PREPARED FOR SUBMISSION TO JCAP

On the matched filter Signal to Noise Ratio in LIGO signals

Akhila Raman[a]

[a]University of California at Berkeley

E-mail: akhila.raman@berkeley.edu

Abstract. This article analyzes the data for the three gravitational wave (GW) events detected in both Hanford(H1) and Livingston(L1) detectors by the LIGO[1] collaboration. It is shown that GW151226 and GW170104 are very weak signals whose amplitude does not rise significantly during the GW event, and they are indistinguishable from non-stationary detector noise. It is shown that LIGO's matched filter, which uses long 32 second windows, misfires with high SNR, coincidentally at H1 and L1 sites, even when correlating a low amplitude sine wave or white gaussian noise with the template. Normalized Cross Correlation Function(CCF) method is implemented using short windows, and it is shown that the normalized CCF is poor for GW151226 and GW170104, when correlating H1/L1 and template, which is indistinguishable from detector noise correlations with the template. It is also shown that normalized CCF is poor, for GW151226 and GW170104, when correlating H1 and L1, which is indistinguishable from detector noise correlations. Hence it is suggested that GW151226 and GW170104 be questioned as candidates for GW signals. The implication of these results for GW150914 is also examined. All the results in this paper are demonstrated using modified versions of LIGO's Python scripts[13].[2]

[1]The Laser Interferometer Gravitational-Wave Observatory

[2]The specific Python script used to generate Fig.1 to Fig.13 in this manuscript, is mentioned in [13].

Contents

1 Introduction

The first GW signal observed was GW150914 [1] which was a relatively strong signal whose amplitude, after whitening and filtering[1], rose significantly, well over detector noise level, during the 0.2 second GW event duration. In comparison, the second signal[2] GW151226[Fig. 2] and the third signal[3] GW170104[Fig. 3] were very weak signals, which look like noise after whitening and filtering and whose signal amplitude does not rise above the detector noise level during the GW event of duration 1 and 0.12 seconds, respectively.(Fig. 1)

This raises the important question of whether GW151226 and GW170104 could have been caused by non-stationary detector noise. Creswell et al.[5] have reported correlations in the detector noise which, at the time of the event, happen to be maximized for the same time lag as that found for the GW event itself.

Because we are more often likely to observe weak signals which look like noise and whose amplitude does not rise during assumed GW event, it is of paramount importance that we should not classify noise as GW events. We must insist on high standards before classifying an observed time series as a GW signal.

[1]Fourth order Butterworth bandpass filters used in frequency range 20-300Hz for GW150914, 43-800Hz for GW170104 and GW151226.

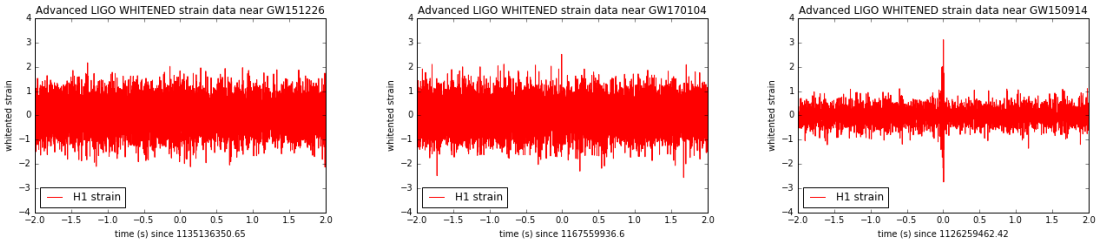

Figure 1. Plots of H1 whitened and filtered strain in GW151226, GW170104 and GW150914.

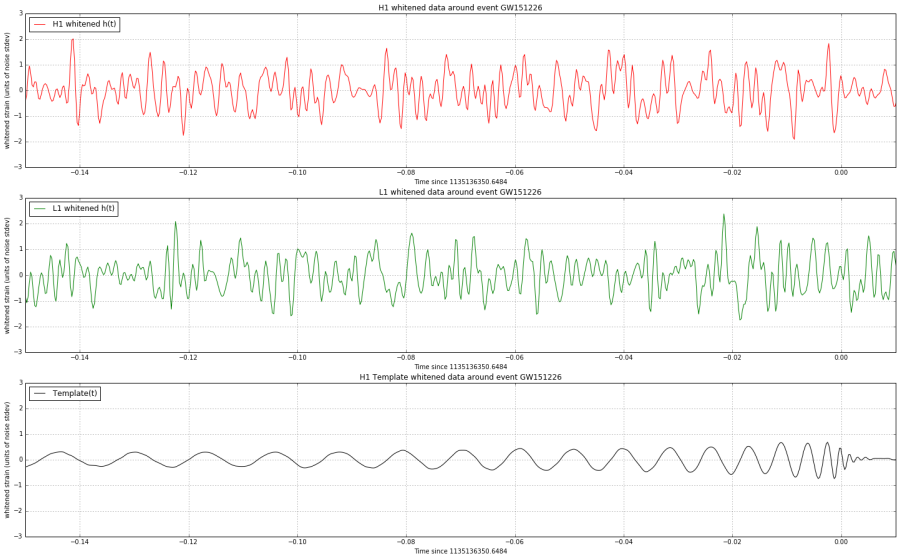

Figure 2. GW151226 whitened and filtered H1 and L1 strain, and the template

The organization of this paper is as follows. In Section 2 we consider the effect of non-stationarity of detector noise on the false alarm rate computation and argue that brief bursts in detector noise coincident at both detectors could produce a high SNR in the matched filter, even though the amplitude of the whitened signal does not rise significantly over detector noise level in the time domain. Such bursts could be mistaken for GW events, particularly for GW151226 and GW170104.

In Section 3, we examine the theory and implementation of LIGO's matched filter which is the **core engine** driving the two independent analyses, PYCBC and GSTLAL. It will be shown that the matched filter misfires with high SNR even for a sine wave and also noise. In Section 4, using Normalized Cross Correlation Function(CCF) method, it will be shown that the normalized CCF of H1/L1 with the template, is indistinguishable from the CCF of detector noise vs template, for GW151226 and GW170104.

In Section 5, using Normalized Cross Correlation Function(CCF) method, it will be shown that the normalized CCF of H1 vs L1, is indistinguishable from detector noise correlations, for GW151226 and GW170104. In Section 6, the case is presented for questioning GW151226 and GW170104 as candidates for GW signals. In Section 7, a case for questioning GW150914 is presented, despite its

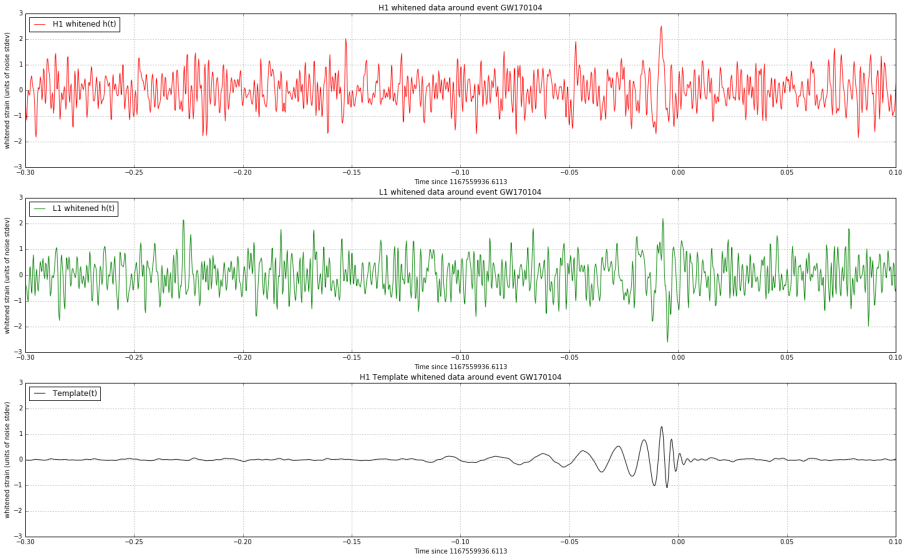

Figure 3. GW170104 whitened and filtered H1 and L1 strain, and the template

relatively strong signal amplitude during the GW event window.

