INSTITUTE OF PHYSICS PUBLISHING CLASSICAL AND QUANTUM GRAVITY

Class. Quantum Grav. **20** (2003) S915–S924 PII: S0264-9381(03)62372-5

Data analysis methods for non-Gaussian, nonstationary and nonlinear features and their application to VIRGO

The VIRGO collaboration

F Acernese[1], P Amico[2,3], N Arnaud[4], D Babusci[5], G Ballardin[6], R Barillé[6], F Barone[1,7], M Barsuglia[4], F Beauville[8], F Bellachia[8], M A Bizouard[4], C Boccara[9], D Boget[8], F Bondu[10], C Bourgoin[6], A Bozzi[6], S Braccini[11], C Bradaschia[11], A Brillet[10], V Brisson[4], L Brocco[12,13], D Buskulic[8], J Cachenaut[10], G Calamai[14,15], E Calloni[1,16], E Campagna[14,17], C Casciano[11], C Cattuto[2,3], F Cavalier[4], S Cavaliere[1], R Cavalieri[11], G Cella[11,18], E Chassande-Mottin[10], F Chollet[8], F Cleva[10], T Cokelaer[10], G Conforto[19], S Cortese[6], J P Coulon[10], E Cuoco[14], V Dattilo[11], P Y Davíd[8], M Davier[4], M De Rosa[1], R De Rosa[1], L Di Fiore[1], A Di Virgilio[11], B Dujardin[10], P Dominici[14,19], A Eleuteri[1], D Enard[6], G Evangelista[1], I Ferrante[11,18], F Fidecaro[11,18], I Fiori[14,19], R Flaminio[8], D Forest[20], J D Fournier[10], L Fournier[8], S Frasca[12,13], F Frasconi[6,11], L Gammaitoni[2,3], P Ganau[20], A Gennai[6], G Gennaro[6], L Giacobone[8], A Giazotto[11], G Giordano[5], C Girard[8], G Gougoulat[8], G M Guidi[14,19], H Heitmann[10], P Hello[4], R Hermel[8], P Heusse[4], L Holloway[11,21], F Honglie[5], M Iannarelli[5], L Journet[8], S Krecklbergh[4], B Lagrange[20], P La Penna[6], M Leliboux[9], B Lieunard[8], T Lomtadze[11], V Loriette[9], G Losurdo[14], M Loupias[6], J M Mackowski[20], E Majorana[11], C N Man[10], F Marchesoni[2,3], F Marion[8], F Martelli[14,19], A Masserot[8], L Massonnet[8], S Mataguez[6], F Menzinger[6], M Mazzoni[14,17], C Michel[20], L Milano[1,16], J L Montorio[20], F Moreau[8], J Moreau[9], N Morgado[20], F Mornet[10], B Mours[8], P Mugnier[8], F Nenci[6], J Pacheco[10], A Pai[10], C Palomba[12], F Paoletti[6,11], A Paoli[6], L Paoli[6], A Pasqualetti[6], R Passaquieti[11,18], D Passuello[11], M Perciballi[12], S Peruzzi[6], B Perniola[14,19], L Pinard[20], R Poggiani[11,18], P Popolizio[6], E Porter[4], S Puccinelli[6], M Punturo[2,3], P Puppo[12], K Qipiani[1], J Ramonet[8], P Rapagnani[12,13], V Reita[9], A Remillieux[20], F Ricci[12,13], F Richard[6], J P Roger[9], P Ruggi[11,18], G Russo[1], S Solimeno[1], R Stanga[14,17], R Taddei[6], J M Teuler[6], E Tournefier[8], F Travasso[2,3], H Trinquet[10], E Turri[5], M Varvella[1,16], D Verkindt[8], F Vetrano[14,19], O Veziant[8], A Viceré[14,19], S Vilalte[8], J Y Vinet[10], H Vocca[2,3], M Yvert[8] and Z Zhang[6]

[1] Istituto Nazionale di Fisica Nucleare, Sez. Napoli, Italy
[2] Istituto Nazionale di Fisica Nucleare, Sez. Perugia, Italy
[3] Dipartimento di Fisica, Università di Perugia, Italy
[4] Laboratoire de l'Accélérateur Linéaire, Orsay, France

[21] Present address: University of Illinois, Urbana, IL, USA.

