aberrations and with losses induced by intrinsic defects of the optics. The sensing is based on

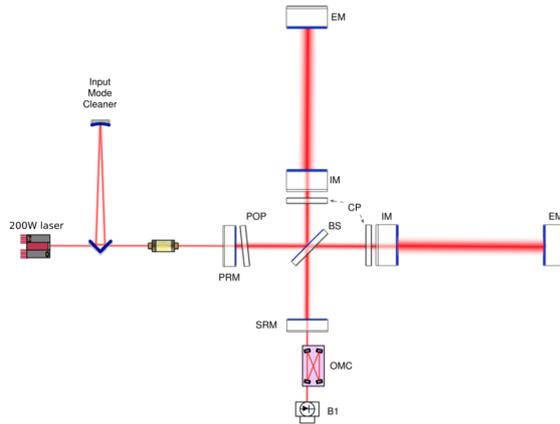


Figure 1.8: Schematic of the optical layout of the Advanced Virgo ifo. Each 3 km-long arm-cavity is formed by an Input Mirror (IM) and an End Mirror (EM). The recycling cavities at the center of the interferometer are 12 meters long and are formed by the Power Recycling Mirror (PRM), the Signal Recycling Mirror (SRM) and the two IM. Figure taken from [1].

Hartmann sensors and phase cameras, while the ring heaters around several suspended optics will be used as actuators to change the radius of curvature. CO<sub>2</sub> laser projectors, which shine on dedicated compensation plates, allow the compensation of thermal lensing and optical defects.

- **Mirrors** To cope with the increased impact of radiation pressure fluctuations the AdV test masses will be twice as heavy (42 kg) as those of Virgo. Fused silica grades with ultra low absorption and high homogeneity have been chosen for the most critical optics. State of the art polishing technology is used to reach a flatness better than 0.5 nm rms in the central area of the test masses. Low-loss and low-absorption coatings are used to limit as far as possible the level of thermal noise and the optical losses in the cavities, which eventually determine the sensitivity in the high frequency range.
- **Stray light control** Scattered light could be a significant limitation on detector sensitivity. In order to limit phase noise caused by part of the light being back-scattered into the interferometer, new diaphragm baffles will be installed. These will be either suspended around the mirrors, or ground-connected inside the vacuum links.
- **Payloads and vibration isolation** A new design of the payloads has been developed. This was triggered mainly by the need to suspend heavier mirrors, baffles and compensation plates, while controllability and mechanical losses have also been improved.
- **General detector infrastructure** Important modifications have been undertaken in the main experimental hall turning them into acoustically-isolated clean rooms. The vacuum has been upgraded by installing large cryotrap at the ends of the 3 km pipes, in order to lower the residual pressure by a factor of about 100.

# Chapter 2

## Methodology: Adaptive and Fractal Algorithms

In this Chapter, the methodology developed in the framework of this Thesis, and adopted for data analysis is firstly outlined in section 2.1. The concepts on which it relies, i.e. adaptive and fractal algorithms, are then described in more detail in subsequent sections. Hilbert Spectral Analysis (HSA) and the concepts of instantaneous frequency (IF), amplitude (IA) and Bandwidth (IB) are described in section 2.2. The Empirical Mode Decomposition algorithm is described in section 2.3. Applying EMD, a given time series is decomposed in Intrinsic Mode Functions (IMFs). If then HSA is applied to the obtained IMFs, EMD and HSA are referred to as Hilbert Huang Spectral Analysis (HHSA). It provides the Hilbert Huang Transform (HHT) i.e. a time-frequency-amplitude representation of the analysed data, which has a higher resolution compared to Fourier based methodologies. In section 2.4, the Hilbert-Huang Transform is described. The issue of mode mixing and Ensemble methodologies like Ensemble EMD (EEMD) used to mitigate it are then described in section 2.5 while a test for significance of IMFs against white noise is described in section 2.6. The recently developed adaptive algorithm time-varying filter EMD (tvf-EMD), on which the adopted methodology is based, is finally described in section 2.7. Finally, in section 2.8 is described the fractal algorithm Detrended Fluctuation Analysis (DFA) and also how to compute the Hurst exponent locally. In section 2.9 the tool developed in the framework of this Thesis is described in more detail.

### 2.1 Outline of adopted Methodology

In this Section the methodology of time series analysis tested and developed in the framework of this Thesis is briefly described. It relies on concepts, like HSA, EMD, HHSA and tvf-EMD which are all described in more details in the following sections. It aims at describing time series of data  $x(n)$  in term of three main components, namely trend, signal and noise

$$x(n) = T(n) + s(n) + r(n) \quad (2.1)$$

where  $T(n)$ ,  $s(n)$  and  $r(n)$  stands for trend, signal and residual components, respectively, and  $n = 1 \dots N$  is the number of data samples. To deal with data that

are possibly affected by nonlinearities and/or nonstationarities, an adaptive approach based on EMD is employed. To extract and characterise the three terms of Equation (2.1), the following techniques are employed

- **Empirical Mode Decomposition (EMD)** is used for adaptive detrending. Last term of EMD is a slowly varying function  $T(n)$ , which can be either the adaptive trend or the baseline wandering of the data, if present[17][18].
- **Time varying filter EMD (tvf-EMD)**, a recently developed algorithm, is employed in order to extract narrow band oscillatory modes embedded in the data[6]. For sake of clarity, such modes are hereafter referred to as Intrinsic Mode Functions (IMFs), as the ones extracted by standard EMD. Due to the completeness property of the EMD and the tvf-EMD algorithms, summing up all the extracted IMFs exactly recovers the original data.
- **Denoising** is performed making use of a thresholding approach estimating persistency of the extracted IMFs, persistency which is quantified by means of their Hurst exponent  $H$ , with  $0 < H < 1.5$ <sup>1</sup>. The Hurst exponent of each IMF is evaluated with DFA and it is used to separate "signal-like" IMFs, having  $H > H_{thresh}$  and that are persistent, from "noisy-like" IMFs, having  $H < H_{thresh}$  and that exhibit anti persistent behavior. Summed together, IMFs above threshold give the extracted signal  $s(n)$  while residuals  $r(n)$ , or noise, are obtained summing up IMFs under threshold. After that, the three terms of Equation 2.1 can be independently characterised, either in term of amplitude and frequency modulations (AM-FM) of the modes constituting  $s(n)$ , or in term of outlier occurrence in the case of  $r(n)$ .
- Having separated the signal term  $s(n)$  from the noise term  $r(n)$ , **denoising performances** can be evaluated. This has been done following the approach of [19], making use of the following parameters
  - Mean squared error (MSE)
  - Mean absolute error (MAE)
  - Signal to noise ratio (SNR)
  - Peak signal to noise ratio (PSNR)
  - Crosscorrelation (xcorr)
 between the extracted signal  $s(n)$  and the input data  $x(n)$ . Low values of MSE, MAE and high values of SNR, PSNR and xcorr indicate good denoising performances, respectively.

In the remaining of this Chapter, the main features of the adopted methodology are introduced and described. First, the HSA and the concepts of IF, IA and IB are introduced. The equations of HSA bear physical information and can be meaningfully applied to monocomponent or at least narrowband oscillatory modes.

---

<sup>1</sup>The Hurst exponent is a useful fractal index quantifying time series long term correlation behaviour. Uncorrelated white noise has  $H = 0.5$  while persistency (antipersistency) is present if  $H > 0.5$  ( $H < 0.5$ ), respectively. Pink noise has  $H = 1$  while for  $H > 1$  the data are nonstationary.  $H = 1.5$  corresponds to Brownian Noise. A value of  $H < 1$  corresponds to noise like time series while  $H > 1$  to random walk like time series.

Hence, when multicomponent signals are to be analysed, they need to be decomposed into their components, or oscillatory modes, before performing HSA. To this end, different algorithms are employed. One of such algorithms is EMD, which is described in section 2.3. Since standard EMD applied to noisy datasets is affected by mode mixing, i.e. the presence of widely different oscillations in the same mode or the presence of two modes with oscillations of the same scale, noise assisted methodologies like Ensemble EMD (EEMD) and Complete-EEMD with adaptive noise (CEEMDAN) have been introduced to mitigate such problem [20][21], and are hence briefly described in section 2.5. Noise assisted methodologies succeed at extracting meaningful IMFs from noisy data, adding different realisations of white noise and then averaging the results of the ensemble of decompositions, hence they improve standard EMD performances. The output of the analysis though, depends on the standard deviation of the added white noise and on the number of trials. The tvf-EMD, is a recently developed algorithm [6], which improves EMD performances against mode mixing and intermittency and it employs B-splines as a time varying filter. It extracts narrow band modes of oscillation characterising the data. It is the algorithm adopted in this Thesis to extract narrowband oscillatory modes and it is extensively described in section 2.7, based on [6] and references therein. Since the adopted methodology is based on empirical algorithms, and performances need to be evaluated a posteriori, in Chapter 3 the algorithm is applied to datasets from different physical systems to test its ability to extract meaningful physical information.

## 2.2 Hilbert Spectral Analysis: The instantaneous frequency, amplitude and bandwidth

In this Section, the concepts of IF, IA and IB, are introduced based on [22][23][24] and references therein. The IF is a useful parameter when describing signals which spectral characteristics, e.g. the frequency of spectral peaks, are time varying. Such signals are referred to as **nonstationary**<sup>2</sup>, a notable example being the chirp signal, a sinusoidal wave which frequency sweeps in time.

The IF is a time varying parameter which defines the location of the signal spectral peak as it varies with time. It can be interpreted as the frequency of a sine wave which locally fits the data [22]. Its definition is application related, and the IF has then to be estimated empirically. It bears physical meaning only if estimated for a signal which is monocomponent or at least narrow band. If instead the analysed signal is multicomponent, it first needs to be decomposed into its components and this is usually achieved using EMD and related algorithms.

The IF can be defined based on the concept of analytic signal, firstly introduced

---

<sup>2</sup>The term stationarity is also applied to the distribution of random variables. If the distribution is not dependent on time, the time series is said to be strictly stationary. Furthermore, a stationary process has the property that mean, variance and autocorrelation do not change with time. In the remaining of this Thesis, the term nonstationary is referred to varying frequency of oscillatory modes. Variation in the mean is instead taken into account in the trend term  $T(n)$ .

by Gabor [23]. Given a signal  $x(t)$ , the analytic signal  $z(t)$  is defined by

$$z(t) = x(t) + iH[x(t)] = x(t) + iy(t) = a(t)e^{i\phi(t)} \quad (2.2)$$

where the Hilbert Transform (HT) of  $x(t)$  is defined by

$$H[x(t)] = y(t) = \frac{1}{\pi}PV \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau, \quad (2.3)$$

and is the convolution product of  $x(t)$  and  $1/\pi t$ . The analytic signal  $z(t)$  can be considered the best local fit of an amplitude and phase varying trigonometric function to  $x(t)$  [17]. An intuitive definition of IF can be given making use of the following general expression for a frequency modulated (FM) signal [22]

$$x(t) = a \cos(\phi(t)) = a \cos \left[ \int_0^t 2\pi f_i(t) dt + \theta \right]. \quad (2.4)$$

The instantaneous frequency is hence defined by

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}. \quad (2.5)$$

A more general definition of IF for an AM-FM signal of the form  $x(t) = a(t)\cos\phi(t)$ , is given in term of the phase of the analytic signal

$$f_i(t) = \frac{1}{2\pi} \frac{d}{dt} [\arg(z(t))], \quad (2.6)$$

where  $z(t)$  is the analytic signal as defined in Equation 2.2. Signals  $x(t)$  and  $H[x(t)]$  are said to be in quadrature, because theoretically they're out of phase by  $\pi/2$ . For real signals though, this is only true if the conditions of the Bedrosian's product theorem (BPT) hold [22], i.e. the spectra of  $a(t)$  and of  $\cos\phi(t)$  are disjoint. In this case it follows that

$$H[a(t)\cos\phi(t)] = a(t)H[\cos\phi(t)] = a(t)\sin\phi(t), \quad (2.7)$$

and Equation 2.2 is verified. The BPT is reported in Appendix D.

The notion of analytic signal was firstly introduced by Gabor in order to define the central moments of the frequency [22][23]

$$\langle f^n \rangle = \frac{\int_{-\infty}^{\infty} f^n |Z(f)|^2 df}{\int_{-\infty}^{\infty} |Z(f)|^2 df}, \quad (2.8)$$

where  $Z(f)$  is the spectrum of the analytic signal  $z(t)$ . If instead the spectrum of the real signal  $x(t)$  is used, all the odd moment would be zero, since  $|X(f)|^2$  is an even function of frequency [22].

Making use of Gabor's results, it can be shown that the average frequency in a signal's spectrum equals the time average of the IF [24]

$$\langle f \rangle = \langle f_i \rangle \quad (2.9)$$

where

$$\langle f \rangle = \frac{\int_{-\infty}^{\infty} f |Z(f)|^2 df}{\int_{-\infty}^{\infty} |Z(f)|^2 df}, \quad (2.10)$$

$$\langle f_i \rangle = \frac{\int_{-\infty}^{\infty} f_i(t) |z(t)|^2 dt}{\int_{-\infty}^{\infty} |z(t)|^2 dt}. \quad (2.11)$$

Based on this relations, Ville formulated the Wigner-Ville distribution, which first moment respect to frequency equals the IF [22][24].

### 2.2.1 Interpretation of instantaneous frequency and bandwidth

Following [22], an insight on the meaning of IF can be given considering a signal  $x(t)$  in the frequency domain. The spectrum of the analytic signal of Equation 2.2 is given by

$$Z(f) = \int_{-\infty}^{\infty} z(t) e^{-i2\pi ft} dt = \int_{-\infty}^{\infty} a(t) e^{i[\phi(t) - 2\pi ft]} dt. \quad (2.12)$$

Applying the stationary phase principle, it follows that the integral have its largest value at  $f_i$  such that

$$\frac{d}{dt} [\phi(t) - 2\pi f_i t] = 0, \quad (2.13)$$

which then gives

$$f_i(t) = \frac{1}{2\pi} \frac{d}{dt} [\phi(t)] \quad (2.14)$$

Hence, if the instantaneous frequency  $f_i(t)$  is time dependent, it provides a measure of how the energy concentration of a signal varies in the frequency domain, as a function of time [22]. The **instantaneous bandwidth (IB)** is then defined as the standard deviation, or spread, of frequencies at that time.

In general the HT and the analytic signal can have meaningful physical interpretation if the spectra of  $a(t)$  and of  $\phi(t)$  are separated in frequency, e.g. are spectrally disjoint, as also stated by the BPT. The more a signal is narrowband, the better will be the estimate of its IF, and also the closer  $iH[x(t)]$  will approach the quadrature signal of  $x(t)$ . For this reason, before performing HSA on a multicomponent signal, its components need to be separated. This is the ultimate goal of the Empirical Mode Decomposition and related algorithms. EMD, noise assisted methodologies like EEMD and CEEMDAN, and the tvf-EMD algorithm are hence described in the next sections.

## 2.3 Empirical Mode Decomposition

EMD is an adaptive algorithm which aims at decomposing nonlinear and/or non-stationary data into Intrinsic Mode Functions (IMF), which are nearly monocomponent oscillatory modes of the data. The EMD algorithm, firstly introduced by

Huang [17][18], relies on the fundamental assumption that any time series of data consists of a superposition of different intrinsic modes of oscillation. The main assumption of EMD, as stated by Huang in [18], is the following: *"Any data consists of different simple intrinsic modes of oscillation. Each intrinsic mode, linear or nonlinear, represents a simple oscillation which will have the same number of extrema and zero-crossings. The oscillation will also be symmetric with respect to the local mean. At any given time, the data may have many different coexisting modes of oscillation, one superimposing on the others. Each of these oscillatory modes is represented by an intrinsic mode function (IMF)"*. The extraction of physically meaningful IMFs is the ultimate goal of the EMD algorithm. According to Huang, any oscillatory function must satisfy the two following **requirements** to be considered an IMF [18]

- In the whole dataset, the number of extrema and the number of zero-crossings must either equal or differ at most by one.
- At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

Considering a time series  $x(t)$ , where  $t = 1 \dots N$ , with  $N$  being the data length, the main steps of the EMD algorithm are [17, 18]

- 1) Identify all the local maxima/minima and connect them by a cubic spline to form an upper/lower envelope, covering all the data between them.
- 2) The envelope mean, designated as  $m_{1,j}(t)$  is subtracted from the data. This difference is the first component and is referred to as proto IMF

$$h_{1,j}(t) = x(t) - m_{1,j}(t). \quad (2.15)$$

- 3) If both requirements described earlier are satisfied

$$h_{1,j} = c_1(t) \quad (2.16)$$

is considered to be the first IMF and it is then subtracted from the data. The algorithm is then repeated on the data minus the first IMF

$$x_1(t) = x(t) - c_1(t). \quad (2.17)$$

- 4) If instead  $h_{1,j}(t)$  doesn't respect the aforementioned conditions, steps 1) and 2) of the EMD algorithm are repeated  $j = 1 \dots S$  times on  $h_{1,j}(t)$  until an IMF is obtained according to the requirements. This procedure is referred to as **sifting**.

The sifting procedure aims at making the wave profiles symmetric eliminating riding waves, so that the HT can give meaningful estimates of the IF when applied to mono-component or narrow band signals [17].

**Stopping criteria** are typically employed in standard EMD and related methodologies. They either limit the number of sifting iterations to  $S$  (typical values are  $S \leq 10$ ) or introduce a threshold on the required symmetry of the upper/lower

envelopes. Stopping criteria are necessary both to help the algorithm to converge but also because oversifting could possibly remove physically meaningful amplitude modulations. One example of stopping criteria is the Cauchy convergence test, in which the sifting is stopped when the normalised squared difference between two successive sifting iterations

$$SD_j = \frac{\sum_{t=0}^T |h_{j-1}(t) - h_j(t)|^2}{\sum_{t=0}^T h_{j-1}^2(t)} \quad (2.18)$$

is lower than a given threshold. Other criterion allows instead to define different thresholds on the IMF mean that must be respected respectively on local and global intervals of the data. Stopping criteria contribute to relax the otherwise very strict restrictions of the aforementioned two conditions for having an IMF [17].

The EMD algorithm extracts IMFs until the residual term is smaller than a selected value or when it becomes a monotonic function, and can then be considered as the adaptive trend or baseline wandering of the data, if present. Such trend component is hereafter referred to as  $T(n)$ , where  $n = 1 \dots N$  are the data samples of a time series of length  $N$ . Hence, after applying EMD, the data are decomposed into  $j = 1 \dots K$  different IMFs,  $c_j(n)$ , plus a term  $T(n)$ , which is the adaptive trend

$$x(n) = \sum_{j=1}^K c_j(n) + T(n). \quad (2.19)$$

From equation (2.19), it can be seen that the EMD algorithm is **complete**, e.g. summing up the extracted oscillatory modes plus the trend exactly recovers the original data. It is found empirically that the number of extracted IMFs depends logarithmically on the length of the data  $K \simeq \log_2 N$ . The step of removing the mean is not needed when using EMD, because it is adaptively extracted by the EMD algorithm itself. Figure 2.1 shows the initial steps of the EMD algorithm and how the envelope mean is obtained and then subtracted providing a proto IMF.

## 2.4 Hilbert-Huang Transform: a time-frequency-energy representation of nonlinear nonstationary data

The EMD algorithm allows to obtain a collection of IMFs,  $c_j(n)$ , with  $n = 1 \dots N$  and  $j = 1 \dots K$  number of samples and of extracted modes, respectively. The expansion of the data  $x(n)$  obtained performing EMD is defined in term of an adaptive base, i.e. the IMFs, which is obtained from the data a posteriori. Following the notation of [17], the data can be expressed as

$$x(t) = Re \left\{ \sum_{j=1}^K a_j(t) e^{j \omega_j(t) dt} \right\}, \quad (2.20)$$

where both frequency and amplitude of the basis functions, i.e. the IMFs, are time varying. The instantaneous frequency of the  $j$ th IMF is given by  $f_j(t) = \omega_j(t)/2\pi$ ,

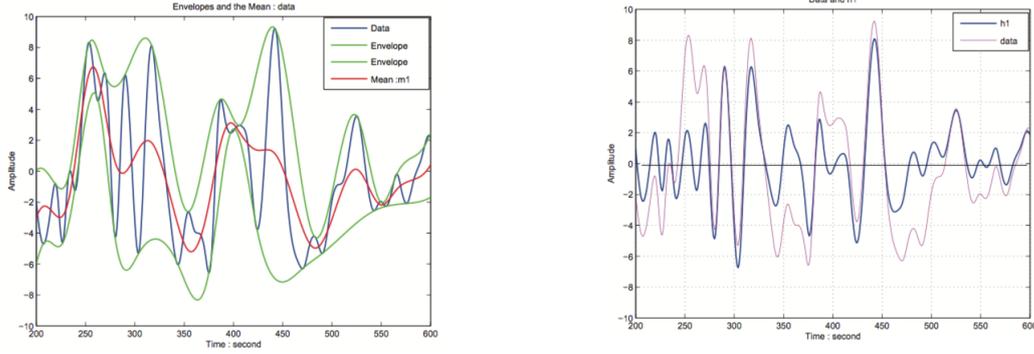


Figure 2.1: Shown is the EMD procedure to obtain oscillatory modes of the data. **Left:** In green the upper and lower envelope of the data, which are shown in blue. Envelopes mean  $m_1$  is shown in red. **Right:** In blue is shown the first oscillatory mode  $h_1$ , or proto IMF, obtained subtracting the mean  $m_1$  from the data, depicted in pink. After the first application of the algorithm,  $h_1$  is not yet an IMF. EMD is repeated until the requirements to be an IMF are met, also referred to as **sifting**. Figure taken from [17][18]

where  $\omega(t)$  is the time derivative of the adaptively determined phase function. It is expressed in Hertz. Differently from frequency, the IF is no longer a mean value over the whole time domain obtained by a given transform. It is instead characterised by instantaneous values at different times [25].

Applying HSA to the  $K$  different modes  $c_j(n)$  extracted by EMD provides the Hilbert-Huang Transform (HHT), or Hilbert spectrum,  $H(\omega, t)$ , which is a time-frequency-amplitude representation of the data. The expansion of Equation 2.20 can be considered a generalisation of the classical Fourier expansion, which would've been

$$x(t) = Re \left\{ \sum_{j=1}^K a_j e^{i\omega_j t} \right\}, \quad (2.21)$$

where both  $a_j$  and  $\omega_j$  are constants. Restrictions on amplitude and frequency of the Fourier components are no longer retained due to the adaptive nature of the algorithm, which allows it to expand nonlinear non stationary datasets in term of amplitude and frequency modulations of its components, which are empirically extracted.

### 2.4.1 Marginal Hilbert Spectra, Instantaneous Energy and Degree of Stationarity

The Hilbert spectrum  $H(\omega, t)$ , can be integrated in time to obtain the so called marginal Hilbert spectrum

$$h(\omega) = \int_0^T H(\omega, t) dt. \quad (2.22)$$

It gives the total amplitude contribution at each frequency, as accumulated during the entire data length.

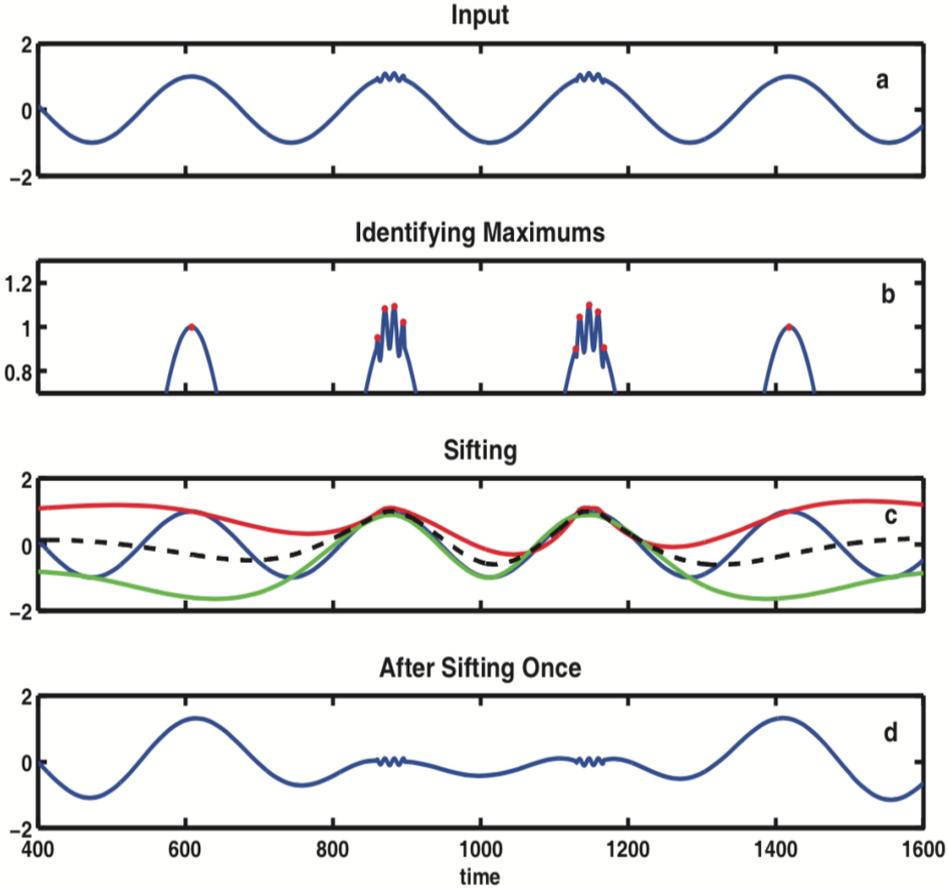


Figure 2.2: a) Data to be decomposed by EMD. b) Extrema identification step. c) Subtraction of envelope mean. d) IMF affected by mode mixing. It can be seen that oscillations of widely different scales are present in the same IMF. Figure taken from [17]

Integrating the Hilbert spectrum over the frequency domain instead, provides the instantaneous Energy [26]

$$IE(t) = \int_{\omega} H(\omega, t) d\omega. \quad (2.23)$$

giving informations on time variations of the energy. Finally, the degree of stationarity  $DS(\omega)$  is defined as

$$DS(\omega) = \frac{1}{T} \int_0^T \left(1 - \frac{H(\omega, t)}{n(\omega)}\right) dt, \quad (2.24)$$

where  $n(\omega) = h(\omega)/T$  is the mean marginal spectrum. The Hilbert spectrum of a stationary process doesn't depend on time, it contains only horizontal lines and  $DS(\omega)$  is in this case zero and the marginal Hilbert spectrum corresponds to the Fourier spectrum [26].