2 Non-stationary Detector Noise and its effect on False Alarm Rate

LIGO detector noise is non-stationary and non-Gaussian and has no analytical model. The number of background noise events which cross a specified SNR threshold, at the output of the matched filter, is determined empirically [4]. Creswell et al.[5] have pointed out that the sources of non-stationary and non-Gaussian detector noise need to be identified and eliminated to ensure reliable GW detection and that analysis methods for stationary noise cannot be used for non-stationary noise in LIGO detectors. It has been pointed out that, LIGO does not assume the noise is Gaussian and stationary for the observation of compact binary mergers.

It should be noted that, when LIGO papers claim high SNR for the three GW events [1–3], they mean SNR observed at the output of the matched filter. This does not necessarily mean that the signal amplitude during the GW event is significantly higher than the background noise amplitude. Let us define Signal Power Ratio (**SPR**) which is defined as the ratio of signal power during the GW event to the signal power outside the GW event (background detector noise). We can see in Fig. 1 that SPR is close to unity for weak signals GW151226(SPR=1.29) and GW170104(SPR=1.08), while GW150914 has relatively higher SPR=4.122. When SPR is close to unity or less than unity for weak signals, assumed GW template is buried in the detector noise, and the probability that non-stationary detector noise simulates the template becomes higher.

LIGO's search software uses two independent analyses, PYCBC and GSTLAL, and false alarm rate computations in Eq.8 and Eq.12 in [4] estimate empirically the number of coincident events above a certain threshold. For example, in the PYCBC analysis, the false alarm rate is given by the

equation below, where T is the observation time of the search, T_b is the background time, $n_b(\hat{\rho}_c)$ is the number of background noise events that cross the candidate event's re-weighted SNR threshold.

$$F(\hat{\rho}_c) = 1 - e^{\frac{-T}{T_b}(1+n_b(\hat{\rho}_c))} \tag{2.1}$$

We may not observe high coincident SNR during 8 days of observation, but during next several months, by **definition** of non-stationarity, there is a possibility that the mean and variance of detector noise **may** change unpredictably and simultaneously in both sites, producing high coincident SNR in the output of the matched filter, even though the signal amplitude does not rise during the GW event, as is the case for GW151226 and GW170104 (Fig. 1). This will result in the estimate of the false alarm rate not applicable, particularly for weak signals GW151226 and 170104.

This is illustrated by an example below, where even low amplitude Additive White Gaussian Noise(AWGN), added to the detector noise, can cause high SNR and re-weighted SNR at the matched filter, coincident at both sites, in mere 32 seconds of observation.

2.1 LIGO's Matched Filter Misfires for AWGN

This point is illustrated by the signal $w_d(t) + a(t)$, where $w_d(t)$ is the **unwhitened** detector noise, obtained from a 32 second block of data which does not contain the GW signal, and $a(t)$ is AWGN of duration 1 second, scaled so that the standard deviation of $a(t)$ is at least 500 times lower than standard deviation of $w_d(t)$. Then we correlate $w_d(t) + a(t)$ with the GW151226 template $h(t)$ in the LIGO matched filter. Fig. 4 shows the results, and we can see from the right panels that, even white gaussian noise added to the detector noise can produce high reweighted SNR in LIGO's matched filter, similar to what actual GW151226 produces (plot). Similar results are observed for GW170104. We observe high SNR and re-weighted SNR at the matched filter, coincident at both sites, in mere 32 seconds of observation.

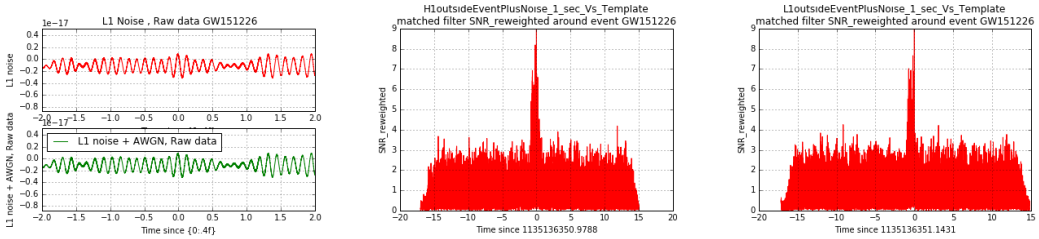

Figure 4. GW151226: Left panel: Raw unwhitened L1 detector noise, with and without AWGN added. Right Panels: LIGO matched filter Reweighted SNR for AWGN + H1/L1 detector noise vs template.

Hence, we cannot be sure that the high SNR is produced by the GW signal and **not** by a low amplitude coincident noise burst. For this reason alone[Reason 1], **independent** of other reasons, GW151226 and GW170104 should be questioned.

2.2 Why False Alarm Rate calculation is unreliable

GW151226 is observed over a duration of 1 second in the window $[t_1, t_2]$. Then LIGO software searches the vicinity for a duration of 10 days before or after the event, say $[t_a, t_b]$, where $t_a = t_1 - (11$ days) and $t_b = t_1 - (1$ day). It is possible that, during observation period $[t_a, t_b]$, there were no **coincident** noise bursts, and measured noise events $n_b(\hat{\rho}_c)$ was low, and computed false alarm rate is 1 in 200,000 years.

Given the non-stationarity of detector noise, it is possible that, **only** during the presumed GW event window $[t_1, t_2]$, there was indeed low amplitude noise burst, as in Fig. 4, causing high reweighted SNR in the matched filter for both H1 and L1 and high detection statistic value(DSV). This noise burst can be $\frac{1}{500}$ times **smaller** than the standard deviation of detector noise.

There is simply **no way** we can make an argument that, just because LIGO team observed no coincident noise bursts in $[t_a, t_b]$, we cannot expect a coincident noise burst in $[t_1, t_2]$, because detector noise is non-stationary. Even a low-amplitude noise burst with standard deviation $\frac{1}{500}$ times **smaller** than the standard deviation of detector noise, can cause **coincident** high SNR in matched filter. We should **include** this event $[t_1, t_2]$, in characterizing detector noise and **not pre-suppose** that these are valid GW events and then look in the vicinity only, **excluding** these events, for detector noise characterization and false alarm rate computation.

The false alarm rate computation associated with the 5σ significance, using the expression $F(\hat{\rho_c}) = 1 - e^{\frac{-T}{T_b}(1+n_b(\hat{\rho_c}))}$, is dependent on the critical parameter $n_b(\hat{\rho_c})$, which is empirically determined, by observations over a number of days, near the vicinity of the GW event. This empirical parameter $n_b(\hat{\rho_c})$ is the **weakest** link in the argument. The above counter-example with coincident low amplitude noise bursts, directly rebuts it. The empirical parameter $n_b(\hat{\rho_c})$ has **no theoretical basis** and is **not** a reliable predictor of false alarm rate.

If the computed false alarm rate is 200,000 years, this only means that the **average time** between 2 successive false coincident detections is 200,000 years. This **does not mean** that, we need to wait for 200,000 years, following the observation period, to observe a false coincident detection. We could get the first false coincident detection a few months after the end of the observation period, and the second detection a few months later, as well. If we observe the detector noise over a period far larger than 200,000 years, we would expect coincident detections due to noise, with an **average interval** of 200,000 years.

This is the reason, normalized CCF of H1 vs L1 test is **crucial**, which will reject this **coincident** false detection. We can **rule out** this case of false coincident detection by the crucial new test of normalized CCF on H1 vs L1, explained in Section 5.

3 LIGO Matched Filter Implementation

The **core engine** of the LIGO software for identification of GW signals is the matched filter as described in Eq.1-4 in [4], which is used in two independent search methods, PYCBC and GSTLAL analysis[2]. For the case of PYCBC analysis, matched filter SNR(MF-SNR) $\rho^2(t)$ is given as follows.

$$\begin{aligned} \rho^2(t) &= \frac{1}{|\langle h|h\rangle|}|\langle s|h\rangle(t)|^2 \\ \langle s|h\rangle(t) &= 4\int_0^\infty \frac{\widehat{s}(f)\widehat{h}^*(f)}{S_n(f)} e^{i2\pi ft} df \end{aligned} \tag{3.1}$$

where $s(t)$ and $h(t)$ are the strain signal and the template respectively and $\widehat{s}(f)$ and $\widehat{h}^*(f)$ are the Fourier Transforms of $s(t)$ and $h^*(-t)$ respectively and $S_n(f)$ is the power spectral density of the

[2] GSTLAL analysis also uses matched filter search, and as per Page 7 in [4], "the data s(t) and templates h(t) are each whitened in the frequency domain by dividing them by an estimate of the power spectral density of the detector noise." and also "By the convolution theorem, $\rho(t)$ obtained in this manner is the same as the $\rho(t)$ obtained by frequency domain filtering in Eq. (1)."[in PYCBC analysis]. These equations are implemented in lines 662-740 in the matched filter section of LIGO's tutorial python script. [11].

detector noise. In the time domain, this is equivalent to the convolution of whitened version of $s(t)$ with real $h(-t)$, which is equivalent to the Cross-Correlation Function(CCF) of the whitened strain signal $s(t)$ and the template $h(t)$, with a normalization scale factor as follows.

$$\langle h|h\rangle = 4\int_0^\infty \frac{\widehat{h}(f)\widehat{h}^*(f)}{S_n(f)}df \tag{3.2}$$

It is noted that the above CCF is **not** normalized to give MF-SNR =1 for an ideal template correlated with itself. The expression for re-weighted SNR is given in Eq.6. in [4].