0264-9381/03/170915+10$30.00

[5] Laboratori Nazionali di Frascati dell'INFN, Frascati, Italy
[6] European Gravitational Observatory, Cascina, Italy
[7] Dipartimento di Scienze Farmaceutiche, Università di Salerno, Italy
[8] Laboratoire d'Annecy-le-Vieux de Physique des Particules, Annecy, France
[9] Ecole Supérieure de Physique et de Chimie Industrielles, Paris, France
[10] Observatoire de la Côte d'Azur, Nice, France
[11] Istituto Nazionale di Fisica Nucleare, Sez. Pisa, Italy
[12] Istituto Nazionale di Fisica Nucleare, Sez. Roma, Italy
[13] Dipartimento di Fisica, Università di Roma 'La Sapienza', Italy
[14] Istituto Nazionale di Fisica Nucleare, Sez Firenze/Urbino, Italy
[15] Osservatorio Astrofisico di Arcetri, Firenze, Italy
[16] Dipartimento di Scienze Fisiche, Università di Napoli 'Federico II', Italy
[17] Dipartimento di Astronomia, Università di Firenze, Italy
[18] Dipartimento di Fisica, Università di Pisa, Italy
[19] Dipartimento di Fisica, Università di Urbino, Italy
[20] Institut de Physique Nucléaire de Lyon, France

E-mail: guidi@fis.uniurb.it and ecm@obs-nice.fr

Received 17 April 2003
Published 18 August 2003
Online at stacks.iop.org/CQG/20/S915

Abstract
The commissioning of the VIRGO central interferometer occasioned the implementation and tests of various algorithms for the characterization of the non-Gaussianity, non-stationarity and non-linearity of the dark fringe data. This library of prototypes will serve as groundwork for the near commissioning of VIRGO (full scale). We make a summary of the activities on that subject including the description of the selected algorithms and some results obtained with the data of the engineering runs.

PACS numbers: 04.80.Nn, 07.05.Kf, 95.75.Wx

(Some figures in this article are in colour only in the electronic version)

Advanced technologies are currently used in the gravitational wave (GW) interferometric detectors [1] to optimize the sensitivity curves or equivalently reduce the noise level to a minimum. Despite this, the noise cannot be entirely suppressed and one can at best expect that it will be stationary and Gaussian. These two requirements greatly simplify the statistical procedures for the detection of GWs. It is therefore important to check if the detector output noise satisfies these nominal properties.

Linearity is another characteristic of importance in the analysis of GW data. The output of the detector is expected to be related linearly to the main noise sources. Internal malfunctions of the apparatus may be detected by checking whether the output data appear to be produced by a nonlinear process.

With these objectives in view, several approaches have been proposed and tested on the commissioning data of the VIRGO central interferometer. We report here on these activities. The paper is divided into three parts concerning the monitoring of the noise stationarity (in section 1) and Gaussianity (in section 2). The last part of section 3 is devoted to a test of the linearity of the data. For each of these items, we give the results of the applications of the proposed algorithms to the data of the fifth engineering run (E4, the first one being labelled E0) of the VIRGO central interferometer [2].

1. Investigating the noise stationarity

We will concentrate here on the 'long-term' non-stationarities of the GW detector noise with typical duration of minutes or hours. Artefacts of shorter duration (of the order of a second) may be detected by looking for transients (see [3] for details on this subject).

1.1. Time–frequency representation with spectrograms

When investigating the long-term stationarity, we are mainly interested in two questions: how much does the noise floor fluctuate? What are the drifting frequency lines if any? Time–frequency methods are natural tools for such problems in the sense that they provide a representation of the variations of the signal spectrum. A large variety of such representations exists [4] but we made the pragmatic choice of solutions based on short-time Fourier analysis because they are computationally tractable when applied to large datasets.

We use the *spectrogram* (squared modulus of the Fourier transform) defined by

$$S(t,\, f) \equiv \left| (1/f_s) \sum x(s) h(t-s)\, \mathrm{e}^{-2\pi \mathrm{i} s f/f_s} \right|^2, \tag{1}$$

where $x(t)$ is the output of the GW detector (i.e., the non-calibrated dark fringe signal as seen after the output mode-cleaner and sampled at frequency f_s) and $h(t)$ is a window (e.g., Hanning type, scaled to unit L^2 norm).