## 2.5 Mode mixing and noise assisted methods

Performances of EMD are diminished when applied to noisy data, due to the fact that the algorithm is sensitive to fluctuations in both extrema locations and magnitude. Intermittent changes due to noise can give rise to **mode mixing**, which is defined either as a single IMF consisting of waves of widely disparate scales or as a signal of similar scales residing in different IMFs [18]. An example of mode mixing due to intermittent noise can be seen in Figure 2.2. To overcome this problem, noise assisted methodologies were developed, ensemble empirical mode decomposition (EEMD) being one of them. When applying EEMD to data  $x(n)$ , an ensemble  $x_i(n)$  is obtained adding different realisations of finite amplitude white noises  $w_i(n)$  to the target data

$$x(n)_i = x(n) + w_i(n), \quad (2.25)$$

with  $i = 1 \dots M$  number of trials. EMD is then applied to each one of such realisations  $x_i(n)$ , and a collection of IMFs are obtained. The ensemble means of the corresponding IMFs are considered to be the true oscillatory modes of the data. In the output of EEMD the contribution from the added uncorrelated white noises is averaged out after a large number of trials  $M$ . A possible drawback of this approach is that, for each trial  $i = 1 \dots M$ , the obtained total number of IMFs,  $K_i$ , can vary slightly. This can lead to noisy outputs, since IMFs of different scales are being averaged. Furthermore, the number of trials and the standard deviation of noise, which has to be of finite amplitude, needs to be defined a priori. Beside these drawbacks though, EEMD usually helps mitigate mode mixing and can provide a meaningful decomposition of noisy datasets, hence being an improvement of standard EMD.

Another example of noise assisted methodology is the Complete EEMD with adaptive noise (CEEMDAN)[27, 28], in which realisations of white noise are added at each step of the sifting process. The obtained average IMF is then subtracted at each step, before continuing with the decomposition.

## 2.6 Statistical significance test of IMFs

The statistical characteristics of Gaussian white noise have been studied intensively using the EMD algorithm [17]. It is found that IMFs of white noise follows normal distributions, and hence their energy follows a  $\chi^2$  distribution. This result have been used to test statistical significance of the extracted IMFs against white noise. This approach can be used to determine whether or not an IMFs contained in the analysed data is possibly due to white noise. For sake of clarity and only in this section the notation of [17] is adopted, and  $j = 1 \dots N$  refers to the number of data samples while  $n$  identifies the different IMFs. The mean energy density of the nth IMF is defined by

$$\bar{E}_n = \frac{1}{N} \sum_{j=1}^N [c_n(j)]^2. \quad (2.26)$$

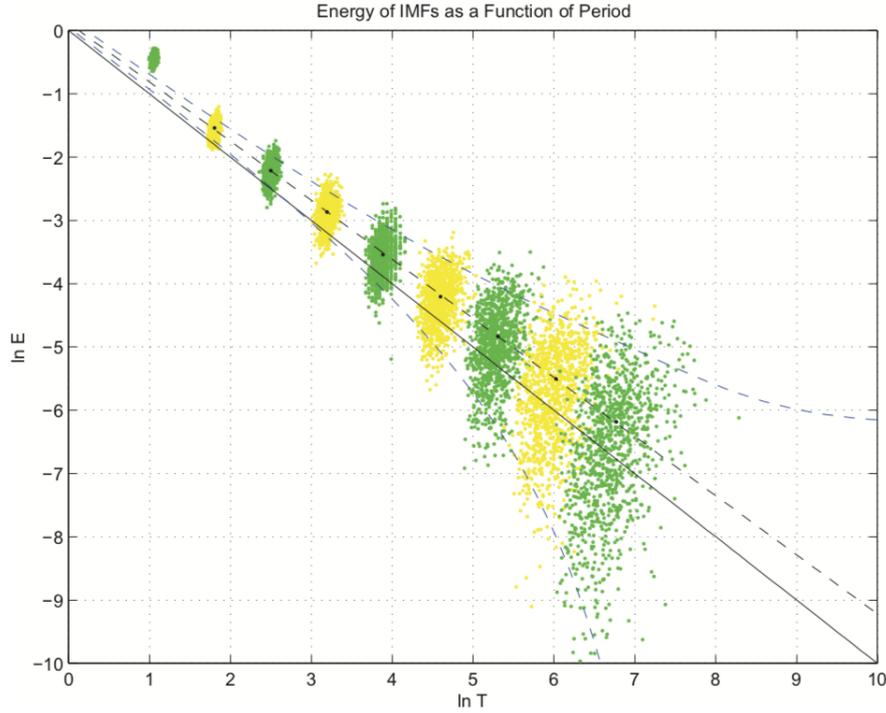


Figure 2.3: Relationship between the mean energy density of IMFs obtained decomposing white noise and their corresponding mean period (black dots). Also shown are the spread lines of energy density. Figure from [17, 18].

Due to the two requirements introduced earlier, IMF mean period  $\bar{T}_n$ , can be estimated either with the extrema counting method or making use of the IMF Fourier spectrum, referred to as spectrum weighted method. Empirically it is found that, for white noise [17]

$$\ln \bar{E}_n + \ln \bar{T}_n = 0. \quad (2.27)$$

Once combined, equations (2.26)(2.27) allows to obtain the spread lines of the energy density of white noise IMFs

$$y = -x \pm k\sqrt{2/N}e^{x/2}, \quad (2.28)$$

where  $k$  is related to the percentiles of a standard normal distribution and  $y = \ln \bar{E}_n$  and  $x = \ln \bar{T}_n$ . The spread lines for confidence against white noise of Equation 2.28 are shown in Figure 2.3, obtained using the Monte-Carlo method analysing a large dataset of synthetically generated long Gaussian white noises [17].

Following this approach, in order for an IMF to be considered a signal against white noise, with a given confidence level, it has to lie outside the thresholds defined by the spread lines for white noise.

## 2.7 Time varying filter EMD (tvf-EMD)

In this section the time varying filter EMD (tvf-EMD) algorithm, used for the data analysis carried out in this Thesis work, is described based on [6] and references therein.

The output of tvf-EMD is a collection of narrow band oscillatory modes, which for clarity are hereafter referred to as IMFs, even if the aforementioned requirements are not strictly met (i.e. equal numbers of extrema and zero crossings, zero envelope mean). The tvf-EMD algorithm makes use of B-splines as a filter with time varying frequency cut off to be able to deal with nonstationary data.

Furthermore, due to a frequency realignment step, the algorithm is able to deal with both mode mixing and intermittency. It has a higher frequency resolution compared to standard EMD and it is able to separate oscillatory modes with frequency ratio up to 0.9 [6].

During the sifting process of tvf-EMD, the analysed data  $x(t)$ , where  $t = 1 \dots N$ , are separated into a local high frequency (LHF) and a local low frequency (LLF) component. The iterative application of the algorithm ensures that the obtained LHF signal to be extracted is local narrow band. To that purpose a threshold criteria based on the instantaneous bandwidth is employed to stop the sifting procedure, once it is achieved that the extracted IMF is local narrow band. In the tvf-EMD algorithm the envelope mean need not to be zero, and the correspondent requirement is relaxed.

In the remaining of this section, the following features of the tvf-EMD algorithm are summarised, based on [6] and references therein

- B-splines approximation of a signal in Section 2.7.1
- Estimation of local frequency cutoff in Section 2.7.2
- Frequency realignment step in Section 2.7.3
- Sifting procedure of the tvf-EMD method in Section 2.7.4

### 2.7.1 B-splines approximation of a signal

B-splines are piecewise polynomials which, once joined together, approximate the desired signal. The joining points of the polynomial segments are referred to as knots. The process of B-spline approximation with equally spaced knots is called uniform B-spline approximation and it consist of constructing polynomial splines that approximate the input signal. Mathematical definition and some properties of spline functions are given in Appendix E, based on [29] [30] [31]. Considering a B-spline function  $\beta^n(t)$  of order  $n$  having step size of the knot sequence  $m$ , a signal in this B-spline space is defined by [6][29].

$$g_n^m(t) = \sum_{k=-\infty}^{\infty} c(k)\beta^n(t/m - k), \quad (2.29)$$

where the  $c(k)$  are the B-splines coefficients. The signal, or approximation result, is determined by the order of the spline  $n$ , its knot distribution  $m$  and the coefficients

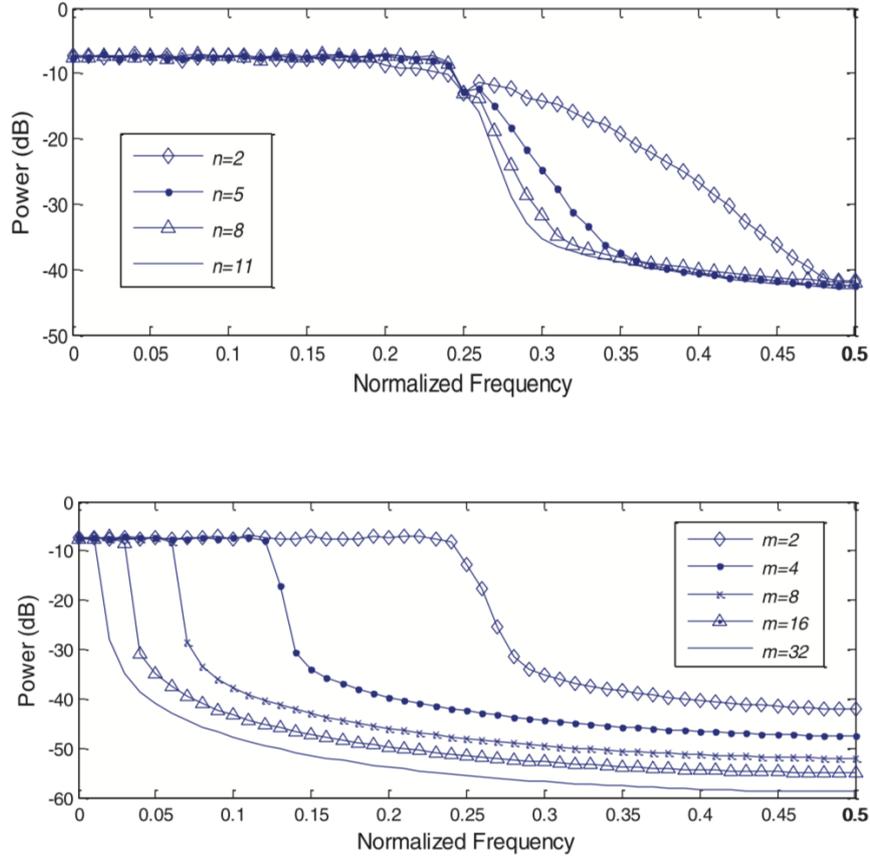


Figure 2.4: Frequency responses of B-spline approximation for various values of  $m$  and  $n$ . Top:  $m = 2$ ,  $n = 2, 5, 8, 11$ . Bottom:  $m = 2, 4, 8, 16, 32$ ,  $n=14$ . Figure taken from [6].

$c(k)$ . Given  $n$  and  $m$ , B-spline approximation aims at obtaining the coefficients  $c(k)$  which minimize an approximation error.

As described in [6], B-spline approximation is a form of low pass filtering. Figure 2.4 shows the frequency response of B-spline based filters of different orders and knot spacing, respectively. In Figure 2.4, top panel, it is shown the frequency response for  $m = 2$  and different spline orders  $n = 2, 5, 8, 11$ . The filter approaches an ideal low pass filter when the spline order  $n$  increases. Figure 2.4, bottom panel, shows instead the frequency response for various knot spacing values  $m = 2, 4, 8, 16, 32$ , and spline order  $n = 14$ . The cut-off frequency of B-spline based filters is  $1/2m$ , function of the knot spacing  $m$  [6].

## 2.7.2 Estimation of local frequency cutoff

In [6], an expression for the local cut-off frequency  $\phi_{bis}$  is derived analytically making use of the notions of instantaneous amplitude IA and frequency IF, which have been introduced in previous sections. The starting point to obtain the local frequency cut-off is the general expression for a multicomponent signal, comprised

of  $N$  narrow-band components

$$x(t) = \sum_{m=1}^N a_m(t) \cos(\phi_m(t)), \quad (2.30)$$

where  $a_m(t)$  is positive and  $\phi_m(t)$  is non-decreasing, and both  $a_m(t)$  and  $\phi_m(t)$  are slowly varying [17]. Following [32], a signal

$$z(t) = A(t)e^{j\phi(t)} \quad (2.31)$$

can be defined as **narrow-band** if  $A(t)$  is band limited signal and its highest frequency is far less than  $\phi'(t)$ . For a band-limited multicomponent signal  $x(t)$ , its instantaneous amplitude  $A(t)$  and frequency  $\phi'(t)$  can be obtained [32]

$$A^2(t) = \sum_{m=1}^N \sum_{n=1}^N a_m(t)a_n(t) \cos[\phi_m(t) - \phi_n(t)] \quad (2.32)$$

$$\phi'(t) = \frac{1}{A^2(t)} \sum_{m=1}^N (\phi'_m(t) \sum_{n=1}^N a_m(t)a_n(t) \cos[\phi_m(t) - \phi_n(t)]) + \sum_{m=1}^N \sum_{n=1}^N \frac{a'_m(t)a_n(t)}{A^2(t)} \sin[\phi_m(t) - \phi_n(t)] \quad (2.33)$$

Since  $x(t)$  is a multi-component signal, the previous relations may not be meaningfully used to obtain its instantaneous amplitude and frequency. In the tvf-EMD algorithm, the previous relations are used for cut-off frequency estimation. The main steps for obtaining such quantity are hereafter summarised based on [6]. A multicomponent signal can be expressed as a combination of two signals as

$$z(t) = A(t)e^{i\phi(t)} = a_1(t)e^{i\phi_1(t)} + a_2(t)e^{i\phi_2(t)} \quad (2.34)$$

If the signal is local narrow-band, both  $a_1(t)$  and  $a_2(t)$  vary slowly, and so do  $\phi'_1(t)$  and  $\phi'_2(t)$ . For a two component signal ( $N = 2$ ) the previous relations are [6][32]

$$A^2(t) = a_1^2(t) + a_2^2(t) + 2a_1(t)a_2(t) \cos[\phi_1(t) - \phi_2(t)] \quad (2.35)$$

and

$$\phi'(t) = \frac{1}{A^2(t)} \left( \phi'_1(t)(a_1^2(t) + a_1(t)a_2(t) \cos[\phi_1(t) - \phi_2(t)]) + \phi'_2(t)(a_2^2(t) + a_1(t)a_2(t) \cos[\phi_1(t) - \phi_2(t)]) \right) + \frac{1}{A^2(t)} \left( a'_1(t)a_2(t) \sin[\phi_1(t) - \phi_2(t)] - a'_2(t)a_1(t) \sin[\phi_1(t) - \phi_2(t)] \right)$$

Supposing that both  $a_1(t)$  and  $a_2(t)$  vary much slower than  $\cos[\phi_1(t) - \phi_2(t)]$ , according to the previous relation the local extrema of  $A(t)$  is approximately determined by the local extrema of  $\cos[\phi_1(t) - \phi_2(t)]$ . If a local minimum or maxima of  $A(t)$  is obtained at  $t_{min}$  or at  $t_{max}$ , respectively, and defining

$$\beta_1(t) = |a_1(t) - a_2(t)| \quad (2.36)$$

$$\beta_2(t) = a_1(t) + a_2(t) \quad (2.37)$$

From the previous relations it follows that

$$\beta_1(t_{min}) = A(t_{min}) = |a_1(t_{min}) - a_2(t_{min})| \quad (2.38)$$

$$\beta_2(t_{max}) = A(t_{max}) = a_1(t_{max}) + a_2(t_{max}) \quad (2.39)$$

Since  $a_1(t)$  and  $a_2(t)$  are both slow varying,  $\beta_1(t)$  and  $\beta_2(t)$  can be estimated by finding the curve going through all the minima and maxima of  $A(t)$ , respectively. After that, assuming  $a_1(t) \geq a_2(t)$ , the two functions  $a_1(t)$  and  $a_2(t)$  are simply obtained by

$$a_1(t) = [\beta_1(t) + \beta_2(t)]/2 \quad (2.40)$$

$$a_2(t) = [\beta_2(t) - \beta_1(t)]/2 \quad (2.41)$$

To obtain the bisecting frequency, also  $\phi_1(t)$  and  $\phi_2(t)$  needs to be estimated. This is achieved defining

$$\eta_1(t) = \phi_1'(t)[a_1^2(t) - a_1(t)a_2(t)] + \phi_2'(t)[a_2^2(t) - a_1(t)a_2(t)] \quad (2.42)$$

and

$$\eta_2(t) = \phi_1'(t)[a_1^2(t) + a_1(t)a_2(t)] + \phi_2'(t)[a_2^2(t) + a_1(t)a_2(t)], \quad (2.43)$$

it holds that

$$\eta_1(t_{min}) = \phi'(t_{min})A^2(t_{min}) \quad (2.44)$$

$$\eta_2(t_{max}) = \phi'(t_{max})A^2(t_{max}) \quad (2.45)$$

Because  $a_1(t)$ ,  $a_2(t)$ ,  $\phi_1'(t)$  and  $\phi_2'(t)$  are slow varying components,  $\eta_1(t)$  can be estimated through interpolation of a set of point given by  $\phi(t_{min})'A^2(t_{min})$ , while similarly  $\eta_2(t)$  can be estimated by interpolation of the set of points given by  $\phi(t_{max})'A^2(t_{max})$ .

Then  $\phi_1'(t)$  and  $\phi_2'(t)$  are obtained by solving the previous relations, and have the following expression

$$\phi_1'(t) = \frac{\eta_1(t)}{2a_1^2(t) - 2a_1(t)a_2(t)} + \frac{\eta_2(t)}{2a_1^2(t) + 2a_1(t)a_2(t)} \quad (2.46)$$

$$\phi_2'(t) = \frac{\eta_1(t)}{2a_2^2(t) - 2a_1(t)a_2(t)} + \frac{\eta_2(t)}{2a_2^2(t) + 2a_1(t)a_2(t)} \quad (2.47)$$

The local cutoff frequency is then simply obtained by

$$\phi'_{bis} = \frac{\phi_1'(t) + \phi_2'(t)}{2} = \frac{\eta_2(t) - \eta_1(t)}{4a_1(t)a_2(t)} \quad (2.48)$$

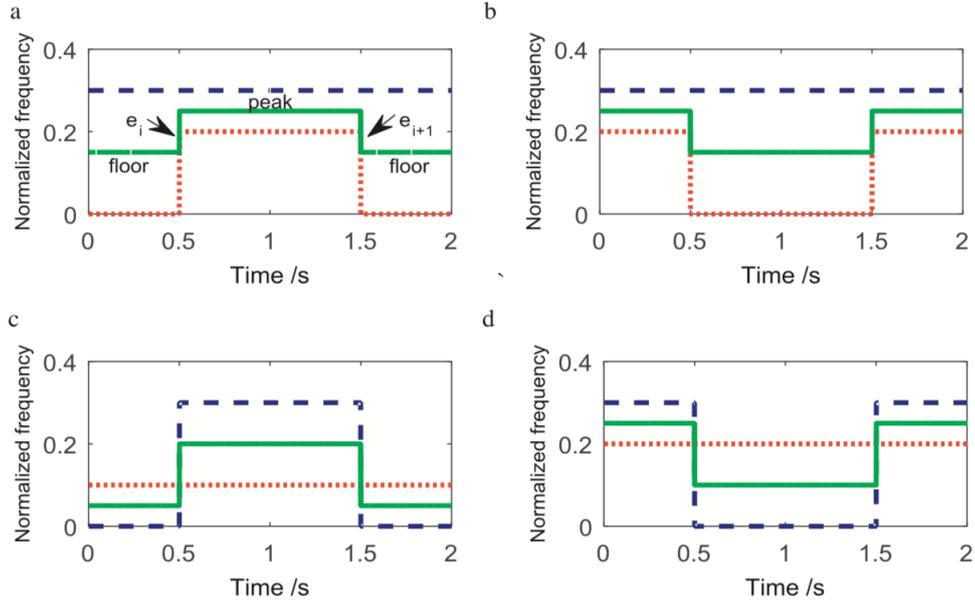


Figure 2.5: Examples of mode mixing in a two-component model. The dotted lines correspond to the LLF component, the dashed lines correspond to the LHF component and the solid lines correspond to the local cut-off frequency. In (a) and (b) the two components can be successfully separated while in (c) and (d) mode mixing occurs because the LLF component is segmented by the local cut-off frequency.

The bisecting frequency doesn't change in case  $a_1(t) \leq a_2(t)$ . Following the notation of [6], given the bisecting frequency  $\phi'_{bis}(t)$ , a signal  $h(t)$  can be retrieved by

$$h(t) = \cos \left[ \int \phi'_{bis}(t) dt \right] \quad (2.49)$$

In the tvf-EMD algorithm, the B-spline approximation is employed as a time varying filter, having a time varying frequency cutoff, where the extrema timings of  $h(t)$  are taken as knots. The approximate result is denoted as  $m(t)$

### 2.7.3 The intermittency problem: frequency realignment to mitigate mode mixing

The estimate of the cutoff frequency is affected by the possible occurrence of intermittence or noise. When this occur, it's value changes abruptly and it needs to be realigned. In the tvf-EMD algorithm this is achieved first determining where intermittency occurs and the cut-off frequency changes abruptly, then setting a threshold on its changing rate within a certain time span, chosen as the time interval between two adjacent maxima. The frequency realignment step of the tvf-EMD algorithm is now briefly described based on [6], Algorithm 2.

First of all, in the time span where intermittency occurs the floor and peak of the frequency cutoff need to be determined: If  $\phi'_{bis}(t)$  is increasing within the time span, it is on its rising edge, then  $\phi'_{bis}(t)$  to the left of the intermittence is consid-

ered to be a floor. Conversely, if  $\phi'_{bis}(t)$  is decreasing within the time span, it is on its falling edge, then  $\phi'_{bis}(t)$  to the right of the intermittence is considered to be a floor. This can be seen in Figure 2.5. The identification of a floor is achieved by examining the neighboring intermittences. The frequency changing rate between two adjacent maxima is limited by a thresholding factor  $\rho$ . More details can be found in [6], Algorithm2.

## 2.7.4 Sifting of tvf-EMD: Stopping criterion based on instantaneous bandwidth and local narrow-band signal extraction

In the following, both the stopping criterion of the tvf-EMD algorithm and the sifting procedure to extract local narrow-band oscillatory modes are described based on [6]. The stopping criterion is based on the concept of instantaneous bandwidth, introduced in a previous section.

The instantaneous bandwidth was firstly defined by Cohen for an analytic signal of the kind  $z(t) = A(t)e^{i\phi(t)}$  as [33, 34]

$$IB_{Cohen} = |A(t)' / A(t)| \quad (2.50)$$

Based on  $IB_{Cohen}$ , a signal is narrow-band if its instantaneous amplitude varies very slowly. In order to overcome the problem of the presence of oscillatory terms present in  $IB_{Cohen}$  for a two tone signal of constant frequencies, an alternative expression for instantaneous bandwidth was obtained in [35] as the standard deviation in the weighted average instantaneous frequency (WAIF) [36]  $\phi_{avg}$  at a given time, which for a two component signal is defined as a weighted average of the instantaneous frequencies of the individual components

$$\phi_{avg} = \frac{a_1^2(t)\phi_1'(t) + a_2^2\phi_2'(t)}{a_1^2 + a_2^2} \quad (2.51)$$

Then, Loughlin IB for a two component signal is defined as

$$IB_{Loughlin}(t) = \sqrt{\frac{a_1'^2(t) + a_2'^2(t)}{a_1^2(t) + a_2^2(t)} + \frac{a_1^2(t)a_2^2(t)(\phi_1'(t) - \phi_2'(t))^2}{(a_1^2(t) + a_2^2(t))2}} \quad (2.52)$$

In the previous expression, if  $|\phi_1'(t) - \phi_2'(t)|$  increases, both the separation between the two components and Loughlin instantaneous bandwidth increases. Also if  $|a_1'(t)|$  or  $|a_2'(t)|$  increase, the amplitude modulation is more pronounced and the instantaneous bandwidth again increases. Loughlin bandwidth quantifies how two components are separated in frequency. A signal can hence be defined as local narrow band based on the Loughlin's definition of instantaneous bandwidth. Since  $IB_{Loughlin}(t)$  is an absolute value given in rad/s, to quantify how much it deviates from the WAIF defined earlier, in [6] a relative criteria is introduced

$$\theta(t) = \frac{IB_{Loughlin}(t)}{\phi_{avg}} \quad (2.53)$$

Given a threshold  $\xi$ , If an oscillatory function  $x(t)$  has  $\theta(t) < \xi$  is referred to as local narrow band, i.e. is already an IMF. Otherwise it is set  $x(t) = x(t) - m(t)$  and the procedure described in previous sections is repeated.

In summary, in the tvf-EMD algorithm the definition of Intrinsic Mode function is replaced by the one of local narrow band signal. For sake of clarity, in the remaining of this Thesis the terms IMFs and local narrow band signal are both referred to the output of the tvf-EMD algorithm.

To summarise, the sifting process of the tvf-EMD algorithm is carried out in three steps [6]

- Estimate the bisecting frequency (local cut-off frequency)  $\phi_{bis}(t)$ .
- Filter the input signal using tvf to obtain the local mean  $m(t)$ .
- Check if the residual signal satisfies the stopping criterion

The complete description of the sifting method of the tvf-EMD algorithm can be found in [6], Algorithm 3, and is here reported for sake of clarity

- 1) Calculate the instantaneous amplitude  $A(t)$  and instantaneous frequency  $\phi'(t)$  of  $x(t)$  using Hilbert transform
- 2) Locate the local minima and maxima of  $A(t)$ , denoted as  $\{t_{min}\}$  and  $\{t_{max}\}$  respectively.
- 3) Interpolate the set of points  $A(\{t_{min}\})$  to obtain  $\beta_1(t)$ . In the same manner, interpolate  $A(\{t_{max}\})$  to obtain  $\beta_2(t)$ . Compute  $a_1(t)$  and  $a_2(t)$ .
- 4) Interpolate  $\phi'(t_{min})A^2(t_{min})$  and  $\phi'(t_{max})A^2(t_{max})$  to obtain  $\eta_1(t)$  and  $\eta_2(t)$ . Compute  $\phi'_1(t)$  and  $\phi'_2(t)$ .
- 5) Calculate the local cut-off frequency  $\phi'_{bis}(t) = [\phi'_1(t) + \phi'_2(t)]/2$ .
- 6) Realign  $\phi'_{bis}(t)$  to deal with the intermittence problem (Algorithm 2 of [6]).
- 7) Compute  $h(t)$  according to 2.49. Then apply B-spline approximation filter on  $x(t)$ , by taking the extrema timings of  $h(t)$  as knots, i.e.,  $\{t_{min}\}$  and  $\{t_{max}\}$ . The approximate result is denoted as  $m(t)$ .
- 8) Calculate the stopping criterion  $\theta(t)$ . If  $\theta(t) \leq \xi$ ,  $x(t)$  is taken to be an IMF, else  $x(t) = x(t) - m(t)$  and repeat steps 1-7.

