Comparison of unknown gravitational wave signals in two detectors as a mechanism of detection

Osvaldo M. Moreschi

Facultad de Matemática, Astronomía, Física y Computación (FaMAF),

Universidad Nacional de Córdoba,

Instituto de Física Enrique Gaviola (IFEG), CONICET,

Ciudad Universitaria, (5000) Córdoba, Argentina.

August 28, 2023

Abstract

We present a new measure that can be used for the detection of an unknown gravitational wave signals in two detectors, without recurring to *a priori* templates or whitening of the strains.

For an evaluation of its properties we apply our measure to the LIGO data of the GW150914 event and detect the existence of a similar signal with 99.99% confidence level. We also use the new measure to study the strains in the gravitational wave observatories for the events 151012_2, GW151012, GW170104 and GW190521.

We compare our measure with other standard measures and find that it is stronger for the studied data. Thus we are presenting a new powerful tool for the systematic study of unknown gravitational wave signals in two or more observatories.

Contents

IF-2023-00750845-UNC-ME#FAMAF

1 Introduction

As the interferometric gravitational-wave observatories increase their sensitivity, the number of detections also increase. In particular, in the O3a run of the Advanced LIGO and Advanced Virgo they were able to increase the number of confident gravitational wave detections (GWTC-2)[1] more than threefold over the first transient catalog GWTC-1[2]. In the presentation of the catalog GWTC-2[1] the authors communicate that they have used two methods to identify candidates; one that searches for minimally modeled sources and other that searches for signals from a bank of template wave forms[1]. They have also mentioned that they have used the Coherent WaveBurst (cWB) algorithm based on the maximum-likelihood-ratio applied to power excesses in the time-frequency domain[3]. More recently the LIGO/Virgo Collaboration has presented the catalog GWTC-3[4]; where they also use the same type of detection techniques.

Every detection of a gravitational wave event captures the attention of a large part of the community which is very interested in the details of the detected signals. For this reason we have been studying the characteristics of the these signals; in order to be able to relate them with corresponding theoretical frameworks. In this process, we have first developed a non-destructive pre-processing filtering technique; that had allowed us to discover more relevant physical signal for the GW150914 event. Afterwards we have started the study of techniques for the comparison of unknown similar signals in the strains of two detectors, and in the process we have constructed an optimized measure that could be used, in post-detection detailed studies, that can even provide more information of the detection process itself. This article is devoted to the presentation of such measure.

When employing templates, most pipelines identify the gravitational-wave signals by matched filtering[5, 6, 7, 8] data; using a bank of filter waveforms with a range of source parameters. Although normally the search is carried out with an assessment in the mass range and the assumption of quasicircular orbits[4]. However, with the increase in sensitivity of the observatories, there appear detections of systems that seem outside of the expected range of parameters. For example, the event GW190521 has been reported to have a remnant of $150M_{\odot}$[9]; which is an unusual high mass. In reference [10] the LIGO/Virgo Collaboration have presented the properties of this system 'under the assumption of a quasi-circular BBH coalescence'. But it has also been suggested[11] that probably this system corresponds to the capture of two non-spinning black holes on hyperbolic orbits. In other words, some simplifying assumptions seem not tenable. Thus, it would be advantageous to have a method of detection that does not depend on assessments and assumptions on the details of the astrophysical system. Our proposed method below is constructed with these ideas in mind.

In the cWB algorithm[3] the detected time series is first treated with a Wilson-Daubechies-Meyer transform and is filtered by a whitening process. These time-frequencies series are then

combined from those coming from all detectors, by their sum of squares; which are maximized for all possible time-of-flight delays in the network. Dealing with time-frequencies series implies that one has to deal with two dimensional arrays, with the consequent demand on computer memory.

Our method, only deals with the time series, and does not use any whitening techniques. Instead we use here the pre-processing method based on FIR filters that have excellent behavior with the phase of the strain, that avoids the attenuation of astrophysical signal, and therefore allows to see 'more' relevant information of the gravitational detected wave; for this reason we call this the pre-processing More19 method, which was described in [12]. In particular by using this method, we discovered in reference [12] that the signal of a gravitational wave has a duration 0.5s; while in their publication[13] the LIGO/Virgo team could only show a signal of a about 0.1s duration.

In spite of the drawbacks of the whitening techniques, associated to huge deformation of the signal and the statistics, those methods have been used in a successful and efficient manner for the detection of transient gravitational waves[14]. The measure we present in this article has been developed as a useful tool in post-detection studies; although, as we will see in this article, it could probably be adjusted to be used in the detection process itself.

We here suggest to carry out the comparison process of similar gravitational-wave signals in two detectors using an optimized measure(OM) that we present below; which is independent of model assumptions, and templates and does not use whitening techniques. We denote the new measure OM with the symbol Λ, which it will be shown to be somehow related to the likelihood ratio[15] calculation for the detection of a known signal in a single detector. Before applying this measure we subject the strains to the pre-processing More19, mentioned above.

While the measure is designed for a pair of detectors; in the case of a network of detectors with three or more of them, one can either apply the measure for each pair independently, or assuming the statistical independence of the measuring process for each detector, one can consider the multiplication of the measures of all pairs

The organization of this article is as follows. In section 2 we present a new measure to compare the content of signals in the strains of two detectors; where for the sake of simplicity of the presentation we relegate to an appendix some arguments in favor of our choice. Also, we include a short description of the correlation coefficient in another appendix. In section 3 , to test its properties, we apply this new measure to the case of the GW150914 event, and present the analysis that conduces us to the detection of similar signals in the two LIGO strains recorded for this event. We carryout some detailed studies of the new measure with this well known event, that has a very strong signal. To test the behavior of the measure OM under different circumstances, we compare the signals in the events 151012_2, GW151012, GW170104 and GW190521 in sections 4 - 7. In particular GW190521 involves three strains, since the Virgo detector was online at that time. In section 8 we include some final comparisons and comments on this work.

2 New measure for the comparison of signals in two detectors

2.1 Preliminaries on measures

In the literature one can find many approaches for the study of data that intends to determine whether a known signal is present in the data; however works considering to determine whether an unknown signal is present in two separate and independent sets of data are rare. In order to build the needed new measure, we have been guided by the usual method of maximum likelihood as much as we could; since: "Although the maximum likelihood principle is not based on any clearly defined optimum considerations, it has been very successful in leading to satisfactory procedures in many specific problems."[16] The choice of a maximum likelihood method is also based on the fact that we are dealing with a non-parametric detection[15]. Due to the fact that from our proposed measure Λ, we will deduce important results, we will go through a detailed presentation of it; but to

avoid distraction of the main content of this article, we present in appendix A a line of arguments that support our choice for the new measure Λ. These arguments are based on trying to adapt the likelihood method, of searching for one signal in one data, to our case. However we have to moderate the first natural definition of likelihood so that the measure becomes useful.

The arguments presented in appendix A, are not intended to be a deduction. They are only presented to connect our definition of the measure Λ with other related constructions that have been used in related works. That is, there is no right or wrong definition of a measure, any definition of a measure is arbitrary and so is ours; the question is if it is useful for some purpose. We claim that our definition is useful since it is the most powerful[16] when compared with two other natural choices and it has allowed us to observe the gravitational wave in the LIGO data of the GW150914 event, with unprecedented level of significance, that we present below.

To give perspective to the strengths of our measure Λ, we here, in section 2.2, present our calculation of the likelihood ratio $\mathbf{L}$ to test the hypothesis that a similar signal is recorded in the strains of two detectors, versus the hypothesis that no similar signal has been recorded in both detectors; we present our measure Λ in section 2.3 and we also recall the correlation coefficient ρ between the two strains in section 2.4. Then, in section 3 we discuss the application of these three measures to the case of the data of the GW150914 event.

2.2 The likelihood ratio for the detection of an unknown signal in two detectors

In appendix A.2 we have deduced the expression for the likelihood ratio to test the hypothesis that a similar signal is recorded in the strains of two detectors, versus the hypothesis that no signal has been recorded in both detectors, which, in terms of the data, is given by:

$$\mathbf{L}(\mathbf{v}_1,\mathbf{v}_2) = \exp\left[\frac{m-1}{2}\Big(\frac{1}{\sum_{k=1}^{m} v_{(1)k}^2}+\frac{1}{\sum_{k=1}^{m} v_{(2)k}^2}\Big)\sum_{k=1}^{n} v_{(1)k}\, v_{(2)k}\right]; \tag{1}$$

where the width of the window to calculate sample variances, that is m, is chosen appropriately depending on the nature of the observations, and we are assuming that the means are zero.

This estimation of the desired measure have some difficulties. The factor $\left(\frac{1}{N_{0_1}}+\frac{1}{N_{0_2}}\right)$(See appendix A.2 for details.) is rather huge, when using LIGO data, and it does not contribute to strengthen the comparison of the data. Expression (1) is the theoretical deduction of the likelihood ratio, but for actual numerical application to the gravitational-wave data, we will use (7); for reasons that we explain below. Huge exponents are undesirable since might lead to unwanted numerical error or even overflow errors. For these reasons, we present our measure Λ next.

2.3 The measure Λ

Due to the difficulties found in using the likelihood ratio, just presented, we constructed a new measure OM that turns out to be useful for our purposes. Our arguments, that led us to this construction, are presented in appendix A.

We here present the measure in synthetic form.

When dealing with data from gravitational-wave observatories, one is confronted with time series, which are supposed to contain signals from gravitational waves; that might last, from a fraction of a second to several seconds; depending on: the astrophysical nature of the source, the intensity, the noise state of the detectors at the time of recording, etc. Then, in designing tools for the analysis of the data, and considering the transitory nature of the expected gravitational-wave signals, it is convenient to introduce windowing techniques that allow for the study of portions of the data. For this reason we introduce an inner product that contemplates this point.

IF-2023-00750843-UNC-ME#FAMAF

We define an inner product for two strains $\mathbf{x}(\tau)$ and $\mathbf{y}(\tau)$. The inner product is calculated through the convolution with an appropriate window $w(t-\tau)$, and is defined by:

$$< \mathbf{x}, \mathbf{y} > (t_j) = \sum_k x(t_k)\, y(t_k)\, w(t_j - t_k); \tag{2}$$

where we use the fact that the data is obtained at discrete time intervals. For the continuum notation below we will use t instead of t_j.

For the specific case of the observation of gravitational waves at two detectors, let $\mathbf{v}_1(\tau)$ represents the strain at one detector with respect to its proper time, and $\mathbf{v}_2(\tau - \delta)$ represents the strain at the other detector, with time shift δ.

Then, we define the time dependent Λ measure from

$$\Lambda(\mathbf{v}_1, \mathbf{v}_2, \delta, t) = \exp\left[\frac{1}{\sigma_{12}^*} \frac{< \mathbf{v}_1, \mathbf{v}_2 >}{< (\mathbf{v}_1 - \mathbf{v}_2), (\mathbf{v}_1 - \mathbf{v}_2) >}\right]; \tag{3}$$

where σ_{12}^* is the standard deviation of $\frac{<\mathbf{v}_1, \mathbf{v}_2>}{<(\mathbf{v}_1 - \mathbf{v}_2), (\mathbf{v}_1 - \mathbf{v}_2)>}$. As already mentioned, in appendix A we present arguments that support this choice for the measure.

In short, this measure can be understood as coming from an adaptation of the likelihood method to our case, with contrast accentuation and overall moderation. Our choice gives reasonable results with actual LIGO data.

Although the arguments presented in appendix A use the idea of having the same signal in two strains, the measure Λ also indicates the existence of similar signals in two strains; namely, when the signals s_1 and s_2 recorded in the detectors satisfy $s_2 = s_1 + \epsilon$, with $\max|\epsilon| \ll \max|s_1|$. This is precisely what happens when we apply the measure OM to the data of the GW150914 event; as shown below.

For the window w we use a Tukey window, from the scipy signal PYTHON library, with width that is chosen for each event, and parameter $alpha = 1/8$; which has excellent behavior. The convolution is performed with the efficient fftconvolve function from the same library.

It is probably worthwhile to mention here that during the first times of gravitational-wave observations, only the Hanford and Livingston detectors were on line; and that due to their orientation, for all practical purposes, the strain of one of them was often compared with minus the strain in the other; as it was the case in the event we are considering here GW150914. However, it is not true that the recorded signals satisfy that one is minus the other, since we know that actually due to a non-exactly opposite alignment of the detectors, they should record two different projections of the spin 2 gravitational wave. When considering other detectors in the network, as for example Virgo or Kagra; this effect becomes more noticeable. But in all cases one would be confronted with the situation that any two of the detectors in the network, would record, let us say for a binary collapse, few oscillations of comparable magnitude and frequencies. Our measure would report on these type of coincidences; that is, even if the morphology of the rest of the gravitational wave have different phase behavior; as expected when recording a spin 2 gravitational wave with detectors having different orientations.

2.4 The correlation coefficient

One could have thought that the natural thing to do was just to consider the correlation coefficient[17, 18, 19, 16] between the two strains, with a time shift added to one of them. Then, taking into consideration the characteristics of the gravitational-wave observation, mentioned above, we define the correlation coefficient in terms of the natural inner product by:

$$\rho_{\mathbf{v}_1, \mathbf{v}_2} \equiv \frac{< \mathbf{v}_1, \mathbf{v}_2 >}{\sqrt{< \mathbf{v}_1, \mathbf{v}_1 >< \mathbf{v}_2, \mathbf{v}_2 >}}. \tag{4}$$

The fact is that the correlation coefficient gives very low response; as we show in appendix B and in discussions below in section 3 and 8.