3.1 LIGO's Matched Filter Misfires for Sine Wave

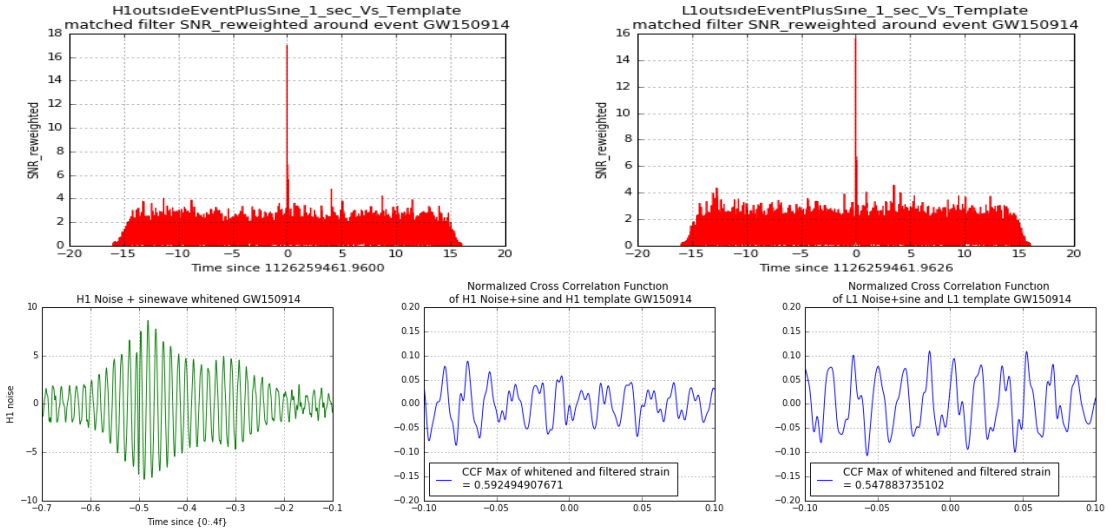

Figure 5. Upper Panels: LIGO matched filter Reweighted SNR for Sinewave + H1/L1 detector noise vs template. Lower left panel: whitened and filtered signal in the time domain. Lower right panels: Normalized CCF of whitened and filtered (Sinewave + H1/L1 detector noise) vs template.

In this section, we will examine whether LIGO's matched filter fires with high SNR, if it is fed with a signal other than the template. Let us test it by replacing the GW150914 signal with the signal $w_d(t) + a(t)$, where $w_d(t)$ is the unwhitened H1/L1 detector noise, obtained from a 32 second block of data which does not contain the GW signal, and $a(t)$ is a sine wave of duration 1 second and frequency 64 Hz at H1 site and 59 Hz at L1 site, scaled so that the standard deviation of $a(t)$ is at least 10 times lower than standard deviation of $w_d(t)$ and has an exponential decay. Then we correlate $w_d(t) + a(t)$ with the template $h(t)$ in the LIGO matched filter.

Fig. 5 shows the results, and we can see from upper panels that, even a sine wave added to the detector noise can produce high reweighted SNR in LIGO's matched filter, similar to high SNR produced by GW150194. This result is not unique to any specific sine wave frequency, but similar results are observed for other frequencies as well and also for GW151226 and GW170104. We observe high SNR and re-weighted SNR at the matched filter, coincident at both sites(64 Hz at H1 and 59 Hz at L1), in mere 32 seconds of observation. This means, we cannot be sure whether GW signal or a sine wave produced high SNR in LIGO's matched filter.

On the other hand, normalized Cross Correlation Function(CCF), described in the next section, of whitened and filtered $w_d(t) + a(t)$ with the template $h(t)$, using short windows for the duration

of the GW signal, does not produce peaky CCF(explained in Section 4) as in the lower right panels in Fig. 5. Hence it is suggested that normalized CCF using short windows be used. Besides, long 32 second windows used in LIGO's matched filter, include noise outside the duration of the GW signal, which can only reduce the sensitivity, as shown in the subsection below.

3.2 Long windows vs short windows for CCF

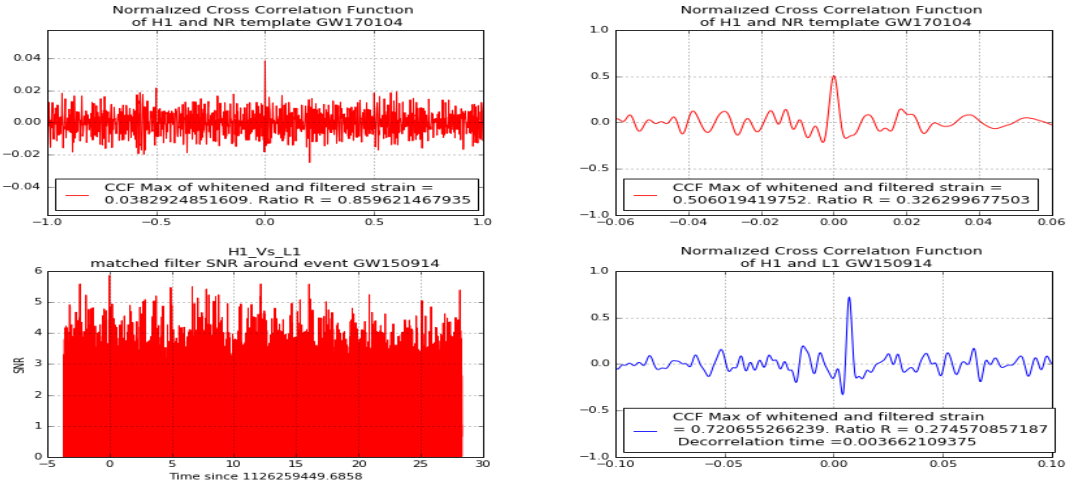

Figure 6. Upper Panels: Normalized CCF for GW170104. Left Panel: done over long 20 second window. Right panel: done over GW event duration of 0.12 seconds. Lower Panels: GW150914: Left Panel: LIGO matched filter SNR H1 vs L1, done over 32 second windows. Right panel:Normalized CCF done over GW event duration of 0.2 seconds. If maximum value of CCF is negative, the plot is inverted. Ratio $R = R_3$.

GW150914, GW151226 and GW170104 have been observed over a window of duration 0.2, 1 and 0.12 seconds respectively[1–3] and more than 95 percent of the template energy is in this window and hence there is **no point** in using windows longer than the duration of these signals.

The upper left panel in Fig. 6 shows normalized CCF done for GW170104, correlating H1 with the template, over long 20 second windows. We can see that the CCF is less peaky than the right panel and the decorrelation ratio(explained in Section 4) is very high ($R > \frac{1}{e}$). On the other hand, the upper right panel in Fig. 6 correlates H1 with the template, over short 0.12 second window, over which the GW event was observed and CCF is sufficiently peaky and the decorrelation ratio $R < \frac{1}{e}$.

Lower panels in Fig. 6 show similar results for GW150914. Lower left panel shows LIGO matched filter SNR, correlating H1 and L1, using long 32 second windows and do not show peaky CCF, while lower right panel performs normalized CCF over short 0.2 second windows, correlating H1 vs L1 and shows peaky CCF. Hence it is suggested that normalized CCF using short windows be used.