In the tvf-EMD algorithm there are two parameters, i.e., the bandwidth threshold  $\xi$  to be used in the stopping criterion, and the B-spline order  $n$ , which determines the time varying filter frequency roll-off.

The bandwidth threshold parameter  $\xi$  is used to determine whether the input signal should be filtered further. It also determines the separation performance. Notably the parameter  $n$  is entirely unrelated to the cut-off frequency estimation. It is assumed that  $A(t)$  have a succession of local minima and maxima. If  $A(t)$  does not contain enough minima and maxima to continue interpolation, it is slowly varying. The input signal is hence considered to be already local narrow-band and is not further decomposed. A comprehensive description of the tvf-EMD algorithm can be found in [6] and references therein.

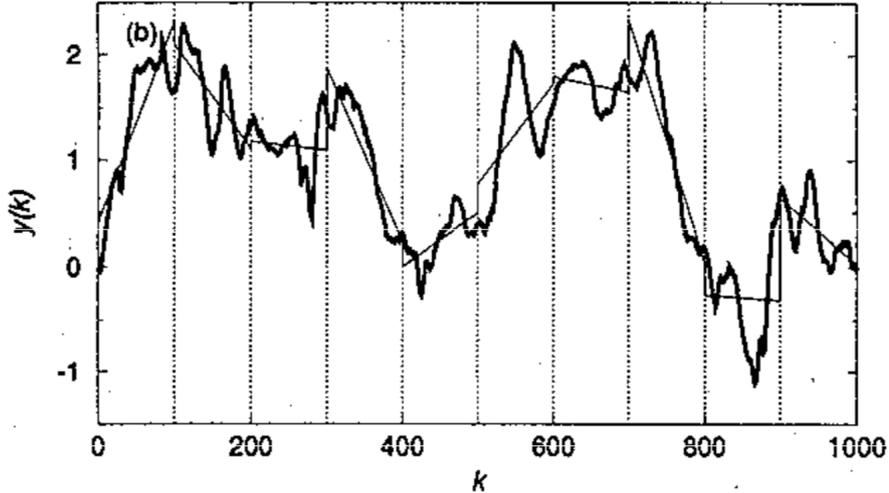


Figure 2.6: In DFA the integrated time series  $y(k)$  is divided in windows of length  $n$ , it is detrended and then the average root mean square computed for windows of different scales. Figure from [37].

## 2.8 Fractal Algorithms: Detrended Fluctuation Analysis and Local Hurst Exponent

The DFA method is here described based on [37]. It has applications in many fields of time series analysis, e.g. geophysics physiology, and economics. It aims at characterising the self similarity behaviour of a given time series. This is achieved by computing fluctuations  $F(n)$  in the data as function of different timescales  $n$ . Figure 2.6 shows the first step of the DFA algorithm, the integrated time series is divided in windows of length  $n$ , linear detrending is performed subtracting the trend  $y_n(k)$  in each window and the root mean square is evaluated by means of the following relation

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (2.54)$$

If the investigated process is self similar, such fluctuations are found to scale in a power law fashion with the length of the time window considered, the exponent being the DFA coefficient  $H$ , or Hurst exponent

$$F(n) \propto n^H \quad (2.55)$$

For uncorrelated white noise  $H = 0.5$  while persistency (antipersistency) is present in the data if  $H > 0.5$  ( $H < 0.5$ ), respectively. Pink noise has  $H = 1$  while for  $H > 1$  the data are unbounded.  $H = 1.5$  corresponds to Brownian Noise.

If the analysed time series can be described by one Hurst exponent, constant over the whole time span, it is referred to as monofractal. If instead the time series is a multifractal, is characterised by local fluctuations with both small and large magnitudes [5]. A multifractal time series can be described by more than one Hurst exponent, which will be different at different time intervals. To highlight this feature,

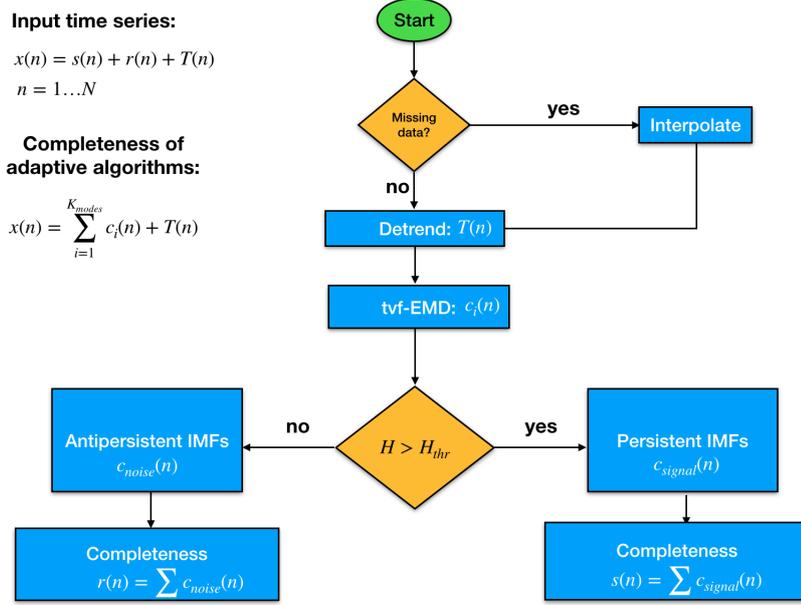


Figure 2.7: Flowchart illustrating the different steps of the tool for adaptive time series analysis used in this Thesis work.

the Hurst exponent can be computed locally, sliding windows of small size, with a procedure similar to DFA. This is described in [5], where a tutorial is also provided in MATLAB. Some applications of the local Hurst exponent for the description of persistency in radionuclide time series can be found in [38, 39, 40, 41, 42], while characterisation of outliers occurrence in radionuclide time series in [43, 44].

## 2.9 Tool for Adaptive time series analysis based on tvf-EMD algorithm

The tool developed for the data analysis of this Thesis work is hereafter described, based on the concepts outlined in previous sections. Figure 2.7 shows a flowchart of the adopted methodology. Once applied to a time series  $x(n)$ , with  $n = 1 \dots N$ , it characterise it according to equation (2.1). The steps of the adopted methodology are the following

- Initially, data are checked for missing values. If missing values are present, interpolation is carried out.
- Standard EMD is then performed and the last term so obtained is a slowly varying function  $T(n)$ , with  $n = 1 \dots N$  number of data samples. It can either be the adaptive trend or the baseline wandering affecting the data, if present. It is subtracted from the data before further analysis.
- Then, tvf-EMD algorithm is performed on  $x(n)$  and a set of narrow-band modes of oscillation,  $c_i(n)$  embedded in the data is obtained.

- The denoising step is then carried out. The set of modes  $c_i(n)$  is firstly classified based on persistent behaviour, calculating their Hurst exponent  $H_{c_i(n)}$  by means of DFA method. The denoising step is carried out making use of a thresholding approach [19]. For this reason a threshold Hurst exponent  $H_{thr}$  is chosen. The thresholding based denoising is then carried out as follows
  - If  $H_{c_i(n)} < H_{thr}$ , noise related IMF
  - If  $H_{c_i(n)} > H_{thr}$ , signal related IMF
- After the denoising step, persistent IMFs are summed up and they constitute the signal  $s(n)$ , which can be characterised in term of amplitude and frequency modulations of its main modes of oscillation, e.g. the persistent IMFs, through the Hilbert Huang Transform  $H(\omega, t)$ .
- Summing up the remaining IMFs gives the residuals  $r(n)$ , which instead can be characterised in term of outliers occurrence.
- Following the approach of [19], the following parameters can be employed for denoising performance evaluation
  - **Mean squared error:**  $MSE = \frac{\sum_{n=1}^N (x(n) - s(n))^2}{N}$
  - **Mean absolute error:**  $MAE = \frac{\sum_{n=1}^N |x(n) - s(n)|}{N}$
  - **Signal to noise ratio:**  $SNR = 10 \log_{10} \frac{\sum_{n=1}^N x(n)^2}{\sum_{n=1}^N (x(n) - s(n))^2}$
  - **Peak signal to noise ratio:**  $PSNR = 20 \log_{10} \frac{\max(x(n))}{RMSE}$
  - **Crosscorrelation**  $xcorr = \frac{E((s(n) - \mu_{s(n)})(x(n) - \mu_{x(n)}))}{\sigma_{s(n)}\sigma_{x(n)}}$ ,

between  $x(n)$  and  $s(n)$ , where  $x(n)$  is the input data,  $s(n)$  is the extracted signal, RMSE is the square root of MSE and  $n = 1 \dots N$  is the number of data samples. Low MSE, MAE and high SNR, PSNR, xcorr are indicative of good denoising performances.

- The Marginal Hilbert spectra  $h(\omega)$  can also be calculated, based on (2.22), giving information regarding the total amplitude accumulated at each frequency in the analysed time span.
- Furthermore, decomposing with tvf-EMD a white noise time series  $w(n)$  confidence lines against white noise can be obtained as described in 2.5. Significance of the obtained oscillatory modes can then be tested against white noise having standard deviation of the same order of the input data.
- The Hurst exponent can be computed locally, to evaluate persistency in the data, making use of the algorithm described in [5].

Even though in the tvf-EMD algorithm there's no restriction on the number of IMFs that can be extracted, in the framework of this Thesis, the number of IMFs to be extracted have been set to  $K = \log_2 N$  where  $N$  is the data length, since this

is commonly the case when applying standard EMD algorithms. Since the adopted methodology is empirical, results need to be evaluated a posteriori, applying the methodology to data from real physical systems. In the next Chapter, the methodology described in this Section is applied for the characterisation of seismometer and radionuclide time series data.

# Chapter 3

## Results

In this Chapter, results from the application of both adaptive and fractal algorithms, described in Chapter 2, on data from different physical systems are presented. The analysed data are the following

- First, the algorithm described in 2.9 was tested on synthetic GW data corrupted by purple noise and results are presented in section 3.1.
- In 3.2 the methodology described in Section 2.9 was used for adaptive denoising of seismometer data monitoring Virgo NEB during four different acoustic noise injections performed for detector characterisation purposes. Separation of the seismic noise, induced by the acoustic noise injections, from the underlying nonlinear nonstationary data is achieved relying on the adaptive nature of tvf-EMD and on the different persistency of the perturbation with respect to the underlying seismometer data.
- In 3.3 the local Hurst exponent was computed for an array of 38 seismometers to characterise the site response of the Virgo ifo WEB. The analysis is able to distinguish among seismometers placed in different parts of the room, namely among the ones on top of a tower platform and not on it, evaluating time series persistency.
- Results from the characterisation of time series of atmospheric  ${}^7Be$  can be found in 3.4. The methodology described in 2.9 can extract the yearly IMFs of  ${}^7Be$ , which phase shift with latitude can then be interpreted in term of features of large scale atmospheric dynamics like e.g. the seasonal shift of the Hadley cell, as also reported in [7]. Data were sampled by the IMS of the CTBTO.

Other results, i.e. the investigation of noises properties with the Burg Maximum Entropy Method and with adaptive algorithms, results from scattered light noise hunting at Virgo, and results from the noise hunt of the  $1/f^{2.5}$  broadband noise affecting Virgo during the commissioning phase are instead reported in Appendixes A C F and G, respectively.

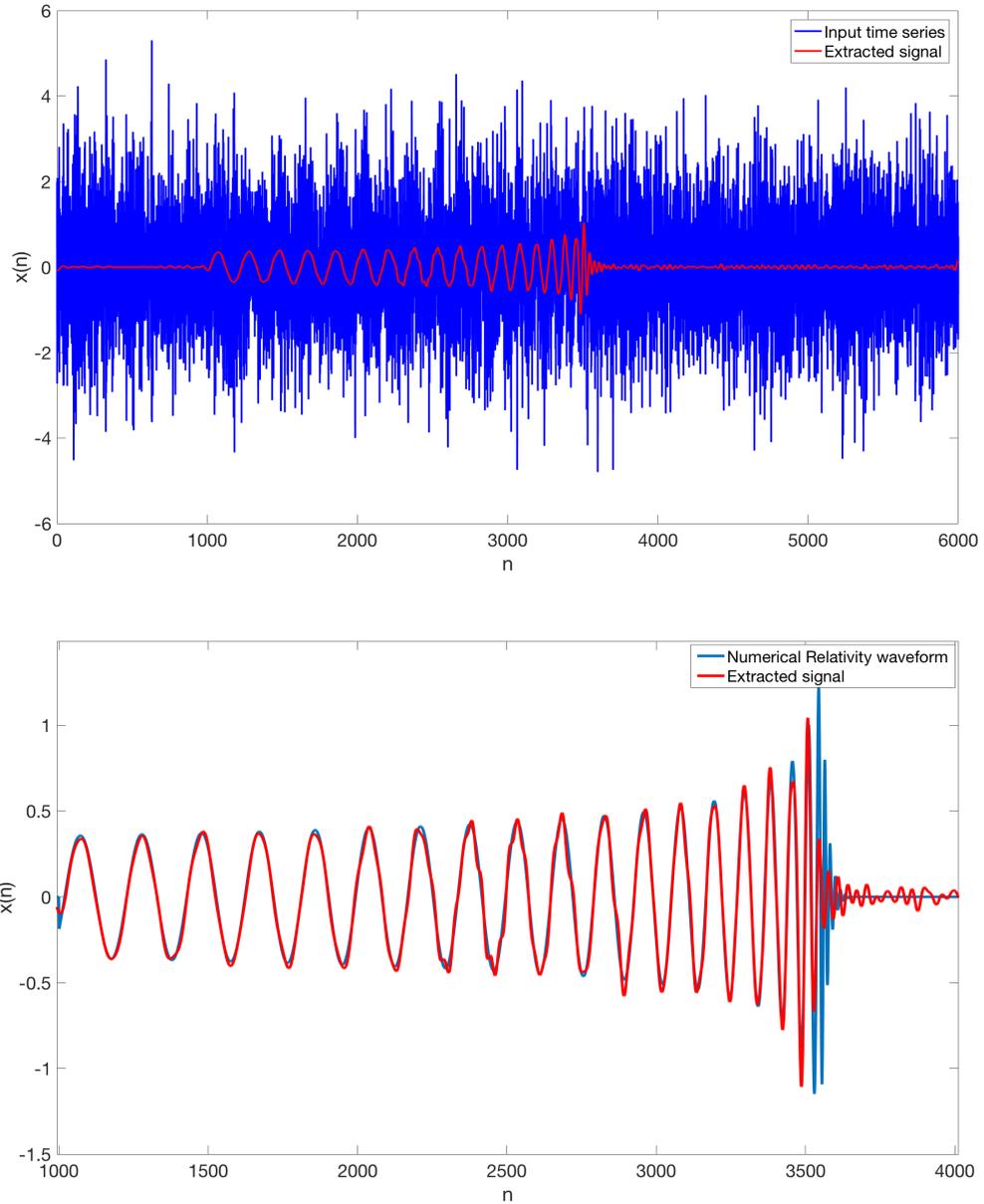


Figure 3.1: Top: NR waveform in purple noise (blue) and waveform extracted with the adaptive algorithm tvf-EMD (red). Parameters used in this case are  $\xi = 0.1$ ,  $n = 35$ ,  $H_{thr} = 0.1$ . Bottom: Comparison of clean NR waveform ( $x(n)$ , in blue) with extracted signal ( $s(n)$  red), after modes of the purple noise  $w_{purple}$  have been classified based on their persistency and subtracted.

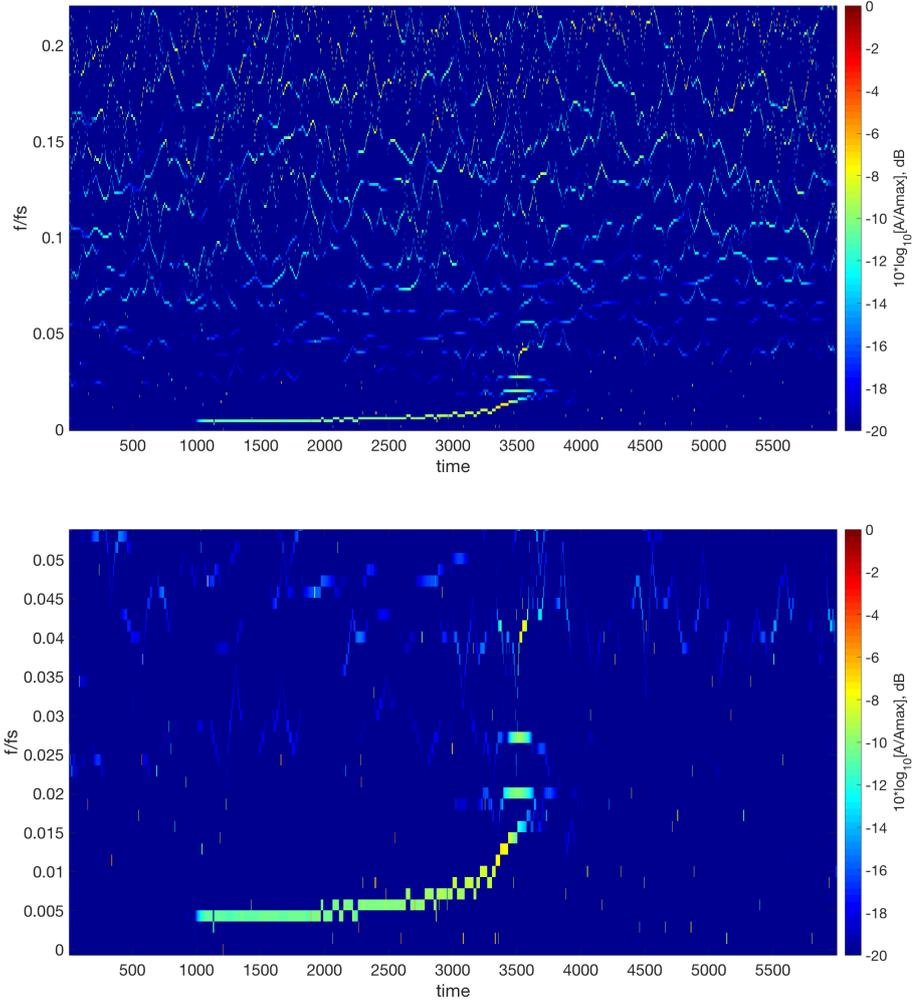


Figure 3.2: Top: HHT of the data, i.e. NR waveform in purple noise. Bottom: HHT zoomed in the frequency range 0 – 1000Hz. The sweep in frequency of the chirp signal is visible. The colorscale gives the amplitude of the oscillatory modes, as described in Equation 3.2.

### 3.1 Testing: NR Waveform in Purple Noise

The algorithm developed in the framework of this Thesis has been firstly tested on a Numerical Relativity (NR) waveform <sup>1</sup>, namely a chirp signal,  $s(n)$  buried in purple noise  $w_{purple}(n)$ .

$$x(n) = s(n) + w_{purple}(n) \quad (3.1)$$

The sampling frequency is  $f_s = 16384Hz$ . In Figure 3.1, top panel, is reported how the chirp waveform  $s(n)$ , in red, is extracted from  $x(n)$  with the adopted methodology, i.e. how it is separated from  $w_{purple}(n)$ . Figure 3.1, bottom panel, compares the original NR waveform,  $x(n)$ , in blue, with the extracted signal  $s(n)$ , in red. Beside some discrepancy toward the last cycles of the chirped signal, the waveform is well reconstructed. This due to the adaptive nature of the tvf-EMD

<sup>1</sup>NR waveform data obtained from <https://www.gw-openscience.org/events/>

algorithm and to the different persistency of the noise compared to the chirp signal. The time-frequency representation of the data is obtained performing HSA on the modes extracted with tvf-EMD, hence obtaining the HHT. Figure 3.2 top panel reports the HHT of the NR waveform in purple noise, while in the lower panel this is zoomed in the frequency band 0 – 1000Hz. The colour scale is normalised to the maximum amplitude obtained, among the oscillatory modes. Is given in logarithmic scale by

$$10\log_{10}A/A_{max}. \quad (3.2)$$

Due to the fact that that the extracted signal is still slightly affected by mode mixing, the IF frequency obtained in the HHT is also affected, showing a slight oscillatory behaviour. The frequency sweep though is well tracked, among with the correspondent amplitude modulation. It should be noted that, the IF being defined for each time, the HHT has higher frequency resolution as opposed to periodograms or wavelet based spectra [17, 18].

Further tests have been carried out with noises of higher spectral indexes, ranging from blue to red ( $\beta = -1$  to  $\beta = 2$ ). Since the denoising step is based on the antipersistency of the noisy IMFs, performances progressively worsen at higher spectral indexes. Hence, the adopted methodology is best applied to strongly nonlinear nonstationary time series corrupted by anticorrelated noise and in general when noise and signal to be extracted exhibit different persistent behavior.

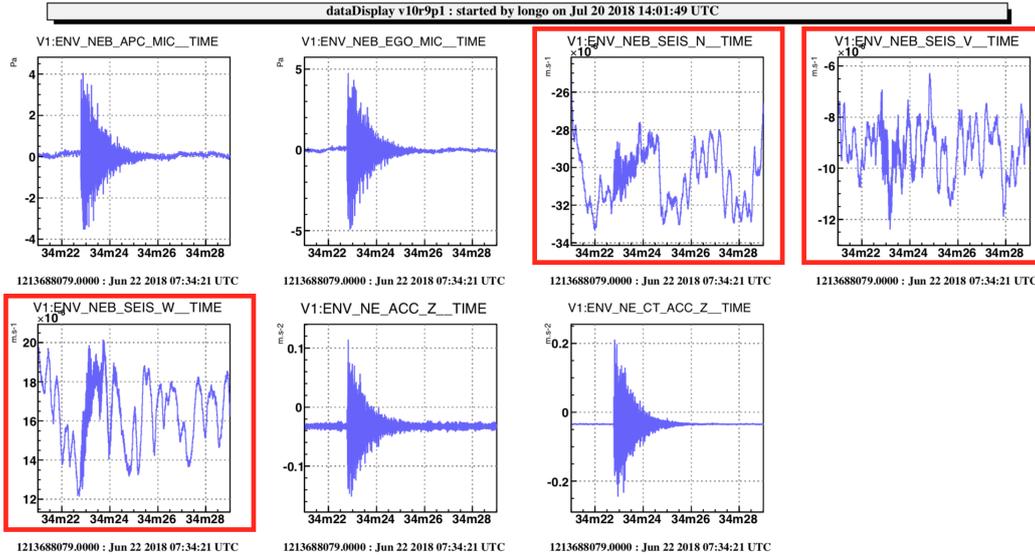


Figure 3.3: Recordings from microphones, seismometer and accelerometers monitoring Virgo NEB during the fourth acoustic noise injection. Highlighted in red are the time series analysed with the methodology described in 2.9. Figure obtained with DataDisplay

## 3.2 Adaptive Denoising of Acoustic Noise Injections

As described in Chapter 1 many sources of noise can affect the sensitivity of ground based interferometers such as Virgo and LIGO in different frequency bands, some relevant examples being thermal noise, quantum noise, seismic noise, Newtonian noise and infrasound atmospheric noise [12, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 8, 55, 56, 57, 1, 58, 3, 59, 2]. In particular, seismic noise can limit interferometers sensitivity and impulsive events such as earthquakes can even lead to unlocks [60]. Hence, to monitor the environment surrounding the detector, many sensors such as microphones, seismometers and accelerometers, deployed both indoor and outdoor the detector facility to monitor natural and anthropogenic sources of noise, continuously sample data in the form of time series.

The methodology developed in the framework of this Thesis, based on the adaptive algorithms EMD [17, 61], tvf-EMD [6] and on fractal analysis [37, 5], and described in Chapter 2 (see Figure 2.7) has been applied to data from a seismometer monitoring the NEB of the Virgo ifo, which is located near Cascina (Pisa), during four different acoustic noise injections performed for detector characterisation purposes. In Figure 3.3 it can be seen the effect of the acoustic perturbation on sensors monitoring NEB, i.e. on microphones, on a seismometer and on accelerometers. It is found that the adopted approach can achieve separation of the seismic perturbation induced by the acoustic noise injection from the underlying nonlinear nonstationary recordings of the seismometer monitoring the NEB. The effect of the acoustic perturbations on the seismometer monitoring NEB have been further quantified computing the local Hurst exponent  $H(t)$  of seismometer data.

For a detailed description of how  $H(t)$  is computed see [5]. Evaluated around the time of the acoustic noise injection,  $H(t)$  tracks changes in the persistency of all components of the triaxial seismometer, as expected. Notably, it is found that the changes in persistency are of the same order of the threshold chosen to achieve denoising, i.e.  $H_{thr}$ . This suggests that fractal analysis could be used to automatically select a threshold value for denoising. Though further research is needed in this regard.

### 3.2.1 Dataset

Analysed data were collected at the NEB of the Virgo ifo site, near Cascina (Pisa) on June 22nd, 2018. The acoustic noise injections were performed for detector characterisation purposes, i.e. to measure the reverberation time of NEB. Measurements were performed with two Brüel & Kjær infrasound microphones model 4193-L-004 [62], that have a flat response down to 0.5 Hz, connected to two amplifiers NEXUS 2690 [63] operating on a bandwidth from 0.1Hz to 100kHz. The two microphones were placed at the two extremities of the NEB. In total, two 14 inches balloons and two firecrackers (Raudi mod. Jupiter, art. 1057) were blew up and each explosion was recorded by the two microphones simultaneously. The balloon explosions were too weak compared to the background noise and only the signals given by the firecrackers were enough high to allow the RT60 estimate [64]. It has been shown in the Virgo logbook entry n. 41940 [65] that the strongest firecracker impulse was also seen by others environmental sensors inside NEB (i.e. accelerometers, seismometers, microphones) which continuously monitor the interferometer environment [47] and are used for the detector noise characterization studies [66].

The data analysed in this Section are sampled by one triaxial broadband seismometer (Guralp 40 T 60 s with flat velocity response from 0.17 Hz to 50 Hz [67]), monitoring along the north/west arm direction (N,W) and in the vertical direction (V), with sampling frequency of  $f_s = 1000$  Hz. The sensor is located on the NEB building concrete slab supporting the test mass suspension and vacuum chamber. The times of the four impulsive events, mentioned above, are:

- First balloon explosion: 7 14' 42" (UTC)
- Second balloon explosion: 7 16' 18" (UTC)
- First firecracker explosion: 7 18' 50" (UTC)
- Second firecracker explosion: 7 34' 21" (UTC)

No missing data were present. A time interval  $t = 8$  s has been chosen for the analysis. i.e. data length is  $N = 8000$ .