2.5 Minimal test for Gaussianity behavior

In the preceding subsection we have presented a new measure to compare similar signals in two detectors. The motivating arguments are presented in the appendix A which make use of the assumption that the noise at both detectors are close to a Gaussian behavior. Just by observing the amplitude spectral density of the strains at Hanford and Livingston, one can see that the data is no perfectly Gaussian. However, there are many ways in which this assumption can be tested, and several articles in the past have addressed this point; but we here instead present graphs of what can be considered the most direct minimal test for Gaussianity behavior, namely the histogram of the data of the GW150914 event. We must emphasize that the behavior shown in the graphs of figures

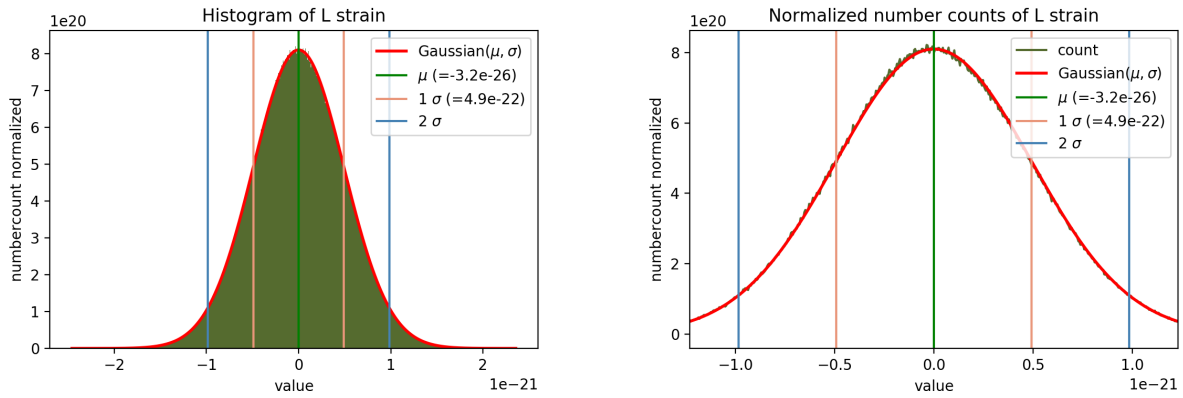

Figure 1: Comparison of the histogram of Livingston data(green) with a Gaussian(red). On the left the complete histogram of 256 seconds of the data, centered at the time of the event. On the right the detail of the graph in the central region.

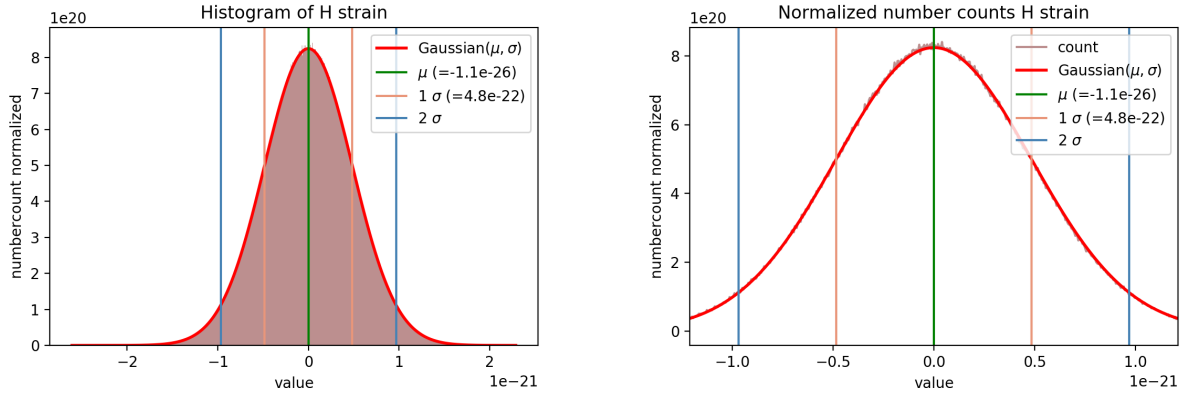

Figure 2: Comparison of the histogram of Hanford data(colored surface) with a Gaussian(red curve). On the left the complete histogram of 256 seconds of the data, centered at the time of the event. On the right the detail of the graph in the central region.

1 and 2 , are obtained after we have applied the pre-processing filtering techniques More19 that we have described in [12]. Figure 1 shows the histograms for the Livingston data of 256 seconds centered at the time of the event, and compared with a Gaussian with $\sigma_L = 4.92212e-22$. Figure 2 shows the histograms for the Hanford data of 256 seconds centered at the time of the event, and compared with a Gaussian with $\sigma_H = 4.83976e-22$. The standard deviations σ_L and σ_H were calculated from the data. For completeness, the Gaussian function is given by: $G(x, \sigma, \mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\big(-\frac{(x-\mu)^2}{2\sigma^2}\big)$. For both detectors the median was essentially zero.

In order to give perspective to the relevance that has this Gaussian behavior as a result of our More19 filtering approach, we show here what is the result on the time series after applying the whitening filtering procedure used by the LIGO/Virgo Collaborations. The histograms of the time series of 256 seconds centered at the time of the event, after applying the whitening filters to the Livingston and Hanford strains of the GW150914 event, are shown in the graphs of Fig. 3 and 4 .

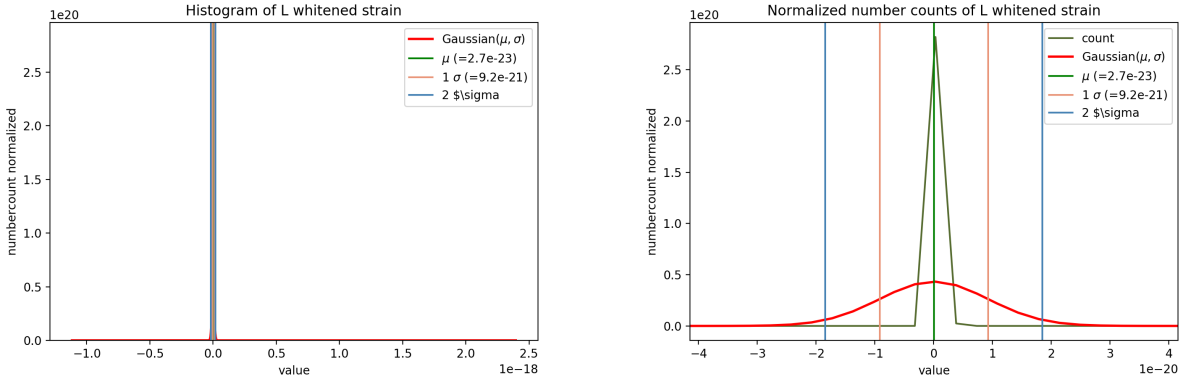

Figure 3: Comparison of the histogram of the Livingston time series after applying the whitening procedure as describe by LIGO tutorials. We use the same graphic scripts as above. On the right the detail of the graph in the central region. The fact that the maximum appears a little to the right, on the detailed graphs, is an artifact of the graphic script based on the notions of bins.

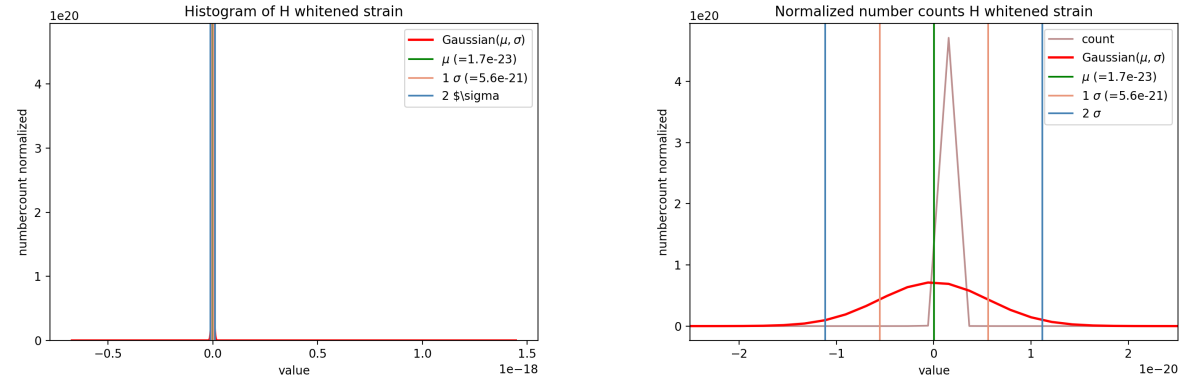

Figure 4: Comparison of the histogram of the Hanford time series after applying the whitening procedure as describe by LIGO tutorials. We use the same graphic scripts as above. On the right the detail of the graph in the central region. The fact that the maximum appears a little to the right, on the detailed graphs, is an artifact of the graphic script based on the notions of bins.

It can be seen in the graphs of Figs. 3 and 4 that the whitening procedure changes completely the statistics of the time series, since now the histograms show a huge departure from a Gaussian behavior. One might ask, what is the reason for this extreme behavior. This can be inferred from the time domain graphs of the strains after whitened and band-pass filters. In Fig. 5 we show the complete 256 seconds strain after whitening and band-pass filters are applied, following LIGO procedures. It can be seen that there are important boundary effects. For this reason we trimmed the strain, by cutting off 3s on each extreme, to obtain a total length of 250s strains. In Fig. 6 we show the remaining 250 seconds strains after whitening, band-pass filters and trimming are applied. It can be seen that now the strain is very quiet with an important spike near the time of the event.

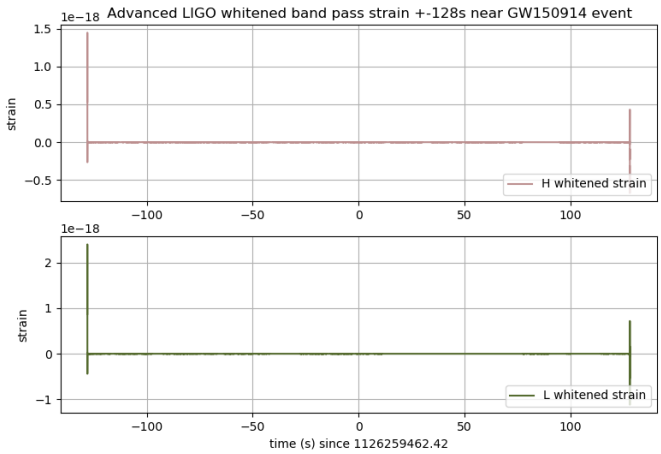

Figure 5: Time domain graph of the whitened and band-pass strains of the complete 256seconds data.

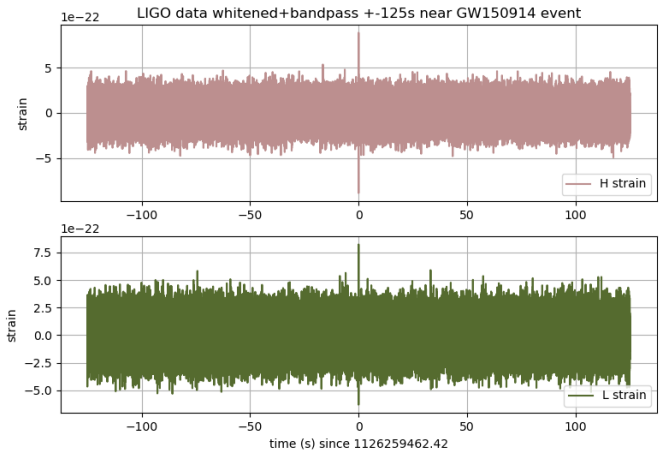

Figure 6: Time domain graph of the whitened, band-pass and trimmed strains of the remaining 250seconds strain.

The corresponding histograms after the trimming are shown in the graphs of Fig.7 ; where now one can see a Gaussian behavior for the remaining whitened noise.

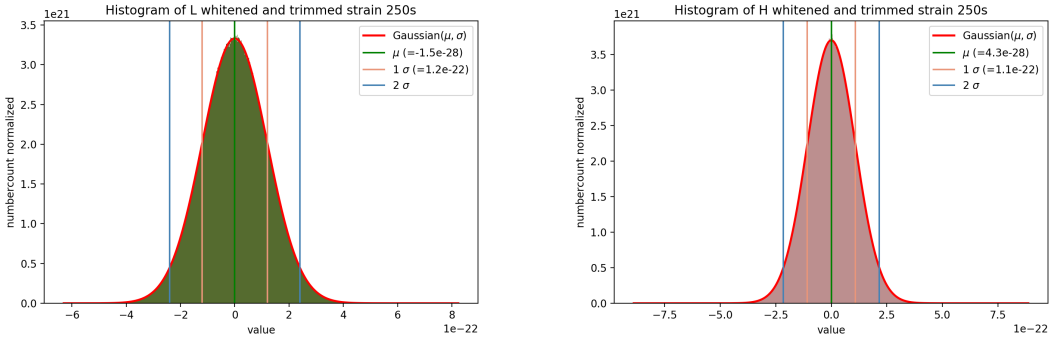

Figure 7: Histograms of the strains after applying whitening and band-pass filters, and after trimming the extremes in order to avoid boundary effects.

This Gaussian behavior for the remaining noise makes one wonder whether our measure Λ would give good results when applied to this whitened strains; which we will study in the next section.

3 Detection of similar signals in the two LIGO observatories for the GW150914 event

3.1 Study of the measure Λ as a function of time

In order to see how this new measure behaves with real data, we apply it to the event GW150914 as a test bed to study its properties.