4 Normalized CCF Method

Let us consider a normalized Cross Correlation Function(CCF) $CCF(t) = \int_{-T}^{T} s(\tau)h(t-\tau)d\tau$, where both $s(t)$ and $h(t)$ are normalized over the time window [-T,T] during which GW signal was observed, such that CCF of each signal with itself gives a result of unity for zero lag. GW150914

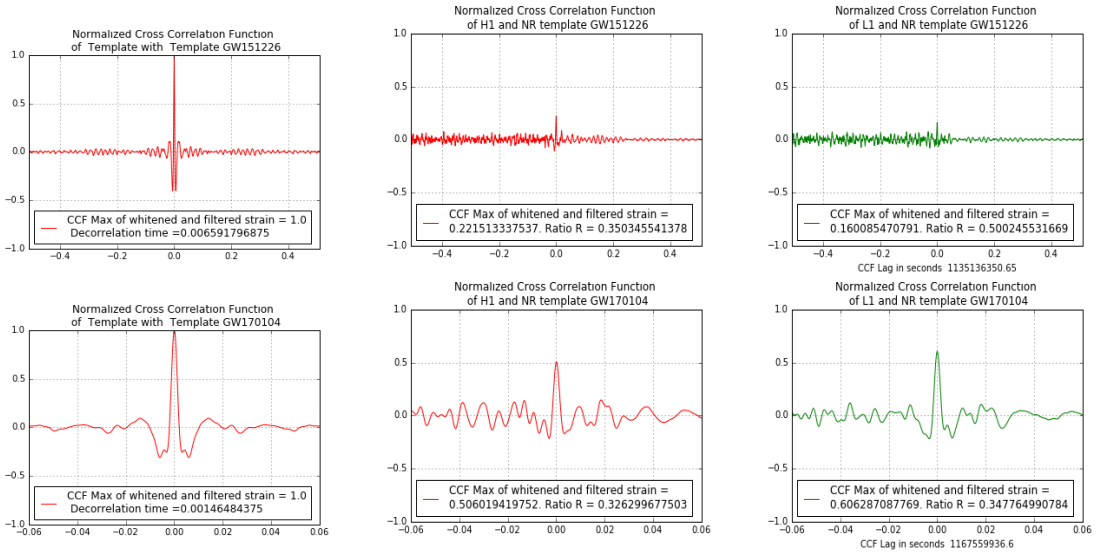

Figure 7. Normalized CCF for GW151226 and GW170104 done over GW event duration of 1 and 0.12 seconds respectively. If maximum value of CCF is negative, the plot is inverted. Ratio $R = R_3$. Left Panels: Reference CCF of Template vs Template. Right Panels: H1 vs template and L1 vs template.

was observed in the window of duration 0.2 seconds, GW151226 in a duration of 1 second and GW170104 in a duration of 0.12 seconds[1–3].[3] We will use reference systems as follows.

Normalized CCF of H1/L1 with template is the same as "matched filtering of H1/L1 with template" using short windows in time domain and is an equally sensitive test. It was shown in Section 3 that matched filter in fact does cross-correlation. The tests listed in this section are **new tests** with short windows.

In this manuscript, the term "Normalized CCF" is used to describe the cross-correlation of **any two signals**, such as H1/L1 with templates **or** H1 with L1, in time domain, using short windows over which GW event was observed, and the signals are normalized such that CCF of a signal with itself gives a value of unity at zero lag. The term "LIGO matched filter" is used to describe the matched filter implemented in LIGO python script[12], which does matched filtering of H1/L1 and the template, using 32 second windows, implemented in frequency domain.

Reference System A: An ideal template for GW151226 or GW170104 $h(t)$ is correlated with itself, after normalization, to give a peaky CCF, the maximum value of CCF = 1 at zero lag, as in left panels in Fig. 7.

By peaky CCF, we mean that the ratio, R, of the the absolute value of CCF for any lag greater than the decorrelation time of the template, to the absolute maximum value of CCF, should be less than a certain threshold. Decorrelation time of the template τ_0 is defined as the time taken for the autocorrelation of the template to fall to $\frac{1}{e} = 0.36$ of the maximum value at zero lag [9] and $\tau_0 = 0.001464$ seconds for the template of GW170104 and $\tau_0 = 0.0066$ seconds for the template of GW151226. We will use the ratio R_3, which is the ratio of the absolute value of CCF at any lag greater than $\tau_0 * 3$ to the absolute maximum value of CCF, and test whether $R_3 < \frac{1}{e}$. Lag greater than $\tau_0 * 3$ is taken to allow for some cushion.

Fig. 7 plots the normalized CCF of GW151226(upper right panels) and GW170104 (lower right

[3]Normalized CCF using running windows method using, say 32 seconds of H1 and 0.2 seconds of L1, gives similar results as this method using short windows for both H1 and L1. See the file LOSC Event tutorial Normalized CCF v1.py

GW signal	R_3 H1 vs template	R_3 L1 vs template	R_3 H1 vs L1
GW150914	0.2396	0.2486	0.2745
GW151226	0.35	0.50	0.88
GW170104	0.3262	0.3477	0.5746

Table 1. Normalized CCF Ratio for GW150914, GW151226 and GW170104. R_3 = Ratio of the absolute value of CCF at any lag greater than $\tau_0 * 3$ to the absolute maximum value of CCF. τ_0 is the decorrelation time of the template. Lag greater than $\tau_0 * 3$ is taken to allow for some cushion.

panels) by correlating H1 with the template and L1 with the template. We can see from the upper right panel in Fig. 7 that GW151226 shows very poor CCF peaks ($R_3 > \frac{1}{e}$) when correlating L1 with the template[4].

Table 1 summarizes the ratio, R_3, of the absolute value of CCF at any lag greater than $\tau_0 * 3$ to the absolute maximum value of CCF, for the three GW signals, where τ_0 is the decorrelation time of the template. GW150914 has good CCF Ratio as in the top panels. GW151226 has poor CCF ratio $R_3 > \frac{1}{e}$ when correlating L1 with the template.(Reason 2)

4.1 Matched Filtering of H1/L1 vs template(using normalized CCF with short windows), indistinguishable from CCF of detector noise vs template

The lower left panel in figure shows Cross Correlation Function (CCF) done over 1 second duration, when correlating L1 with the template of GW151226 and the lower right panel shows CCF peaks obtained by correlating L1 detector noise with the template, using running windows, over 4096 second block of data, excluding the GW event portions and edges. We can see that GW151226 shows CCF peaks **comparable** with CCF peaks corresponding to detector noise correlations, and **do not** rise well above detector noise correlations. If we observe the detector noise for a longer period of 4096*N seconds, we may observe high CCF peaks for H1 detector noise vs template as well.

So, we concatenate 4096 second data for GW151226, GW170104, GW150914 and LVT151012, and exclude the GW event portions and edges and form 4096*4 second block of data. Given the public availability of only 4 blocks of 4096 second data, they are concatenated together and considered as adjacent 4 blocks, for this analysis. Fig. 8 shows CCF done over 1 second duration, when correlating H1/L1 with the template of GW151226. We can compare CCF peaks obtained by correlating H1/L1 with the template(left panel), with CCF peaks obtained by correlating H1/L1 detector noise with the template(right panel), using running windows.

We can see that GW151226 shows CCF peaks(left panel) **comparable** with CCF peaks corresponding to detector noise correlations (right panel), and **do not** rise well above detector noise correlations. This means that we **cannot be sure** if the CCF peaks in GW151226 were caused by GW signals or detector noise correlations. This behaviour is observed just in 4096*4 seconds of H1 and L1 detector noise.

Figure shows Cross Correlation Function (CCF) done over 0.12 seconds duration, when correlating H1/L1 with the template of GW170104. We can compare CCF peaks obtained by correlating H1/L1 with the template(left panel), with CCF peaks obtained by correlating H1/L1 detector noise with the template(right panel), using running windows, over 4096*4 second block of data, excluding the GW event portions and edges.

[4]In Fig. 7, we use whitened and filtered versions of H1, L1 strains and the templates, for the 2 GW signals. In LIGO tutorial file LOSC Event Tutorial.py, the variables "strain H1 whitenbp" and "strain L1 whitenbp", "template match" are used.