### 3.2.2 Results of Adaptive Denoising

Figures 3.4 3.5 3.6 and 3.7 show the obtained results for the denoising of seismometer data during the four different acoustic noise injections. From top to bottom

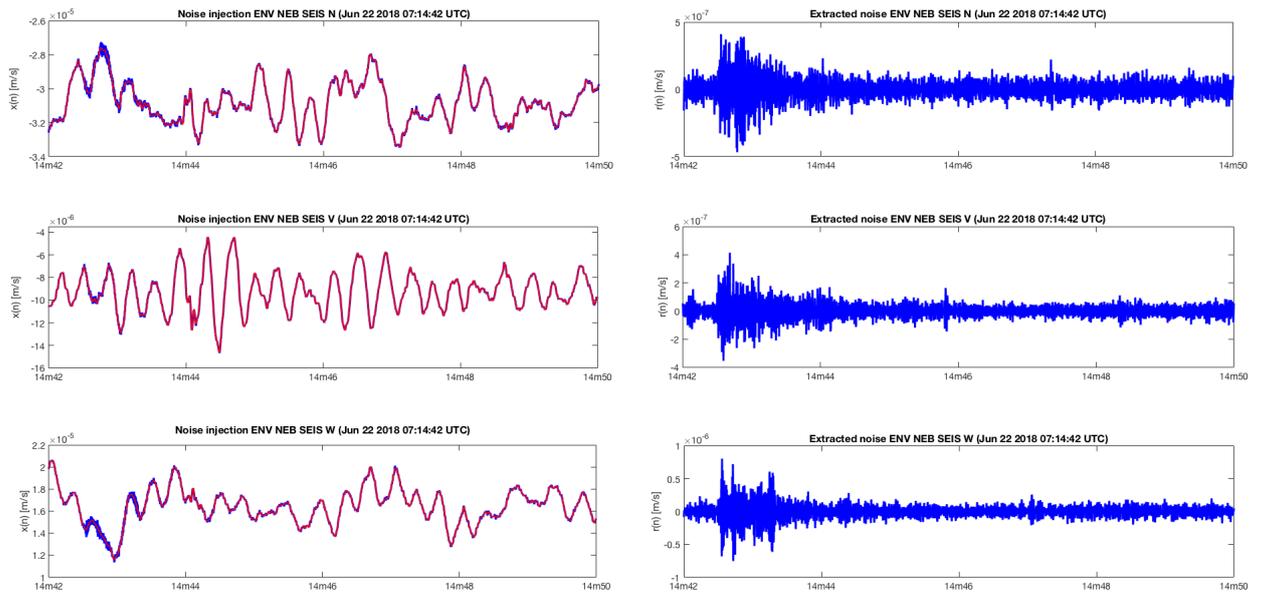


Figure 3.4: First acoustic noise injection: In the left panel, the original seismometer data  $x(n)$  are shown in blue while in red is shown the sum of extracted signal  $s(n)$  and trend terms  $T(n)$ . In the right panel is instead shown the extracted seismic perturbation  $r(n)$ .

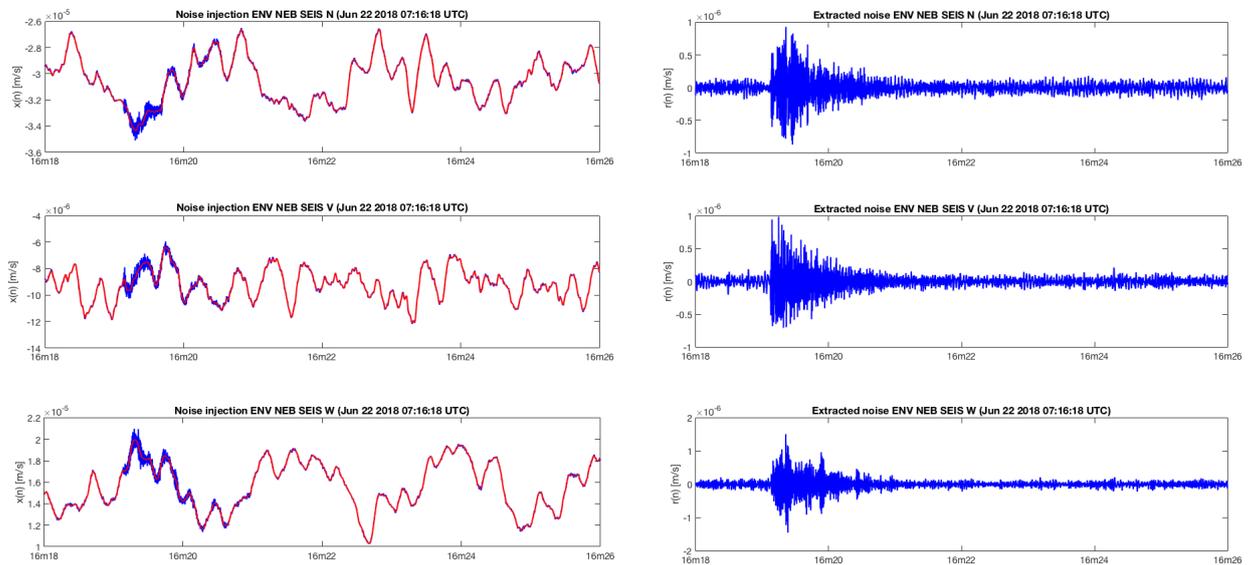


Figure 3.5: Second acoustic noise injection: In the left panel, the original seismometer data  $x(n)$  are shown in blue while in red is shown the sum of extracted signal  $s(n)$  and trend terms  $T(n)$ . In the right panel is instead shown the extracted seismic perturbation  $r(n)$ .

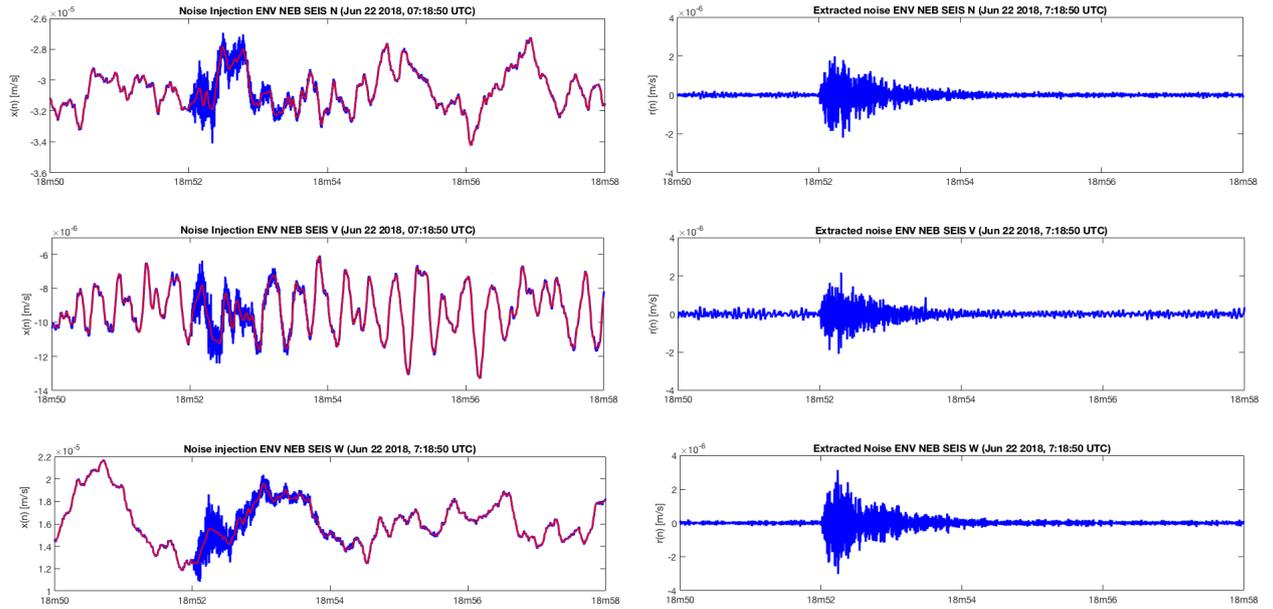


Figure 3.6: Third acoustic noise injection: In the left panel, the original seismometer data  $x(n)$  are shown in blue while in red is shown the sum of extracted signal  $s(n)$  and trend terms  $T(n)$ . In the right panel is instead shown the extracted seismic perturbation  $r(n)$ .

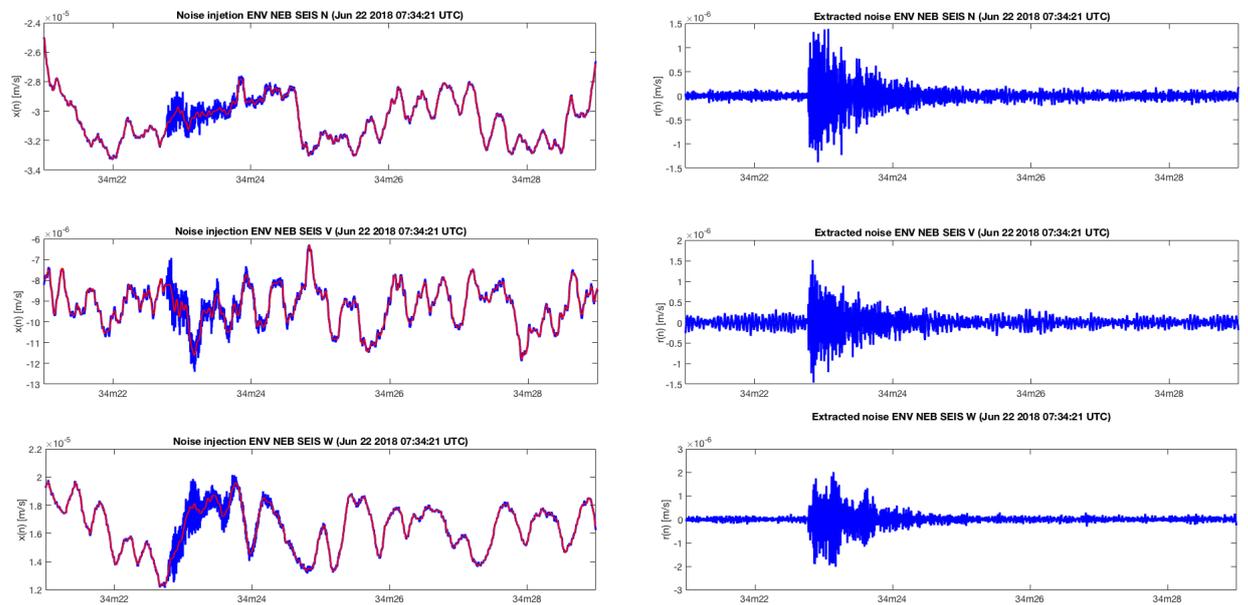


Figure 3.7: Fourth acoustic noise injection: In the left panel, the original seismometer data  $x(n)$  are shown in blue while in red is shown the sum of extracted signal  $s(n)$  and trend terms  $T(n)$ . In the right panel is instead shown the extracted seismic perturbation  $r(n)$ .

are shown results for the three different components of the Triaxial seismometer sampling around the time of the acoustic noise injection. In the left panel, the original seismometer data  $x(n)$ , where  $N = 1 \dots 8000$ , are shown in blue while superimposed in red is shown the sum of extracted signal  $s(n)$  and trend terms  $T(n)$ . In the right panel is instead shown the extracted seismic perturbation  $r(n)$ . The signal  $s(n)$  and noise  $r(n)$  term are the sum of persistent/antipersistent modes, where persistent is relative to the chosen value of  $H_{thr}$ . The parameters used are: bandwidth threshold  $\xi = 0.1$ , B-spline order  $n = 26$ , threshold on modes Hurst exponent  $H_{thr} = 0.3$ . The adopted methodology allows to separate the nonlinear and nonstationary perturbation  $r(n)$ , induced by the acoustic noise injection, from the nonlinear nonstationary seismic signal  $s(n)$ .

### 3.2.3 Local Hurst Exponent of Seismometer Data

From Figures 3.4 3.5 3.6 and 3.7 it can be seen how, in unperturbed conditions, seismometer data  $x(n)$  shows persistent behaviour while antipersistent behaviour is characteristic of perturbed conditions, i.e. during the noise injection. To further investigate and quantify such behaviour, the local Hurst exponent,  $H(t)$ , of  $x(n)$  have been evaluated. A local Hurst exponent  $H(t)$  can be obtained with a procedure similar to DFA, using sliding windows of small size  $\delta$  [5]. Evaluated at windows of suitable size,  $H(t)$  allows to monitor changes of persistency in the seismometer data  $x(n)$ . The scale chosen is  $\delta = 10$ , corresponding to a sliding window of  $\delta t = 10ms$  over which  $H(t)$  is evaluated. Linear detrending have been performed in each window. Obtained results are shown in Figure 3.8 for the three components of the seismometer monitoring NEB during the loudest noise injection, i.e. the third one. It can be seen that local Hurst exponent  $H(t)$ , evaluated using small sliding windows of  $\delta t = 10ms$ , allows to track changes of persistency in the seismometer data. Figure 3.8 shows a change in persistency in all three components N,V,W of the seismometer data, indicative of red noise in non perturbed conditions and of pink noise during the acoustic noise injection, recovering to red noise after the perturbation has passed.

### 3.2.4 Completeness of tvf-EMD in the Frequency Domain

Finally, to characterise the completeness property of tvf-EMD algorithm, i.e. the fact that summing up extracted modes exactly recovers the original data, in the frequency domain, Figure 3.9 and 3.10 shows the Fourier spectra of seismometer recordings (only vertical component shown) during the third noise injection. Superimposed are instead the spectra of extracted modes. It can be seen how the extracted modes accounts for different portions of the Fourier spectra. Since in the tvf-EMD algorithm the maximum number of modes to be obtained and the threshold on their bandwidth can be selected as input, parameters could be tuned to obtain modes explaining a portion of the Fourier spectra in a given band of frequency.

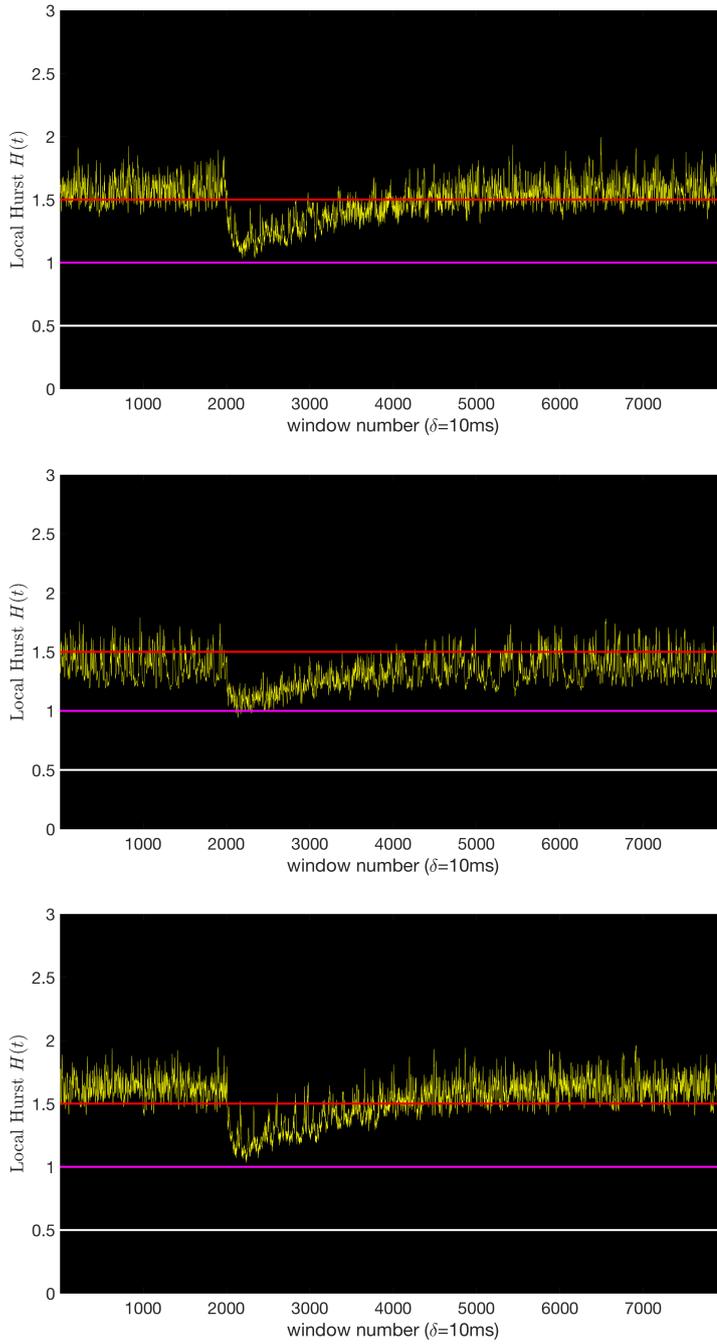


Figure 3.8: Local Hurst exponent  $H(t)$  of the three components of the triaxial seismometer, recording around the loudest acoustic noise injection, i.e. the third one, performed at NEB. From top to bottom are shown results for the N,V and W component respectively. Red pink and white lines represent values of  $H$  typical of red, pink and white noise. Change in persistent behaviour is evident at the time of the injection, where a sudden drop  $\Delta H(t)$  occurs. It is interesting to notice that  $\Delta H(t)$  is of the same order of  $H_{thr}$  selected to achieve denoising.

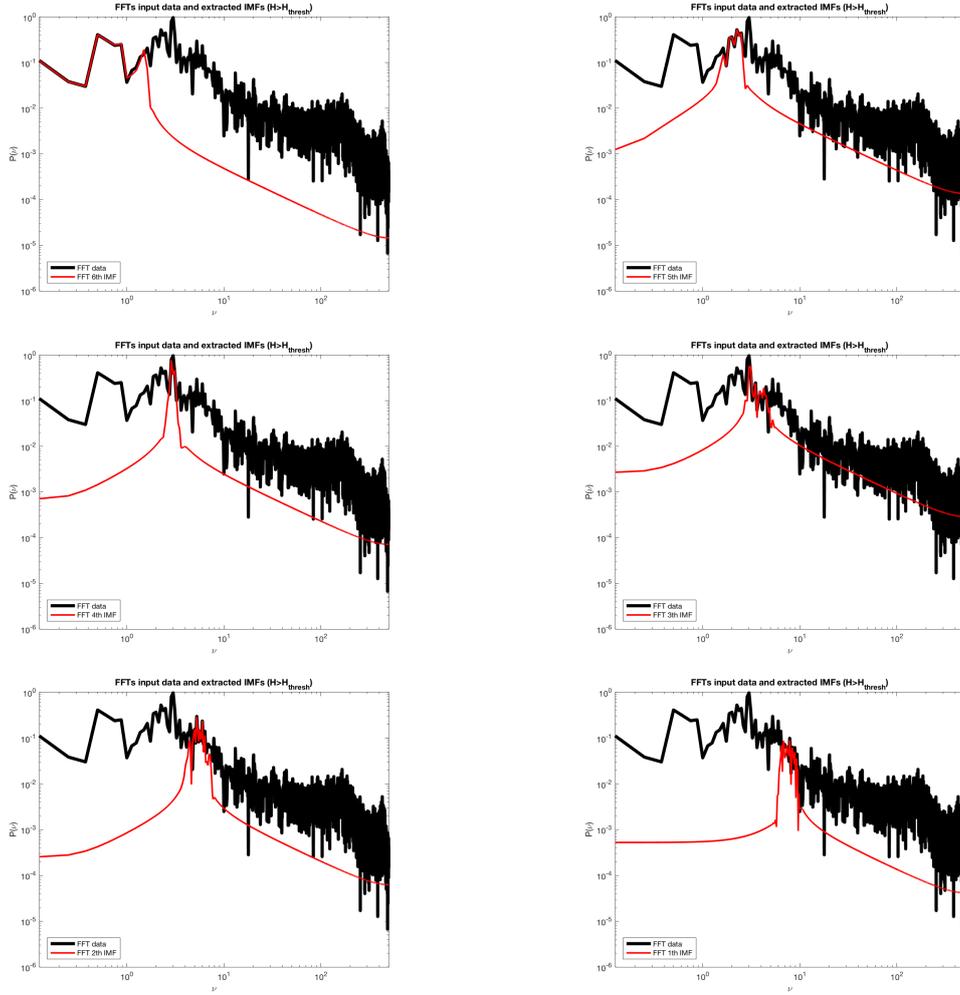


Figure 3.9: Shown are the spectra of obtained persistent oscillatory modes, i.e. having a  $H > H_{thr}$  while in black is the Fourier spectra of the input time series  $x(n)$ .

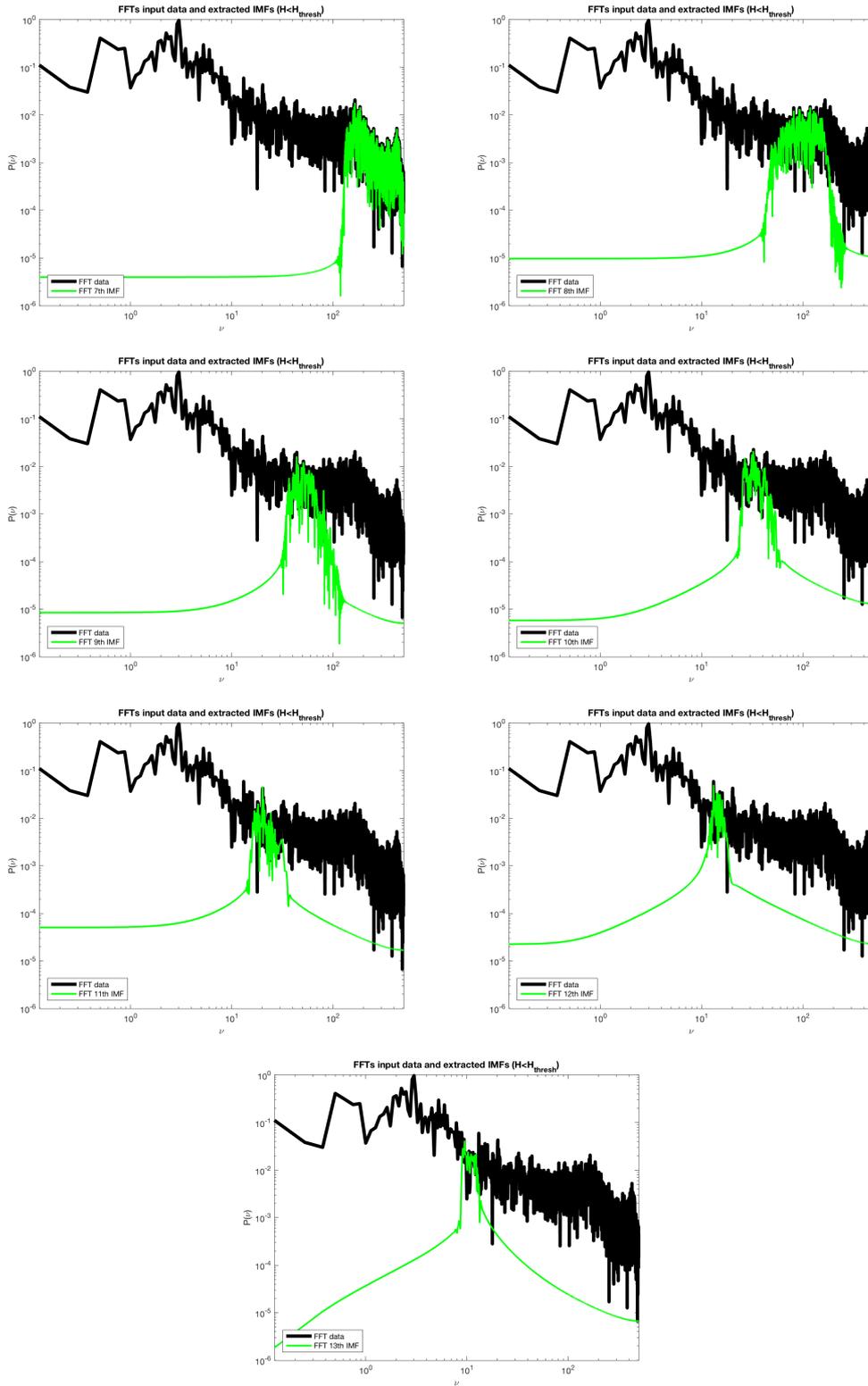


Figure 3.10: Shown in green are the Fourier spectra of obtained antipersistent oscillatory modes, i.e. having a  $H < H_{thr}$  while in black is the Fourier spectra of the input time series  $x(n)$

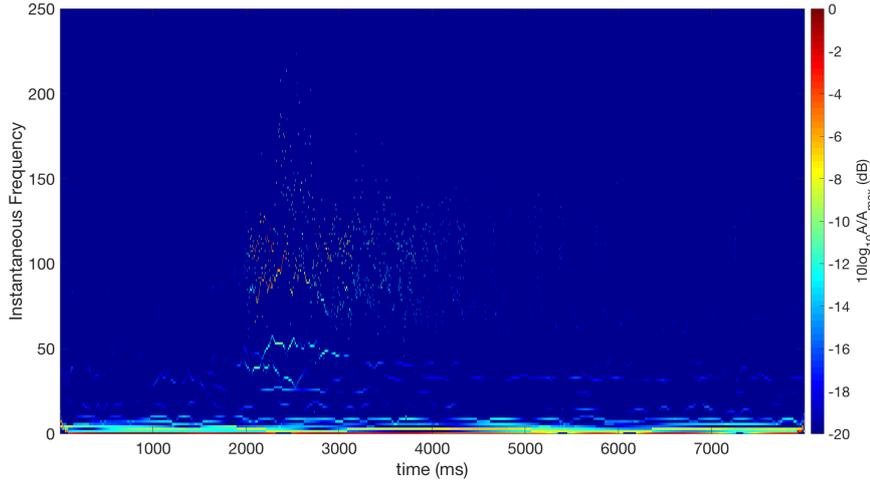


Figure 3.11: Hilbert Huang transform of triaxial seismometer data recorded during the third noise injection. Shown are results for the West arm direction component. High frequency modes are excited due to the injection, lasting for approximately  $\delta t = 2s$ . The colour scale is given in logarithmic scale by the relative amplitude  $A$  of the modes with respect to the maximum amplitude  $A_{max}$ .

### 3.2.5 Hilbert Huang Transform and Denoising Threshold Parameter

Figure 3.11 shows the HHT of seismometer data from the third acoustic noise injection (only W component shown). HHT is obtained applying HSA to the modes extracted with tvf-EMD. The obtained IF show that high frequency oscillatory modes are excited due to the injection, lasting approximately  $\delta t = 2s$ .

The adopted methodology was useful to adaptively extract the seismic noise waveform  $r(n)$ , sum of antipersistent modes, induced by the acoustic noise injection, from the underlying nonlinear nonstationary seismometer data  $s(n)$  given instead by the sum of persistent modes. This is achieved relying both on the adaptive nature of the tvf-EMD algorithm and on the different persistency of the data at the time of the injection compared to the overall seismometer recordings. Such a change in persistent behaviour is evident at the time of the injection and it was further quantified computing the local Hurst exponent  $H(t)$ , as can be seen in Figure 3.8. Further testing is needed to determine which input parameters,  $H_{thr}$ ,  $n$ ,  $\xi$  and number of modes achieve the best separation of perturbation signals from background signals. This could be done evaluating the output in term of standard denoising performance estimators such as the ones defined in [19], namely signal to noise ratio, mean squared error, mean absolute error, crosscorrelation, and is left to future work.