In our set-up, we compute $S(t,\, f)$ using non-overlapping data chunks of about 13.2 s (i.e., chosen so that the number of samples is a power of 2 and the frequency resolution is better than 0.1 Hz, namely 76 mHz) and every $T = 5$ min (i.e., duration of the data in a file), we evaluate the average (which we denote by $\hat{S}(t,\, f)$) of the 23 computed spectra. The resulting two-dimensional matrix can be displayed in the form of a grey-level image as illustrated in figure 1. This gives a panorama of the spectral changes over the whole test-run duration (72 h in the case of E4) where further investigations can be conveniently made.

At various stages of the commissioning of the central interferometer of VIRGO, information extracted from the spectrogram helped us to understand its internal functioning. For instance, the monitoring of the vertical position of the mirror [5] or for discriminating the contributions of seismic origin in the low-frequency (below 10 Hz) noise contents [6] (both contributions used E2 data). Using candidate frequency values (given by a measurement done during a non-operational phase of E4), it was possible [2] to identify and track the thermal modes of the mirror in the spectrogram of the dark-fringe data during operation and thus to have an estimate of the contribution of the thermal noise to the E4 dark-fringe spectrum (see [2] for more details on this analysis). Finally, the fluctuations of the noise floor have been estimated from the averaged spectrogram to be 7.5×10^{-17} m Hz$^{-1/2} \pm 5\%$. We made this measurement using the bandwidth 1.1 kHz $\pm$ 20 Hz because this frequency interval does not contain spurious lines and it is close to the best sensitivity reached during E4. We used the longest period of stable operation which lasts for $\sim$10 h.

1.2. Automated analysis using the two-factor test

In those examples, the useful information has been extracted 'manually' from the time-frequency images. This is not well suited when processing the signal in real-time, motivating the design of a more automated approach based on the work presented in [7]. This approach consists in the computation of a number of statistics which are used to test the non-stationarities of the time series. As the value of these statistics can be updated and their level of significance computed time by time, the test we describe below will give the possibility to detect and

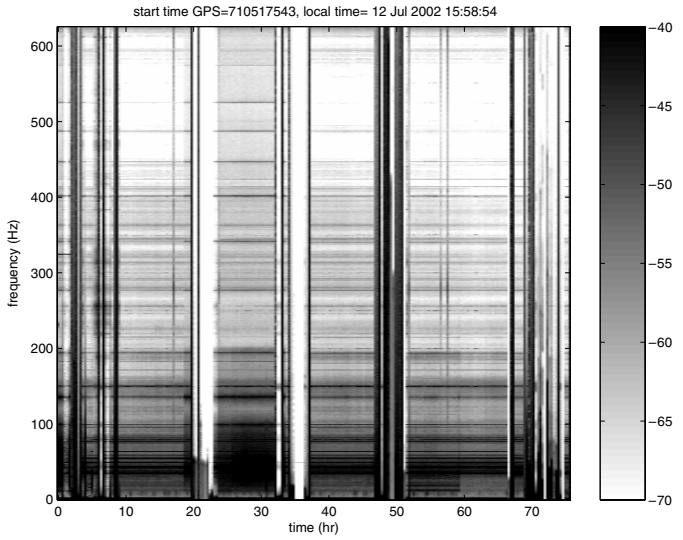

Figure 1. Spectrogram of the dark-fringe signal using E4 data. The time–frequency representation (in dB units) as defined in equation (1) is displayed here as a grey-scaled image. Events appearing as 'vertical' lines at about hours 0–8.4, 19.5–24, 32–37.5, 47–51.7 and 67–end are related to losses of apparatus control (see [2, 12] for details). Outside these periods, the interferometer is kept on the dark fringe in various configurations. For instance, the increase of the noise floor during 24–32 is due to the test of the automatic alignment system (which turned out to be defective) of the input mode cleaner. The transients of large amplitude appearing clearly at hours 17 and 58 have been found to be related to the use of picometers for the output mode-cleaner realignment.

follow the non-stationarities in real-time and eventually act on the instrument if the cause of the abnormal behaviour has been identified. Once the confidence level is fixed by the user, also the false alarm probability is fixed, as the test is based on the χ^2 test.