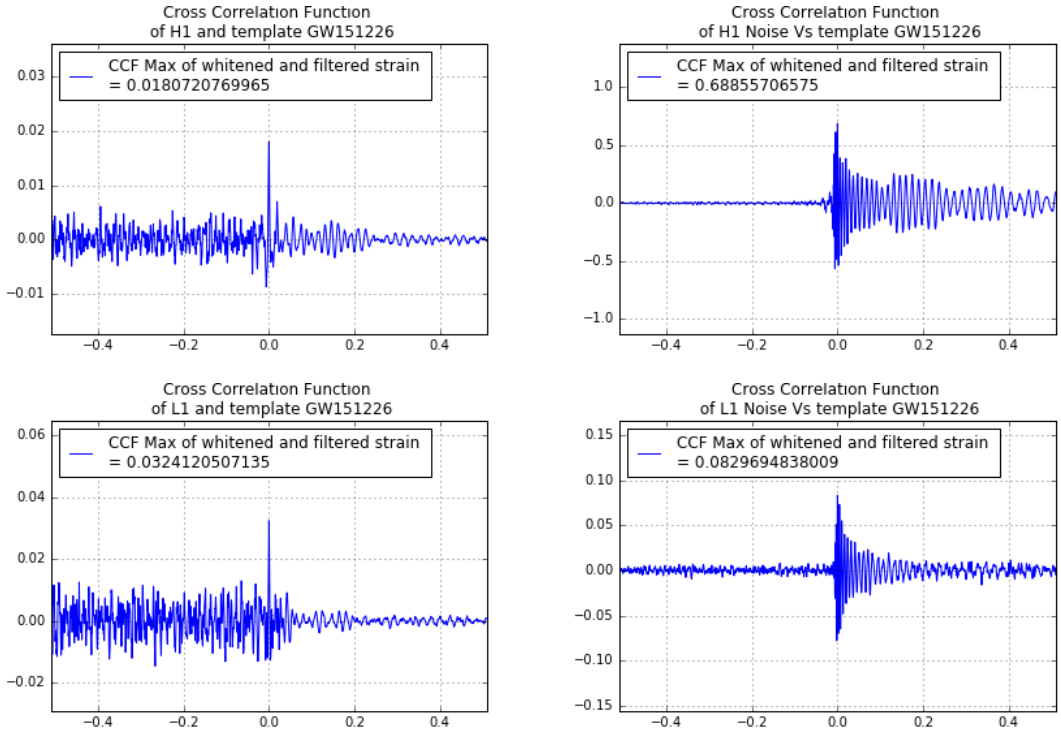

Figure 8. GW151226: CCF Plots done over GW event duration of 1 second, using running windows for the right panel. Upper Left: H1 vs template. Upper Right: H1 Noise vs template. Lower Left: L1 vs template. Lower Right: L1 Noise vs template. H1/L1 detector noise obtained from a 4096*4 second block of data and excluding GW event portions and edges.

We can see that GW170104 shows CCF peaks (left panel) **comparable** with CCF peaks corresponding to detector noise correlations (right panel). This means that we **cannot be sure** if the CCF peaks in GW170104 were caused by GW signals or detector noise correlations. This behaviour is observed just in 4096*4 seconds of H1 and L1 detector noise.

For coincident observation at H1 and L1, we can show that detector noise can produce comparable CCF peaks, when correlated with 250,000 templates, approximately once in every 51 days[$\frac{(4096*4*4096)^2}{(4096*3600*24*250000)} = 51$], thus simulating 1 out of 250,000 templates. Hence **GW170104** and **GW151226** should be questioned, based on this test alone[Reason 3].

We can reproduce these results for **unwhitened longer templates** of GW151226 and GW170104 as well. Fig. 9 shows CCF done over 14 second duration, when correlating H1/L1 with the template of GW170104. Figure shows CCF done over 14 seconds duration, when correlating H1/L1 with the template of GW151226. We can see that GW151226 and GW170104 show CCF peaks (left panel) **comparable** with CCF peaks corresponding to detector noise correlations (right panel).

LIGO team has modified python scripts in August 2017 and they use modified whitened and filtered templates. We can reproduce above results for **modified templates** of GW151226 and GW170104 as well. Figure shows CCF done over 0.12 second duration, when correlating H1/L1 with the template of GW170104. Figure shows CCF done over 1 second duration, when correlating H1/L1 with the template of GW151226. We can see that GW151226 and GW170104 show CCF peaks (left panel) **comparable** with CCF peaks corresponding to detector noise correlations (right panel).

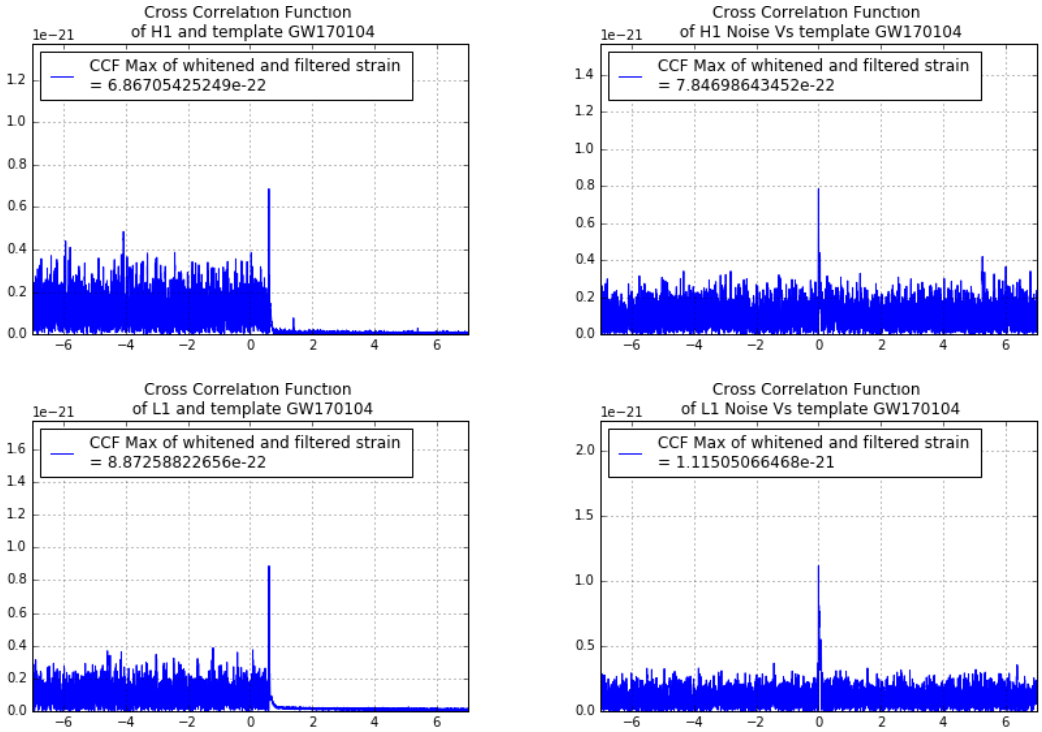

Figure 9. GW170104: CCF Plots done over unwhitened template duration of 14 seconds, using running windows for the right panel. Upper Left: H1 vs template. Upper Right: H1 Noise vs template. Lower Left: L1 vs template. Lower Right: L1 Noise vs template. H1/L1 detector noise obtained from a 4096*4 second block of data and excluding GW event portions and edges.

4.2 False coincident detection with detector noise and fake template

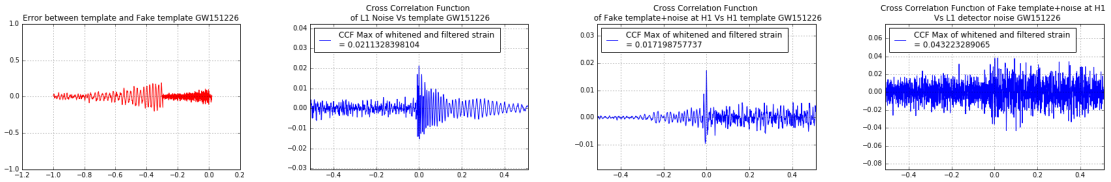

Figure 10. Left Panel:Error between Template vs Fake Template. CCF Plots for GW151226 done over GW event duration. Second: L1 Noise vs template. Third:(Fake Template+H1 noise) vs template .Right Panel: L1 noise vs (fake template+H1 Noise). H1/L1 detector noise obtained from a 4096 second block of data and excluding GW event portions and edges.