Regarding the choice of  $H_{thr} = 0.3$ , it is interesting to notice that the variation  $\Delta H$  of the local hurst exponent computed for the seismometer recordings, i.e. for  $x(n)$ , around the time of occurrence of the noise injection, is similar to the value of  $H_{thr}$  chosen to achieve denoising (see Figure 3.8 ).

Further studies are needed to test if choosing an  $H_{thr} \simeq \Delta H$  gives the best denois-

ing performances according to parameters such as the ones defined in [19]. If this is the case, fractal analysis could be carried out around noisy periods in the data, e.g. around periods of glitches of the instrument, providing a  $\Delta H$ . This value could be then given in input to the tvf-EMD based algorithm to adaptively extract and characterise the waveform of the given nonlinear nonstationary perturbation, i.e. for glitch characterisation and removal.

Regarding detector characterisation, the local Hurst exponent  $H(t)$  could be in general used to monitor persistent behaviour of relevant channels, e.g. around unlocks, or to monitor the properties of noises injected for detector characterisation purposes, hence monitoring stable conditions of the interferometer over time. An example in this regard is reported in appendix G, where the Hurst exponents and their standard deviations have been computed for time series from various sensors monitoring the Virgo ifo during a period of unusually high low frequency noise.

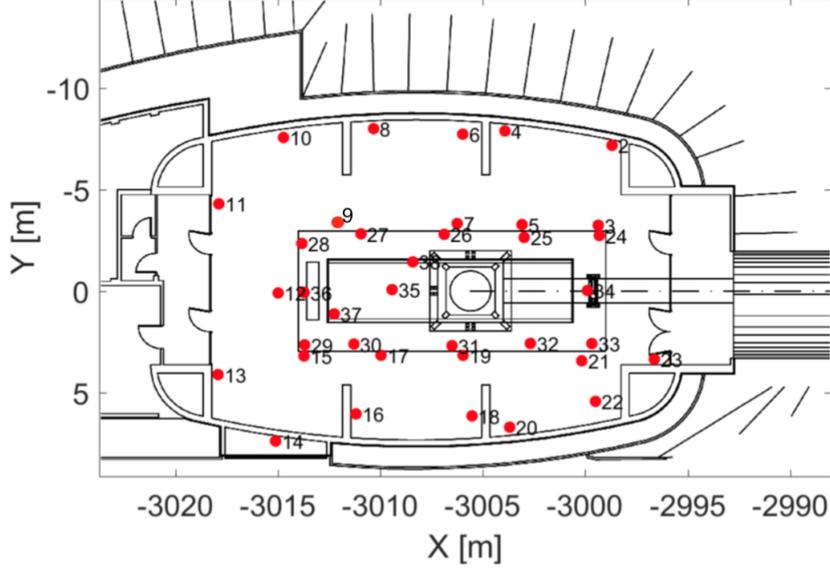


Figure 3.12: Coordinates of the array of 38 seismometer monitoring Virgo WEB for Newtonian Noise characterisation purposes. Figure taken from <https://tds.virgo-gw.eu/?content=3&r=14016>.

### 3.3 Local Hurst Exponent for Seismometer Array Monitoring WEB

The local Hurst exponent  $H(t)$  has been computed for data from an array of 38 seismometers, deployed at the Virgo West End Building for Newtonian Noise characterisation purposes. The analysed period is from January 31st, 2018 to February 5th, 2018. As described in Chapter 2, the Hurst exponent  $H$  is a fractal index quantifying the persistent behaviour of a time series, higher  $H$  corresponding to higher persistency, i.e. slower oscillations around the mean. The Hurst exponent can be computed using DFA and it can be computed locally using small sliding windows. It has been computed in order to characterise the properties of the seismometers data. For a detailed description on how to compute the Hurst exponent locally see [5]. Hourly averages and averages of  $H(t)$  have been computed over the whole analysed period. Results show that seismometers placed on a concrete slab closer to the centre of the room systematically exhibit higher persistency than the ones that are not placed on it. Seismometers placed next to the outer walls also exhibit higher persistency. The one seismometer placed on the ceiling of the tower exhibits instead very low values of persistency during the analysed period, compared to the rest of the array.

#### 3.3.1 Array of seismometers for Newtonian Noise Characterisation

As also described in Chapter 1, mass density fluctuations generated either by seismicity and microseismicity or by density fluctuations of atmospheric air masses

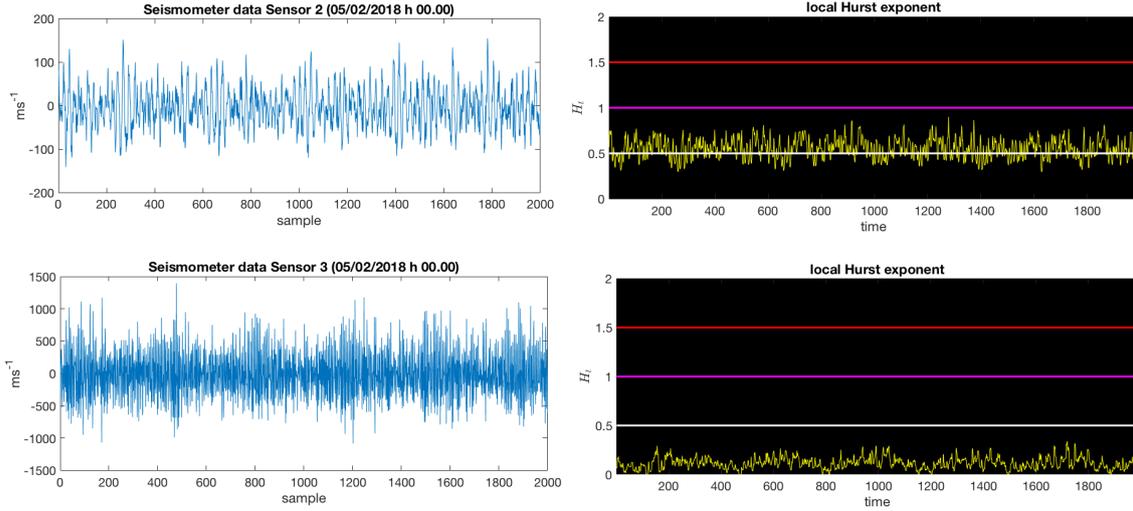


Figure 3.13: Left: Time series of 4s of data recorded from sensors 2 and 3 on February 5th, 2018 at 00:00. Right: local Hurst exponent  $H(t)$  for the two time series. Horizontal line represents values (starting from below) for white, pink and red noise respectively. Sensor 2 is placed next to the outer wall of WEB while sensor 3 is closer to the center of the room but not on the tower platform.

can induce a gravitational field which couples directly to the test masses of the interferometer, giving rise to a noise referred to as Newtonian noise [3, 59]. This gravitational field couples to each stage of the attenuation chain and also directly to the mirror. As described in Chapter 1, in the Virgo interferometer the test masses are isolated from the direct influence of seismic vibrations using a chain of oscillators. Such chain mitigates the amplitude of vibrations of frequency  $f$  at the suspension point by a factor  $(f/f_0)^2$  at each stage, where  $f_0$  is the frequency of a single oscillator [68]. Newtonian Noise has been briefly described in Section 1.6. For Newtonian noise cancellation studies, array of seismometers are typically deployed around test masses.

### 3.3.2 Local Hurst Exponent of Seismometer Array

The local Hurst exponent has been computed for the 38 seismometers of the array, monitoring WEB for six consecutive days from January 31st, 2018 to February 5th, 2018 with sampling frequency of  $f_s = 500 \text{ Hz}$ . In Figure 3.12 the coordinates of the seismometers are shown. In particular, it can be seen how some sensors are placed next to the outer wall of the room, while some are placed closer to the center of the room but not on the platform hosting the tower in which the superattenuator is placed. Finally some sensors are placed on top of the platform hosting the tower. Sensor 35 is instead placed on the tower ceiling. The size of the window used for  $H(t)$  computation is of 10 data points, corresponding to a time interval of  $\delta t = 20 \text{ ms}$ . Figure 3.13 shows 2000 samples from sensors 2 and 3, recorded on

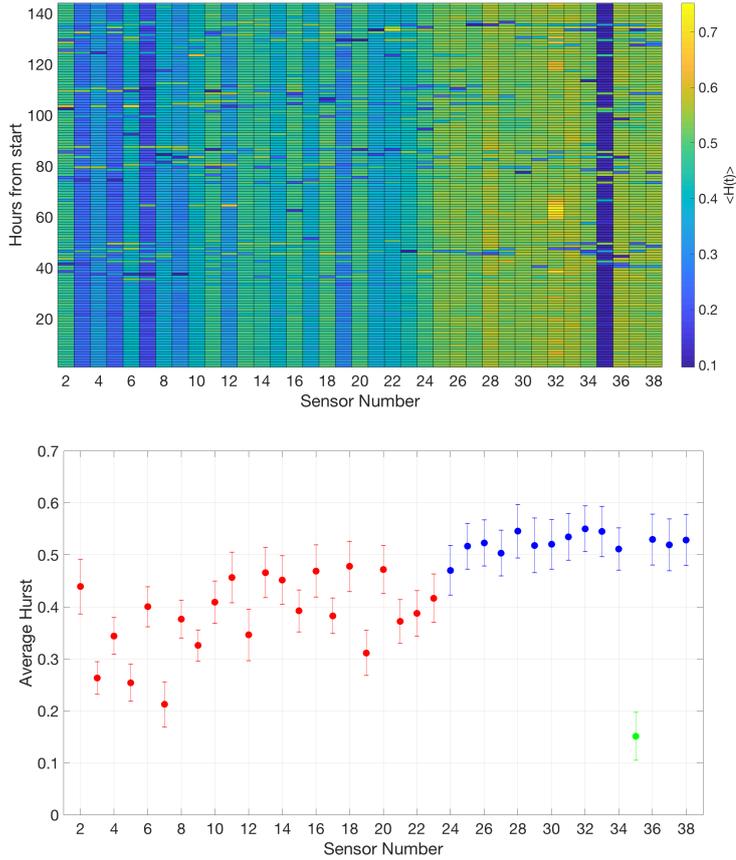


Figure 3.14: Top: Hourly averages of the Hurst exponents for the 38 seismometer of the array over the whole sampling period. Bottom: Average values of the hourly average Hurst exponents. Red dots represent sensors which are not on the tower platform, while blue dots sensors that are on the tower platform. It can be seen how sensors closer to the outer walls of the room exhibit higher persistency. Sensor 35, shown in green, is instead placed on the ceiling of the tower and is found to be strongly antipersistent over the whole analysed period.

February 5th, 2018 at 00:00. Since the sampling frequency is  $f_s = 500 \text{ Hz}$  and the samples are 2000, the time series in Figure 3.13 span a period of 4 s. Sensor 2 is placed next to the outer wall of WEB while sensor 3 is closer to the center of the room but not on the tower platform. The local Hurst exponent is able to discriminate the different persistency of seismometer data, sensor 3 oscillating more rapidly around the mean and exhibiting therefore an antipersistent behaviour. Due to this, sensor 2 exhibits higher persistency compared to sensor 3 during the analysed period, as can be also seen on the right panel of figure 3.13.

Figure 3.14 shows the hourly average values of  $H(t)$  (top) and the average of such hourly values (bottom) for the 38 seismometers of the array. It can be seen how  $H(t)$  is able to discriminate the different positions of the seismometer in the room, namely higher persistency is found in data from sensors closer to the outer walls and placed on top of the tower platform of Virgo WEB compared to seismometers placed in the middle of the room which are not on the tower platform.

This is possibly due to the platform, on which the tower is placed, attenuating the high frequency content of seismometer time series data, giving hence a higher Hurst exponent for sensors on the platform.

Why also sensors closer to the outer walls of WEB have a higher Hurst exponent compared to seismometers at the centre of the room but not to those on the platform is not clearly understood. Seismometer number 35, which is instead placed on the ceiling of the tower, exhibits a lower persistency compared to all the other seismometers of the array.

## 3.4 Yearly Modulation of surface $^7\text{Be}$ activity concentration

In this Section are reported results from the analysis of data collected on a daily basis by 28 stations of the International Monitoring System <sup>2</sup> (IMS), a worldwide distributed network set up and maintained by the Comprehensive Nuclear-Test-Ban Treaty Organisation (CTBTO), whose goal is to monitor over the CTBT compliance. Beside monitoring atmospheric concentrations of Treaty relevant radionuclides, the IMS also detects naturally occurrent radionuclides. A notable example in this regard is  $^7\text{Be}$ , a cosmogenic radionuclide which, due to its physical and chemical characteristics, is a useful tracer of air masses of tropospheric and stratospheric origin [69].

### 3.4.1 Dataset

Every monitoring station of the IMS is equipped with a filter and an air sampler. The flow rate of the air sampler should be at least  $500\text{ m}^3\text{h}^{-1}$ , measured at standard temperature and pressure (STP) during the sampling periods. It follows that at least a total air volume of  $10800\text{ m}^3$  ( $500\text{ m}^3\text{h}^{-1} \times 24\text{ h}(-10\%)$ ) at STP will pass through each filter sample. The total air volume collected is then normalised to STP. The measurement system is based on high resolution gamma spectrometry using a high purity germanium detector (HPGe) with a minimum relative efficiency of at least 40%, and a peak resolution better than  $2.5\text{ keV}$  at full width at half maximum (FWHM) at the gamma ray energy of  $1332\text{ keV}$  under operational conditions [70]. Time series of  $^7\text{Be}$  and surface temperature considered in this Section have been selected based on data availability.  $^7\text{Be}$  time series are sampled on a daily basis ( $\mu\text{Bq}/\text{m}^3$ ), while temperature time series refers to the daily means ( $^{\circ}\text{C}$ ). The length of a given time series depends on when the corresponding station became operative, and the starting year of the time series varies from 2003 to 2009, while all series end on March 2016. Time series length ranges from 2746 to 4711 daily samples. Other informations on the analysed  $^7\text{Be}$  dataset can be found in [39]. In Table 3.1 the names and the exact location of the analysed stations are listed.

### 3.4.2 Obtained Results: Trend, Yearly cycles and Outliers

Results of the analysis are hereafter summarised for the 28 stations of the IMS network. In Figure 3.15 the trends in  $^7\text{Be}$  time series extracted adaptively are shown. Trends are normalised to zero mean and unit variance for a better comparison. A unique behaviour in terms of latitude cannot be discerned, possibly due to the widely different locations and altitudes of the different stations of the network. Monotonic trends, represented by a gradual change of the colour scale, are prevalent in both Northern and Southern hemispheres (divided by the horizontal black line), while there are also non-monotonic trends that express an overall change of  $^7\text{Be}$  concentrations. It is important to notice that trend behaviour is not completely

---

<sup>2</sup><https://www.ctbto.org/map/>

ID	Location	Lat [ $^{\circ}N$ ]	Long [ $^{\circ}E$ ]	Altitude [ $m$ a.s.l.]
RN76	Salchaket, Alaska, USA	64.67	-147.10	150
RN16	Yellowknife, N.W.T., Canada	62.48	-114.47	206
RN63	Stockholm, Sweden	59.41	17.95	0
RN61	Dubna, Russian Federation	56.74	37.25	118
RN71	Sand Point, Alaska, USA	55.34	-160.49	60
RN60	Petropavlovsk, Russian Federation	53.05	158.78	65
RN33	Schauinsland/Freiburg, Germany	47.92	7.91	1208
RN45	Ulaanbaatar, Mongolia	47.89	106.33	1729
RN17	St. John's, N.L., Canada	47.59	-52.74	133
RN75	Charlottesville, VA, USA	38.00	-78.40	104
RN74	Ashland, KS, USA	37.17	-99.77	603
RN72	Melbourne, FL, USA	28.10	-80.65	10
RN37	Okinawa, Japan	26.50	127.90	106
RN79	Oahu, Hawaii, USA	21.52	-157.99	427
RN43	Nouakchott, Mauritania	18.14	-15.92	10
RN50	Panama City, Panama	8.98	-79.53	90
RN64	Dar Es Salaam, Tanzania	-6.78	39.20	104
RN08	Cocos Islands, Australia	-12.19	96.83	5
RN09	Darwin, NT, Australia	-12.43	130.89	32
RN26	Nadi, Fiji	-17.76	177.45	31
RN06	Townsville, QLD, Australia	-19.25	146.77	7
RN23	Rarotonga, Cook Islands	-21.20	-159.81	5
RN10	Perth, WA, Australia	-31.93	115.98	31
RN01	Buenos Aires, Argentina	-34.54	-58.47	31
RN47	Kaitaia, New Zealand	-35.07	173.29	86
RN68	Tristan da Cunha, United Kingdom	-37.07	-12.31	64
RN04	Melbourne, VIC, Australia	-37.73	145.10	31
RN46	Chatham Island, New Zealand	-43.82	-176.48	22

Table 3.1: ID code, exact location and altitude of CTBTO International Monitoring System stations analysed.

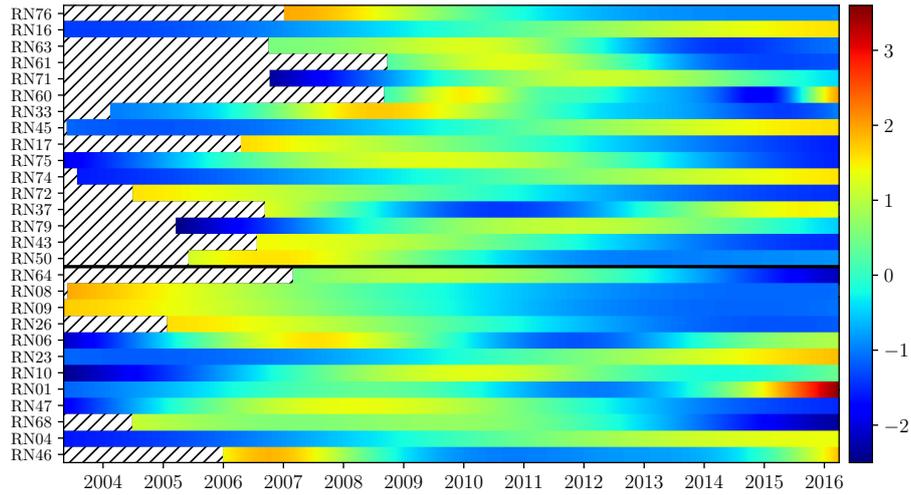


Figure 3.15:  ${}^7\text{Be}$  adaptive trends as obtained from the residual term of EMD for the 28 stations of the IMS network. Stations are ordered by latitude and data have been normalised to zero mean and unit variance for a meaningful comparison. The horizontal black line divides the Northern hemisphere from the Southern one, while the shaded areas represent periods where data are not available. The colour scale is the height of the adaptive trend.

determined for those time series that have a shorter length, since their behaviour before the date they became operative is not known. Furthermore, beside choosing a high number of sifting iteration no attempt was made to mitigate end effects during the trend estimation step with EMD [18]. End effects are clearly visible in Figure 3.15 for the stations RN60 and RN01, where a sudden change in the trend occurs at the right end of the time series.  ${}^7\text{Be}$  trends have been also cross-correlated with the temperature ones. Even though the majority of correlations are high, a clear pattern cannot be evinced. In Figure 3.16, the annual IMF of  ${}^7\text{Be}$  activity concentration is shown for all the analysed stations. To meaningfully compare the different yearly oscillations, they have been normalised to zero mean and unit variance. Maxima and minima alternate regularly and appear to be shifted in time both in the Northern and Southern hemispheres. Peaks of the annual oscillation are almost regularly delayed going from the equator to high latitudes, and the same occurs going from the equator to low latitudes. Such behaviour is possibly related to shifting of the Hadley cell and of the intertropical convergence zone (ITCZ), as also noted in [7]. It should be noted that the considered stations are not uniformly distributed around the globe, and latitudes are not continuous from the top to the bottom of Figure 3.16. Furthermore, stations RN79, RN26, RN06, RN23, and RN04 exhibit the highest value of the yearly peak in 2010, while RN47 in 2009. Due to their locations, this could be possibly related to El Niño event occurred in 2009-2010. To better quantify though, further studies are needed, employing both atmospheric transport models and data of  ${}^7\text{Be}$  production, e.g. cosmic rays variability due to the solar cycle. Furthermore, while the analysis carried out in this Thesis primarily focused on characterising only the yearly cycle of  ${}^7\text{Be}$  activity concentration, other periodicities can be found, see for instance [71]. In Figure 3.17

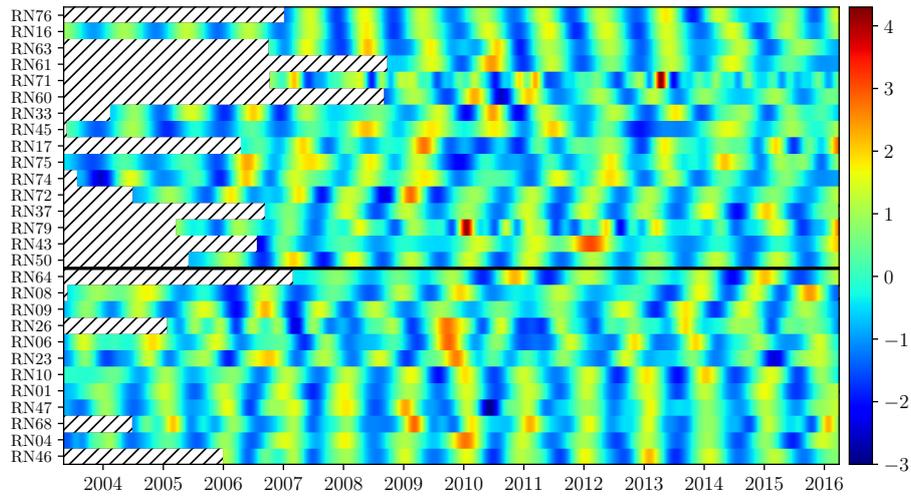


Figure 3.16:  ${}^7\text{Be}$  annual IMF of the 28 stations of the IMS network. Stations are ordered by latitude and data have been normalised to zero mean and unit variance for a meaningful comparison. The horizontal black line divides the Northern hemisphere from the Southern one, while the shaded areas represent periods where data are not available. The colour scale is the amplitude of the oscillations.

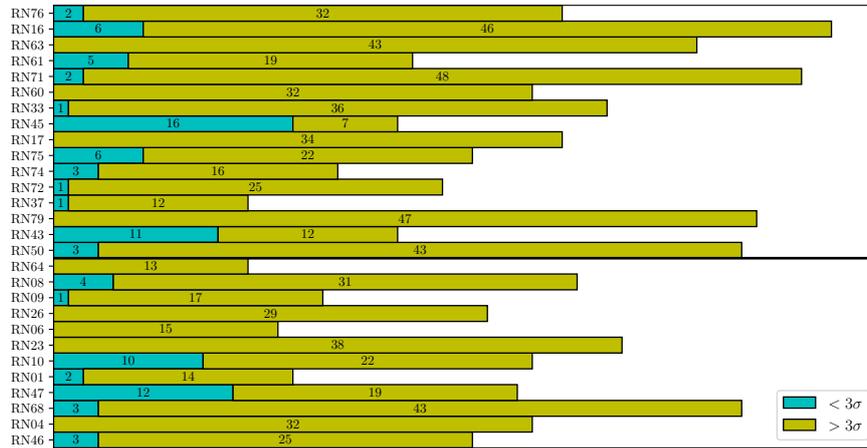


Figure 3.17: Occurrence of outliers higher than  $3\sigma$  (yellow) and lower than  $-3\sigma$  (cyan) in  ${}^7\text{Be}$  residuals, for the 28 stations of the IMS CTBT network. Stations are ordered by latitude. The horizontal black line divides the Northern hemisphere from the Southern one.

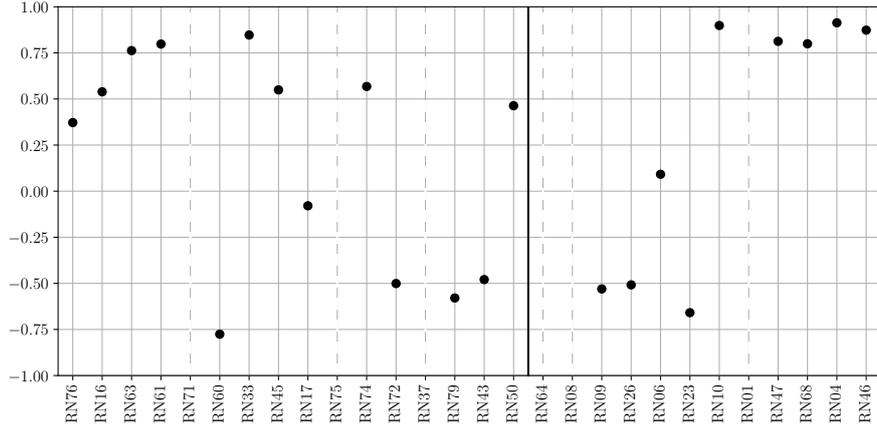


Figure 3.18: Cross-correlations between  ${}^7\text{Be}$  and temperature IMFs of the annual cycle. Dashed vertical lines represent stations with bad quality data.

the occurrence of outliers higher than  $3\sigma$  (yellow) and lower than  $-3\sigma$  (cyan) in  ${}^7\text{Be}$  residuals is shown. In general low values are not frequent, meaning a very rare occurrence of high drops in  ${}^7\text{Be}$  concentrations. Furthermore, the number of outliers in the Northern hemisphere represents the 61.4% of the total outliers. Almost 20% of the Northern hemisphere's outliers are above  $4\sigma$ , and 4% is above  $5\sigma$ . Percentages are lower for the Southern hemisphere, and the number of outliers above  $4\sigma$  and  $5\sigma$  is almost half the number of the corresponding Northern outliers, even though Southern hemisphere stations are the three quarters of the Northern hemisphere ones. Residuals correlations have been estimated via the Hurst exponent  $H$ . All values of  $H$  ranges between  $\sim 0.8$  and  $\sim 1.0$ , indicating strong long-range autocorrelations of residuals time series. Moreover, denoising performance parameters introduced in Section 2.9 are evaluated. The performance of the denoising step resulted to be quite good, with  $MSE$  and  $MAE$  values less than one for all the time series.  $SNR$  and  $PSNR$  values are not very high (but greater than one), while  $xcorr$  value is in the range 0.4-0.7. These values are possibly due to the time series being very noisy and short. Finally, the yearly IMFs extracted from temperature data have been analysed. The number of stations is now less than 28 since some stations had a poor quality temperature time series, i.e. wide gaps and/or unphysical values. A simple seasonal pattern is observed in this case with a six-month shift between the Northern and Southern hemispheres maximum temperature, as expected. Cross-correlations between temperature and  ${}^7\text{Be}$  annual IMFs have also been evaluated and are shown in Figure 3.18. High correlations have been found in all but two stations, namely RN06 in the Southern hemisphere and RN17 in the Northern hemisphere. These two stations are characterised by a non-perfect annual oscillation, as can be seen in Figure 3.16, where with a perfect yearly oscillation a sinusoid with period of one year is intended, not affected neither by non-linearity nor by non-stationarity. Dashed vertical lines represent stations with bad quality data. In summary, data of  ${}^7\text{Be}$  and temperature, sampled daily at 28 different stations of the IMS of the CTBTO and worldwide distributed, have been analysed and characterised using adaptive and fractal methodologies. Oscillatory modes both on shorter and longer scales than the yearly ones were obtained,

according to the denoising criteria. Such modes had periods ranging from three months to up to three years and were removed in order to obtain the residual time series, from which outlier occurrence is evaluated. The analysis then focused on characterising phase and amplitude of the extracted yearly cycles, which showed a different pattern at different latitudes, the annual peak being delayed for stations at higher latitudes, both in the Northern and Southern Hemispheres, and this is probably due to Hadley cell and the Intertropical Convergence Zone dynamics. A higher value of the yearly cycle is also found in those stations influenced by the 2009-2010 El Niño event. Finally yearly cycles of  $^7\text{Be}$  and temperature data have been cross-correlated, yielding high correlation values in general.