We show in figure 8 the graph of our measure Λ as a function of time, for the Livingston strain(L) against the nominal Hanford(H) strain shift of $\delta = -0.007$s, in the interval -10s, 10s from the time of the event t_e; for a Tukey window of 0.5s width, where it can be seen that there is a sharp peak at about 0.0s, that is at the time t_e; where we have advanced the time axis by the width of the window. The sign of the H strain is obviously chosen so that the natural inner product with the L strain gives a positive result when comparing. The actual statistic is calculated in the lapse of time $\pm$11s around the event time, which at this moment is a preliminary arbitrary choice and will be studied further in the next subsection.

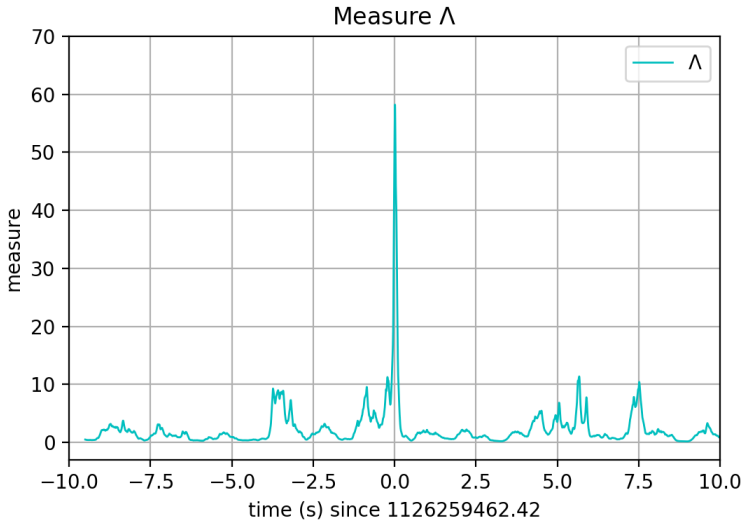

Figure 8: The measure $\Lambda(t)$ for the shift $\delta = -0.007$s for the Hanford (-)strain, in the range $\pm$10s, with respect to the Livingston strain.

Figure 8 shows that close to the time of the event there is a sharp peak, indicating that most probably both detectors have recorded a similar signal in the window w before this time.

The sharp behavior of this measure invites us to calculate a coarse estimate of the level of significance by using directly Chebyshev inequality[19], that we recall next:

Theorem 3.1 *Let X be a random variable and let $\lambda(x)$ be a non-negative function. Then, for any $r > 0$,*

$$P(\lambda(X) \geq r) \leq \frac{\mathbf{E}(\lambda(X))}{r}. \tag{5}$$

Here P means probability and $\mathbf{E}$ expectation value.

The first estimate of the level of significance α_0 can be calculated from identifying $\lambda = \Lambda$ and taking $r = \max(\Lambda) = 58.21$. Then, from the numerical calculation $\mathbf{E}(\Lambda) = 2.338$, we obtain $\alpha_0 = 0.0402$.

With the knowledge of the standard deviation, we can use the customary form of the Chebyshev inequality that sets:

$$P(|X - \mu| \geq t\sigma_X) \leq \frac{1}{t^2}; \tag{6}$$

where $\mu = \mathbf{E}(X)$ and σ_X is the standard deviation. Then, by identifying X with Λ, and using the calculated value of $\sigma_\Lambda = 4.4406$, we obtain the second estimate of the level of significance α_1 given by $\alpha_1 = 0.0063$. It can be seen that this second estimate improves on the first one; since we are using more information on the statistics of Λ.

The Chebyshev inequality is normally applied when the probability distribution for the situation under study is not known. Let us try to infer more information on the statistical properties of Λ.

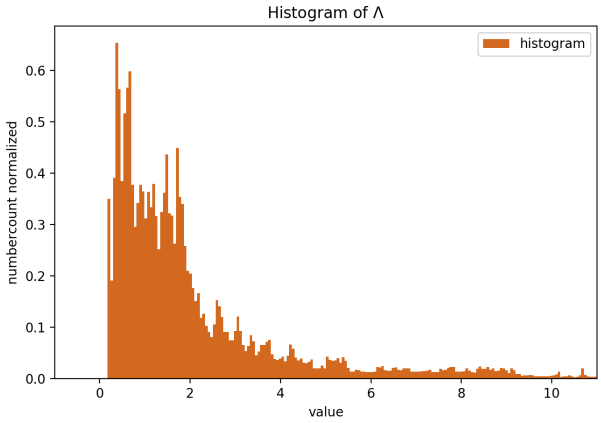

Figure 9: Histogram of Λ measure for a limited part of its domain.

In figure 9 , in order to show some detail, we present the graph of the histogram of our measure for a limited part of its domain, that does not include the maximum value, close to sixty. The sample corresponds to a strain of 22 seconds centered at the time of the event. It can be seen, since the distribution only involves positive values, that the histogram does not show a Gaussian behavior. Instead it does resemble a log-normal behavior.

This, in turn invites us to see what is the behavior of the histogram of the exponent, that we show next.

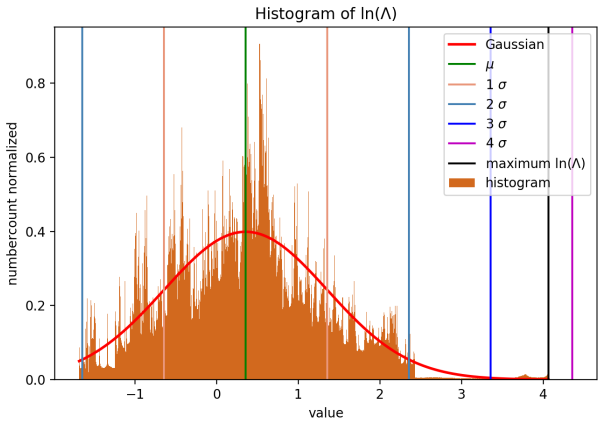

Figure 10: Histogram of logarithm of Λ measure. It is also shown with verticals lines the median, 1, 2, 3 and 4 sigmas, and with a black vertical line, the position of the maximum observed.

In figure 10 we show the graph of the histogram of the logarithm of our measure for the 22s of the strains; where sigma is the standard deviation. The Gaussian curve is calculated from the mean and the standard deviation. It can be observed that the Gaussian red curve is a good smoothed approximation of the behavior of the histogram. The black vertical line shows the position of the maximum of the logarithm of the measure, observed at $z\sigma$, with $z = 3.70749$.

Identifying a Gaussian behavior for the logarithm of Λ, we can assign the level of significance[16] $\alpha = (1/2)[\mathrm{erfc}(z/\sqrt{(2)})]$; where erfc is the complementary error function. Equivalently we can also define the confidence level $\gamma = (1 - \alpha)$. The values so obtained are: a level of significance $\alpha = 0.000105$, and a confidence level, or confidence coefficient[18], $\gamma = 0.99989$.

This is a remarkable strong behavior of the measure OM; since for the data we are analyzing it gives us 99.99% confidence that there are similar signals in both detector strains, for the chosen window. We will see next that this measure is stronger than the other two we are considering in this article.

Let us note that the LIGO/Virgo Collaborations[20] use for this event a measure based on a matched-filter signal-to-noise ratio, which employs templates, and can not be compared directly with our method.

3.2 Adjustment of the lapse of time to be used

Since the measure OM works without a priori assumed templates, but just comparing the two strains of the distinct gravitational wave detectors, it is subject to the possibility of finding strong random noise of seismic origin with similar frequencies and phases. For this reason, one is interested in maintaining the lapse of time, for the strain comparison to a minimum, so as to avoid the previous inconvenience and obtain a reasonable statistic behavior. Because of this, we first carry out a preliminary study on the behavior of the measure OM in order to estimate a reasonable working lapse of time.

In [13] the authors present an study designed to operate without a specific waveform model, for signal frequencies up to 1 kHz and durations up to a few seconds. In reference [21] the LIGO/Virgo Collaboration analyzed coherently 8s of data with a uniform prior. In the article [22] the same Collaboration studied correlations on the order of the duration of transient astrophysical signals; a fraction of a millisecond to a few seconds; and they also added: "noise transients with a large amount of broadband power can corrupt the analyzed data up to the duration of the strain-equivalent noise PSD estimate, ±8s from the time of the noise transient."; which corroborates our view above.

Taking the range of lapse of times used by LIGO/Virgo team, we next study the range for lapse of times: [8, 16, 24, 32, 40, 48, 56, 64] seconds. In Fig.11 we show the graph that studies the behavior of the level of significance α with the length of the lapse of time.

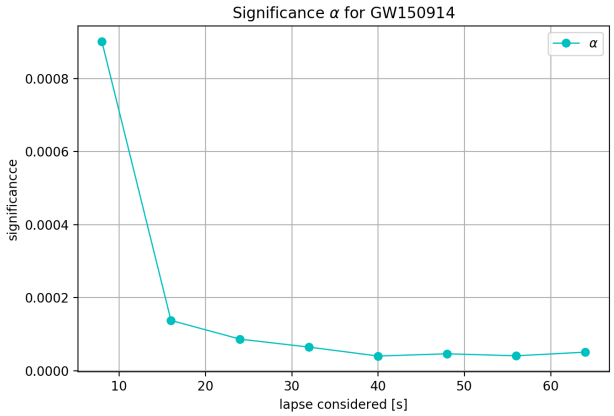

Figure 11: Graph of the level of significance as a function of the length of the lapse of time considered.

It can be seen from Fig.11 that the significance for this event, shows an starting decreasing tendency, up to a width of about 40 seconds, when it gets into a stationary phase. This suggests that we should consider the lapse of time of 40s for our studies.

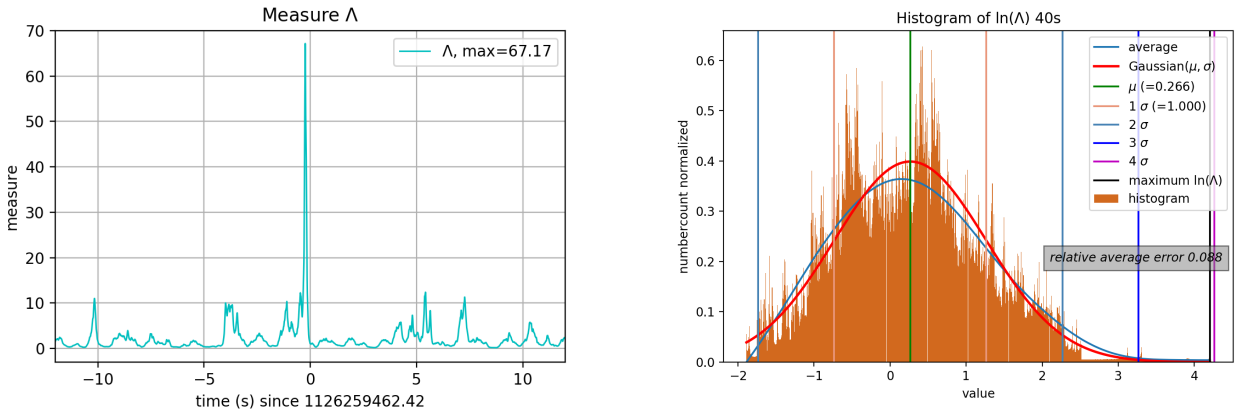

Figure 12: Graph of the measure and histogram of $\ln(\Lambda)$ for the 40s lapse of time.

In Fig.12 we show the graphs of Λ and of the histogram of $\ln(\Lambda)$ for the 40s lapse of time. Let us note that all the histograms are calculated with 1000 bins; so that, as in this case, the amount of data might not be sufficient to produce a smoothed graph. In order to quantify how close the histogram is to a Gaussian behavior, we have also included the curve of the appropriate average, that was calculated with a Kaiser window. One can see that the Gaussian curve, shown in red, is fairly close to the average of the histogram, shown in blue. We have also estimated the relative average error between these two curves; which in this case gives 0.088, that is, just few percents.

With the chosen length for the lapse of time, we now have $z = 3.94165$ and therefore significance of $\alpha = 4.05\text{e-}05$ and confidence coefficient $\gamma = 0.9999595$; which improves by a factor of 2.6 on our previous estimate using the ± 11s interval.

3.3 Applying the optimal measure to the whitened strain

Since we have seen that the noise of the whitened strain also shows a Gaussian behavior; it is natural to consider then the action of the measure OM to the whitened strain. We have also considered the same range for lapse of times and we show the histogram of $\ln(\Lambda)$ for three cases in Fig.13.

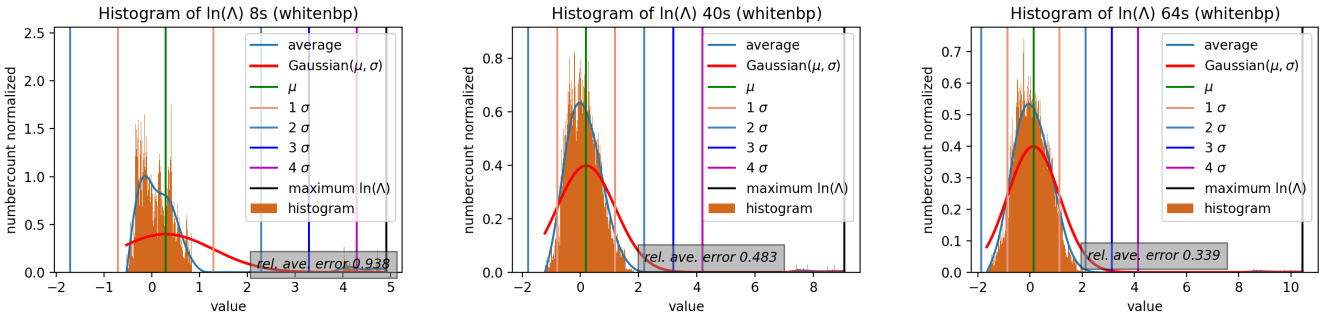

Figure 13: Graph of the histogram of $\ln(\Lambda)$ for whitened strains.