The averaged spectrogram is computed as before (see equation (1)) but with a slightly different set-up: we use overlapping data chunks, and different values for the width of the window $h(t)$ (chosen such that the process can be considered stationary on it) and for the averaging time T (fixed to a value larger than the width of $h(t)$) [7]. The values of the averaged spectrogram taken at two points of the time-frequency plane (t_1, f_1) and (t_2, f_2) are then uncorrelated if the points are separated sufficiently, namely if $|f_2 - f_1| \gg \Delta_h f$ and $|t_2 - t_1| \gg T$ where $\Delta_h f$ is the frequency bandwidth of the window $h(t)$ [7]. Let us consider a set of time–frequency points $(t_i, f_j)_{i=1\ldots I, j=1\ldots J}$ where the distance condition is satisfied.

In a nutshell, the method consists in testing whether the log-spectrogram values $Y_{ij} \equiv \log(\hat{S}(t_i, f_j))$ are uniformly distributed along the time axis. (The motivation for the logarithm transformation is that it stabilizes [7] the variance σ^2 of the spectral estimate to a known value depending on $\Delta_h f$ and T.)

The test uses the techniques of a two-factor analysis [7] of the variance model. Precisely, we test the null hypothesis (nonstationary signal) $H_0 : Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ij}$ against $H_1 : Y_{ij} = \mu + \beta_j + e_{ij}$ i.e. a model where the log-spectrogram values do not depend upon the time index as expected for stationary signals. Here α_i, β_j and γ_{ij} are dummy variables modelling the variations of the spectral estimate Y_{ij} versus time only, frequency only and time *and* frequency, respectively.

For instance, if the process is uniformly modulated, i.e. if the log-spectrogram Y_{ij} can be written in the form $Y_{ij} = c(i)\, Y_j$, where $c(i)$ is a modulation depending only on time, we

expect the γ_{ij} to be all zero. The e_{ij} are the random residuals due to estimation error which can be assumed i.i.d., Gaussian of variance σ^2, the same variance of the spectral estimate we already know.

To do the test, we compute the mean $\overline{Y_{i\,.}}$ with respect to the frequencies for each time, then we compute the mean square deviation Σ of the overall mean $\overline{Y_{..}}$:

$$\Sigma = J \sum_{i=1}^{I} (\overline{Y_{i\,.}} - \overline{Y_{..}})^2. \tag{2}$$

Under H_1, Σ/σ^2 follows a $\chi^2_{(I-1)}$ distribution. This can be tested with a standard χ^2 test given a confidence level chosen by the user.

This test offers the possibility to detect a nonstationary behaviour due to instabilities in a limited part of the spectrum or to a noise floor drift which interests all the spectrum, because in this case we have an uniform modulation. Indeed, if it turns out that the process is nonstationary, we can test if it is uniformly modulated using the statistic:

$$\Sigma_{IR} = \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \overline{Y_{i\,.}} - \overline{Y_{.\,j}} + \overline{Y_{..}})^2, \tag{3}$$

which behaves as $\sigma^2 \chi^2_{(I-1)(J-1)}$ under H_1. Here $\overline{Y_{.\,j}}$ is the mean with respect to the times for each frequency. When the test on Σ_{IR} proves that the interaction between time and frequencies is significant, the process is non-uniformly modulated, and we can test if the non-stationarity is mainly due to some frequency components. We thus select the index of the frequencies of interest $K = \{j_1, j_2, \ldots, j_k\}$ and use the statistic: $\Sigma' = \sum_{j \in K} \sum_{i=1}^{I} (Y_{ij} - \overline{Y_{.\,j}})^2/\sigma^2$ which behaves like a $\chi^2_{k(I-1)}$ under H_1. A similar χ^2 test is then applied to give the decision result.

This approach has been successfully used on simulated data and applied to some E4 data chunks, for whom we obtained good results when the width of $h(t) \geqslant 2.5 \times 10^{-2}$ s and the time of average $T \geqslant 0.5$ s. By choosing different values for T, we can discriminate between different non-stationarities: for example, the amplitude fluctuations of lines and the slow drift of noise floor. If the drift is slower than the amplitude fluctuations, the former is still detected with a large value of T whereas the latter is invisible.

In figure 2 we present the analysis of an E4 data block where non-stationarity is detected between 10^2 and 10^4 Hz using $T = 0.5$ s. We conclude from the distortion of the level curves at the detection point that the duration of this non-stationarity is about 10 s. To find the duration for each non-stationarity we can perform different tests with different T, as we said, or check the rate of change of the statistics of the test.