# Chapter 4

## Discussion

In the framework of this Thesis a methodology of time series analysis based on fractal and adaptive algorithms such as EMD and tvf-EMD has been developed and applied to time series data from different physical systems in order to validate its output based on features extracted from the analysed data. Beside adaptive methodologies, the local Hurst exponent has been computed to evaluate persistency in the data, the Hurst exponent being a fractal index used to estimate time series persistency. Main part of the analysis focused on seismometer data from seismometers monitoring the Virgo ifo both for detector characterisation purposes and for Newtonian Noise cancellation purposes. In the latter case data from an array of 38 seismometers were analysed. For this reason, in Chapter 1 an overview of the operational principles of the Virgo ifo were given, with particular focus on seismic noise, newtonian noise and on the methodologies adopted to mitigate them, both from the experimental point of view and from the data analysis point of view. After that, in Chapter 2, the methodologies adopted for data analysis were widely described. First, HSA has been described, and the concepts of IA, IF and IB were introduced. After that, the EMD algorithm was described widely. Advantages and drawbacks, such as mode mixing, of this algorithm were discussed. Ensemble methodologies to mitigate mode mixing were also described, along with a test for significance of extracted IMFs against white noise. After that the tvf-EMD algorithm, which is the one used in this Thesis for the extraction of oscillatory modes, was described in detail. The main advantage of this algorithm is that, stopping the sifting procedure based on a threshold on modes IB, the mode mixing problem is significantly mitigated. Relevance of extracted IMF was established computing their Hurst exponent by means of Detrended Fluctuation analysis algorithm.

Thanks to the completeness property of adaptive algorithms, namely that summing the extracted modes exactly recovers the original data, adaptive denoising could be performed on the data. Denoising performance parameters were also introduced in Chapter 2 and were used to evaluate denoising performance in  ${}^7\text{Be}$  data.

Both the Hurst exponent and DFA were defined and described in Chapter 2. Having established mode's relevance allowed to separate relevant modes of oscillation from noise related ones, which are typically found to be antipersistent and rapidly oscillating around their mean value.

Applying HSA to IMFs extracted by adaptive algorithms allows to obtain a time frequency representation of the data called HHT. The HHT shows how the IF and

the IA of the extracted modes varies in time. It has higher resolution compared to Fourier based periodograms, since frequency is defined for each time. No windowing nor averages of spectra is needed in this case.

Finally, in Chapter 3, the methodologies described in Chapter 2 were applied to data from real physical systems. The aforementioned methodologies have been applied to

- Denoising of GW chirp signal in purple noise for testing purposes.
- Seismometer data monitoring Virgo NEB during four acoustic noise injections performed for detector characterisation purposes
- Seismometer data from an array of 38 seismometers monitoring Virgo WEB for Newtonian Noise characterisation purposes
- ${}^7\text{Be}$  data acquired from the IMS of the CTBTO.

Regarding the testing phase, it is found that denoising based on persistency is best applied when the signal to be extracted, the GW chirp, exhibit different persistent behavior compared to the noise that corrupts the time series. Performances decrease in fact for noises of higher spectral index. For purple noise though, the chirp signal is correctly extracted as can be seen from Figure 3.1 where it is compared with the noise free NR waveform.

The adopted methodology has been then applied for the denoising of seismometer data recording during four different acoustic noise injections. In this case the seismic perturbation due to the injection is again extracted correctly from the time series data, which are nonlinear and nonstationary, due to the fact that the perturbation has a different persistent behavior compared to the underlying seismometer data. This is also highlighted in Figure 3.8 where a drop in Hurst exponent is visible around the time of the injection. This suggest that computation of the local Hurst exponent could be used to select an optimal threshold  $H_{thr}$  for the denoising, but further testing is needed in this regard.

Regarding the array of 38 seismometers only fractal analysis was performed, and the local Hurst exponent has been employed to characterise persistency of the data. It is found that sensors in different parts of the room have different persistency depending on whether they're placed or not on the concrete slab hosting the platform where the tower containing the superattenuator is placed or if they are in proximity of WEB outer wall.

Then, the yearly component of  ${}^7\text{Be}$  time series was extracted using tvf-EMD and denoising based on persistency. It is found that the yearly component of the 28 analysed stations of the IMS have a phase shift which is latitudinally dependent, feature that can be reconciled with large scale features of atmospheric dynamics such as the seasonal shift of the ITCZ and of the Hadley cell. High values of the yearly component were found for stations possibly influenced by the 2009-2010 El Niño event, though this need further research. Furthermore, outlier occurrence has been characterised and denoising performance parameters evaluated for this dataset, showing good performances of persistency based denoising.

Further studies regarding synthetic noise properties and investigations of noises affecting the Virgo ifo sensitivity were also conducted.

In this regard, additional results are reported in Appendices regarding

- Spectral index evaluation with the Burg Maximum Entropy Method: In Appendix A is shown how Fourier and Lomb Scargle periodograms underestimate spectral indexes  $\beta > 2$  and how a possible solution to the problem is to use the Burg Maximum Entropy Method for spectral index estimation.
- Hurst exponent estimation of fractional noises with EMD: In Appendix C it is shown how EMD can be used to estimate the Hurst exponent of noises, thanks to the similarity of the spectra of extracted IMFs.
- Scattered light noise hunting in Virgo with EMD: In Appendix F it is shown how EMD algorithm can be applied to Virgo data in the attempt to quickly identify sources of scattered light noise, due to movements of reflecting surfaces causing a phase noise in the laser beam.
- $1/f^{2.5}$  Noise Hunting with Multifractal Detrended Fluctuation Analysis Characterisation of Local Hurst Exponent of C11 data: In G some results are shown from a characterisation of the  $1/f^{2.5}$  broadband noise, affecting Virgo during the commissioning phase, computing the local Hurst exponent of downsampled time series from auxiliary channels monitoring Virgo during the Commissioning run C11.

# Appendices

# Appendix A

## Estimate of spectral indexes $\beta > 2$ with the Burg Maximum Entropy Method

In this Appendix is described how periodogram based methods, such as the FFT and the Lomb periodogram, underestimate spectral indexes  $\beta > 2$ , and a possible way to solve this issue, based on the Burg Maximum Entropy Method (MEM), is tested on synthetic time series.

Results obtained with a software for time series analysis, referred to as Multi-step Time Series Analysis (MTsA) [39] which is based on the Lomb periodogram method, are compared with the ones obtained following the approach, based on MEM, from [72]. MTsA allows to separate time series deterministic components from random fluctuations, even when data are affected by missing data. It does so making use of the Generalised Lomb-Scargle Periodogram (GLS) [73] which naturally provide a threshold to separate signal from noise in presence of uncorrelated Gaussian noise. Once the harmonic components have been identified and filtered out, residual time series is obtained. Then, MTsA performs detrended crosscorrelation analysis [74] of residual time series and estimates their Local Hurst exponent  $H_t$ , allowing to obtain quantitative information on long term correlation [5]. A more detailed description of MTsA software, along with results from its application to real geophysical datasets, can be found in [38, 42, 39, 40, 41]. A flowchart of such algorithm is reported in Figure A.1.

A process having power spectrum that follows a power law of the type

$$S(\nu) \sim \nu^{-\beta} \tag{A.1}$$

is said to have spectral index  $\beta$ , where  $\beta = 0, 1, 2$  corresponds to white, pink, and red noise, respectively. One exemple of physical system where different spectral indexes are present in geophysical data can be found in [72], where Burg-MEM technique is applied to ionospheric amplitude scintillation data taken from the MARISAT satellite. In this case, the smoothness of the Burg spectra allows to better track, compared to periodograms results, the onset of moderate scintillation and its development into fully saturated scintillation ([72] Figure 7a) in terms of changes of spectral slope ([72] Figure7b).

One possible way to estimate the power spectrum of a time series is from the

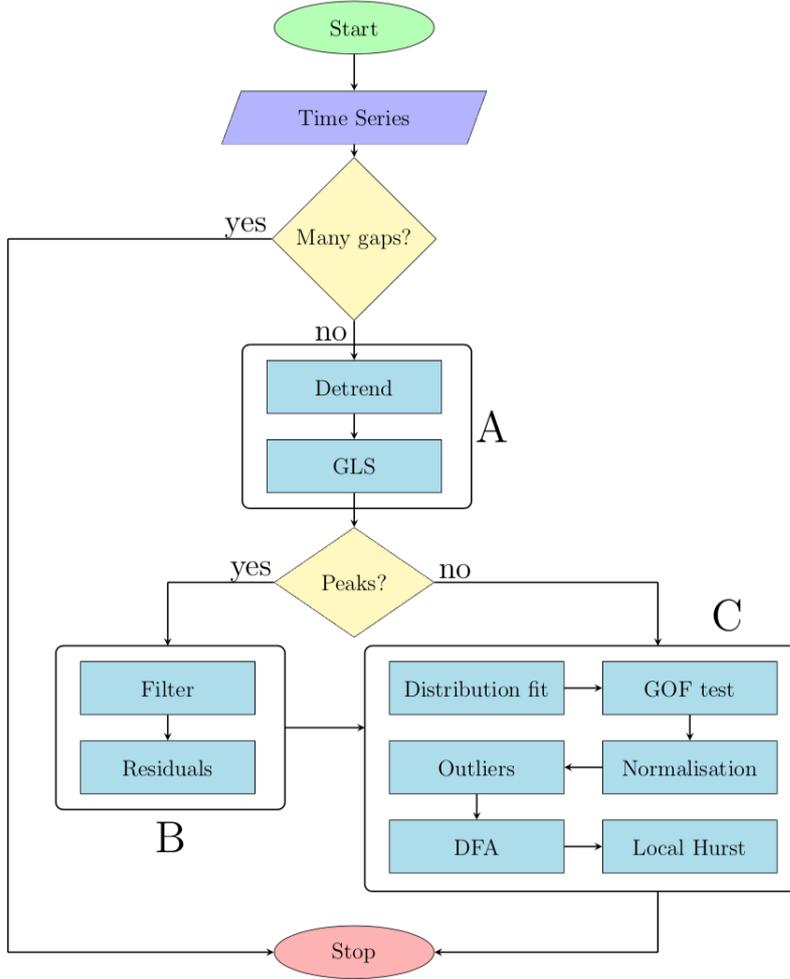


Figure A.1: Flowchart illustrating the different steps of MTsA software [39].

squared magnitude of the Fourier Transform (FT) of the data, i.e. the periodogram. Any finite data set is equivalent to an infinite data set multiplied for a rectangular window  $W$  which is unity inside the interval of data availability and zero outside. Since the Fourier transform of a rectangular window  $W$  has the form [72]

$$FT[W] \sim \sin(\pi f)/\pi f \quad (\text{A.2})$$

and has both a narrow central lobe and high side lobes, this increase the power at points away from the central lobe. It is found that spectral indexes  $\beta > 2$  can't be obtained using a rectangular window, which is the one that provides the narrowest main lobe and hence the best spectral resolution, limited by the width of the main lobe, given by  $(N\Delta t)^{-1}$  Hz [72].

Many other windows have been designed to offset this high side lobe behavior. The rectangular window gives the narrowest main lobe while other windows designed to reduce side lobes do so at the expense of increasing main lobe width [72]. Nevertheless, some windowing is essential and spectral estimation of red noise processes using periodograms requires the use of some nonrectangular window [72]. To eval-

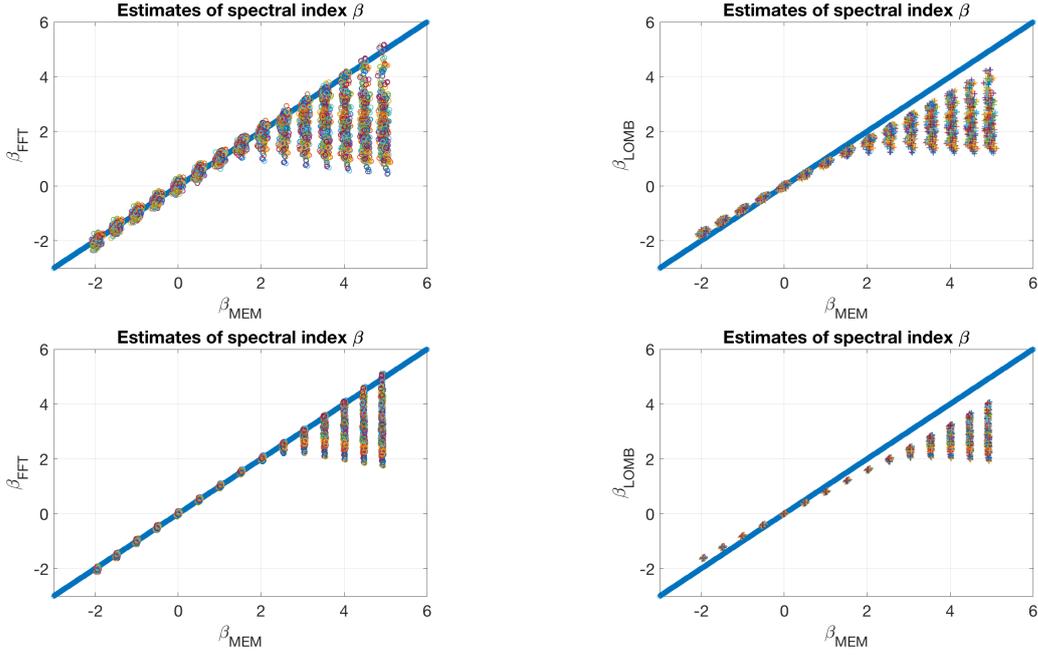


Figure A.2: Comparison of spectral index estimation with different methods. Top: Comparison of Fourier and Burg estimates of spectral indexes (left) and Lomb-Burg estimates (right) for dat length  $N = 10^3$ . Bottom: Same as before, but for data length  $N = 10^4$ . Periodogram based methods tend to underestimate spectral indexes greater than two.

uate such effects also on the Lomb spectrum, which can be considered both as a Fourier Method and a Least Square Method [75], two ensemble of 400 time series having length of  $10^3$  and  $10^4$  respectively, and having spectral index ranging from  $[-2, \dots, 5]$  at intervals of 0.5, has been synthetically generated following the procedure described in [72], that is low-pass filtering gaussian white noise with filters having frequency response tuned in order to obtain as output time series with power-law spectrum in a selected frequency range [76].

The shyntetic time series were than normalised to zero mean and unit variance. The so obtained "coloured" noise were analysed through their Fourier, Lomb and Burg Maximum Entropy (MEM) spectra. The spectral index  $\beta$  for each case is obtained from the slope of the straight lines fitted to the spectra versus frequency in logarithmic coordinates.

As expected, Burg spectrum provides the most accurate results among the three techniques, yielding precise estimation of spectral index up to  $\beta = 5$ , while the two Fourier Based methods applied can't estimate spectral index greater than  $\beta = 2$  when no windowing is performed. Results from the simulation are shown in Figure A.2. It is found that for  $N = 10^4$ , the estimates of  $\beta$  are less scattered around their true value compared to the case  $N = 10^3$ . Some of the synthetic time series of length  $N = 1000$  used for this study are shown Figure A.3 and as can be seen they exhibit different persistency.

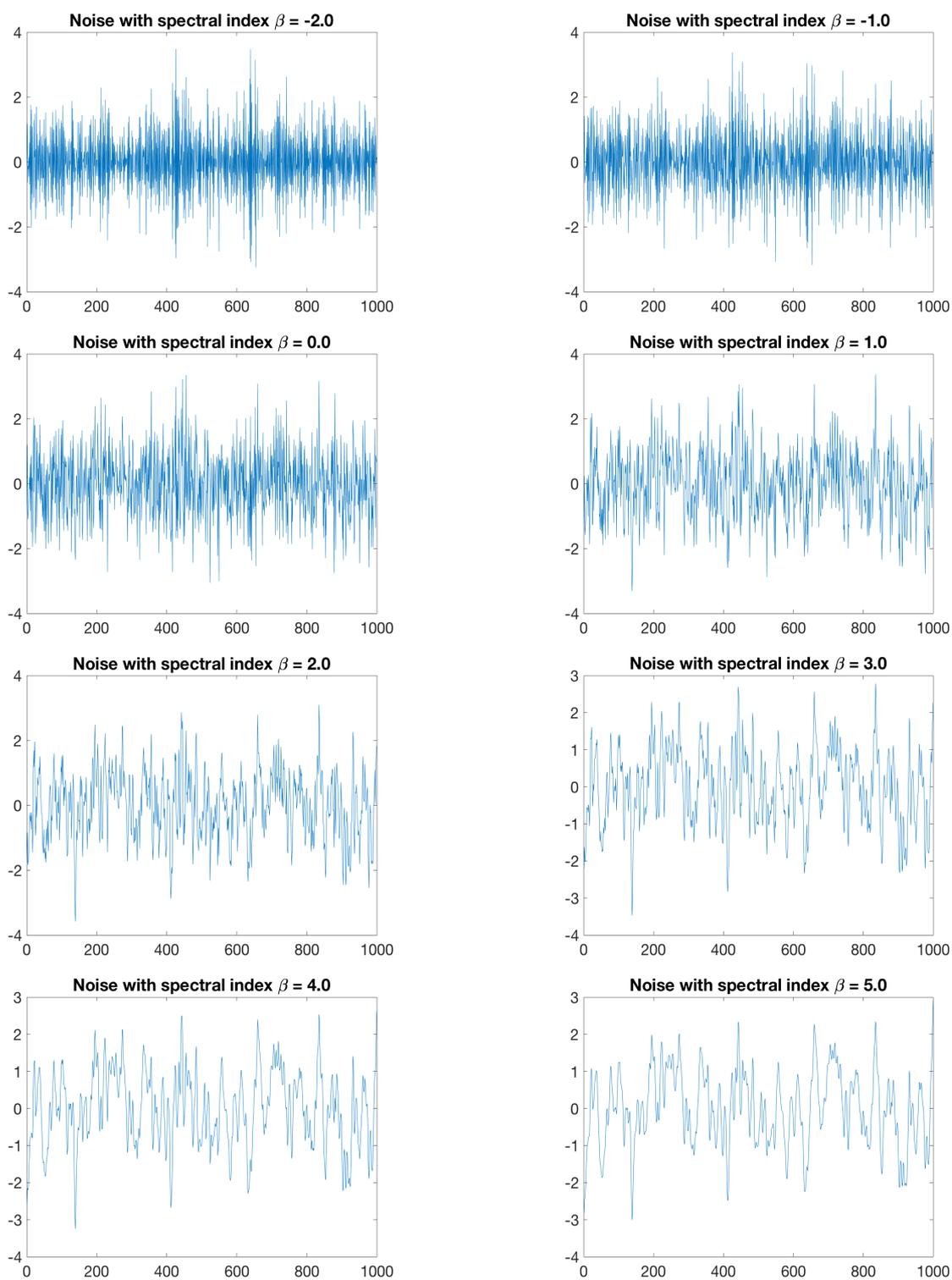


Figure A.3: Synthetically generated noises of length  $N = 10^3$ , having zero mean and unit variance. Spectral index  $\beta$  ranges from -2 (top left) to 5 (bottom right), increasing by 1.

## Appendix B

# Fractional Gaussian noise (fGn) and fractional Brownian motion (fBm)

In this Thesis work the Hurst exponent have been used as a threshold parameter for denoising purposes. Generally though, the Hurst exponent  $H$  is a parameter widely used to quantify long term correlation, or "memory" in a time series. In this case, the exact meaning of  $H$  depends on whether a signal is fractional Gaussian noise (fGn) or fractional Brownian motion (fBm), since while  $H$  values indicate the correlation structure of a fGn signal, for a fBm signal the correlation structure described by  $H$  values refers to the increments resulting from differencing the time series [77],[78]. To classify a signal as fGn or fBm before proceeding with its fractal analysis is hence needed, as described for example in [79][80]. Fractional Gaussian noise (fGn) is a stationary long-memory process while fBm is a non-stationary, long-memory process. A more precise definition of fGn involves its autocorrelation function. As can be found in [81] [18], the process

$$x_H[n], \quad n = \dots, -1, 0, 1, \dots \quad (\text{B.1})$$

is a fGn of index  $H$  if and only if it is a zero-mean Gaussian stationary process whose autocorrelation sequence is

$$r_{H[k]} = \frac{\sigma^2}{2}(|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}). \quad (\text{B.2})$$

Summing data of a fGn time series will produce a fBm time series. Conversely the increments of a fBm process, obtained by subtracting each value from the prior, form a fGn signal [80]. As described earlier, different values of  $H$  indicate different types of long-memory. Following [77]:

- $H = 0.5$  indicates absence of long-memory and data points are uncorrelated with each other (the process is random) or are correlated only at very small scales.
- When  $0 < H < 0.5$  the Hurst exponent indicates an anti-correlated or anti-persistent process (negative long-range correlation) for cases of fGn and fBm, respectively. In particular increases in the signal (for fGn) or in the increments of the signal (for fBm) are likely to be followed by decreases and conversely decreases are likely to be followed by increases

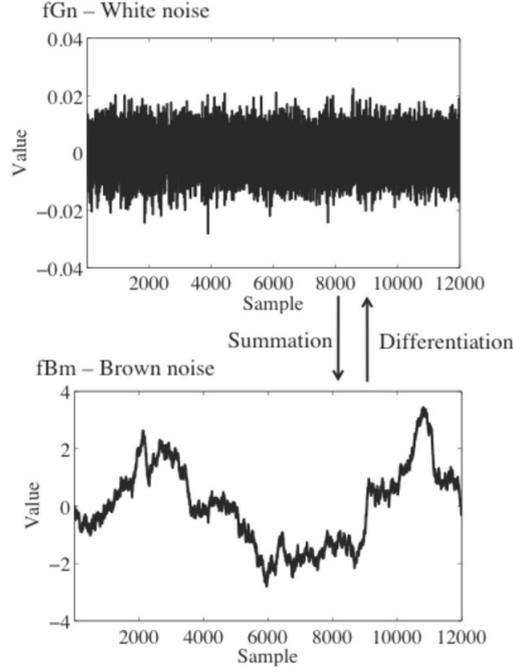


Figure B.1: In this figure an example of how Fractional gaussian noise and fractional Brownian motion can be obtained from one another summing or differentiating is shown. Figure taken from [79].

- Finally when  $0.5 < H < 1$  this indicates a correlated process for fGn or a persistent process for fBm. In this case, increases/decreases in the signal (for fGn) or in the increments of the signal (for fBm) are likely to be followed by further increases/decreases (positive long-range correlation).

Anti-persistent and persistent processes contain structure that distinguishes them from truly random sequences of data. Taking the Fourier transform of the auto-correlation function of a fGn yields its power spectrum density

$$S_H(f) = C\sigma^2 |e^{i2\pi f} - 1|^2 \sum_{k=-\infty}^{\infty} \frac{1}{|f+k|^{2H+1}} \longrightarrow C\sigma^2 |f|^{1-2H}, \quad (\text{B.3})$$

for  $H \neq 1/2$  when  $f \rightarrow 0$  [18]. fGn is then a model for power-law spectra at low frequencies. From its spectral properties, the particular value  $H = 1/2$  delineates two domains with distinct behaviours.

- For  $0 < H < 1/2$ , the power spectral density  $S_H(0) = 0$ , and the spectrum is high-pass.
- For  $1/2 < H < 1$ , we have  $S_H(0) = \infty$  with a  $1/f$ -type spectral divergence.

Over the Nyquist frequency band it holds the following relation in log-log coordinates

$$\log S_H(f) \sim (1 - 2H) \log |f| + C, \quad (\text{B.4})$$

for most frequencies  $-1/2 \leq f \leq 1/2$  [18].

# Appendix C

## Hurst Exponent of fGn and fBm noises: EMD based estimation

An ensemble of 4000 synthetically generated time series of a given Hurst exponent was generated, following the procedure described in [82] and [18], with the purpose of evaluate the characteristics of IMFs obtained decomposing such processes with EMD and to test methods for Hurst exponent estimation.

Differently from [82], where fGn processes were generated via the Wood and Chen algorithms [83], in this simulation fBm like time series of a given Hurst exponent were generated using the algorithm described in [84] and then differentiated to obtain fGn time series, as explained in Appendix B.

An example of the generated fractional noises, with Hurst exponent spanning from 0.1 to 0.9 is shown in Figure C.1.

The IMFs showing band pass behaviour (index  $k > 2$ ) have spectra with similar shape, apart from some shift in abscissa and ordinate [82]. This allows to look for self similarity in IMFs spectra, according to the relation

$$S_{k',H}(f) = \rho_H^{\alpha_H(k'-k)} S_{k,H}(f \rho_H^{k'-k}), \quad (\text{C.1})$$

where  $2 \leq k \leq k_{min}$ ,  $k_{min}$  being the smallest number of IMFs common to all the time series of the ensemble.

The average spectra of the ensemble generated time series and of their pass band IMFs are shown in Figure C.2, where the self-similarity property is evident.