It can be seen from the graphs in Fig.13 that when we apply the measure OM to the whitened strains, we can not sustain a behavior that is close to Gaussian. In particular we can see that the relative error for a 8s width is 0.938, for a 40s width the relative error is 0.483s while for the 64s width the relative error is 0.339. In all these cases this quantitative estimation of the relative error

is too high to allow us to consider the behavior of the measure OM on the whitened strains to be close to Gaussian. For this reason we do not consider our measure on whitened data any further.

Of course one could still use the language of 'signal to noise' ratios, applied to our measure on whitened strains; but on one hand we do not want to compete with LIGO techniques, and on the other hand we are trying to stay close to standard statistical treatments based on the concept of probability.

3.4 Study of the other measures as a function of time

For the numerical calculation of the likelihood ration $\mathbf{L}$, we make use of the natural inner product defined above, so that the detailed expression of $\mathbf{L}$ to be used with gravitational-wave data is:

$$\mathbf{L}(\mathbf{v}_1, \mathbf{v}_2) = \exp\left[\frac{1}{2}\left(\frac{1}{<\mathbf{v}_1, \mathbf{v}_1>} + \frac{1}{<\mathbf{v}_2, \mathbf{v}_2>}\right) < \mathbf{v}_1, \mathbf{v}_2 >\right]. \tag{7}$$

When we try to use the likelihood ration $\mathbf{L}$, that we calculated above, to the data of the GW150914 event, we obtain an overflow error using python. From the previous discussion, it is suggested to also study the behavior of the logarithm of $\mathbf{L}$. In figure 14 we shown the histogram of the logarithm of the likelihood ratio, along with the position of its median, sigmas, and the maximum of the logarithm of $\mathbf{L}$ with vertical lines and the theoretical Gaussian calculated from the median and the standard deviation. The maximum of the logarithm of $\mathbf{L}$ is located at $z_{\mathbf{L}}\sigma_{\mathbf{L}}$, with $z_{\mathbf{L}} = 2.85729$ (Shown in figure 23.). Identifying a Gaussian behavior for the logarithm of $\mathbf{L}$, we can assign the level of significance $\alpha_{\mathbf{L}} = 0.00213$. This means that this measure gives a signal about 53 times weaker than our optimized measure.

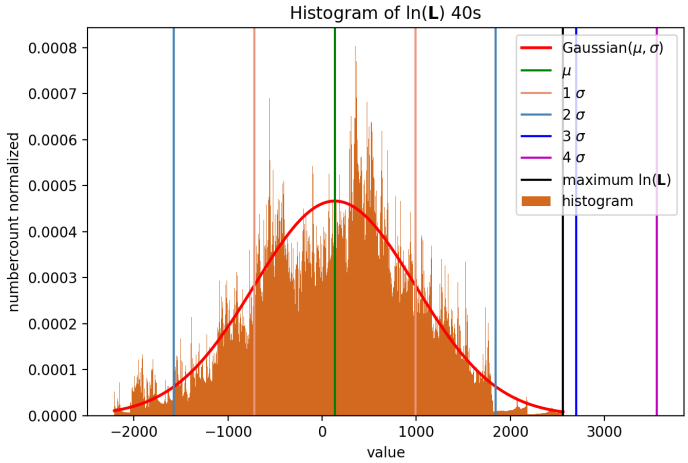

Figure 14: Histogram of the logarithm of likelihood ratio $\mathbf{L}$. It is also shown with verticals lines the median, 1, 2, 3 and 4 sigmas, and with a black vertical line, the position of the maximum observed.

The behavior of the correlation coefficient is studied in appendix B ; instead, here in figure 15 we shown the histogram of the correlation coefficient, with the position of its median, sigmas, and the maximum of ρ with vertical lines and the theoretical Gaussian calculated from the median and the standard deviation. The maximum of ρ is located at $z_\rho\sigma_\rho$, with $z_\rho = 2.85978$. Identifying a Gaussian behavior for ρ, we can assign the level of significance $\alpha_\rho = 0.00212$. This means that the correlation measure gives a signal about 52 times weaker than our measure.

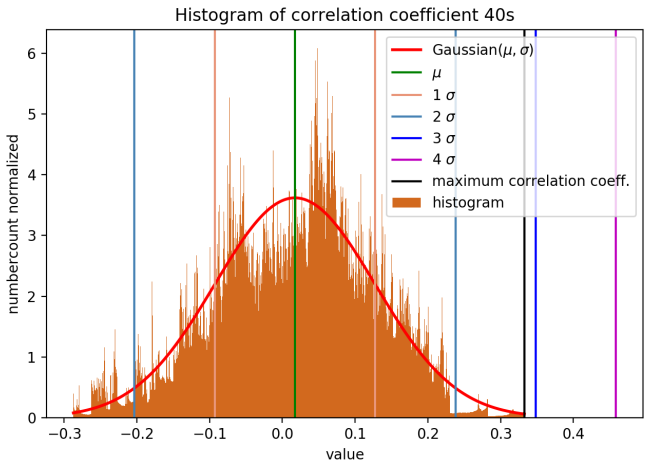

Figure 15: Histogram of the logarithm of the correlation coefficient measure ρ. It is also shown with verticals lines the median, 1, 2, 3 and 4 sigmas, and with a black vertical line, the position of the maximum observed. They show a remarkable similarity in shape.

Summarizing, the problem of detecting a gravitational-wave signal in noise can be posed as a statistical *hypothesis testing problem*[23, 24], and we have studied here the problem to test the hypothesis that a similar signal is recorded in the strains of two detectors, versus the hypothesis that no similar signal has been recorded in both detectors with three different statistical measures. Our optimized measure OM, the likelihood measure and the correlation coefficient measure all show a statistical behavior that is very closed to Gaussian. In particular for the measure OM, we have quantified its behavior. The use of each of these three measures gives us very high confidence levels for the detection of similar signal in both detectors.

By comparing the level of significance that we can give to the detection of a similar signal in the two LIGO observatories data, for the GW150914 event, using the three measures, we conclude that the measure OM, that we have introduced, is the strongest one.

In the next sections we apply the measure OM to a variety of different events, presented in chronological order.

4 Applying the measure OM to the 151012_2 event

The event 151012_2 was released as a 'GWTC-1-marginal' kind, and assigned the GPS time 1128666662.2; equivalently the UTC Time: 2015-10-12 06:30. It was also assigned the network SNR of 9.6. We could not found any suggestion for the arrival time at Hanford relative to Livingston. We have chosen $\Delta t_{HL} = 0.0099$s.

We have first applied the same band-pass filter of 22-1024Hz, but a preliminary study of the strains showed a strong presence of low frequency noise. For this reason we decided to move the low frequency of the pass band from 22 to 32Hz.

With this choice for the low limit of the pass band filter, the strains have a Gaussian behavior and the 20s graph of Λ, with a 0.5s window, is shown in Fig. 16.

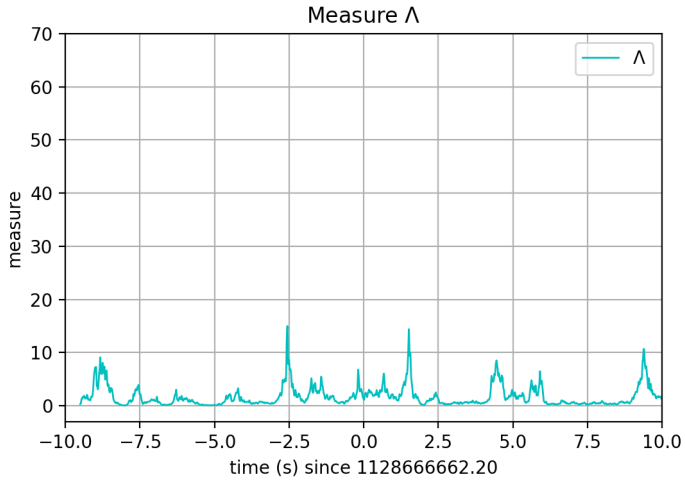

Figure 16: Graph of Λ for event 151012_2 near the event time.

Although there are small peaks, they do not seem to be associated to a signal, but to the state of the observatories; since noise of low frequencies is present around the time of the event; as it can be seen in Fig. 17.

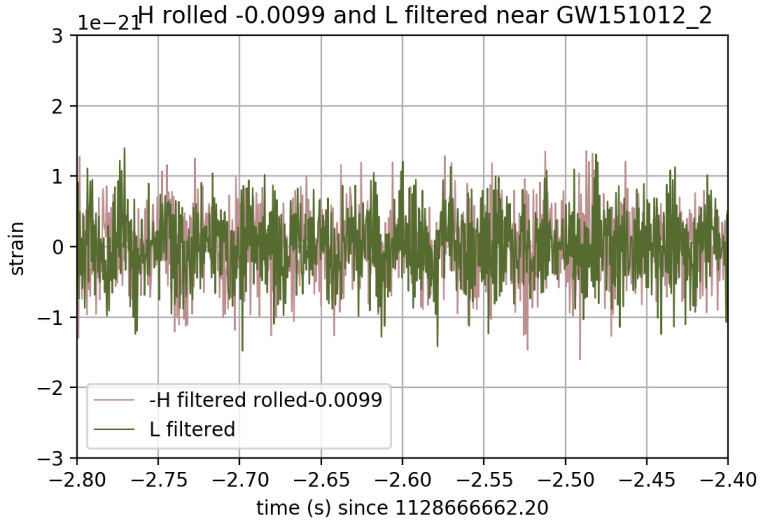

Figure 17: Closer look to the strains in time domain of 151012_2, near the peak around at -2.5s before the event time.

In a work dedicated to 151012_2 one should consider further filters; but in our case, we are just studying the behavior of our measure to this marginal event. Then, from Fig. 16 , we conclude that the measure OM does not show a confident signal near the time of the event for the duration considered.

5 Applying the measure OM to the GW151012 event

The event GW151012 was released as a 'GWTC-1-confident' kind, and assigned the GPS time 1128678900.4; equivalently the UTC Time: 2015-10-12 9:54; but we use in our work the time-event 1128678900.45. It was also assigned the network SNR of 9.3. This event shows low levels of signals.

Due to the proximity to the previous event, in this case, it shares the characteristic of the noise for this study.

In particular in reference [25] the authors suggest that the arrival time at Hanford relative to Livingston is $\Delta t_{HL} = -0.0006 \pm 0.0006$s. From the same reference, one can infer that the authors estimate a signal duration of about 0.5seconds. Since the pre-processing filtering techniques More19 normally shows longer signals than those found by the LIGO team, we tried several longer durations and finally settle for a signal duration of 0.8s. The preliminary application of the measure OM to this event with $\Delta t_{HL} = -0.0006$s and duration of 0.8s gave no signal close to the time of the event.

Due to this negative result, we studied in detail the time delay and detected that the value $\Delta t_{HL} = -0.0012$s produces a signal in the Λ which we show in Fig. 18 .

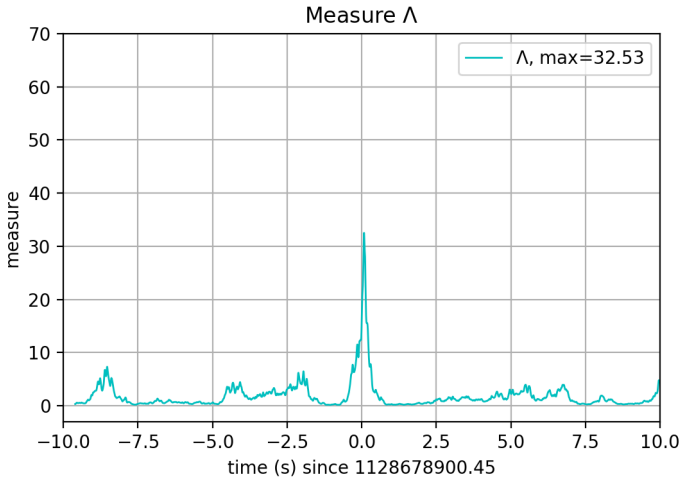

Figure 18: Behavior of the measure Λ near the event GW151012; using the time delay $\Delta t_{-HL} = -0.0012$s, and a window of 0.8s.

We concludes from Fig. 18 that there is a weak signal corresponding to the comparison of the Hanford and Livingston strains close to the GPS time 1128678900.45, corresponding to the arrival time at Hanford relative to Livingston of $\Delta t_{-HL} = -0.0012$s and with a signal duration of at least 0.8s.

We have seen then that even in this problematic case, with low level signals, the measure OM can assign a weak signal for the GW151012 event. Note however that this measure can not clarify whether the signals are of astrophysical origin; since it only deals with the similarities.

Here we just study the behavior of the measure OM when applied to the GW151012 event. It is not our intention to proceed further with a detailed study of the event, and instead concentrate on the behavior of the measure on different kinds of events. For further properties of GW151012 see for example [20, 26, 25, 27].

6 Applying the measure OM to the GW170104 event

The event GW170104 was released as a 'GWTC-1-confident' kind, and assigned the GPS time 1167559936.6; equivalently the UTC Time: 2017-01-04 10:11. It was also assigned the network SNR of 13.8. We have used version v2 of the corresponding LIGO strains.

According to reference [28] the relative Hanford to Livingston arrival time shift is $\Delta t = -0.003$s; and we have used $\Delta t_{-HL} = -0.00305$s in our work.