A package to implement these procedures on real-time in the next engineering runs is currently being developed.

2. Measuring discrepancies to Gaussianity with the kurtosis

The optimality of the matched filter statistic is intrinsically related to the hypothesis of a Gaussian noise. The various analysis of the current experimental data show that this assumption is not always satisfied and therefore needs to be monitored.

There are many ways for the noise to depart from Gaussianity, however not all of them are relevant for the problem of GW detection. One of the possibilities is to have a noise probability density function (PDF) with heavier tails than the Gaussian bell curve. This particular discrepancy is a problem because it causes an increase of the false detection rate (i.e., the large values in the tails make the matched filter trigger more often).

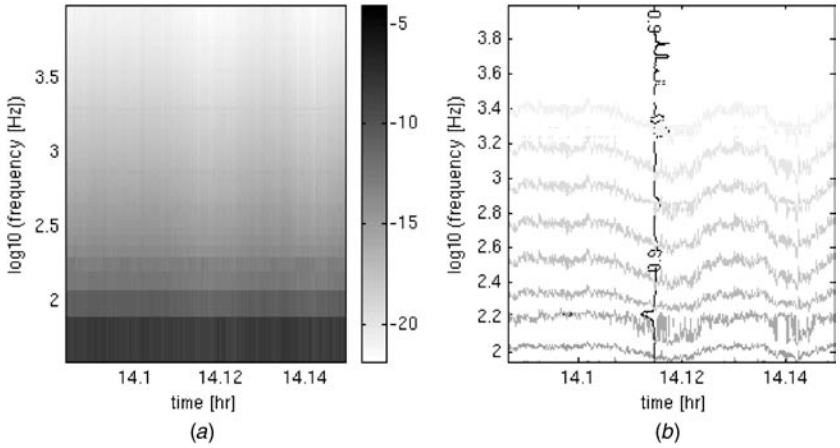

Figure 2. Non-stationarity monitoring index. We illustrate the use of the index presented in section 1 focusing on a few minutes of the E4 run where a nonstationary behaviour is detected. (*a*) Log-spectrogram of the period examined (time resolution of 0.5 s, frequency resolution of 39 Hz). (*b*) Probability resulting from the χ^2 test applied to the statistic defined in equation (2) (calculated with a coarser frequency resolution of 79 Hz). Here only the level curve at the confidence level value fixed to 0.9 is displayed, showing the limit where this probability exceeds the confidence level. It is superimposed onto the contour plot (in grey scale) of the log-spectrogram presented in (*a*). The test detects a nonstationarity occurring around time ≈14.12 h. Note that this spectral change is not obvious in the representation of panel (*a*), but it is clearly in coincidence with fluctuations of the contour plots shown in (*b*).

The kurtosis [8] is a well-known measurement of the decay rate of the PDF in the tails. This motivates us to use the kurtosis as an index measuring normality. The mean defined by $\mu_1(t) \equiv \mathbb{E}[x(t)]$ where $\mathbb{E}[\cdot]$ denotes the expectation operator is the first-order moment of the signal $x(t)$. At higher order $r \geqslant 2$, the central moments [8] of $x(t)$ are given by $\mu_r(t) \equiv \mathbb{E}[(x(t) - \mu_1(t))^r]$. The kurtosis is defined by the following ratio $\kappa_4(t) \equiv \mu_4(t)/(\mu_2(t))^2$. In the nominal assumption where the signal is Gaussian, κ_4 equals 3. If the signal PDF has heavier tails, κ_4 is larger than 3. We are therefore interested in locating excess of kurtosis.

In [9], we propose an estimate $\hat{\kappa}_4(t)$ of the kurtosis using the data selected by a time-sliding exponential window $w(n) = (1 - C_1)^{t-n}$ defined for $n \leqslant t$, and where $0 < C_1 < 1$ sets the size of the window. This estimate can be efficiently computed with the following recursive expression:

$$\hat{\kappa}_4(t) = (1 + C_1 - 2C_1\delta^2 x(t))\,\hat{\kappa}_4(t-1) + C_1\delta^4 x(t), \tag{4}$$

where $\delta^2 x(t) \equiv (x(t) - \hat{\mu}_1(t-1))^2/\hat{\mu}_2(t-1)$ computes a normalized distance of the current signal sample to the mean. The short-time and unbiased estimates of the mean and variance can also be computed recursively with

$$\hat{\mu}_1(t) = (1 - C_1)\hat{\mu}_1(t-1) + C_1 x(t) \tag{5}$$

$$\hat{\mu}_2(t) = (1 - C_1)\hat{\mu}_2(t-1) + C_2(x(t) - \hat{\mu}_1(t-1))^2, \tag{6}$$

where $C_2 = C_1(1 + C_1/2)$. If $C_1 \ll 1$, $\hat{\kappa}_4(t)$ can be shown [9] to have a small bias.