Due to the self-similarity property, for a given  $H$  the following relationships can be derived among the mode number  $k$ , the mode empirical variance  $V_H[k]$  and its mean period  $\bar{T}_H[k]$

$$V_H[k] = C \rho_H^{2(H-1)k} \simeq C 2^{2(H-1)k} \quad (\text{C.2})$$

$$V_H[k] = C' (\bar{T}_H[k])^{2(H-1)} \quad (\text{C.3})$$

even though  $\rho_H$  is reported to be slightly increasing with the Hurst exponent. The empirical variance of an IMF of index  $k$  and Hurst exponent  $H$ , namely  $d_{k,H}[n]$  is given by

$$V_H[k] = \frac{1}{N} \sum_{n=1}^N (d_{k,H}[n])^2. \quad (\text{C.4})$$

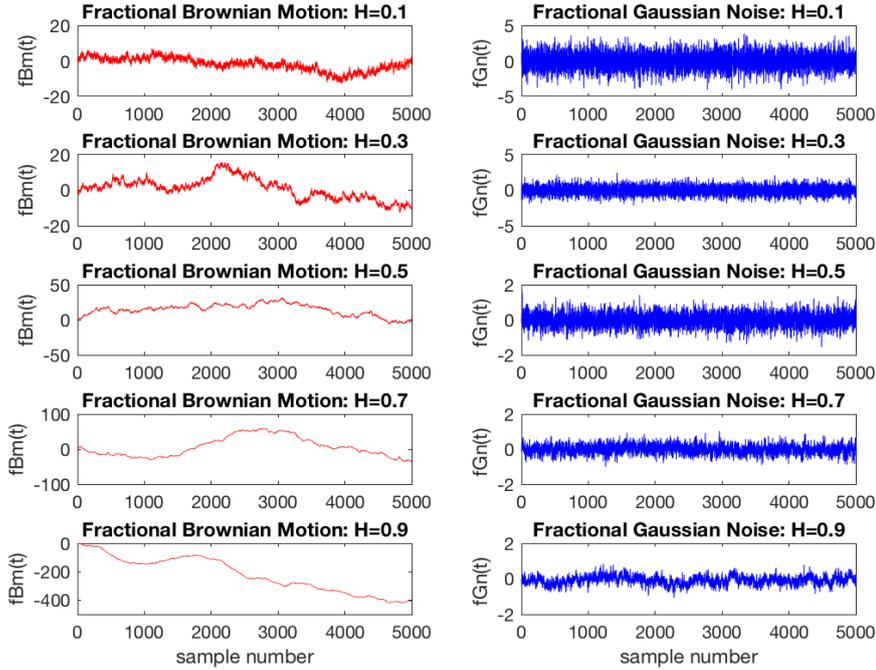


Figure C.1: Fractional Brownian motions (fBm) and fractional Gaussian Noises (fGn), with Hurst exponent spanning from 0.1 to 0.9, at steps of 0.2. Data length is  $N = 5 \cdot 10^3$

As described in [82] and [18], due to the peculiar properties of IMFs, namely that extrema appear as alternating local minima and local maxima separated by only one zero-crossing, their mean frequency (period) can be calculated counting the number of zero crossings or alternatively counting the number of their extrema. The latter approach has been chosen in this study, hence the period of an IMF has been evaluated as the mean time difference occurring in between adjacent maxima. The empirical relations introduced above, as a function of a given Hurst exponent, have been used to test the ability of the EMD algorithm to estimate  $H$ . Results are collected in Figure C.2, where the average spectra of the noises, for different values of  $H$ , are reported in blue and are compared with the ensemble average spectra of the IMFs, in which EMD decomposes them.

The EMD algorithm, although based on no "a priori" assumption, like Fourier basis or wavelet mother function, correctly estimates Hurst exponent for  $H \geq 0.5$ , while it overestimates it for  $H \leq 0.5$ . This is due to the fact that for  $H \leq 0.5$  some of the IMFs act as an "active filter", amplifying the lower frequency band [82]. This behaviour was also found in the simulation presented in this study, since the IMFs spectra (green) raise above the corresponding spectra of the full fGn process as can be seen in Figure C.2 for  $H < 0.5$ .

This behaviour should be taken into account when performing EMD analysis of signals containing mostly high frequency, as remarked in [82].

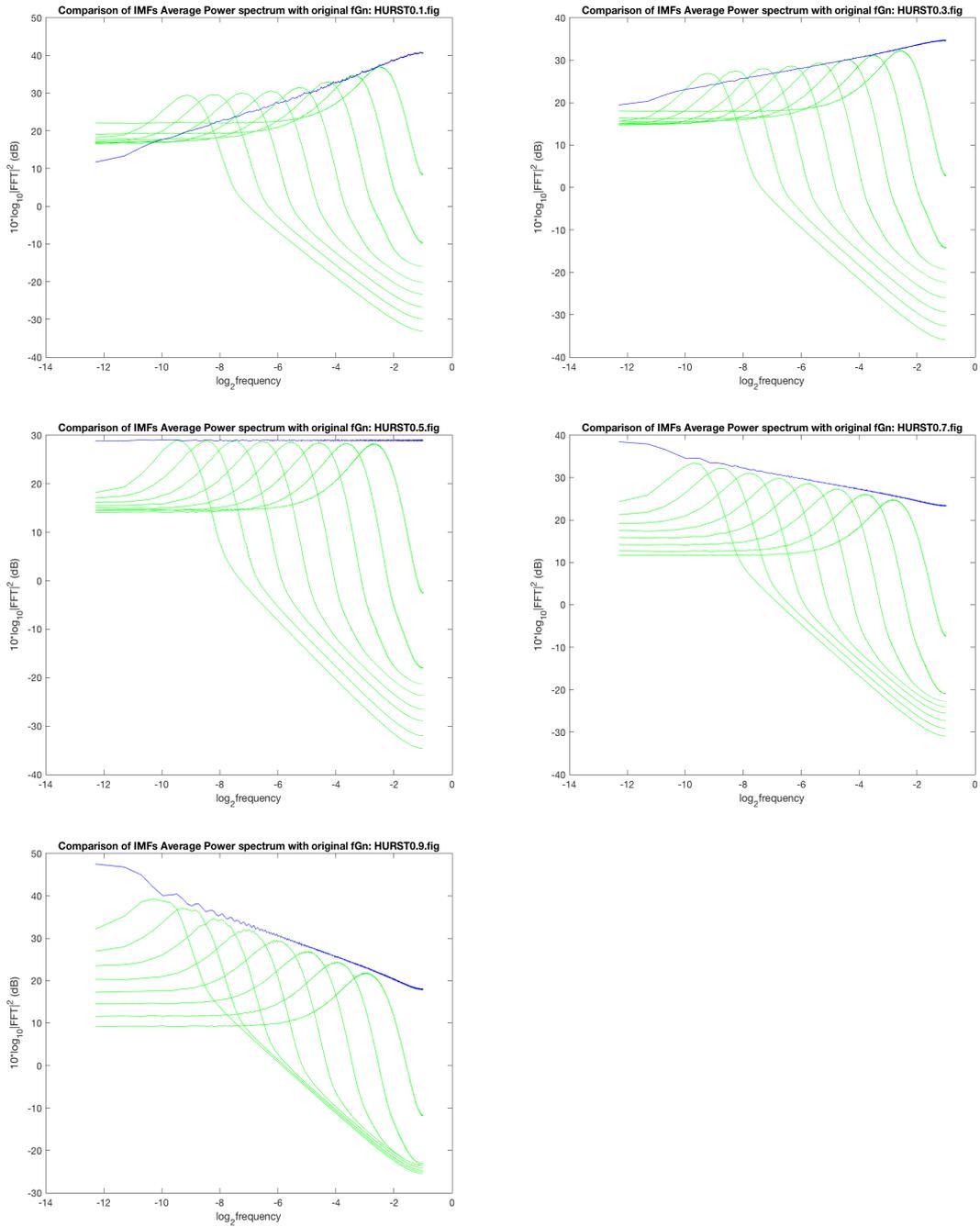


Figure C.2: Average PSD of fGn time series, in blu, compared to average spectra of band pass IMFs, in green. The selfsimilarity is in this case evident and except from the case  $H = 0.1$  and  $H = 0.3$  in all other cases the correct  $H$  can be obtained from the spectra of the extracted IMFs.

# Appendix D

## Bedrosian Product Theorem

Although the Hilbert transform exists for any function of  $L^p$  class, the phase function of the transformed function will not always yield a physically meaningful instantaneous frequency.

The EMD algorithm helps in this regard since reducing the data into IMFs, i.e. oscillatory functions having equal number of extrema and zero crossings and zero mean, improves the chance of getting a meaningful instantaneous frequency. Though, obtaining IMFs satisfies only a necessary condition, since additional limitations are given by two theorems, the Bedrosian Product Theorem (BPT) and the Nuttall theorem [18]. The BPT theorem is briefly described in this Appendix while more details can be found in [85]. The BPT has important implications in the computation of the instantaneous frequency, since this has to be computed from the phase function of the analytic signal, already defined in Equation 2.2 and here reported for clarity

$$z(t) = x(t) + iH[x(t)] = a(t)e^{i\phi(t)} = a(t)\cos[\phi(t)] + ia(t)\sin[\phi(t)] \quad (\text{D.1})$$

For Equation D.1 to hold the Bedrosian Product Theorem [85, 18] must be respected.

The Bedrosian theorem states that the Hilbert transform for the product of two functions  $f(t)$  and  $h(t)$  can be written as

$$H[f(t)h(t)] = f(t)H[h(t)], \quad (\text{D.2})$$

where  $H[\ ]$  is the Hilbert Transform, only if the Fourier spectra for  $f(t)$  and  $h(t)$  are totally disjoint in frequency space. Furthermore the frequency range of the spectrum for  $h(t)$  should be higher than that of  $f(t)$ .

Hence according to the BPT, D.1 is true only if the amplitude  $a(t)$  is varying so slowly that the frequency spectra of the envelope and the carrier waves are disjoint. In this case

$$H[x(t)] = H[a(t)\cos\phi(t)] = a(t)H[\cos\phi(t)] = a(t)\sin\phi(t) \quad (\text{D.3})$$

The proof of the BPT is reported in [85].

# Appendix E

## B-splines

In the tvf-EMD algorithm, used for data analysis in this Thesis, B-splines are used as a filter with time varying frequency cut off. B-splines and their properties are hereafter briefly described.

Polynomial splines are especially useful to consider image data as a continuum rather than a discrete array of pixels. Considering an image as a continuously defined function is also often desirable for feature extraction such as contour detection [86].

Schoenberg, who invented splines, also proved that any uniform spline can be represented in terms of a B-spline expansion, which uses shifted B-spline basis functions. This B-spline representation turns out to be most convenient for performing signal and image processing [86].

Splines were first described in 1946 by Schoenberg [87]. He showed how splines can be used to interpolate equally spaced samples of a function and B-splines are introduced as the basic functions by which polynomial splines are constructed [87, 88]. Splines are piecewise polynomials with pieces that are smoothly connected together. The joining points of the polynomials are called knots. For a spline of degree  $n$ , each segment is a polynomial of degree  $n$ . Then,  $n + 1$  coefficients are needed to describe each piece. However, since there is an additional smoothness constraint that imposes the continuity of the spline and its derivatives up to order  $(n - 1)$  at the knots, there is only one degree of freedom per segment [88].

Hereafter only splines with uniform knots and unit spacing are considered.

The remarkable result, due to Schoenberg [87], is that polynomial splines  $s(x)$  are uniquely characterised in terms of a B-spline expansion

$$s(x) = \sum_{k \in \mathbb{Z}} c(k) \beta^n(x - k) \tag{E.1}$$

which involves the integer shifts of the central B-spline of degree  $n$  denoted by  $\beta^n(x)$ . The parameters of the model are the B-spline coefficients  $c(k)$ .

As described in [88], B-splines are symmetrical, bell-shaped functions constructed from the  $(n + 1)$ -fold convolution of a rectangular pulse  $\beta^0(x)$

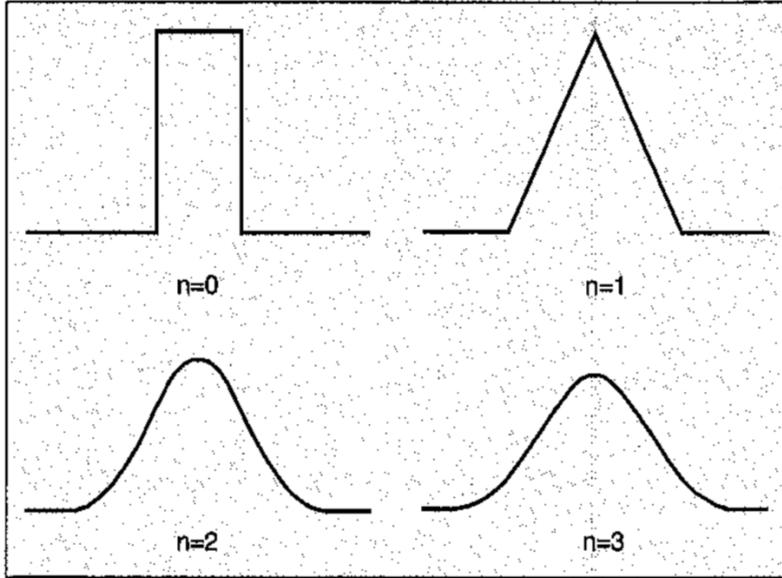


Figure E.1: The B-splines of degrees 0 to 3 are shown. Figure taken from [88]

$$\beta^0(x) = \begin{cases} 1, & \text{if } -\frac{1}{2} < x < \frac{1}{2} \\ \frac{1}{2}, & \text{if } |x| = \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

$$\beta^n(x) = \beta^0(x) * \beta^0(x) * \dots * \beta^0(x) \quad (\text{E.2})$$

where the convolution is carried out  $n + 1$  times. Each spline is uniquely described by its sequence of B-spline coefficients  $c(k)$ , which has the structure of a discrete signal, even though the model is continuous [88]. A time domain expression for the B-spline of order  $n$  is obtained in [88]

$$\beta^n(x) = \frac{1}{n!} \sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k \left(x - k + \frac{n+1}{2}\right)_+^n \quad (\text{E.3})$$

where the  $+$  indicated the one sided power function. This result shows that  $\beta^n(x)$  is a piecewise polynomial of degree  $n$  and that is differentiable up to order  $n$ .

# Appendix F

## Scattered Light Noise Hunting at Virgo with Empirical Mode Decomposition

In the framework of this Thesis, Empirical Mode Decomposition was applied to Virgo data also for scattered light noise hunting, i.e. for detector characterisation purposes, following the approach described in [89].

Scattered light noise diminish the sensitivity of interferometric detectors and can occur when a fraction of laser light is diffused by moving reflective surfaces, e.g. mirrors located along the beam path Z, and then couples back to the main laser beam. This effect can be seen in Figure F.1, where the Fourier spectra of DARM, the differential arm motion, with and without scattered light are shown.

The recombined scattered light forms arch-shaped figures or fringes visible in DARM spectrograms. The scattered light phase angle after reflecting once from the scattering surface is given by [89]

$$\phi_{scattering}(t) = 2\frac{2\pi}{\lambda}(x_0 + \delta x_{surface}(t)) \quad (\text{F.1})$$

where  $x_0$  is the static optical path, as sensed by Position Sensing Devices,  $\lambda$  is the laser wavelength and  $\delta x_{surface}(t)$  is the position of the moving object.

From the phase angle the predictors can be obtained, i.e. the fringes frequency, computed with time derivative of Equation F.1

$$f_{fringe}(t) = \frac{2}{\lambda}|v_{surface}(t)| \quad (\text{F.2})$$

where  $v_{surface}(t)$  is the velocity at which scattering surface is moving.

In case of multiple reflections, arches in DARM spectrograms at frequencies multiple of  $f_{fringe}(t)$  appear since when light is reflected  $N$ -times before recombining the phase angle is

$$\phi_{scattering}(t) = 2N\frac{2\pi}{\lambda}(x_0 + \delta x_{surface}(t)) \quad (\text{F.3})$$

And this generates  $N$  fringes in the DARM spectrograms, with frequencies

$$f_{fringeN}(t) = Nf_{fringe}(t) \quad (\text{F.4})$$

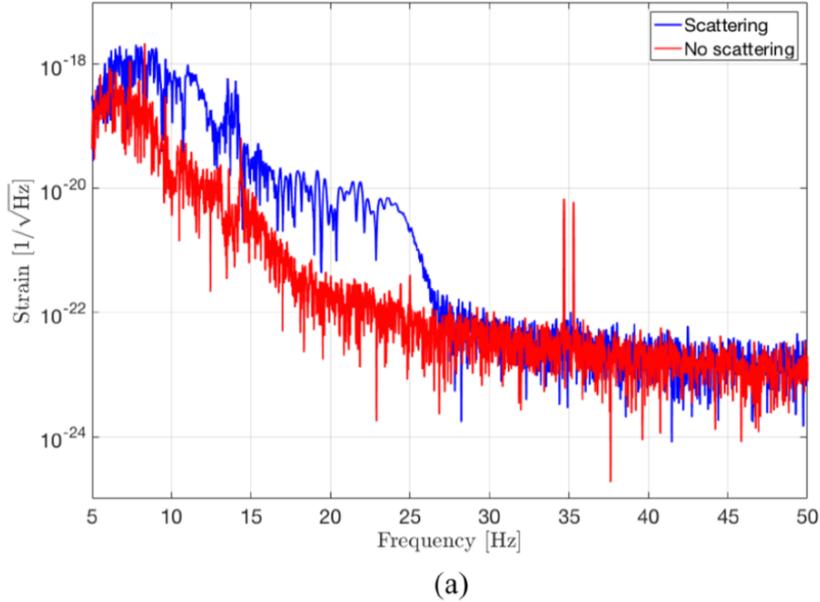


Figure F.1: Fourier spectra of DARM degree of freedom, with and without scattered light noise affecting the sensitivity. Figure from [89]

As described in [89] Decomposing DARM time series with EMD and selecting the appropriate IMF allows to correlate the predictors, which are computed based on Equation F.2, from auxiliary channels sensing the positions of the reflecting surfaces with the IMF's IA, since these are both functions of time. When applied to a list of possible scatterers this allows to rank the most probable source of scattered light noise.

In this regard, the methodology described in [89] was tested during multiple hardware injections.

- An hardware injection was performed on August 22 2019 h 12:12:30 s when a sinusoidal line having frequency of  $0.1Hz$  was injected on auxiliary channel Sa IB F0 X. Details regarding this injection can be found in logbook entry n. 46744 of Virgo logbook<sup>1</sup>.
- Other noise injections were performed on October 03 2019 h.19:55:30s and 20:04:00s, when a sinusoidal line having frequency of  $0.1Hz$  was injected on auxiliary channel SNEB LC Z. In this case 200s and 120s od data were analysed, respectively.
- Finally, a sinusoidal line having frequency of  $0.1Hz$  was injected on auxiliary channel SWEB LC Z on October 03 at 20:26:30s.

The methodology was tested on all Virgo channels having units of  $\mu m$  and having sampling frequency of  $f_s = 500Hz$ . As can be seen in Figure F.2, the hardware injection, shown in the top panel, causes arch shaped figures to appear in the spectrogram of one of the main degrees of freedom of the ifo, namely the power

<sup>1</sup>Virgo logbook can be found at <https://logbook.virgo-gw.eu/virgo/?c=1>

recycling cavity length (PRCL) as is shown in the bottom panel. The methodology described in this Appendix was applied to test whether in this case is possible to identify the auxiliary channel Sa IB F0 X as the most probable source of scattered light. In this regard, results are shown in figure F.3 where in this case the most probable source of scattered light is correctly identified to be Sa IB F0 X. The predictor from this channel is found to have the highest correlation with the IA of the first IMF obtained decomposing DARM time series with EMD.

In Figure F.4 the other hardware injections on which the tool has been tested are shown. In Figure F.5 the results obtained from the application of the methodology are shown for the hardware injections of Figure F.4. The tool correctly ranks as most probable cause of scattered light the channel SNEB LC Z ground corrected, regarding the injections of October 03 2019 h.19:55:30s and 20:04:00s as can be seen in top and middle panel. Bottom panel shows instead, that for the hardware injection performed on October 03 at 20:26:30s, the most probable scatterer, the one having the highest correlation, is found to be SBE SWEB diff bench MIR Z. High correlations,  $\rho > 0.4$  were found also for channels

- SWEB LC Z err ground corrected
- SWEB LC Z err
- SWEB LC Z
- SBE SWEB SA F0 diff LVDT Z
- SBE SWEB SA F7 diff LVDT Z
- SBE SWEB diff bench MIR Z.

Further tests are needed, in particular testing during periods in which scattered light is affecting the sensitivity and the cause of scattered light noise is not known.

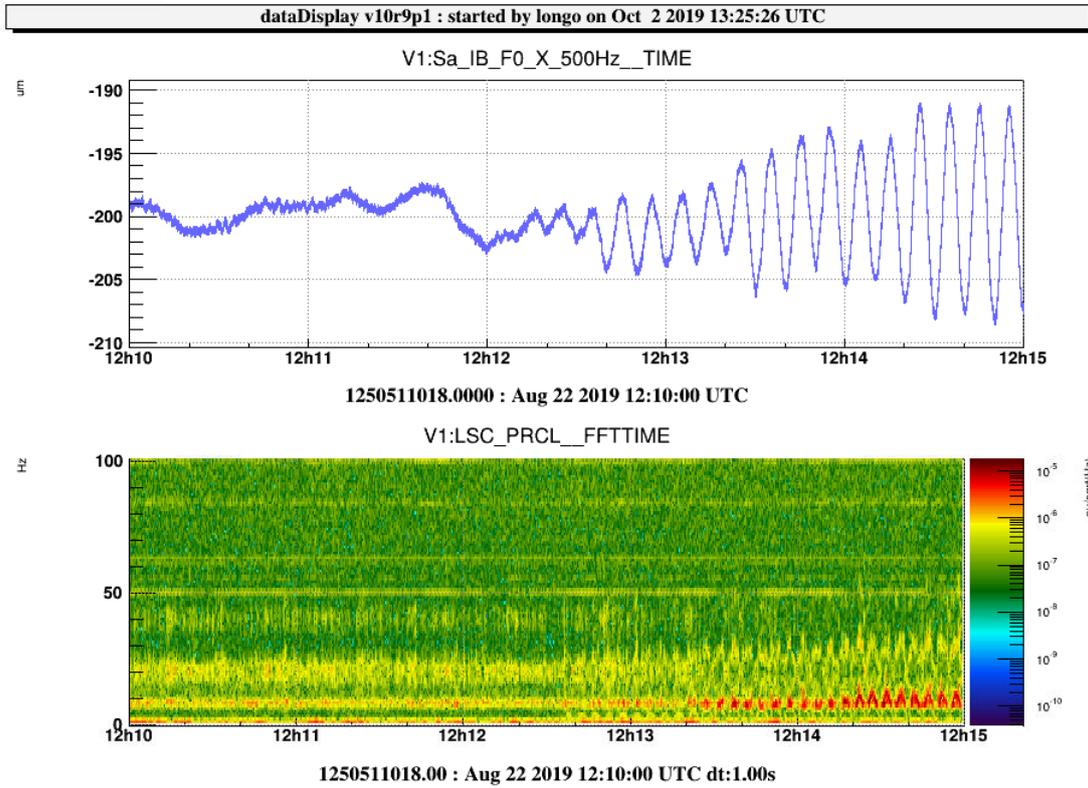


Figure F.2: Hardware injection shown in the top panel where a sinusoidal line is injected in Sa IB F0 X. PRCL spectrogram is shown in the bottom panel, arch shaped fringes are visible in red.