The authors of [28] also assert that the two strains are well calibrated in the the frequency range [20,1024]Hz. For the pre-processing filtering we used a pass band of [27.0,1003.0]Hz.

The graphs appearing in [28] show strains with signals lasting approximately 0.11s. From observing the strains after applying the pre-processing filtering techniques More19, we detect signals that last approximately 0.28s; which we use in the settings of the measure OM.

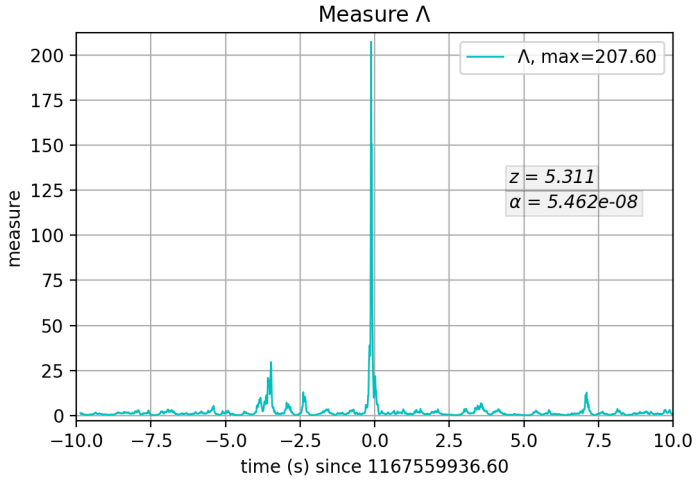

Figure 19: Behavior of the measure Λ near the event GW170104; using the time delay $\Delta t_{-HL} = -0.00305$s and a window of 0.28s.

It can be seen in the graph of Fig. 19 a very strong Λ signal near the event time. In fact, the statistic of the optimize measure OM allows us to assign an outstanding level of significance of $\alpha = 5.5 \times 10^{-8}$ to the detection of similar signals in the two LIGO observatories for the event GW170104.

The impressive confident level that this measure gives for this event, corroborates the extension of the physical signal to about 0.28s. This obviously deserves a detailed study of the event, but as mentioned before, we are here concentrating in presenting the behavior of the measure OM to different events; and we will proceed with in-depth studies on another occasion.

7 Applying the measure OM to the GW190521 event

At the Gravitational Wave Open Science CenterGWOSC web page, the GW190521 event is presented in its fourth version as a release of type 'GWTC-2.1-confident'. However, since we have noticed that the v4 version has been subjected to unknown kind of filters, we work with version v1; which is described as a release of type 'O3_Discovery_Papers'.

The LIGO/Virgo Collaboration has assigned the value 14.3 for their network SNR.

During this event the Virgo observatory was in operation, so that we also consider here its strain.

According to [9], GW190521 is a short transient signal with a duration of approximately 0.1s. Instead, after applying the pre-processing filtering techniques More19, we have found that the transient signal has a duration of approximately 0.35s; which we use in our work.

We found no information in the literature[9, 10] on the relative times of arrival of the signal among the detectors. The comparison of the minus Hanford strain with the Livingston strain, gives a time shift of arrival of approximately $\Delta t_{-HL} = 0.0025$s. When studying the comparison of the Virgo strain with that of Livingston, and using the minus Virgo strain we found an acceptable time shift of arrival of approximately $\Delta t_{-VL} = -0.0005$.

While for Hanford and Livingston strains we have used the pass band of [25,995]Hz, for the Virgo strain we have to use the pass band of [35.3,883]Hz; due to the fact that Virgo had more instrumental and seismic noise present.

In Fig.20 it is shown the graph of $\Lambda(L, -H)$ near the time of the event GW190521.

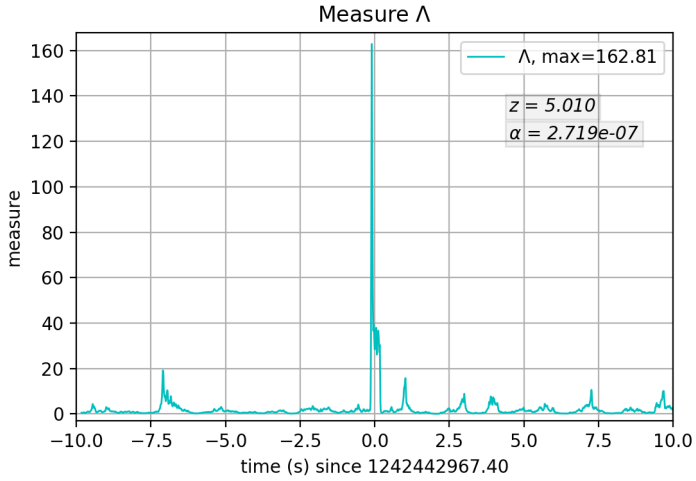

Figure 20: Behavior of the measure $\Lambda(L, -H)$ near the time of the event GW190521; using the relative Hanford to Livingston arrival time of $\Delta t_{-HL} = 0.0025$s and a window of 0.35s. For both detectors the pass band was [25,995]Hz.

From the information of the graph in 20 , we infer a strong Λ signal for the pair $(L, -H)$ near the time of the event GW190521.

When comparing the Virgo signal, one has to take into account, that this detector had a higher noise level, so that the duration of the signal has to be estimated; for which we found the value of 0.18s. Then, using the appropriate delay time, we found no significant Λ signal; as it can be seen in 21 .

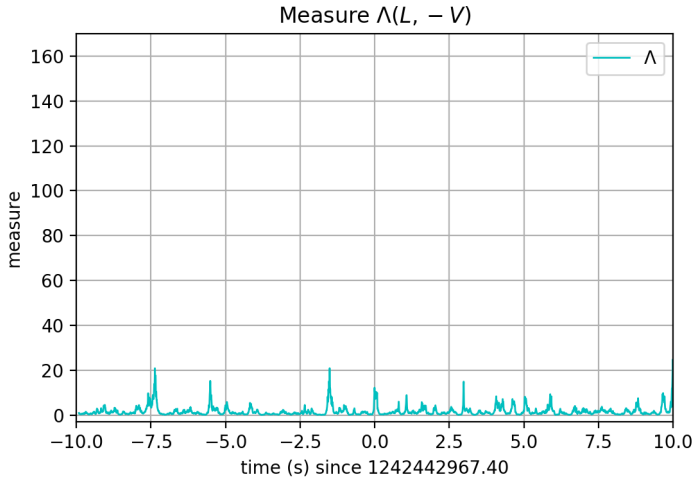

Figure 21: Behavior of the measure $\Lambda(L, -V)$ near the event GW190521; using the relative Virgo to Livingston arrival time of $\Delta t_{-VL} = -0.0005$s and a window of 0.18s. For the Livingston detector the pass band was [25,995]Hz. For the Virgo detector the pass band was [35.3,883]Hz.

As already mentioned, the Virgo strain had higher levels of noise than the LIGO detectors; so that it was somehow expected to have some difficulties when comparing the Virgo strain with the others. In order to mitigate this issue, we have applied a low pass filter of 350Hz to all strains, before doing further comparisons. In Fig. 22 we show the graph of $\Lambda(L, -V)$ near the time of the event GW190521, after applying a low pass 350Hz filter.

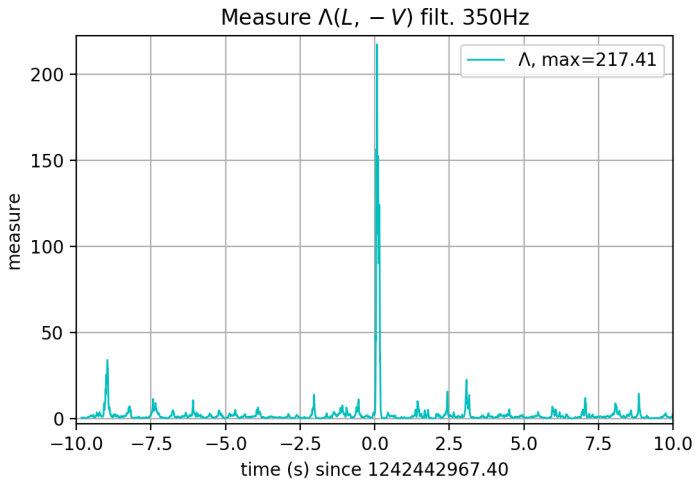

Figure 22: Behavior of the measure $\Lambda(L, -V)$ near the time of the event GW190521; using the arrival time at Virgo relative to Livingston of $\Delta t_{-VL} = -0.0005$s and a window of 0.18s; and an extra low pass filter of 350Hz applied to both strains.

One can now see a strong signal for the measure $\Lambda(L, -V)$ relative to this new statistic. It must be stressed that when applying the extra low pass 350Hz filter, one is changing the statistics, so that the maximum value of Λ in Fig. 22 can not be compared directly with that shown in Fig. 20; which refers to the previous statistics. In fact, the corresponding maximum value of $\Lambda(L, -H)$ after applying the 350Hz filter; is much higher that the maximum shown for $\Lambda(L, -V)$ in the same statistic.

This event of course deserves further study, which will contribute to the localization in the sky of the source of the signal; but here we are concerned with presenting the behavior of the measure OM to a variety of situations, and individual event detailed studies will be carried out elsewhere.

8 Final comments

The difficulties encountered in the identification of transient gravitational signals in the strains recorded by the interferometric gravitational-wave observatories, as LIGO/Virgo, have various facets. We have already mentioned some of them. For example, that some of the detection technique use templates, which in turn use physical assumptions that are probably not tenable for every situation. We have also recalled the problems involved when using whitening techniques. Another difficulty found in the detection process is the appearance of noise transients (glitches) that could trigger false detections. For example, in reference [29] they propose a way to identify and characterize transient noise. Also in [30] the authors present a method to identify glitches in gravitational-wave data. Another effort for glitch classification is presented in [31]. In relation to the problems presented by the appearance of glitches, it can be seen from the nature of our measure Λ, as expressed in (3), that glitches, which occurs at one observatory, will be attenuated by design.

In this work we have presented a new optimized measure OM, denoted by Λ, which we have shown is useful for the detection of similar signals in two strains of gravitational-wave observatories. Our measure works with the data in the time domain, so that it makes efficient use of computer memory resources. With the measure OM we have been able to show the detection of similar signals in the two LIGO strains of the GW150914 event, close to the time of the event, with a duration of about 0.5s. This is more than the 0.1s lapse of time reported in LIGO articles. It is essential to stress that first we have used the pre-processing filtering techniques More19, presented before[12], as a starting point in the treatment of the observed data at LIGO observatories. We have demonstrated here that after applying our filters the strains show a behavior very close to Gaussian.

We have also presented the calculation of the likelihood ratio **L** for the detection of two similar signals in two strains. Due to the fact that several authors use correlation coefficients[32, 33, 34, 35, 36, 37, 38] to analyze the observed data, we have also considered in section 2 the corresponding correlation coefficient ρ for the two strains. In appendix B we have presented the result of applying the correlation coefficient to the problem we have treated here; and it was shown that it gives information with too much noise. We have also shown that it is a weaker measure, when compared with Λ measure; see the table below also.

The application of the measure Λ, in section 3 , to the two LIGO strains of 22s around the GW150914 event, with a nominal time shift of -0.007s, for (-)Hanford data, and with a window of 0.5s length, gives a clean, sharp peak as a function of time; which can be assigned an unrefined level of significance $\alpha_1 = 0.0063$, or a level of significance $\alpha = 0.00010$, by recognizing a Gaussian behavior for the logarithm of Λ. Another way of saying this is that we can trust the hypothesis that there is a similar signal in both strains with a 99.99% confidence level. When comparing our measure with the likelihood ratio **L** statistics and the correlation coefficient ρ statistics, we have shown, that among the three measures studied in section 2 , our measure Λ can be used with a level of significance, for this problem, which is more than 30 times stronger than those than can be used with the likelihood ratio **L** or the correlation coefficient ρ. When applying the measure Λ to the two LIGO strains of 40s around the GW150914 event, the behavior of the measure OM improves; because now the level of significance is $\alpha = 4.0e-05$; and the relation with the other two measures also improves, since now they give signals that are more than 50 times weaker than our measure, as shown en subsection 3.4 . This suggests the preference for the use of the measure OM.

We have also take the opportunity in section 3 to study the effects on the statistics after whitening the strains. We have seen that if one follows the procedures, as suggested in LIGO tutorials, one is left with a strain that is far from a Gaussian aspect. Only after trimming the extremes of the strains, one is left with data that is close to a Gaussian behavior; as shown in Figs. 5 , 6 and 7 . However we have also shown, in subsection 3.3, that the statistics of the optimized measure applied to the whitened data, does not show a Gaussian behavior, and therefore one can not apply this methods in this case. This unwanted behavior is also observed with the other two measures.

In the detailed study of event GW150914, we have shown that it is possible to fine tune the choice of the lapse of time considered, in order to maximize the response for the level of significance; as we demonstrated in subsection 3.2 . Naturally this fine tuning can be done for each event; which we avoid here in order to present examples of first approach to different events.

To check the behavior of the measure OM with the strains of other events, we have also applied a 22s study to the events: 151012_2, GW151012, GW170104 and GW190521 in sections 4 , 5 , 6 and 7 respectively. For each of these other cases we have also carried out preliminary studies in which we have checked minimal Gaussianity, basic behavior in the time domain, phase behavior, amplitude spectral density behavior and others; that we do not include in the text, in order to keep the presentation focused on the measure OM response in different situations.