We use $\hat{\kappa}_4(t)$ to estimate the kurtosis of the output of the interferometer, thus generating a data stream of same size as the original signal. Because the storage space allocated to the monitoring data is limited, we need to 'summarize' the information given by $\hat{\kappa}_4(t)$. As

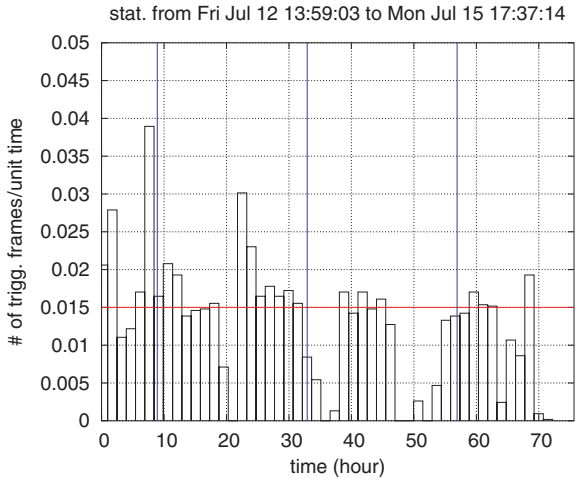

Figure 3. Rate of data frames presenting at least one excess of kurtosis. This rate introduced in section 2 is an indicator of the Gaussianity of the considered data (here, the dark-fringe signal resampled at 50 Hz and cleared of the known lines with a multiple notch filter). Each bar represents the rate estimated using ∼1:30 h of thresholded kurtosis. If the signal is Gaussian, this rate should not exceed 1.5% (a simple calculation shows that this corresponds to a max number of frames presenting an excess of kurtosis of about 80 out of 5400 s (=1:30 h) under Gaussian assumptions). We see that it is roughly the case when the interferometer is locked onto the dark fringe (see figure 1).

mentioned above, we are essentially interested in the excess of kurtosis, i.e., $\hat{\kappa}_4(t) > \eta$ with $\eta > 3$ an arbitrary threshold. Because of the estimation error, there is a non-zero probability that a false excess of kurtosis is observed for a given threshold. Instead of recording the time locations of all excesses of kurtosis (which would therefore include uninteresting false events), our approach is to compute the rate r of frames (i.e., 1 frame is equivalent to 1 GPS second of data) where there is at least one excess of kurtosis and compare this rate with the (maximum) one r_0 obtained by applying the same procedure to a simulated Gaussian noise with a similar spectrum to the processed data (see section 3). For a threshold $\eta = 4$, the maximum rate can be estimated to be $r_0 \approx 1.5\%$ (see [9] for an abacus relating η to the maximum rate). We decide that the signal presents a significant departure from Gaussianity when the measured rate r exceeds the maximum rate r_0. Figure 3 shows the application of this approach to the dark-fringe signal of the E4 run.

Note that the final result is somewhat independent of the threshold (provided that the r_0 is updated when the threshold is changed). However, the requirement that r is sufficiently well estimated put some restrictions on the possible values of η: if η is large, there are too few threshold crossings which lead to a bad estimate of r. Conversely if η is too small, almost all frames present an excess of kurtosis even if the data are Gaussian, i.e., $r_0 \approx 1$, preventing us from distinguishing the Gaussian from the non-Gaussian case. The value $\eta = 4$ ($r_0 = 1.5\%$) is a good compromise.

3. Non-linearity test with surrogate data

As non-linearity cannot be characterized by analysing the noise autocorrelation, which retains only the linear characteristic of a process, it is necessary to apply new concepts and algorithms.