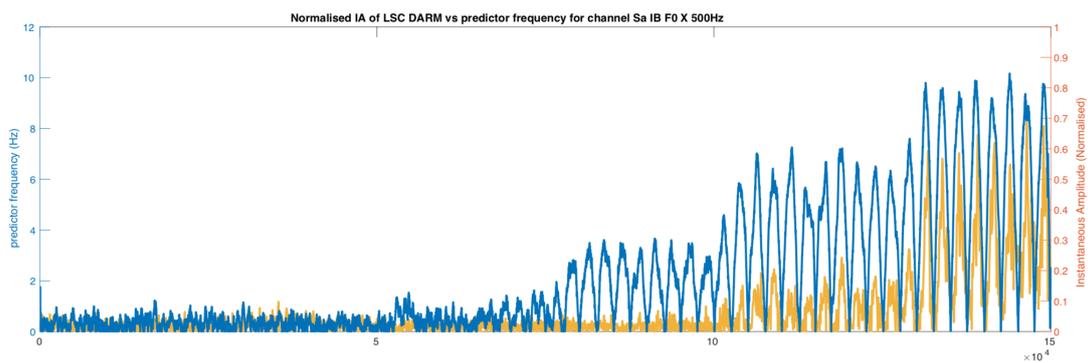
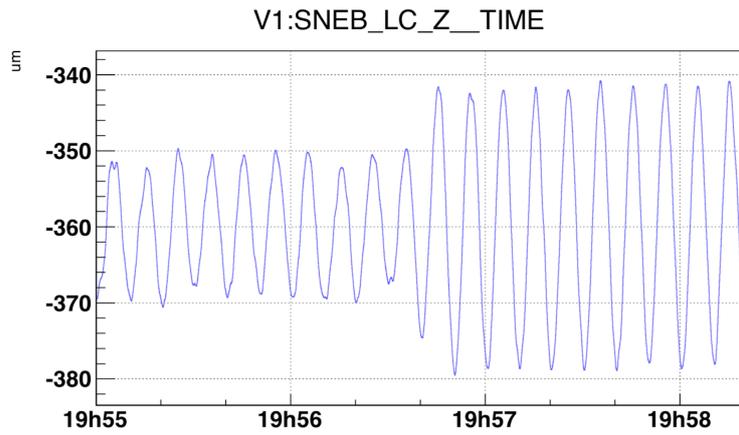


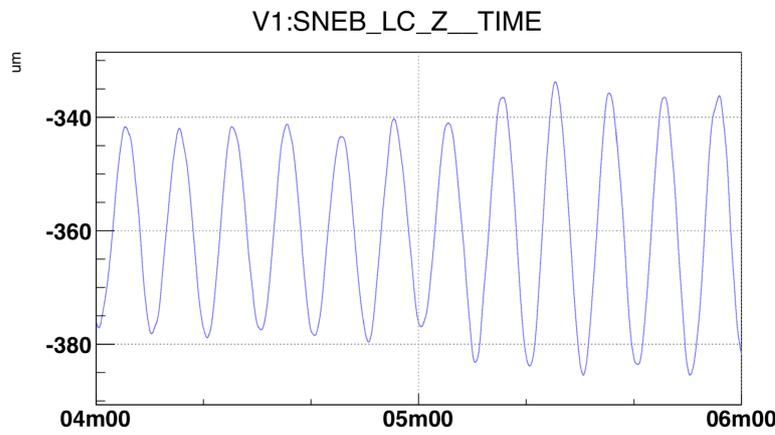
Figure F.3: Shown in yellow is the IA of the first IMF obtained decomposing DARM with EMD while in blue the predictor for channel Sa IB F0 X is shown, which resulted to have the highest correlation among all the auxiliary channels tested.

dataDisplay v10r9p1 : started by longo on Oct 10 2019 08:18:58 UTC



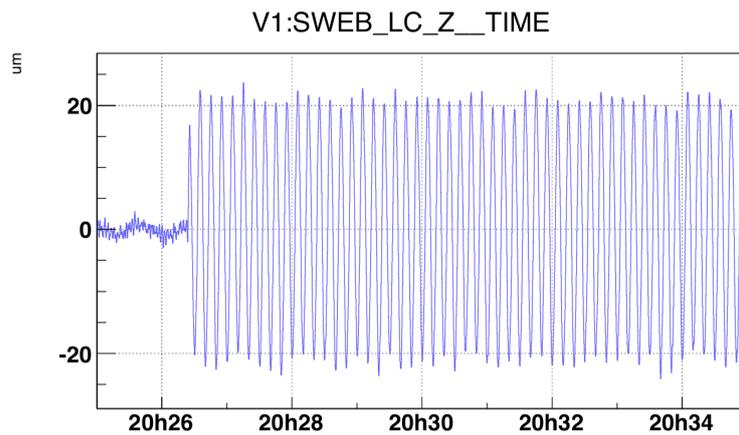
1254167718.0000 : Oct 3 2019 19:55:00 UTC

dataDisplay v10r9p1 : started by longo on Oct 10 2019 08:20:06 UTC



1254168258.0000 : Oct 3 2019 20:04:00 UTC

dataDisplay v10r9p1 : started by longo on Oct 9 2019 16:29:17 UTC



1254169518.0000 : Oct 3 2019 20:25:00 UTC

Figure F.4: The hardware injections performed on 3 October are shown. A sinusoidal line of frequency  $f = 0.1Hz$  was injected in auxiliary channel SNEB LC Z and SWEB LC Z

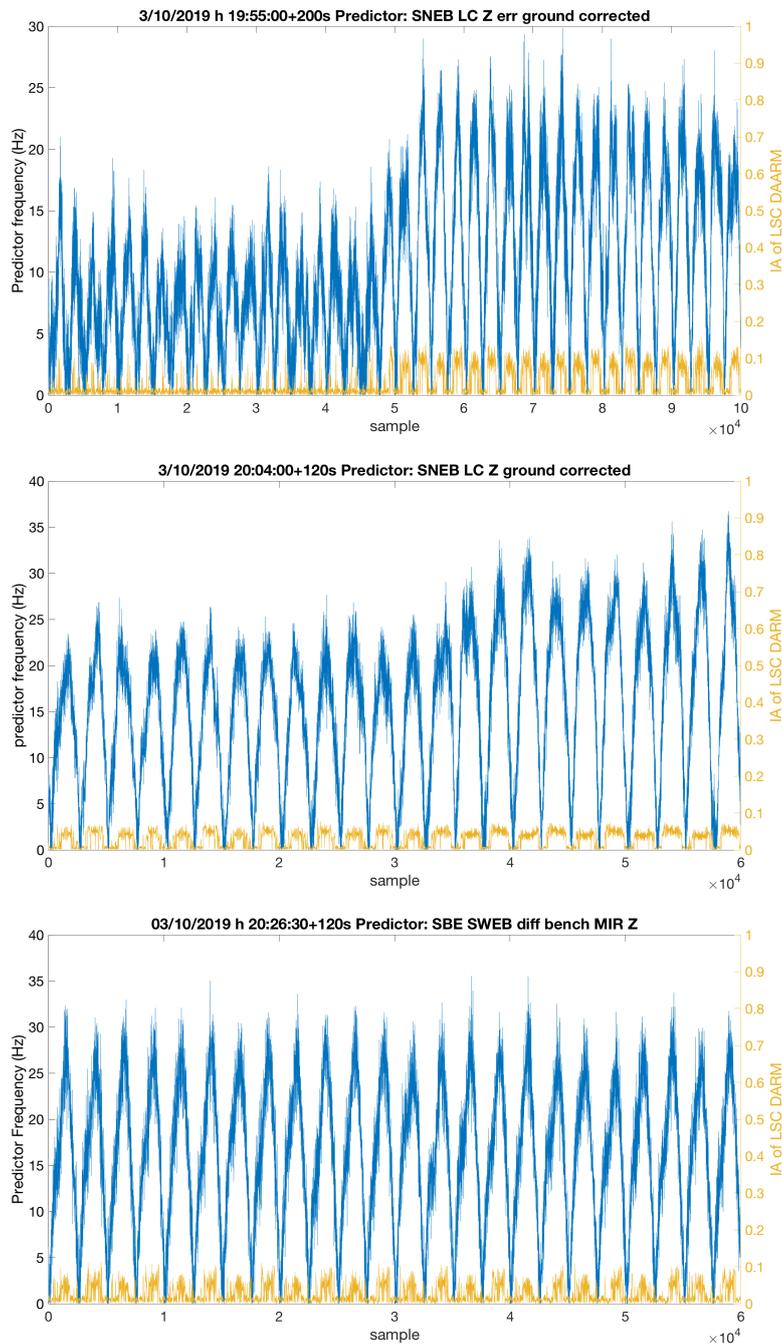


Figure F.5: Results from application of the tool for scattered light noise hunting are shown. In yellow, the IA of the first IMF obtained decomposing DARM with EMD is shown, while in blue the predictor for channels SNEB LC Z ground corrected and SBE SWEB diff bench MIR Z, which resulted to be the most correlated among the ones tested, are shown.

# Appendix G

## $1/f^{2.5}$ Noise Hunting with Multifractal Detrended Fluctuation Analysis: Characterisation of Local Hurst Exponent of C11 data

Fractal analysis was applied in the attempt to identify possible sources of the  $1/f^{2.5}$  noise which was affecting Virgo ifo during the commissioning run C11, started 12/10/2018 h 22:40:55 and lasting until /15/10/2018 h 08:00:00. The duration of C11 is  $\delta t = 206345\text{s}$ . The  $1/f^{2.5}$  noise was a broadband noise severely affecting Virgo sensitivity in the  $20-100\text{Hz}$  frequency band and reducing the range of the ifo of almost  $15\text{Mpc}$ . More details about such noise can be found at <https://wiki.virgo-gw.eu/Commissioning/OneOverFCube>. In the framework of this Thesis, fractal analysis was applied to all auxiliary channels from the standard detector characterisation channel list during the period of C11, with the aim of flagging either noises having a high value of Hurst exponent or that exhibit high standard deviation of  $H$  and that hence have significantly different values of persistency at different times. Data downsampled at  $1\text{Hz}$  were analysed. A simple relation in between the Hurst exponent and the spectral index of a selfsimilar noise of spectral index  $\beta$  holds [79]

$$H = \frac{\beta + 1}{2} \quad (\text{G.1})$$

The procedure of how to compute the local Hurst exponent can be found in [5], where it is explained how the Hurst exponent can be computed locally with a procedure similar to DFA, but sliding windows of small scale, as also described in Chapter 2. Figure G.1 shows some results of the analysis. The culprit are indicated to be sensing and control channels of PR NE WE NI and WI mirrors, which had a variable persistent behaviour during the analysed period, having a high standard deviation in their Hurst exponent. Comprehensive results can be instead found at <https://tds.virgo-gw.eu/?content=3&r=14837>. Further research is needed in order to test whether the local Hurst exponent could be used to monitor stable conditions of the ifo, possibly testing the algorithm on many auxiliary channels at higher sampling rates, during stable conditions and around unlocks of the detector.



# Bibliography

- [1] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, *et al.*, “Advanced Virgo: a second-generation interferometric gravitational wave detector,” *Classical and Quantum Gravity*, vol. 32, no. 2, p. 024001, 2014.
- [2] D. Fiorucci, J. Harms, M. Barsuglia, I. Fiori, and F. Paoletti, “Impact of infrasound atmospheric noise on gravity detectors used for astrophysical and geophysical applications,” *Physical Review D*, vol. 97, no. 6, p. 062003, 2018.
- [3] J. Harms, “Terrestrial gravity fluctuations,” *Living reviews in relativity*, vol. 18, no. 1, p. 3, 2015.
- [4] F. Robinet, “Omicron: an algorithm to detect and characterize transient events in gravitational-wave detectors,” tech. rep., Tech. Rep, 2015.
- [5] E. A. F. E. Ihlen, “Introduction to multifractal detrended fluctuation analysis in Matlab,” *Frontiers in physiology*, vol. 3, p. 141, 2012.
- [6] H. Li, Z. Li, and W. Mo, “A time varying filter approach for empirical mode decomposition,” *Signal Processing*, vol. 138, pp. 146–158, 2017.
- [7] C. Doering and P. Saey, “Hadley cell influence on  $^7\text{Be}$  activity concentrations at Australian mainland IMS radionuclide particulate stations,” *Journal of environmental radioactivity*, vol. 127, pp. 88–94, 2014.
- [8] T. Accadia, F. Acernese, M. Alshourbagy, P. Amico, F. Antonucci, S. Aoudia, N. Arnaud, C. Arnault, K. Arun, P. Astone, *et al.*, “Virgo: a laser interferometer to detect gravitational waves,” *Journal of Instrumentation*, vol. 7, no. 03, p. P03012, 2012.
- [9] B. Caron, A. Dominjon, C. Drezen, R. Flaminio, X. Grave, F. Marion, L. Massonnet, C. Mehmél, R. Morand, B. Mours, *et al.*, “The Virgo interferometer,” *Classical and Quantum Gravity*, vol. 14, no. 6, p. 1461, 1997.
- [10] M. Abernathy, F. Acernese, P. Ajith, B. Allen, P. Amaro Seoane, N. Andersson, S. Aoudia, P. Astone, B. Krishnan, L. Barack, *et al.*, “Einstein gravitational wave Telescope conceptual design study,” 2011.
- [11] A. Einstein, “Näherungsweise Integration der Feldgleichungen der Gravitation, 22 Jun 1916,” 1916.

- [12] R. Weiss, “Electromagnetically coupled broadband gravitational antenna. Quarterly Progress Report 105, Research Laboratory of Electronics,” 1972.
- [13] G. Ballardini, L. Bracci, S. Braccini, C. Bradaschia, C. Casciano, G. Calamai, R. Cavalieri, R. Cecchi, G. Cella, E. Cuoco, *et al.*, “Measurement of the VIRGO superattenuator performance for seismic noise suppression,” *Review of Scientific Instruments*, vol. 72, no. 9, pp. 3643–3652, 2001.
- [14] V. Collaboration *et al.*, “Advanced Virgo Technical Design Report, VIR0128A12.”
- [15] M. Beccaria, M. Bernardini, E. Bougleux, S. Braccini, C. Bradaschia, C. Casciano, G. Cella, E. Cuoco, E. D’Ambrosio, G. De Carolis, *et al.*, “Extending the VIRGO gravitational wave detection band down to a few hz: metal blade springs and magnetic antisprings,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 394, no. 3, pp. 397–408, 1997.
- [16] M. Beker, G. Cella, R. DeSalvo, M. Doets, H. Grote, J. Harms, E. Hennes, V. Mandic, D. Rabeling, J. van den Brand, *et al.*, “Improving the sensitivity of future GW observatories in the 1-10 hz band: Newtonian and seismic noise,” *General Relativity and Gravitation*, vol. 43, no. 2, pp. 623–656, 2011.
- [17] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, and H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [18] N. Huang, *Hilbert-Huang transform and its applications*, vol. 16. World Scientific, 2014.
- [19] A. Sundar, V. Pahwa, C. Das, M. Deshmukh, and N. Robinson, “A comprehensive assessment of the performance of modern algorithms for enhancement of digital volume pulse signals,” *International Journal of Pharma Medicine and Biological Sciences*, vol. 5, no. 1, p. 91, 2016.
- [20] P. Luukko, J. Helske, and E. Räsänen, “Introducing libeemd: A program package for performing the ensemble empirical mode decomposition,” *Computational Statistics*, vol. 31, no. 2, pp. 545–557, 2016.
- [21] J. Lu, X. Chen, and S. Feng, “A GPS Time Series Prediction Model Based on CEEMD,” *Journal of Advances in Computer Networks*, vol. 4, no. 1, 2016.
- [22] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal. i. Fundamentals,” *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, 1992.
- [23] D. Gabor, “Theory of communication. Part 1: The analysis of information,” *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.

- [24] J. Ville, "Theorie et application de la notion de signal analytique," *Câbles et transmissions*, vol. 2, no. 1, pp. 61–74, 1948.
- [25] N. E. Huang, K. Hu, A. C. Yang, H.-C. Chang, D. Jia, W.-K. Liang, J. R. Yeh, C.-L. Kao, C.-H. Juan, C. K. Peng, *et al.*, "On Holo-Hilbert spectral analysis: a full informational spectral representation for nonlinear and non-stationary data," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150206, 2016.
- [26] A. R. Rao and E.-C. Hsu, *Hilbert-Huang transform analysis of hydrological and environmental time series*, vol. 60. Springer Science & Business Media, 2008.
- [27] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomedical Signal Processing and Control*, vol. 14, pp. 19–29, 2014.
- [28] S. G. F. P. Torres ME, Colominas MA, "A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise,"
- [29] Q. Chen, N. Huang, S. Riemenschneider, and Y. Xu, "A B-spline approach for empirical mode decompositions," *Advances in computational mathematics*, vol. 24, no. 1-4, pp. 171–195, 2006.
- [30] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing. i. Theory," *IEEE transactions on signal processing*, vol. 41, no. 2, pp. 821–833, 1993.
- [31] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing. ii. Efficiency design and applications," *IEEE transactions on signal processing*, vol. 41, no. 2, pp. 834–848, 1993.
- [32] B. Xuan, Q. Xie, and S. Peng, "EMD sifting based on bandwidth," *IEEE signal processing letters*, vol. 14, no. 8, pp. 537–540, 2007.
- [33] L. Cohen and C. Lee, "Instantaneous frequency, its standard deviation and multicomponent signals," in *Advanced Algorithms and Architectures for Signal Processing III*, vol. 975, pp. 186–209, International Society for Optics and Photonics, 1988.
- [34] L. Cohen and C. Lee, "Instantaneous bandwidth for signals and spectrogram," in *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2451–2454, IEEE, 1990.
- [35] P. J. Loughlin and K. L. Davidson, "Modified Cohen-Lee time-frequency distributions and instantaneous bandwidth of multicomponent signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1153–1165, 2001.
- [36] G. Jones and B. Boashash, "Instantaneous frequency, instantaneous bandwidth and the analysis of multicomponent signals," in *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2467–2470, IEEE, 1990.

- [37] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, “Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 82–87, 1995.
- [38] A. Longo, S. Bianchi, and W. Plastino, “Xenon and radon time series analysis: A new methodological approach for characterising the local scale effects at CTBT radionuclide network,” *Applied Radiation and Isotopes*, vol. 139, pp. 209–216, 2018.
- [39] S. Bianchi, A. Longo, and W. Plastino, “A new methodological approach for worldwide beryllium-7 time series analysis,” *Physica A: Statistical Mechanics and its Applications*, vol. 501, pp. 377–387, 2018.
- [40] S. Bianchi, A. Longo, W. Plastino, and P. Povinec, “Evaluation of  $^7\text{Be}$  and  $^{133}\text{Xe}$  atmospheric radioactivity time series measured at four CTBTO radionuclide stations,” *Applied Radiation and Isotopes*, vol. 132, pp. 24–28, 2018.
- [41] S. Bianchi and W. Plastino, “Uranium time series analysis: A new methodological approach for event screening categorisation,” *Journal of environmental radioactivity*, vol. 183, pp. 37–40, 2018.
- [42] S. Bianchi, W. Plastino, E. Brattich, V. Djurdjevic, A. Longo, M. A. Hernández-Ceballos, D. Sarvan, and J. Ajtić, “Analysis of trends, periodicities, and correlations in the beryllium-7 time series in Northern Europe,” *Applied Radiation and Isotopes*, vol. 148, pp. 160–167, 2019.
- [43] W. Plastino, R. Plenteda, G. Azzari, A. Becker, P. Saey, and G. Wotawa, “Radioxenon time series and meteorological pattern analysis for CTBT event categorisation,” *Pure and applied geophysics*, vol. 167, no. 4-5, pp. 559–573, 2010.
- [44] A. Longo, S. Bianchi, and W. Plastino, “tvf-EMD based time series analysis of  $^7\text{Be}$  sampled at the CTBTO-IMS network,” *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 908–914, 2019.
- [45] M. Holst, O. Sarbach, M. Tiglio, and M. Vallisneri, “The emergence of gravitational wave science: 100 years of development of mathematical theory, detectors, numerical algorithms, and data analysis tools,” *Bulletin of the American Mathematical Society*, vol. 53, no. 4, pp. 513–554, 2016.
- [46] E. Daw, J. Giaime, D. Lormand, M. Lubinski, and J. Zweizig, “Long-term study of the seismic environment at LIGO,” *Classical and Quantum Gravity*, vol. 21, no. 9, p. 2255, 2004.
- [47] F. Barone, R. De Rosa, A. Eleuteri, L. Milano, and K. Qipiani, “The environmental monitoring system of VIRGO antenna for gravitational wave detection,” *IEEE Transactions on Nuclear Science*, vol. 49, no. 2, pp. 405–410, 2002.

- [48] J. Van Den Brand, M. Beker, M. Doets, E. Hennes, and D. Rabeling, “Einstein Telescope site selection: Seismic and gravity gradient noise,” vol. 203, no. 1, p. 012076, 2010.
- [49] C. Bradaschia, R. Del Fabbro, A. Di Virgilio, A. Giazotto, H. Kautzky, V. Montelatici, D. Passuello, A. Brilliet, O. Cregut, P. Hello, *et al.*, “The VIRGO project: a wide band antenna for gravitational wave detection,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 289, no. 3, pp. 518–525, 1990.
- [50] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, M. Adamo, C. Adams, T. Adams, P. Addesso, *et al.*, “Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914,” *Classical and Quantum Gravity*, vol. 33, no. 13, p. 134001, 2016.
- [51] T. Accadia, F. Acernese, F. Antonucci, P. Astone, G. Ballardin, F. Barone, M. Barsuglia, A. Basti, T. S. Bauer, M. Beker, *et al.*, “Automatic Alignment system during the second science run of the Virgo interferometer,” *Astroparticle Physics*, vol. 34, no. 6, pp. 327–332, 2011.
- [52] T. Accadia, F. Acernese, F. Antonucci, P. Astone, G. Ballardin, F. Barone, M. Barsuglia, A. Basti, T. S. Bauer, M. Beker, *et al.*, “Performance of the Virgo interferometer longitudinal control system during the second science run,” *Astroparticle Physics*, vol. 34, no. 7, pp. 521–527, 2011.
- [53] T. Accadia, F. Acernese, P. Astone, G. Ballardin, F. Barone, M. Barsuglia, A. Basti, T. S. Bauer, M. Bebronne, M. Beker, *et al.*, “Characterization of the Virgo seismic environment,” *Classical and Quantum Gravity*, vol. 29, no. 2, p. 025005, 2011.
- [54] T. Accadia, F. Acernese, F. Antonucci, P. Astone, G. Ballardin, F. Barone, M. Barsuglia, A. Basti, T. S. Bauer, M. Beker, *et al.*, “Calibration and sensitivity of the Virgo detector during its second science run,” *Classical and Quantum Gravity*, vol. 28, no. 2, p. 025005, 2010.
- [55] F. Acernese, M. Alshourbagy, P. Amico, F. Antonucci, S. Aoudia, K. Arun, P. Astone, S. Avino, L. Baggio, G. Ballardin, *et al.*, “Lock acquisition of the Virgo gravitational wave detector,” *Astroparticle Physics*, vol. 30, no. 1, pp. 29–38, 2008.
- [56] F. Acernese, F. Antonucci, S. Aoudia, K. Arun, P. Astone, G. Ballardin, F. Barone, M. Barsuglia, T. S. Bauer, M. Beker, *et al.*, “Performances of the Virgo interferometer longitudinal control system,” *Astroparticle Physics*, vol. 33, no. 2, pp. 75–80, 2010.
- [57] F. Acernese, P. Amico, M. Alshourbagy, F. Antonucci, S. Aoudia, P. Astone, S. Avino, D. Babusci, G. Ballardin, F. Barone, *et al.*, “Status of Virgo detector,” *Classical and Quantum Gravity*, vol. 24, no. 19, p. S381, 2007.

- [58] E. Cuoco, G. Calamai, L. Fabbroni, G. Losurdo, M. Mazzoni, R. Stanga, and F. Vetrano, “On-line power spectra identification and whitening for the noise in interferometric gravitational wave detectors,” *Classical and Quantum Gravity*, vol. 18, no. 9, p. 1727, 2001.
- [59] J. Driggers, J. Harms, and R. Adhikari, “Subtraction of Newtonian noise using optimized sensor arrays,” *Physical Review D*, vol. 86, no. 10, p. 102001, 2012.
- [60] N. Mukund, M. Coughlin, J. Harms, S. Biscans, J. Warner, A. Pele, K. Thorne, D. Barker, N. Arnaud, F. Donovan, *et al.*, “Predicting surface wave velocities at gravitational wave observatories using archival seismic data,” *arXiv preprint arXiv:1812.05185*, 2018.
- [61] Z. Wu and N. Huang, “A study of the characteristics of white noise using the empirical mode decomposition method,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.
- [62] “Brüel & Kjær microphone model 4193-l-004.” <https://www.bksv.com/en/products/transducers/acoustic/microphones/microphone-preamplifier-combinations/4193-L-004>.
- [63] “Brüel & kjær amplifier NEXUS 2690.” <https://www.bksv.com/media/doc/bp1702.pdf>.
- [64] M. Falxa, D. Fiorucci, I. Fiori, F. Paoletti, J. Harms, and M. Barsuglia, “Acoustic characterization of Advanced Virgo buildings.” <https://tds.virgo-gw.eu/?content=3&r=14737>.
- [65] F. Paoletti, I. Fiori, D. Fiorucci, and M. Falxa, “Reverberation Time (RT60) measurements in NEB.” <https://logbook.virgo-gw.eu/virgo/?r=41940>.
- [66] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, T. Adams, *et al.*, “The characterization of Virgo data and its impact on gravitational-wave searches,” *Classical and Quantum Gravity*, vol. 29, no. 15, p. 155002, 2012.
- [67] “Guralp triaxial broadband seismometer 40T60s.” <http://www.guralp.com/products/instruments/guralp-40-series>.
- [68] M. Beccaria, M. Bernardini, S. Braccini, C. Bradaschia, A. Bozzi, C. Casciano, G. Cella, A. Ciampa, E. Cuoco, G. Curci, *et al.*, “Relevance of Newtonian seismic noise for the VIRGO interferometer sensitivity,” *Classical and Quantum Gravity*, vol. 15, no. 11, p. 3339, 1998.
- [69] C. Doering and R. Akber, “Describing the annual cyclic behaviour of  $^7\text{Be}$  concentrations in surface air in Oceania,” *Journal of environmental radioactivity*, vol. 99, no. 10, pp. 1703–1707, 2008.
- [70] “Certification of IMS particulate radionuclide stations (with guidelines for station installation),” 2007.

- [71] D. Sarvan, Đ. Stratimirović, S. Blesić, V. Djurdjevic, V. Miljković, and J. Ajtić, “Dynamics of beryllium-7 specific activity in relation to meteorological variables, tropopause height, teleconnection indices and sunspot number,” *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 813–823, 2017.
- [72] P. F. Fougere, “On the accuracy of spectrum analysis of red noise processes using maximum entropy and periodogram methods: Simulation studies and application to geophysical data,” *Journal of Geophysical Research: Space Physics*, vol. 90, no. A5, pp. 4355–4366, 1985.
- [73] N. R. Lomb, “Least-squares frequency analysis of unequally spaced data,” *Astrophysics and space science*, vol. 39, no. 2, pp. 447–462, 1976.
- [74] B. Podobnik and H. E. Stanley, “Detrended cross-correlation analysis: a new method for analysing two nonstationary time series,” *Physical review letters*, vol. 100, no. 8, p. 084102, 2008.
- [75] R. Vio, P. Andreani, and A. Biggs, “Unevenly-sampled signals: a general formalism for the Lomb-Scargle periodogram,” *Astronomy & Astrophysics*, vol. 519, p. A85, 2010.
- [76] J. McClellan, T. Parks, and L. Rabiner, “A computer program for designing optimum FIR linear phase digital filters,” *IEEE Transactions on Audio and Electroacoustics*, vol. 21, no. 6, pp. 506–526, 1973.
- [77] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao, “A tutorial introduction to adaptive fractal analysis,” *Frontiers in physiology*, vol. 3, p. 371, 2012.
- [78] M. J. Cannon, D. B. Percival, D. C. Caccia, G. M. Raymond, and J. B. Basingthwaighte, “Evaluating scaled windowed variance methods for estimating the Hurst coefficient of time series,” *Physica A: Statistical Mechanics and its Applications*, vol. 241, no. 3-4, pp. 606–626, 1997.
- [79] A. Eke, P. Herman, L. Kocsis, and L. Kozak, “Fractal characterization of complexity in temporal physiological signals,” *Physiological measurement*, vol. 23, no. 1, p. R1, 2002.
- [80] A. Eke, P. Herman, J. Basingthwaighte, G. Raymond, D. Percival, M. Cannon, I. Balla, and C. Ikrényi, “Physiological time series: distinguishing fractal noises from motions,” *Pflügers Archiv*, vol. 439, no. 4, pp. 403–415, 2000.
- [81] B. B. Mandelbrot and J. W. Van Ness, “Fractional Brownian motions, fractional noises and applications,” *SIAM review*, vol. 10, no. 4, pp. 422–437, 1968.
- [82] G. Rilling, P. Flandrin, and P. Gonçalves, “Empirical mode decomposition, fractional Gaussian noise and Hurst exponent estimation,” in *Proceedings.(ICASSP’05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 4, pp. iv–489, IEEE, 2005.

- [83] G. Chan and A. T. Wood, “Simulation of stationary Gaussian vector fields,” *Statistics and computing*, vol. 9, no. 4, pp. 265–268, 1999.
- [84] P. Abry and F. Sellan, “The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation,” 1996.
- [85] B. E., “A product theorem for Hilbert transform,” *Proceedings of IEEE*, vol. 51, pp. 868–869, 1963.
- [86] M. A. Unser, “Splines: a perfect fit for medical imaging,” in *Medical Imaging 2002: Image Processing*, vol. 4684, pp. 225–237, International Society for Optics and Photonics, 2002.
- [87] I. J. Schoenberg, “Contributions to the problem of approximation of equidistant data by analytic functions. part B. On the problem of osculatory interpolation. A second class of analytic approximation formulae,” *Quarterly of Applied Mathematics*, vol. 4, no. 2, pp. 112–141, 1946.
- [88] M. Unser, “Splines: A perfect fit for signal and image processing,” *IEEE Signal processing magazine*, vol. 16, pp. 22–38, 1999.
- [89] G. Valdes, B. O’Reilly, and M. Diaz, “A Hilbert-Huang transform method for scattering identification in LIGO,” *Classical and Quantum Gravity*, vol. 34, no. 23, p. 235009, 2017.