The 151012_2 has been presented as a release of type 'GWTC-1.marginal', and our study also did not show any signal in the Λ graph.

Event GW151012 has been presented as a release of type 'GWTC-1-confident' with a network SNR of 9.3; which showed very low levels of signals. From the preliminary comparison of the strains with the Δt_{-HL} value of -0.0006s suggested by LIGO, we could not found a signal in the Λ graph. After further study we discovered that the value $\Delta t_{-HL} = -0.0012$s yields a weak signal of the measure OM, using a signal duration of 0.8s; which is longer than the 0.5s LIGO suggested duration.

The event GW170104 has been presented as 'GWTC-1-confident' of the second observing run, with a network SNR of 13.8. In this case, the signal was very strong and our optimize measure OM was able to assign a level of significance of $\alpha = 5.46 \times 10^{-8}$. For this statistics we have used a time duration of 0.28s, instead of the 0.1s shown in the published data[28]. Thus, the optimize measure OM can be used to test the duration of the signal for each event.

The case of event GW190521 is of special interest because the Virgo detector was also in operation at that time. This event was released as of type 'GWTC-2.1-confident', with its version v4 of data. In comparing the different data version we found unexplained filtering, so that we have decided to work with version v1 of the strains, which was released as of type 'O3_Discovery_Papers'. The GW190521 event was assigned the network SNR of 14.3 and a duration of approximately 0.1s[9]. Our study instead shows a duration of approximately 0.35s; which we have used in our work. Although the $\Lambda(L,-H)$ gave a strong signal, the initial settings did not show any signal for the $\Lambda(L,-V)$. We found a signal after applying a low pass filter of 350Hz to the strains, since the spectrograms show that the signals where below this threshold. This case showed that the optimize measure OM can be used even in problematic situations; in which more care has to be taken to analyze the strains.

We have not pursued here any further the deserved individual detailed studies for each event, since we are only concerned here with the presentation of the optimized measure OM. For this reason we have only presented here the first approach of the application of the measure OM to a variety of events with different characteristics and of all three first LIGO/Virgo runs.

Note that in this work we have not attempted to determine the exact functional form of the gravitational wave; since all the information comes from the comparison of the strains through the use of our measure Λ; after preparation with the above mentioned pre-processing filtering technique.

In table 1 we show the comparison of our measure OM with the other two widely known and used measures, the likelihood ratio $\mathbf{L}$ and the correlation coefficient ρ. Since one might feel uncomfortable with recognizing a Gaussian behavior for the measures, we also include the indisputable weaker level of significance α_1; calculated in terms of the direct Chebyshev inequality[1], and having an undeniable probabilistic interpretation. It can be seen that even using only this more conservative and so weaker index, the measure OM is much stronger than the other two considered. In fact, only using the conservative version one obtains excellent confidence levels, since all confidence coefficients are of the order of 99%; as it can be seen in table 2 .

events	level of significance α_1 (Chebyshev inequality)			level of significance α (Gaussian statistics)		
	Λ	$\mathbf{L}$	ρ	$\ln(\Lambda)$	$\ln(\mathbf{L})$	ρ
GW150914	6.32e-03	nan	1.36e-01 (22)	1.047e-04	3.708e-03 (35)	3.384e-03 (32)
GW151012	1.21e-02	nan	1.40e-01 (12)	3.418e-04	4.031e-03 (12)	3.796e-03 (11)
GW170104	1.67e-03	nan	5.67e-02 (34)	5.462e-08	1.489e-05 (273)	1.347e-05 (245)
GW190521	2.11e-03	nan	6.46e-02 (31)	2.719e-07	5.416e-05 (199)	4.184e-05 (154)

Table 1: Comparison of the three measures for different events and two levels of significance. The significance α_1 denotes the unrefined level of significance calculated from Chebyshev inequality; and the significance α is calculated with the Gaussian statistics. In parenthesis we show how weak is the measure with respect to the measure OM. The 'not a number' sign 'nan' is due to the appearance of overflows in standard calculations.

[1]We show in Appendix C that one can trust three significant figures for the strain size considered. But for a more precise bound, we also provide in the appendix a corresponding inequality for a sample.

events	confidence coefficient γ (Chebyshev inequality)
GW150914	0.99368
GW151012	0.9879
GW170104	0.99833
GW190521	0.99789

Table 2: Conservative confidence coefficients for detections for different events, from the measure OM.

Summarizing our work; we have started by developing the likelihood ratio strategy applied to the detection of a similar signal in two time series, and found that this measure has some difficulties involving very big exponents, and undesirable numerical overflow errors. Then, in order to circumvent the likelihood ratio problematic, we have defined the optimized measure OM; which turned out to be stronger than the likelihood measure and the correlation coefficient signals. The comparison has been done in two ways; by direct application of Chebyshev inequality on the bare quantities and, recognizing a Gaussian behavior of the exponents, using the standard normal level of significance. The superiority of the optimized measure OM is demonstrated in tables 1 and 2 above. The measure OM depends on several characteristics of the strains, and its values have to be understood in terms of the local statistic of the strains. We have also shown that it can be applied to a variety of real data, by covering five events, with very different signal characteristics, belonging to the observing runs O1, O2 and O3. In all these cases the measure OM gives reasonable and sensitive signals, including the detection in the Virgo strain of a very low level signal. Therefore, it has been shown in this article that the measure OM is a powerful versatile tool in post-detection studies.

We plan to use this measure in a variety of situations, and we expect that it will become useful in checking possible gravitational lensed black hole mergers by providing a new approach to this problem.

Acknowledgments

We owe great thanks to Carlos Briozzo and Dante Paz for a careful reading of the manuscript and for comments and discussions on statistics and data analysis, and to Emanuel Gallo for criticism and discussions on the physics of gravitational-wave signals. The pre-processing for event GW170104 was done in collaboration with Florencia Urrutia, and that of event GW190521 was done with Camilo Crisman. We thank the criticism of an anonymous Referee that has persuaded us to extend the statistical study of the measure.

We are very grateful to the LIGO/Virgo Collaboration for making available the data and the python scripts on data analysis at https://gwosc.org.

We acknowledge support from SeCyT-UNC, CONICET and Foncyt.

A Arguments to build the measure

A.1 Detection of a known signal

We here recall the basics of the likelihood method as applied to one set of data following the notation of [15].

Samples of white Gaussian noise taken by an instrument having a bandwidth $W = 2\pi\Delta\nu$ will

be Gaussian random variables with the probability density function (p.d.f.)

$$p_1(x) = \frac{1}{\sqrt{2\pi N_0}} e^{-\frac{x^2}{2N_0}}, \tag{8}$$

with

$$N_0 = \frac{NW}{2\pi} = N\Delta\nu; \tag{9}$$

where N is the unilateral spectral density. The joint p.d.f. of samples $x_1, x_2, ..., x_n$ at times $t_1, t_2, ..., t_n$ separated by intervals much longer than $2\pi/W$ will be statistically independent, given by

$$p(x_1, t_1; x_2, t_2; ...; x_n, t_n) = \prod_{k=1}^{n} p_1(x_k) = \frac{1}{(2\pi N_0)^{n/2}} e^{\left(-\frac{1}{2N_0}\sum_{k=1}^{n} x_k^2\right)}. \tag{10}$$

In treating the detection of signals in the presence of this kind of noise, we shall imagine sampling the random processes by an instrument whose bandwidth is much greater than that of any of the signals involved. We can then apply eq. (10) to the values of the noise at times $t_1, t_2, ..., t_n$ that are arbitrarily close.

Let $s(t)$ be a signal superimposed on a Gaussian noise $\mathsf{n}(t)$; so that one observes

$$v(t) = \mathsf{n}(t) + s(t), \tag{11}$$

in the interval $0 < t < T$. The hypotheses H_0 is that the signal is not present, and the hypotheses H_1 is that the signal is present in the observation v.

The observations are made at n uniformly spaced times $t_k = k\Delta t = k\frac{T}{n}$, with $k = 1, 2, ..., n$; with values $v_k = v(t_k)$. The observations for the two possibilities are described by the joint probability density functions $p_0(\mathbf{v}) = p_0(v_1, v_2, ..., v_n)$ and $p_1(\mathbf{v}) = p_1(v_1, v_2, ..., v_n)$ under the hypotheses H_0 and H_1 respectively. The observer's decision is best made on the basis of the likelihood ratio,

$$\mathbf{L}(\mathbf{v}) = \mathbf{L}(v_1, v_2, ..., v_n) = \frac{p_1(\mathbf{v})}{p_0(\mathbf{v})}. \tag{12}$$

Its value for the data at hand is compared with a fixed decision level $\mathbf{L}_0$; if $\mathbf{L}(\mathbf{v}) < \mathbf{L}_0$ the observer decides that there is no signal present.

It is assumed that the measurements of $v(t)$ at times t_k are made by an instrument of such a large bandwidth that however small the intervals Δt between them, their outcomes have statistically independent noise components. Then, under hypothesis H_0 their joint probability density function is

$$p_0(\mathbf{v}) = \frac{1}{\left(2\pi N_0\right)^{-n/2}} \exp\Big(-\sum_{k=1}^{n} \frac{v_k^2}{2N_0}\Big), \tag{13}$$

with $N_0 = \frac{NW}{2\pi}$.

When the signal is present, the part of the observed v_k due to the noise is $v_k - s_k$, with $s_k = s(t_k)$ a sample of the signal. Therefore, the data v_k should behave as independent Gaussian random variables with mean values s_k and variances N_0, namely, under hypothesis H_1 the joint p.d.f. of the data is,

$$p_1(\mathbf{v}) = \frac{1}{\left(2\pi N_0\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_k - s_k)^2}{2N_0}\right). \tag{14}$$

The likelihood ratio, eq. (12), now becomes

$$\mathbf{L}(\mathbf{v}) = \exp\left(\sum_{k=1}^{n} \frac{2s_k\, v_k - s_k^2}{2N_0}\right). \tag{15}$$

The observer chooses hypothesis H_0 if $\mathbf{L}(\mathbf{v}) < \mathbf{L}_0$ or, because of the monotone character of the exponential function, if

$$\Delta t \sum_{k=1}^{n} s_k \, v_k < \frac{1}{2} \Delta t \sum_{k=1}^{n} s_k^2 + N_0 \Delta t \ln \mathbf{L}_0. \tag{16}$$

Hence the observer can base the decision on the value of the quantity

$$G_n = \Delta t \sum_{k=1}^{n} s(t_k) \, v(t_k), \tag{17}$$

comparing it with some fixed amount G_{n0} determined by some criterion. In the n-dimensional Cartesian space with coordinates v_k, the decision surface D is a hyperplane

$$\sum_{k=1}^{n} s(t_k) \, v(t_k) = \text{constant}, \tag{18}$$

which is perpendicular to the vector with components s_k.

It can be seen that the natural inner product of the expected signal and the strain, that we denote by $< \mathbf{v}, \mathbf{s} >= \sum_{k=1}^{n} s(t_k) \, v(t_k)$, is the basic quantity of the likelihood calculation. And if $< \mathbf{s}, \mathbf{s} >$ can be neglected in front of $< \mathbf{v}, \mathbf{s} >$, or if one only concentrates in the functional dependence on the data, one arrives at the working expression for the likelihood to be

$$\mathbf{L}(\mathbf{v}) = \exp\left(\frac{< \mathbf{v}, \mathbf{s} >}{N_0} \right); \tag{19}$$

where for the sake of simplicity in this presentation, we are assuming a constant N_0, although the expressions can easily be generalized. Normally, N_0 is measured from the local properties of the data, close to the time of the event under study.

In actual situations, in which the expected signal has some characteristic length in time, one does not use the natural inner product but a convolution of it with an appropriately chosen window w, with a width of the order of the characteristic length of the signal. So that one actually works with the definition:

$$< \mathbf{v}, \mathbf{s} > (t_j) = \sum_{k} v(t_k) \, s(t_k) \, w(t_j - t_k). \tag{20}$$

In this way one samples the data with an appropriate width.

The likelihood method is used by the LIGO/Virgo Collaboration as a standard way to obtain the matched templates to the observed signals[14, 21].

A.2 The case of the same or similar unknown signal in two detectors

To simplify the notation we are going to omit when possible the index denoting the time variation. Let one detector observe the data v_1, which is supposed to contain the signal s_1 in the presence of the noise n_1, and similarly for the other detector so that

$$v_1 = \mathsf{n}_1 + s_1, \tag{21}$$

and

$$v_2 = \mathsf{n}_2 + s_2; \tag{22}$$

but for a moment let us consider first the basic assumption that both detectors contain the same signal (Although it also applies to similar signals $s_2 = s_1 + \epsilon$, for some small ϵ, as we will show below.)

$$s_1 = s_2 = s. \tag{23}$$

Then one can express

$$\mathsf{n}_1 = v_1 - s = v_1 - v_2 + \mathsf{n}_2; \tag{24}$$

so that instead of (14) now we will have

$$p_1(\mathbf{v}_1, s_1) = \frac{1}{\left(2\pi N_{0_1}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(1)k} - v_{(2)k} + \mathsf{n}_{(2)k})^2}{2N_{0_1}} \right); \tag{25}$$

where $\mathbf{v}_1 = (v_{(1)1}, v_{(1)2}, ..., v_{(1)k}, ...)$ denotes the complete strain of detector 1. In this situation, hypothesis 1 is that detector 1 have recorded the same signal as detector 2, and hypothesis 0 is that detector 1 has not recorded the signal.