Given that detector noise can produce comparable correlations with the template, as the observed GW151226 signal, let us consider the case where a) detector noise at L1 produces CCF peaks, when correlated with the template, in 4096 second block of data, as in the second panel of Fig. 10, and b) a noisy signal roughly resembling the template but generated by frequency modulation(FM) and amplitude modulation(AM) [5] in the last 0.3 seconds and zero in the first 0.7 seconds, is observed

[5]The template for GW151226 is represented by a FM+AM signal $h(t) = A(t) * cos(2\pi f_0 t + m(t))$ where $f_0 = 56$ Hz. Fake template is given by $h_f(t) = A_f(t) * cos(2\pi f_0 t + m_f(t))$ where $A_f(t) = A(t) + w_a(t), m_f(t) = m(t) + w_m(t)$ and $w_a(t), w_m(t)$ are white gaussian noise.

at H1, which also produces CCF peaks, when correlated with the template, as in the third panel of Fig. 10. This means that we **cannot be sure** if the CCF peaks in GW151226 were caused by GW signals or detector noise correlations and fake templates. Similar results hold for the case of GW170104. Hence **GW170104** and **GW151226** should be questioned, based on this test alone[Reason 3b].

4.3 Need for H1 vs L1 cross-correlation test: Ruling Out False coincident detection

We do not want to accept above false coincident detection in section 4.2 as a valid GW signal, because L1 detector noise simulated the template at L1 detector and fake template was observed at H1 detector. We can rule out this case of false coincident detection by correlating H1 with L1, which gives poor CCF peaks, as in the right panel in Fig. 10.

H1 vs L1 cross-correlation test is also necessary to rule out false coincident detections in sections 4.1 and 2.1.

5 Normalized CCF by Correlating H1 with L1

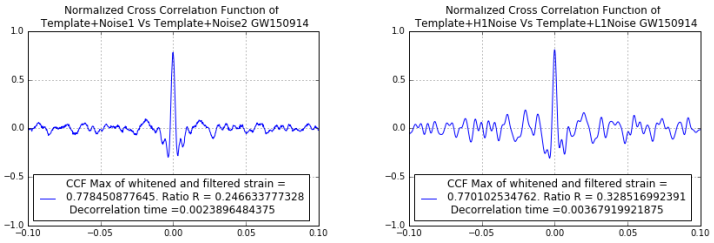

Figure 11. Reference Systems with Normalized CCF over 0.2 second windows for GW150914. Left: Reference System 1. Right: Reference System 2. If maximum value of CCF is negative, the plot is inverted.

In the previous section 4.3, it was shown that it is important to cross-correlate H1 and L1, to rule out false coincident detections. Given that the same GW signal is expected to be received in both sites, these signals should give a high CCF when correlated with each other. This is a **crucial** test which must be performed.

We **can** cross-correlate two noisy signals and expect a peaky CCF, if the signals are correlated. In fact, wireless communication with sensors routinely use cross-correlation of two noisy signals [8]. We wish to correlate two noisy signals $s(t) = h_H(t) + w_H(t)$[H1] and $h^{'}(t) = h_L(t) + w_L(t)$[L1] where $w_H(t)$ and $w_L(t)$ represent the detector noise, and $h_H(t)$ and $h_L(t)$ represent the templates of H1 and L1. This is equivalent to correlating $s(t) = h^{'}(t) + (w_H(t) - w_L(t)) + (h_H(t) - h_L(t))$ and $h^{'}(t)$ and hence correlating $s(t) = h^{'}(t) + w(t)$ and $h^{'}(t)$, where $w(t) = (w_H(t) - w_L(t)) + (h_H(t) - h_L(t))$. $h^{'}(t)$ is the noisy template. The theory of matched filter or cross-correlation imposes no constraints on the characteristics of the template.

The same template $h(t)$ is used in the right panels CCF plots of Fig. 11, because the templates received at H1 and L1 detectors are nearly identical, which differ only by a time shift and a scale factor, and this does not matter for the ratio R, in the the normalized CCF plots. We observe that $R < \frac{1}{e}$. Given that GW signals are assumed to be the sum of an ideal template and detector noise, this comparison is reasonable.

Reference System 1: Fig. 11 shows this system in the second panel, which correlates two noisy signals, $s(t)$ and $h^{'}(t)$, where $s(t)$ is the sum of the template for GW150914 and AWGN, and $h^{'}(t)$ is the sum of the same template and an independent AWGN. Then it computes the normalized

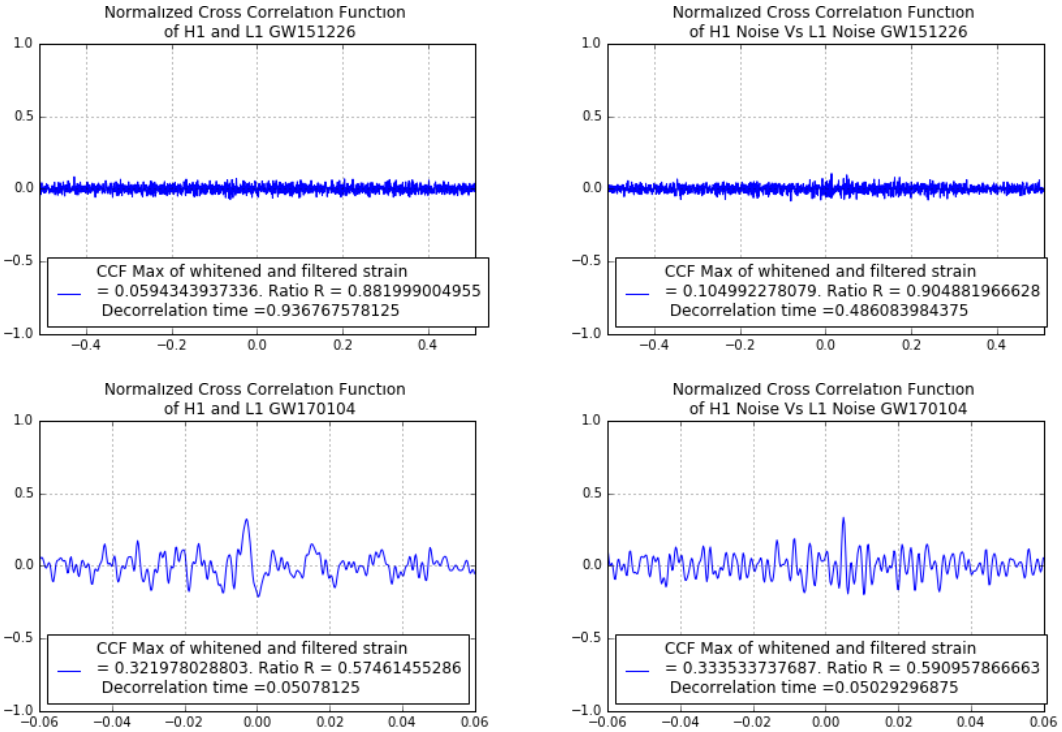

Figure 12. Normalized CCF Plots done over GW event duration. Top Row: GW151226. Bottom Row: GW170104. Left panels: H1 vs L1. Right panels: H1 noise vs L1 noise. H1 and L1 detector noise obtained 1 second after the end of GW151226 and 10 seconds after the end of GW170104.

$CCF(t) = \int_{-T}^{T} s(\tau)h^{'}(t-\tau)d\tau$. We can see that the CCF shows strong peaky behaviour. Average decorrelation time for this system is $\tau_1 = 0.0024$ seconds.

Reference System 2: Fig. 11 shows a reference system 2 in the third panel, which correlates two noisy signals, $s(t)$ and $h^{'}(t)$, where $s(t)$ is the sum of the template for GW150914 and H1 detector noise outside the GW event, and $h^{'}(t)$ is the sum of the same template and and L1 detector noise outside the GW event, and computes the normalized CCF. We can see that the CCF shows strong peaky behaviour. Average decorrelation time for this system is $\tau_2 = 0.0037$ seconds.

Reference systems 1 and 2 are shown only for the purpose of demonstrating the fact that we **can** cross-correlate two noisy signals and expect a peaky CCF. Given the need for high standards required in classifying GW signals, we will use only the decorrelation time τ_0 of the reference system A, which correlates the template of each GW signal with itself, when we compare the decorrelation times and the ratio R, of the three GW signals.

We observed in the panels of Fig. 11 that the two reference systems 1 and 2, have normalized CCF with the ratio $R_3 < \frac{1}{e} = 0.36$. It is reasonable to expect a similar ratio $R_3 < \frac{1}{e}$ from GW151226 and GW170104, when correlating H1 with L1.

Fig. 12 plots the normalized CCF for GW151226(top row) and GW170104 (bottom row) by correlating H1 with L1 in the left column and correlating H1 detector noise with L1 detector noise in the right column. We can see that GW151226 and GW170104 show very poor CCF peaks ($R_3 > \frac{1}{e}$) when correlating H1 with L1, and we can see that CCF peaks are **indistinguishable** from CCF peaks corresponding to detector noise in the right column. Hence, we cannot be sure whether detector noise or GW signal caused the CCF peaks.

