LIGO-P080042-00-Z

Detection Confidence Tests for Burst and Inspiral Candidate Events

Romain Gouaty for the LSC

Louisiana State University, Baton Rouge, LA 70803, USA

E-mail: `romain@phys.lsu.edu`

Abstract. In order to detect transient gravitational-wave signals such as "bursts" or "inspirals", the LIGO scientific collaboration is using analysis pipelines which aim to reduce the false alarm rate while keeping optimal the chances of performing a detection. However, because of the non-Gaussian, non-stationary noise exhibited by the LIGO detectors [1], residual false alarms (called background triggers) are found at the end of the pipelines. A critical aspect of the search is then to assess our confidence for gravitational waves and to distinguish them from background triggers. Both the "Compact Binary Coalescence" and the "Burst" working groups have been developing a detection checklist for the validation of candidate-events, consisting of a series of tests including data quality checks, analysis of the candidate appearance, parameter consistency studies, coherent analysis, which aim to corroborate a detection or to eliminate a false alarm. In this paper, the general methodology used for candidate validation is presented. The method is illustrated with an example of simulated gravitational wave signal and a background trigger.

1. Introduction

The Compact Binary Coalescence (CBC) working group of the LIGO-Virgo joint collaboration is a data-analysis group looking for gravitational-wave signals emitted by inspiralling Compact Binary Systems. The duration that such signal spends in the frequency bandwidth of the current ground-based gravitational-wave interferometers lies in a range which goes from a few tenths up to several tens of seconds. The LIGO-Virgo Burst group is a data-analysis group looking for fast transient gravitational-wave signals of typical duration of the order of the milli-second or the tenth of second, without specific assumptions on the expected waveform. Both of the working groups previously mentioned lead analysis that are sensitive to the non-stationary and non-gaussian noise of the detectors. Because of the time-frequency properties of the signals that are being looked for, noise transients can actually induce background events in the analysis pipelines. Although these pipelines have been designed to minimize the rate of false alarm triggers while keeping optimal the probability to detect gravitational-wave signals, some background triggers can be found at the end of the analysis pipelines as accidental coincidences between the interferometers. It is therefore crucial to submit each gravitational-wave candidate found by the analysis to a detection checklist which aims to estimate confidence in this candidate.

The detection checklist is made of a list of standard tests in different stages of development that are used to review the gravitational-wave candidates. A summary of this checklist is provided in section 2. Section 3 summarizes the method implemented by the Burst and CBC groups to estimate the statistical significance of the candidates. In section 4, the paper will describe with more details a few items of the detection checklist, using an example of simulated gravitational-wave signal and background trigger for illustration purposes.

2. Overview of the detection checklist

A detection checklist to review candidate-events has been developed by each of the CBC and Burst groups. Despite some specificities inherent to the kind of signals that are being looked for by the CBC and Burst searches, the method implemented to estimate confidence in a gravitational-wave candidate is very similar between the two data-analysis groups. Thus the summary of the checklist provided in this section applies both to the CBC and Burst groups. The list presented below is a short synthesis of the tests implemented in the CBC and Burst detection checklists. As many of these tests are still under development or refinement, the checklist is rapidly evolving, and the following list should not be considered as exhaustive. Here we outline the main bullets that are currently part of the detection checklist for candidate-events or in the process of implementation:

- **Statistical significance** The first step of the candidate validation procedure consists in determining the statistical significance of the candidates found by the analysis pipeline, that is to say the probability of coincident triggers arising from random coincidences of noise triggers (background triggers). The general method is described in section 3, which will also explain how the candidate's false alarm probability affects the way the other tests of the checklist are addressed.