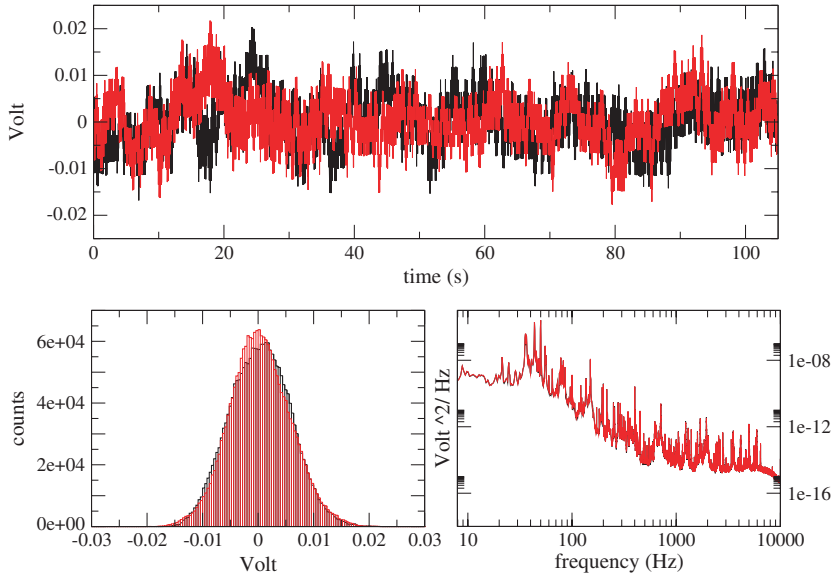

Figure 4. Comparison between real data and a surrogate: in black/dark grey the time evolution, the histogram and the power spectrum of the real data during E4. In red/light grey, the same for a surrogate of the real data.

Our approach to the problem is to test for non-linearity using as statistics the measure of the non-invertibility of a time series:

$$\Phi(\tau) = \frac{1}{N-\tau} \sum_{t=\tau+1}^{N} (x(t) - x(t-\tau))^3. \tag{7}$$

$\Phi(\tau)$ is a nonlinear observable for which we do not know how to find analytically the confidence intervals for a linear process, as we do not know this PDF. We thus have to estimate the PDF, and hence the confidence intervals, generating many realizations of the data that retain only their linear characteristics.

Surrogate data are in fact constructed to have the same linear characteristics as the real data—and hence the same power spectrum [10]. They have no dynamical non-linearities and are Gaussian and stationary: they can thus be useful to estimate the PDF of a statistic in the absence of nonlinearities. The algorithm to produce surrogates is illustrated below:

- calculate the discrete Fourier transform (DFT) of the test data;
- randomize the phase (uniformly distributed in $[0, 2\pi]$);
- take the inverse DFT.

In figure 4, we compare a real dataset for the E4 run and one of the surrogates. We conclude that the surrogate retains the same time characteristics as the real data (upper panel) and that they have the same histogram and power spectrum (lower panels).

The null hypothesis of our test is that the data are generated by a stationary Gaussian linear stochastic process. We perform a two-sided rank non-parametric test using the surrogates [11]:

- we choose the confidence level α of the test;
- we create $S = 2/\alpha - 1$ surrogates from the real data;
- we calculate $\Phi(\tau)$ on the real-data block and on the S surrogate sets;

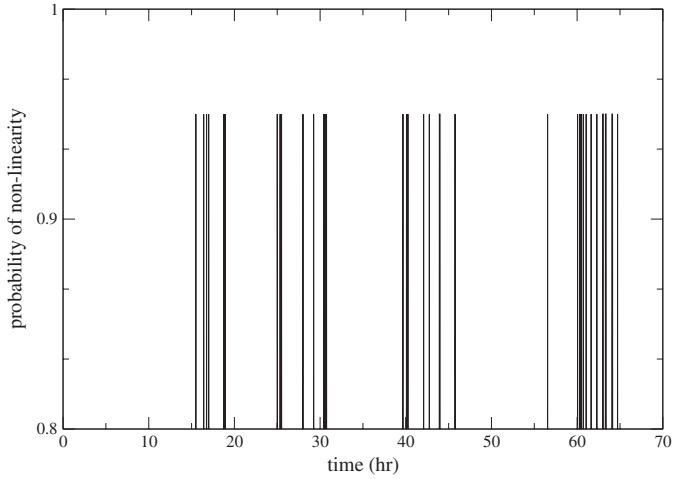

Figure 5. Test of nonlinearity of the E4 data: the bars indicate the time instants during E4 where the test finds nonlinear effects at 95% confidence level. This computation was made only in periods where the interferometer was in the lock state. In the long periods during which the test does not trigger, the lock state is not acquired (see figure 1).