Similarly for detector 2 one also has

$$p_2(\mathbf{v}_2, s_2) = \frac{1}{\left(2\pi N_{0_2}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(2)k} - s_{(2)k})^2}{2N_{0_2}} \right). \tag{26}$$

Assuming the statistical independence of the measuring process in the two detectors we arrive at the joint probability from the product of the probabilities calculated for each detector; namely

$$\begin{aligned} p(\mathbf{v}_1, s_1, \mathbf{v}_2, s_1) = & \frac{1}{\left(2\pi N_{0_1}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(1)k} - v_{(2)k} + \mathsf{n}_{(2)k})^2}{2N_{0_1}} \right) \\ & \frac{1}{\left(2\pi N_{0_2}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(2)k} - v_{(1)k} + \mathsf{n}_{(1)k})^2}{2N_{0_2}} \right). \end{aligned} \tag{27}$$

Note that also (25) can be understood as the conditional probability that detector 1 has observed signal s given that detector 2 has observed signal s. (See section 2.4 of [16].) Then, using Bayes' rule[39] one would also arrive at (27).

The likelihood ratio is calculated from the quotient of $p(\mathbf{v}_1, s_1, \mathbf{v}_2, s_1)$ with the corresponding $p_0(\mathbf{v}_1, \mathbf{v}_2)$; where in p_0 there is no contribution from any signal and is given by

$$\begin{aligned} p(\mathbf{v}_1, \mathbf{v}_2) = & \frac{1}{\left(2\pi N_{0_1}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(1)k})^2}{2N_{0_1}} \right) \\ & \frac{1}{\left(2\pi N_{0_2}\right)^{-n/2}} \exp\left(-\sum_{k=1}^{n} \frac{(v_{(2)k})^2}{2N_{0_2}} \right). \end{aligned} \tag{28}$$

This quotient has the form of a product $\mathbf{L}_1\,\mathbf{L}_2$.

Let us start by calculating the likelihood $\mathbf{L}_1$ considering just one of these factors, and, as above, using $s_1 = s_2$, so that

$$\begin{aligned} \mathbf{L}_1(\mathbf{v}_1) = & \frac{p_1(\mathbf{v}_1)}{p_0(\mathbf{v}_1)} \\ = & \exp\left[\sum_{k=1}^{n} \frac{2v_{(1)k}\, v_{(2)k} - v_{(2)k}^2 - 2v_{(1)k}\, \mathsf{n}_{(2)k} + 2v_{(2)k}\, \mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2}{2N_{0_1}} \right]. \end{aligned} \tag{29}$$

It is convenient to manage the algebra in the following way:

$$
\begin{aligned}
&\sum_{k=1}^{n} 2v_{(1)k}\, v_{(2)k} - v_{(2)k}^2 - 2v_{(1)k}\, \mathsf{n}_{(2)k} + 2v_{(2)k}\, \mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2 \\
&= \sum_{k=1}^{n} 2v_{(1)k}\, v_{(2)k} - v_{(2)k}(v_{(1)k} + \mathsf{n}_{(2)k} - \mathsf{n}_{(1)k}) - 2v_{(1)k}\, \mathsf{n}_{(2)k} + 2v_{(2)k}\, \mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2 \\
&= \sum_{k=1}^{n} 2v_{(1)k}\, v_{(2)k} - v_{(2)k} v_{(1)k} - v_{(2)k}\mathsf{n}_{(2)k} + v_{(2)k}\mathsf{n}_{(1)k} \\
&\qquad - 2(s_{(1)k} + \mathsf{n}_{(1)k})\, \mathsf{n}_{(2)k} + 2(s_{(2)k} + \mathsf{n}_{(2)k})\, \mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2 \\
&= \sum_{k=1}^{n} v_{(1)k}\, v_{(2)k} - (s_{(2)k} + \mathsf{n}_{(2)k})\mathsf{n}_{(2)k} + (s_{(2)k} + \mathsf{n}_{(2)k})\mathsf{n}_{(1)k} \\
&\qquad - 2(s_{(1)k} + \mathsf{n}_{(1)k})\, \mathsf{n}_{(2)k} + 2(s_{(2)k} + \mathsf{n}_{(2)k})\, \mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2 \\
&= \sum_{k=1}^{n} v_{(1)k}\, v_{(2)k} - s_{(2)k}\mathsf{n}_{(2)k} - \mathsf{n}_{(2)k}^2 + s_{(2)k}\mathsf{n}_{(1)k} + \mathsf{n}_{(2)k}\mathsf{n}_{(1)k} \\
&\qquad - 2s_{(1)k}\mathsf{n}_{(2)k} - 2\mathsf{n}_{(1)k}\, \mathsf{n}_{(2)k} + 2s_{(2)k}\mathsf{n}_{(2)k} + 2\mathsf{n}_{(2)k}^2 - \mathsf{n}_{(2)k}^2 \\
&= \sum_{k=1}^{n} v_{(1)k}\, v_{(2)k} + s_{(2)k}\mathsf{n}_{(1)k} + s_{(2)k}\mathsf{n}_{(2)k} - 2s_{(1)k}\mathsf{n}_{(2)k} - \mathsf{n}_{(1)k}\, \mathsf{n}_{(2)k}.
\end{aligned}
\tag{30}
$$

In this algebraic manipulation we have kept the identities of s_1 and s_2 to allow for the situation $s_2 = s_1 + \epsilon$, with $\max|\epsilon| \ll \max|s_1|$.

Since by assumption the noises n_1 and n_2 are considered to have independent Gaussian behavior; we can take the size of the interval big enough so that we attain

$$
\sum_{k=1}^{n} s_{(2)k}\mathsf{n}_{(1)k} \approx \sum_{k=1}^{n} s_{(2)k}\mathsf{n}_{(2)k} \approx \sum_{k=1}^{n} s_{(1)k}\mathsf{n}_{(2)k}, \tag{31}
$$

and therefore

$$
\sum_{k=1}^{n} s_{(2)k}\mathsf{n}_{(1)k} + \sum_{k=1}^{n} s_{(2)k}\mathsf{n}_{(2)k} - 2\sum_{k=1}^{n} s_{(1)k}\mathsf{n}_{(2)k} \approx 0, \tag{32}
$$

and we also have

$$
\sum_{k=1}^{n} \mathsf{n}_{(1)k}\, \mathsf{n}_{(2)k} \approx 0. \tag{33}
$$

So that we arrive at

$$
\mathbf{L}_1(\mathbf{v}_1) = \exp\left[\sum_{k=1}^{n} \frac{v_{(1)k}\, v_{(2)k}}{2N_{0_1}}\right]; \tag{34}
$$

for the likelihood of having the signal s in v_1 that is contained in v_2. From the considerations above, we deduce that the joint likelihood of having the same signal in both detectors is then

$$
\mathbf{L}_a(\mathbf{v}_1, \mathbf{v}_2) = \exp\left[\frac{1}{2}\Big(\frac{1}{N_{0_1}} + \frac{1}{N_{0_2}}\Big)\sum_{k=1}^{n} v_{(1)k}\, v_{(2)k}\right]; \tag{35}
$$

where in practical situation we evaluate N_0 from the sample variance so that the final expression is

$$
\mathbf{L}(\mathbf{v}_1, \mathbf{v}_2) = \exp\left[\frac{m-1}{2}\Big(\frac{1}{\sum_{k=1}^{m} v_{(1)k}^2} + \frac{1}{\sum_{k=1}^{m} v_{(2)k}^2}\Big)\sum_{k=1}^{n} v_{(1)k}\, v_{(2)k}\right]; \tag{1}
$$

where the width of the window to calculate σ^2, that is m, is chosen appropriately depending on the nature of the observations v_k, and we are assuming that the means are zero.

This estimation of the desired measure have some difficulties. The exponent is huge, when using actual LIGO data, and it does not emphasize the comparison of the data in both strains. To understand this in detail, let us see what is the behavior of the logarithm of this measure, when it is applied to the data of GW150914 after post-processing[12], which we shown in figure 23 ; where we have advanced the time axis by the width of the window.

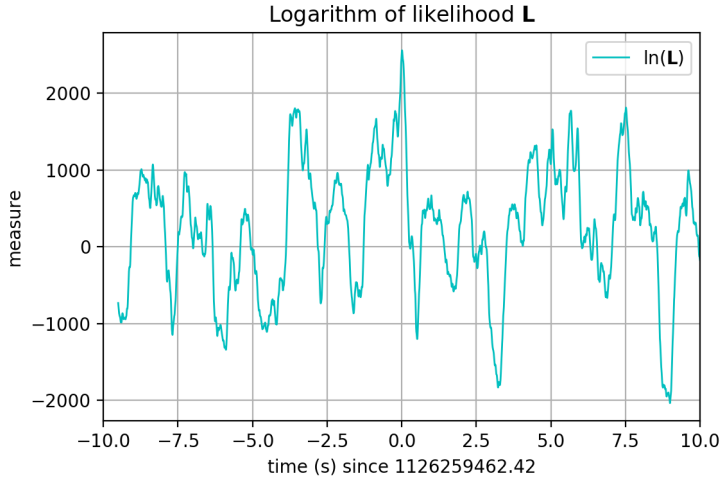

Figure 23: Logarithm of the likelihood ratio applied to the strain at Livingston and the one at Hanford from -10s to 10s around the time of the event GW150914 with the nominal shift of -0.007s.

A.3 The new measure

Due to the unpleasant behavior of the likelihood ratio discussed above, we decide to define a new measure and so we introduce some changes to the likelihood ratio to strengthen the comparison and to moderate the amplitude, so that it becomes useful for the system we have in mind.

To motivate our choice, let us note that for pure and independent noise one has

$$\overline{(\mathsf{n}_1 - \mathsf{n}_2)^2} = \overline{(\mathsf{n}_1)^2} + \overline{(\mathsf{n}_2)^2} - 2\overline{(\mathsf{n}_1\mathsf{n}_2)} = \overline{(\mathsf{n}_1)^2} + \overline{(\mathsf{n}_2)^2} = \sigma_1^2 + \sigma_2^2. \tag{36}$$

Also note that when both variances are similar, then one has that $\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$ is approximately $\frac{4}{\sigma_1^2+\sigma_2^2}$. So that instead of $\frac{1}{N_{0_1}} + \frac{1}{N_{0_2}}$ we can use the estimation $\frac{4}{(v_1-v_2)^2}$; which has the advantage that $v_1 - v_2$ would cancel the information of the signal; and therefore help in accentuating the measure at the time of coincidence.

Then, in order to control this too sensitive behavior, we moderate the measure to be

$$\Lambda_a(\mathbf{v}_1, \mathbf{v}_2) = \exp\left[\frac{1}{\sigma_{12}^*}\left(\frac{1}{\sum_{j=1}^m (v_{(1)j} - v_{(2)j})^2}\right)\sum_{k=1}^n v_{(1)k}\, v_{(2)k}\right], \tag{37}$$

where σ_{12}^* is the standard deviation of $\left(\frac{1}{\sum_{j=1}^m (v_{(1)j} - v_{(2)j})^2}\right)\sum_{k=1}^n v_{(1)k}\, v_{(2)k}$ in the lapse of time of interest.

Employing the notation of the inner product as in (20), which makes use of the window w, we arrive at the final expression for the measure given by

$$\Lambda(\mathbf{v}_1, \mathbf{v}_2) = \exp\left[\frac{1}{\sigma_{12}^*}\frac{< \mathbf{v}_1, \mathbf{v}_2 >}{< (\mathbf{v}_1 - \mathbf{v}_2), (\mathbf{v}_1 - \mathbf{v}_2) >}\right]; \tag{38}$$

where σ^*_{12} is calculated with the window w also.

This measure gives reasonable results with actual LIGO data.

B Behavior of the correlation coefficient

A natural question is whether the correlation coefficient between the two set of data coming from both detectors, is enough to have a well behaved measure for determining if there is a common signal in the strains. For this reason we here present the graphs that show the behavior of the coefficient:

$$\rho_{\mathbf{v}_1,\mathbf{v}_2} \equiv \frac{< \mathbf{v}_1, \mathbf{v}_2 >}{\sqrt{< \mathbf{v}_1, \mathbf{v}_1 >< \mathbf{v}_2, \mathbf{v}_2 >}}; \tag{4}$$

where we are assuming zero average for both strains. This measure is also called the sample correlation coefficient[15].

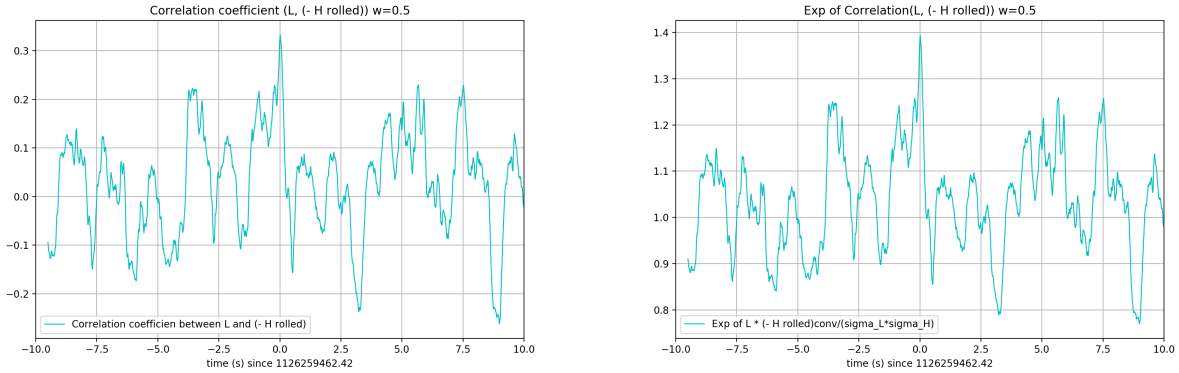

Figure 24: On the left the correlation coefficient between the strain at Livingston and the one at Hanford from -10s to 10s around the time of the event. On the right the exponential of the correlation coefficient in the same lapse of time.