We can reproduce these results for **longer H1/L1 strains** of GW151226 and GW170104 as well. Figure shows CCF done over 14 second duration, when correlating H1 with L1 for GW151226(top panel) and GW170104(bottom panel). We can see that GW151226 and GW170104 show CCF peaks (left panel) **indistinguishable** from CCF peaks corresponding to detector noise correlations (right panel).

Hence **GW170104** and **GW151226** should be questioned, based on this test alone[Reason 4].

6 Case for Questioning of GW151226 and GW170104

GW151226 and GW170104 belong to a class of signals with Signal Power Ratio(SPR) around unity. We need to have much higher standards for classifying such signals as GW signals, than signals like GW150914 which have $SPR > 4$.

Given the high standards required for declaration of GW signal, it is suggested that GW151226 and GW170104 be questioned, due to each of the following Reason 1, 3 and Reason 4 **independently**.

Reason 1: Being weak signals, non-stationary and non-Gaussian detector noise at the two sites could have caused the weak GW signal. When high SNR is observed at the matched filter during the GW event, while signal amplitude does not rise significantly, as is the case for GW151226 and GW170104, it may be due to non-stationary detector noise having high SNR at the output of the matched filter briefly at both sites as well.

Reason 3: In Section 4.1, we observed in Fig. 8 and Fig. 9 that the normalized CCF is very poor ($R_3 > \frac{1}{e}$) when correlating H1 or L1 with the template, and is indistinguishable from detector noise correlations, for GW151226 and GW170104.

Reason 4: In Section 5, we observed in Fig. 12 that the normalized CCF is very poor ($R_3 > \frac{1}{e}$) when correlating H1 and L1, and is indistinguishable from detector noise correlations, for GW151226 and GW170104.

7 Implications of Results for GW150914

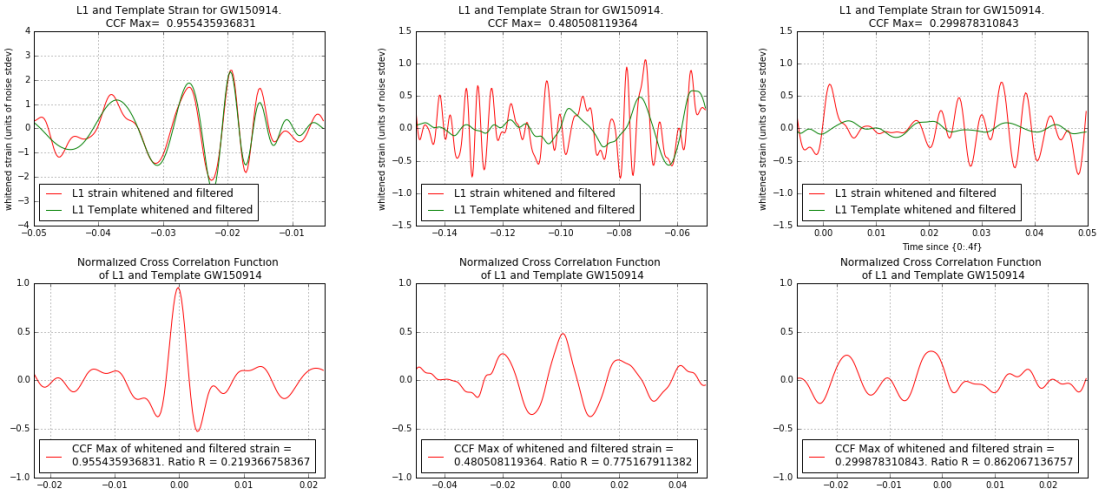

Figure 13. GW150914 whitened and filtered strains: Top Row: Plots of L1 vs Template. Bottom Row: Normalized CCF of L1 vs Template. 3 time windows: [-0.05,-0.005], [-0.15,-0.05] and [-0.005,0.05] seconds.

Fig. 13 shows the visual correlations of whitened and filtered L1 and template strains over the 3 windows and we can see that the correlations are satisfyingly strong only in a short 0.05 second window in the first plot. Correlations in the second and third windows in the rest of $0.2-0.05=0.15$ seconds are not convincing.

This agrees with the maximum values of the normalized CCF, given by CCF_{max} computed for these 3 windows. We can see that $CCF_{max} > 0.9$ for the short 0.05 second window in the first plot, but $CCF_{max} < 0.5$ for the other 2 windows and the Ratio $R_3 > \frac{1}{e}$.[Reason 5]. This is attributed to the high levels of detector noise which is comparable to the template power in these 2 windows. So, we **cannot be sure** if template is present in these 2 windows for GW150914[Reason 5]. This problem can be resolved only by reducing detector noise in future.

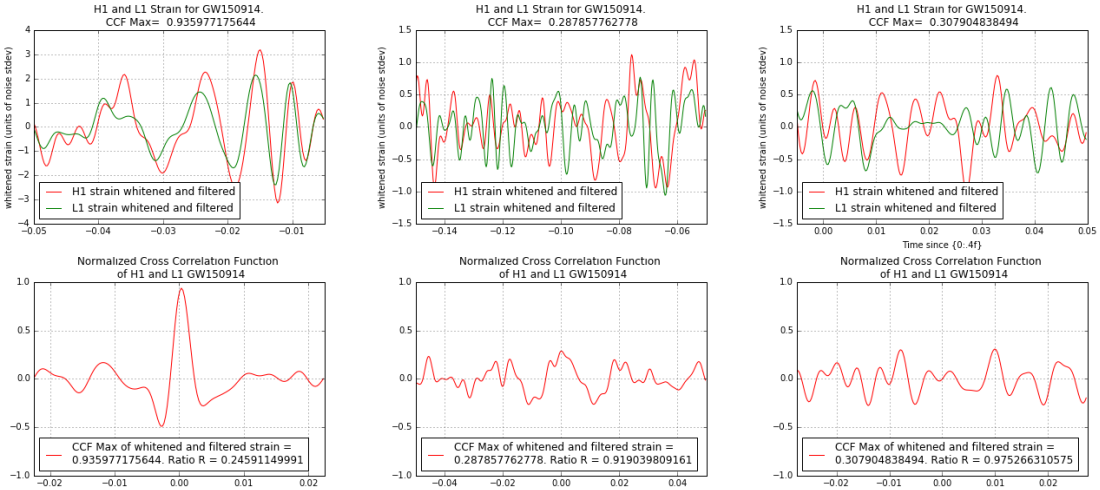

Figure 14. GW150914 whitened and filtered strains: Top Row: Plots of H1 vs L1. Bottom Row: Normalized CCF of H1 vs L1. 3 time windows: [-0.05,-0.005], [-0.15,-0.05] and [-0.005,0.05] seconds.

Fig. 14 shows the visual correlations of whitened and filtered H1 and L1 strains over the 3 windows and we can see that the correlations are satisfyingly strong only in a short 0.05 second window in the first plot. Correlations in the second and third windows in the rest of $0.2-0.05=0.15$ seconds are not convincing.

This agrees with the maximum values of the normalized CCF, given by CCF_{max} computed for these 3 windows. We can see that $CCF_{max} > 0.9$ for the short 0.05 second window in the first plot, but $CCF_{max} < 0.31$ for the other 2 windows and the Ratio $R_3 > \frac{1}{e}$.[Reason 5]. This is attributed to the high levels of detector noise which is comparable to the template power in these 2 windows. So, we **cannot be sure** if template is present in these 2 windows for GW150914[Reason 5]. This problem can be resolved only by reducing detector noise in future.

7.1 Effect of removal of 20-300Hz Bandpass filter

Fourth order butterworth bandpass filters are used in the frequency range 20-300 Hz in plotting whitened and filtered H1 and L1 strains for GW150914 in Fig. 14. The template for GW150914 has 86 percent of its energy in this 20-300 Hz range and **14 percent** of energy in the remainder of the total frequency range of 0-2048 Hz, with sampling frequency of 4096 Hz. The upper left panel in Fig. 15 compares the templates inside 20-300 Hz band and the total range of 0-2048 Hz.