- we check if $\Phi(\tau)$ for real data is outside the interval of the values calculated using surrogates.

In figure 5 we show the result of a test fixing $\alpha = 0.05$ for the E4 run. It is obtained by generating $S = 39$ surrogates for data blocks of 100 s taken each 5 min. The test seems to indicate that nonlinear processes are quite common during the run, but it appears that it is difficult to distinguish them from non-stationarities since the test is also sensitive to these as explained by the following comment.

It is known that the phase randomization of the Fourier transform 'stationarizes' the time series. The statistic of equation (7) can then have a value for surrogate data very different from the value for nonstationary data, so that the test can be positive even if real data are linear. For this reason, the test for non-linearity should be used after having analysed the stationarity of the time series.

4. Concluding remarks

Any methods aiming to analyse/characterize non-stationarity, non-Gaussianity or non-linearity unavoidably suffer from the under-determination of these characteristics. In contrast to their associated and antinomic properties which are well defined, the non-properties are intrinsically multiple and so are the analysis tools. We presented and tested three approaches to address some of the facets of the problem.

Each approach aims to give information about a single departure from one of the three considered non-properties (e.g., check the tails of the PDF in section 2), focusing on discrepancies of interest for the GW detection problem. The final objective is to help the experimentalists in the identification and characterization of the malfunctions of a complex instrument such as VIRGO. Incidentally, it is possible that the results might turn out to be useful to set a veto on false detection in the output of detection algorithms (such as a matched filter bank).

Because of the large dataflow we have to deal with in interferometric GW detectors, the choices we made were mainly oriented towards simplicity, yielding to algorithms that are easy to implement and that have a low computational cost.

The commissioning of the VIRGO central interferometer gave us the opportunity to test the individuated approaches and provide a 'quick' analysis of the engineering data with satisfactory contributions to the understanding of certain parts of the apparatus. We had to adapt the algorithms to the continuously changing configuration of the instrument as new functionalities were commissioned, thus changing significantly the nature of dominating noises. We plan to continue and refine the investigations on the data of (full scale) VIRGO.

Acknowledgments

We would like to thank all the contributors (T Cokelaer, J-Y Vinet and H Vocca) of the VIRGO group working on aspects related to stationarity, Gaussianity and linearity.

References

[1] Here is a list of Internet sites where more information can be found on the main interferometric GW detectors: LIGO (http://www.ligo.caltech.edu), TAMA300 (http://tamago.mtk.nao.ac.jp), GEO600 (http://www.geo600.uni-hannover.de), VIRGO (http://www.virgo.infn.it)

[2] The VIRGO Collaboration 2003 Commissioning of the central interferometer of the gravitational wave detector Virgo webpage http://wwwcascina.virgo.infn.it/commissioning

[3] Verkindt D (The VIRGO Collaboration) 2003 *Proc. 7th Gravitational Wave Data Analysis Workshop (Kyoto, 2002) Class. Quantum Grav.* **20** S623

[4] Flandrin P 1999 *Time–Frequency/Time-Scale Analysis* (London: Academic)

[5] Pai A *et al* 2002 *Virgo Project* VIR-NOT-OCA-1390-199
(Virgo notes can be obtained upon request)

[6] Fiori I 2002 *Virgo Project* VIR-NOT-FIR-1390-212
(Virgo notes can be obtained upon request)

[7] Priestley M B 1989 *Non-Linear and Non-Stationary Time Series Analysis* (London: Academic) pp 140–83

[8] Johnson N L and Kotz S 1970 *Distribution in Statistics. Discrete Distributions—Continuous Univariate Distributions* (New York: Wiley)

[9] Chassande-Mottin E 2003 *Proc. PSIP'2003 (Grenoble, France)* p 157
Chassande-Mottin E 2003 *Preprint* gr-qc/0212010

[10] Schreiber T and Schmitz A 2000 *Physica* D **142** 346

[11] Schreiber T and Schmitz A 1997 *Phys. Rev.* D **55** 5443

[12] Acernese F *et al* (The VIRGO Collaboration) 2003 *Proc. 7th Gravitational Wave Data Analysis Workshop (Kyoto, 2002) Class. Quantum Grav.* **20** S609