In figure 24 it is shown the behavior of the correlation coefficient, where we have made use of the same window employed for the Λ measure, and where we have advanced the time axis by the width of the window. We also show the behavior of the exponentiation of the correlation coefficient, to see if the relation was augmented; but it can be seen that although there is a local maximum close to the time of the event, both graphs give information with too much noise. So, comparing this with the cleaner behavior of our measure Λ, as shown in figure 8 , we choose Λ; which provides a much better tool for analysis.

C Chebyshev inequality for a sample

The case of a sample has been considered in reference [40], and based on this, in [41] it was presented the following inequality. Let $r \geq 2$ a fixed integer, $Y_1, Y_2, ..., Y_r$ and X a weakly exchangeable sample (e.g. identically and independently distributed (i.i.d.), but not necessarily) from some unknown distribution such that $P(Y_1 = Y_2 = ... = Y_r = X) = 0$, and $\lambda \geq 1$. Denote $\bar{Y} = \frac{1}{r}\sum_{i=1}^{r} Y_i$ and $\overline{Var(Y)} = \frac{1}{r-1}\sum_{i=1}^{r}(Y_i - \bar{Y})^2$ the sample mean and variance respectively, and $Q^2 = \frac{r+1}{r}\overline{Var(Y)}$. Then,

$$P(|X - \bar{Y}| \geq \lambda Q) \leq \frac{1}{r+1}\left\lfloor \frac{r+1}{r}\left(\frac{r-1}{\lambda^2} + 1\right)\right\rfloor; \tag{39}$$

where the notation is: $\lfloor x \rfloor$ is the largest integer less than x.

In order to relate to equation 6, let us define t by $\lambda Q = t\sigma$; where σ is the sample standard deviation. Then, one has $\lambda\sqrt{\frac{r+1}{r}} = t$; so that, using

$$\frac{1}{r+1}\left\lfloor \frac{r+1}{r}\left(\frac{r-1}{\lambda^2}+1\right)\right\rfloor = \frac{1}{r+1}\left\lfloor \frac{r+1}{r}\left(\frac{(r-1)(r+1)}{rt^2}+1\right)\right\rfloor ; \tag{40}$$

one has

$$P(|X-\bar{Y}| \geq t\sigma) \leq \frac{1}{r+1}\left\lfloor (r+1)\left(\frac{1}{t^2}\left(1-\frac{1}{r^2}\right)+\frac{1}{r}\right)\right\rfloor . \tag{41}$$

Let us note that for 22 seconds at a sample rate of $fs = 16384$, one would have $r = 360448$; so that $\frac{1}{r} = 2.77432\times10^{-6}$ and $\frac{1}{r^2} = 7.69688\times10^{-12}$. Let us define $g(r,t) = \frac{1}{r+1}\left\lfloor (r+1)\left(\frac{1}{t^2}\left(1-\frac{1}{r^2}\right)+\frac{1}{r}\right)\right\rfloor$. In Fig. 25 we shown the graph of the relative difference of inequality 41 with 6 for the range of t in $(2, 30)$; which covers the values for the cases studied here.

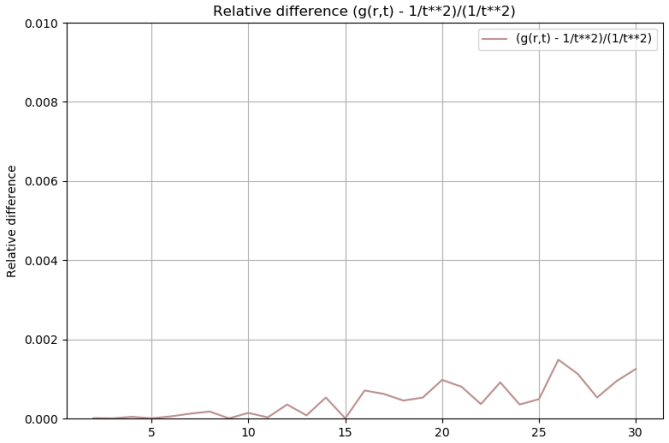

Figure 25: Relative difference of sample inequality and original inequality.

It is deduced that in the range $t \in (2, 30)$ one can safely trust at least three significant figures of the original inequality for this size of sample; which we have tested numerically for the events considered.

References

[1] **LIGO Scientific, Virgo** Collaboration, R. Abbott *et al.*, "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," *Phys. Rev. X* **11** (2021) 021053, arXiv:2010.14527 [gr-qc].

[2] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X* **9** no. 3, (2019) 031040, arXiv:1811.12907 [astro-ph.HE].

[3] S. Klimenko *et al.*, "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors," *Phys. Rev. D* **93** no. 4, (2016) 042004, arXiv:1511.05999 [gr-qc].

[4] **KAGRA, VIRGO, LIGO Scientific** Collaboration, R. Abbott *et al.*, "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3," *Phys. Rev. X* **13** no. 1, (2023) 011048, arXiv:2111.03634 [astro-ph.HE].

[5] B. S. Sathyaprakash and S. V. Dhurandhar, "Choice of filters for the detection of gravitational waves from coalescing binaries," *Phys. Rev.* **D44** (1991) 3819–3834.

[6] S. V. Dhurandhar and B. S. Sathyaprakash, "Choice of filters for the detection of gravitational waves from coalescing binaries. 2. Detection in colored noise," *Phys. Rev.* **D49** (1994) 1707–1722.

[7] B. J. Owen and B. S. Sathyaprakash, "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement," *Phys. Rev. D* **60** (1999) 022002, arXiv:gr-qc/9808076.

[8] I. W. Harry, B. Allen, and B. S. Sathyaprakash, "A Stochastic template placement algorithm for gravitational wave data analysis," *Phys. Rev. D* **80** (2009) 104014, arXiv:0908.2090 [gr-qc].

[9] **LIGO Scientific, Virgo** Collaboration, R. Abbott *et al.*, "GW190521: A Binary Black Hole Merger with a Total Mass of 150 M$\odot$," *Phys. Rev. Lett.* **125** no. 10, (2020) 101102, arXiv:2009.01075 [gr-qc].

[10] **LIGO Scientific, Virgo** Collaboration, R. Abbott *et al.*, "Properties and Astrophysical Implications of the 150 $M_\odot$ Binary Black Hole Merger GW190521," *Astrophys. J. Lett.* **900** no. 1, (2020) L13, arXiv:2009.01190 [astro-ph.HE].

[11] R. Gamba, M. Breschi, G. Carullo, S. Albanesi, P. Rettegno, S. Bernuzzi, and A. Nagar, "GW190521 as a dynamical capture of two nonspinning black holes," *Nature Astron.* **7** no. 1, (2023) 11–17, arXiv:2106.05575 [gr-qc].

[12] O. M. Moreschi, "Convenient filtering techniques for LIGO strain of the GW150914 event," *JCAP* **1904** (2019) 032, arXiv:1903.00546 [gr-qc].

[13] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** no. 6, (2016) 061102, arXiv:1602.03837 [gr-qc].

[14] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals," *Class. Quant. Grav.* **37** no. 5, (2020) 055002, arXiv:1908.11170 [gr-qc].

[15] C. W. Helstrom, *Statistical theory of signal detection*. Pergamon Press, second ed., 1975.

[16] E. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. Springer, third ed., 2005.

[17] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, 1967.

[18] R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise*. Academic Press, second ed., 1995.

[19] G. Casella and R. L. Berger, *Statistical Inference*. DUXBURY, second ed., 2002.

[20] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "GW150914: First results from the search for binary black hole coalescence with Advanced LIGO," *Phys. Rev.* **D93** no. 12, (2016) 122003, arXiv:1602.03839 [gr-qc].

[21] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Properties of the Binary Black Hole Merger GW150914," *Phys. Rev. Lett.* **116** no. 24, (2016) 241102, arXiv:1602.03840 [gr-qc].

[22] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914," *Class. Quant. Grav.* **33** no. 13, (2016) 134001, arXiv:1602.03844 [gr-qc].

[23] P. Jaranowski and A. Krolak, "Gravitational-Wave Data Analysis. Formalism and Sample Applications: The Gaussian Case," *Living Rev. Rel.* **8** (2005) 3, arXiv:0711.1115 [gr-qc]. [Living Rev. Rel.15,4(2012)].

[24] P. Jaranowski and A. Krolak, *Analysis of gravitational-wave data.* Cambridge, UK: Univ. Pr. 257 p, 2009. http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521864596.

[25] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Binary Black Hole Mergers in the first Advanced LIGO Observing Run," *Phys. Rev. X* **6** no. 4, (2016) 041015, arXiv:1606.04856 [gr-qc]. [Erratum: Phys.Rev.X 8, 039903 (2018)].

[26] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence," *Phys. Rev. Lett.* **116** no. 24, (2016) 241103, arXiv:1606.04855 [gr-qc].

[27] C. Talbot and E. Thrane, "Gravitational-wave astronomy with an uncertain noise power spectral density," *Phys. Rev. Res.* **2** no. 4, (2020) 043298, arXiv:2006.05292 [astro-ph.IM].

[28] **VIRGO, LIGO Scientific** Collaboration, B. P. Abbott *et al.*, "GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2," *Phys. Rev. Lett.* **118** no. 22, (2017) 221101, arXiv:1706.01812 [gr-qc].

[29] J. Ding, R. Ng, and J. McIver, "UniMAP: model-free detection of unclassified noise transients in LIGO-Virgo data using the temporal outlier factor," *Class. Quant. Grav.* **39** no. 13, (2022) 135011, arXiv:2111.09465 [gr-qc].

[30] L. Vazsonyi and D. Davis, "Identifying glitches near gravitational-wave signals from compact binary coalescences using the Q-transform," *Class. Quant. Grav.* **40** no. 3, (2023) 035008, arXiv:2208.12338 [astro-ph.IM].

[31] G. Ashton, S. Thiele, Y. Lecoeuche, J. McIver, and L. K. Nuttall, "Parameterised population models of transient non-Gaussian noise in the LIGO gravitational-wave detectors," *Class. Quant. Grav.* **39** no. 17, (2022) 175004, arXiv:2110.02689 [gr-qc].

[32] M. A. Green and J. W. Moffat, "Extraction of black hole coalescence waveforms from noisy data," *Phys. Lett.* **B784** (2018) 312–323, arXiv:1711.00347 [astro-ph.IM].

[33] H. Liu and A. D. Jackson, "Possible associated signal with GW150914 in the LIGO data," *JCAP* **1610** no. 10, (2016) 014, arXiv:1609.08346 [astro-ph.IM].

[34] J. Creswell, H. Liu, A. D. Jackson, S. von Hausegger, and P. Naselsky, "Degeneracy of gravitational waveforms in the context of GW150914," *JCAP* **1803** no. 03, (2018) 007, arXiv:1803.02350 [gr-qc].

[35] H. Liu, J. Creswell, S. von Hausegger, A. D. Jackson, and P. Naselsky, "A blind search for a common signal in gravitational wave detectors," *JCAP* **1802** no. 02, (2018) 013, arXiv:1802.00340 [astro-ph.IM].

[36] H. Liu, J. Creswell, S. von Hausegger, P. Naselsky, and A. D. Jackson, "The role of redundancy in blind signal estimation for multiple gravitational wave detectors," *Int. J. Mod. Phys.* **D28** no. 04, (2018) 1930009.

[37] A. B. Nielsen, A. H. Nitz, C. D. Capano, and D. A. Brown, "Investigating the noise residuals around the gravitational wave event GW150914," *JCAP* **1902** (2019) 019, arXiv:1811.04071 [astro-ph.HE].

[38] A. D. Jackson, H. Liu, and P. Naselsky, "Noise residuals for GW150914 using maximum likelihood and numerical relativity templates," *JCAP* **1905** (2019) 014, arXiv:1903.02401 [astro-ph.IM].

[39] N. V. Kampen, *Stochastic Processes in Physics and Chemistry.* Elsevier, third ed., 2007.

[40] J. Saw, M. Yang, and T. Mo, "Chebyshev inequality with estimated mean and variance," *Am. Stat.* **38(2)** (1984) 130–132.

[41] A. Kabán, "Non-parametric detection of meaningless distances in high dimensional data," *Stat Comput* **22** (2012) 375–385.

Universidad Nacional de Córdoba
1983/2023 - 40 AÑOS DE DEMOCRACIA

Hoja Adicional de Firmas
Informe Gráfico

Número: IF-2023-00750845-UNC-ME#FAMAF

CORDOBA, CORDOBA
Viernes 1 de Septiembre de 2023

Referencia: MORESCHI- Documentacion informe año sabatico parte 3//
EX-2021-00478970- -UNC-ME#FAMAF

El documento fue importado por el sistema GEDO con un total de 31 pagina/s.

Digitally signed by GDE UNC
DN: cn=GDE UNC, c=AR, o=Universidad Nacional de Cordoba, ou=Informatica, serialNumber=CUIT 30546670623
Date: 2023.09.01 12:27:33 -03'00'

SILVIA ANAHI VERON
Administrativa
Mesa de Entradas Facultad de Matemática, Astronomía, Física y Computación
Universidad Nacional de Córdoba

Digitally signed by GDE UNC
DN: cn=GDE UNC, c=AR, o=Universidad Nacional de Cordoba, ou=Informatica, serialNumber=CUIT 30546670623
Date: 2023.09.01 12:27:39 -03'00'