The upper right panel in Fig. 15 compares the whitened H1 strain without filtering and compares it with the unfiltered template. The lower left panel in Fig. 15 compares the whitened H1 and L1

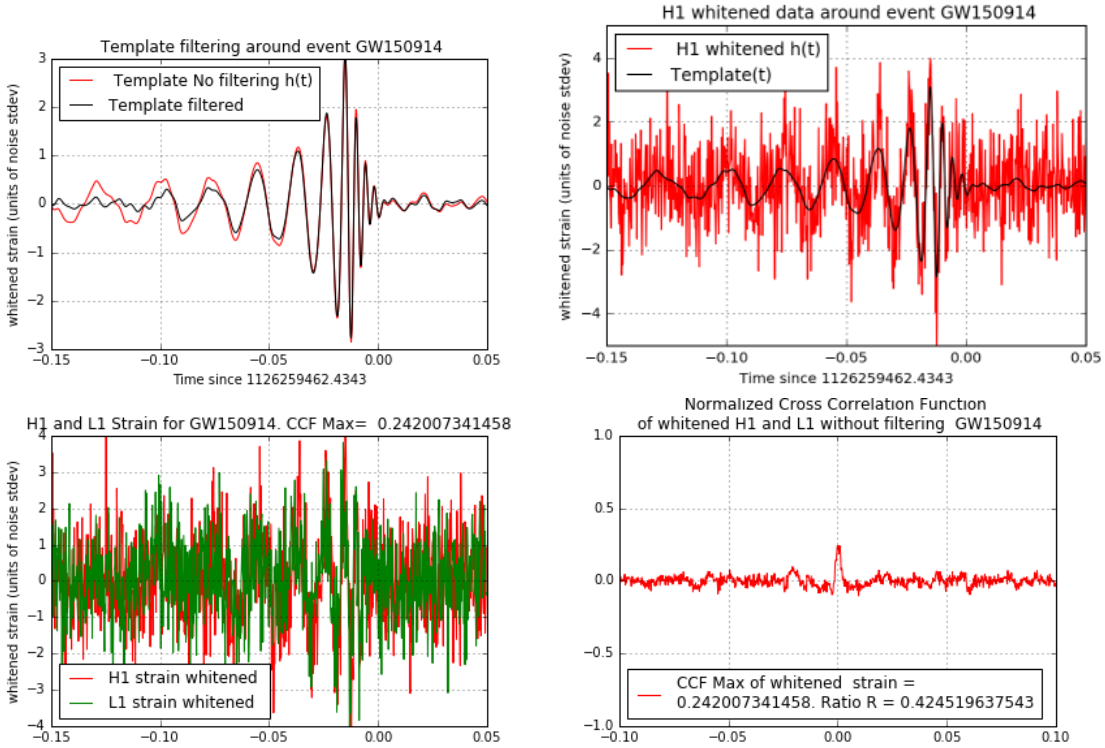

Figure 15. Top Left: Plots of templates with and without filtering. Top Right: Plots of whitened H1 vs template without filtering. Lower Left:Plots of H1 vs L1 whitened strain in GW150914 without filtering. Lower Right: Normalized CCF of H1 vs L1 whitened strain in GW150914 without filtering

strains without filtering. The lower right panel in Fig. 15 compares the Normalized CCF of whitened H1 and L1 strains without filtering. We can see that visual correlations and the normalized CCF values become much poorer if we do not filter the strains with 20-300 Hz filter, so that we can capture the **14 percent** template energy outside 20-300 Hz. We see that the maximum value of CCF falls from 0.72 in the upper right panel in Fig. 7, to 0.24 in the lower right panel in Fig. 15 and also that the Normalized CCF Ratio $R_3 > 1/e$[Reason 6]. This problem can be resolved only by reducing detector noise in future.

7.2 Summary of results

Given the high standards required for declaration of GW signal, it is suggested that GW150914 be studied further, due to each of the following Reason 5 and Reason 6 **independently**.

Reason 5: We observed in Fig. 13 and Fig. 14 that the normalized CCF is high only for a short 0.05 second window when correlating L1 with the template, and also correlating H1 and L1, for GW150914. It is poor for the other 2 windows ($R_3 > 1/e$).

Reason 6: We observed in Fig. 15 that the normalized CCF is poor and the Normalized CCF Ratio $R_3 > 1/e$, when correlating H1 and L1, for GW150914, without using 20-300 Hz filter, so that we can capture the 14 percent template energy outside 20-300 Hz.

8 Concluding remarks

Section 6 gives the reasons why weak signals GW151226 and GW170104 should be questioned. Section 7 gives the reasons why GW150914 should be studied further. Reiterating the point made earlier, because we are more often likely to observe weak signals which look like noise and whose amplitude does not rise during assumed GW event, it is of paramount importance that we should not classify noise as GW events. We need **high** standards to classify an observed time series as a GW signal. LIGO detectors have very strong impulsive interference in 60Hz and harmonics and also in 300Hz-2000Hz range. This may affect the performance of the matched filter. It is very important to clean up the impulsive interference in the detectors.

Acknowledgments

I am grateful to Andrew D. Jackson of Niels Bohr International Academy for his encouragement, suggestions, discussions and review of this manuscript. I would like to thank Sebastian Domenico von Hausegger and Arunava Chaudhuri for review of my Python scripts and helpful suggestions. I would like to thank M.A. Srinivas for discussions and helpful suggestions. I would like to thank LIGO Open Science Center for making the data and Python scripts available online. I would like to thank LIGO scientists who answered my many questions in detail.

References

[1] Abbott, B. P., Abbott, R., Abbott, T. D., et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, Physical Review Letters, 116, 061102 (2016) Online version of paper.

[2] Abbott, B. P., Abbott, R., Abbott, T. D., et al., *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*, Physical Review Letters, 116, 241103(2016) Online version of paper.

[3] Abbott, B. P., Abbott, R., Abbott, T. D., et al.*GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2*, Physical Review Letters, 118, 221101 (2017) Online version of paper.

[4] Abbott, B. P., Abbott, R., Abbott, T. D., et al., *GW150914: First Results from the Search for Binary Black Hole Coalescence with Advanced LIGO*, Phys. Rev. D 93, 122003 (2016) [see pp.4-9] Arxiv copy of paper. see pp.8-13.

[5] Creswell, J., Hausegger,S., Jackson, A. D., Liu,H., Naselsky, P., *On the time lags of the LIGO signals*, 2017, arxiv:1706.04191 Arxiv copy of paper.

[6] Naselsky, P., Jackson, A. D., & Liu, H., *Understanding the LIGO GW150914 event*, Journal of Cosmology and Astroparticle Physics, 8, 029 (2016)

[7] Liu, H., & Jackson, A. D., *Possible associated signal with GW150914 in the LIGO data*, Journal of Cosmology and Astroparticle Physics, 10, 014 (2016)

[8] Garnier, J., Papanicolaou,G., *Passive Sensor Imaging Using Cross Correlations of Noisy Signals in a Scattering Medium*, SIAM J. Imaging Sci., 2(2), 396437.(2009). Online version of paper

[9] Kenneth Y. Jo, Satellite Communications Network Design and Analysis, Page 415.

[10] LOSC LIGO GW150914 Tutorial

[11] LOSC LIGO GW151226 Tutorial matched filter equations are implemented in the section ”matched filtering to find the signal”

[12] LIGO Tutorials

[13] Zip Files of Modified LIGO Tutorials which demonstrate the results in this paper.
Note: The tutorials have changed since August 2017. The tutorial for all 3 GW signals, published in June 2017 is archived here(LOSC Event tutorial.zip) and is used and modified to produce scripts below.
.
Fig.1, Fig.2 , Fig.3: LOSC Event tutorial zoom.py
Fig.4: LOSC Event tutorial MF reweighted SNR noise test v1.py
Fig.5: LOSC Event tutorial MF reweighted SNR sinewave test.py.
Fig.6, Fig 10: LOSC Event tutorial Normalized CCF.py
Fig.7, Fig.8: LOSC Event tutorial Normalized CCF H1 template 4 blocks.py
Fig.9: LOSC Event tutorial Normalized CCF H1 template fake.py
Fig.11: LOSC Event tutorial Normalized CCF H1 L1.py
Fig.12: LOSC Event tutorial Normalized CCF GW150914 v2.py
Fig.13, Fig 14: LOSC Event tutorial Normalized CCF.py
.
Above results can be demonstrated for the August 2017 version of LIGO's python scripts and also longer unwhitened templates. Please see the readme file in Zip Files.
.
Matched Filter SNR results are demonstrated with 32 second block of data, used in the original LIGO tutorial scripts. Results hold for 512 second block of data also.