- **Data integrity:** sanity checks to verify that the data set containing the candidate is not corrupted.
- **Status of the interferometers:** The state of the interferometers and their sensitivity near the time of the candidate are checked. This test also includes a verification of the data quality flags recorded in the database. Section 4.1 will show how this test can allow the identification of noisy data segment containing background triggers.
- **Environmental or instrumental causes:** We analyze the auxiliary channels of the interferometers, such as the environmental sensors or the signals involved in the mirror control loops, to check for the presence of possible noise transients (called "glitches") which could be the cause of an accidental trigger found by the analysis. This effort is also part of the "glitch group" activities [2]. More details on this part of the detection checklist are provided in section 4.2.
- **Candidate's appearance:** Part of the detection checklist consists of qualitative tests of the candidate's appearance. A variety of tools are used to examine the data containing the candidate-event, such as time series, time-frequency spectrograms, or the outputs of the search pipeline, such as the *Signal-To-Noise Ratio* or the χ^2 time series in the case of the CBC search. The appearance of a simulated gravitational-wave signal will be compared to an example of background trigger in section 4.3.
- **Correlation between interferometers:** Other tests of the detection checklist consist in checking for the correlation between the signal measured in the different interferometers of the network. The Burst group has developed many network or coherent analysis, one example of which is *Coherent Waveburst*, based on a constraint likelihood method described in [4]. The CBC group is developping a coherent analysis specific to the search for inspiral waveforms and is currently testing it as part of the detection checklist.
- **Consistency of the candidate's estimated parameters:** Automated tools are being developed in order to establish a likelihood ranking of the candidates given their estimated physical parameters. This relies on the comparison of parameters distribution for simulated gravitational-wave signals and for known accidental coincident triggers.
- **Detection robustness:** It is foreseen to implement in the checklist a test to verify the robustness of the detection taking into account possible errors in the calibration of the detectors.
- **Coincidence with external searches:** The LIGO Scientific Collaboration is implementing procedures to check for possible coincidences between gravitational-

wave candidates and triggers from external searches, such as Gamma-Ray Bursts, optical transients or neutrinos observations.

3. Statistical significance of the candidates

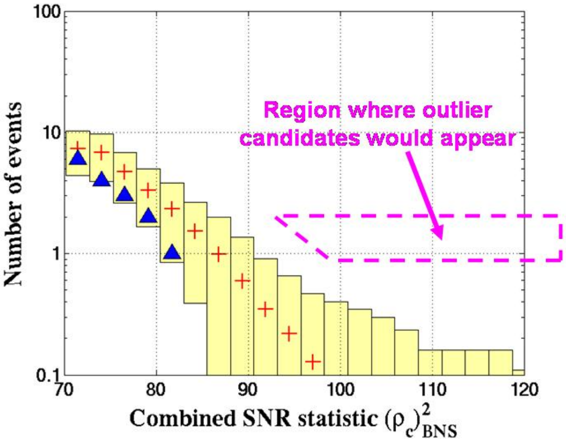

Figure 1. Cumulative histogram of *Signal-To-Noise Ratio* for the S4 Binary Neutron Star search: for the in-time coincident candidate events (triangles), and for the estimated background of accidental coincidences (crosses and one standard-deviation range). This plot has been extracted from [3]. All candidates were found consistent with the background. The dotted box show the region where a statistically significant candidate would be expected.

In order to estimate the false alarm probability of the candidate-events, the CBC and Burst searches compare the in-time coincident triggers to an expected background. The latest is estimated by repeating the analysis after time-shifting the data of each interferometer with respect to each other. This method, called the time-slides analysis, has already been described in previous publications such as [3]. Fig. 1 shows an example of comparison between in-time triggers and expected background for the Binary Neutron Star search run over the data taken during the LIGO S4 run [3]. The loudest candidates of a search are submitted to the detection checklist even if they are not statistically significant. However the goal of the candidate's follow-up with the checklist slightly differs whether the candidate has a low probability of being an accidental coincidence or not. If the candidate is consistent with the estimated background, then the goal of the follow-up is to perform a sanity check to make sure that there is no obvious gravitational-wave signal hidden in the background. Such candidate shall not lead to a detection claim unless strong evidences of a gravitational-wave signal can be found, despite the high probability of accidental coincidence. In case the candidate has a low probability of accidental coincidence, the goal of the follow-up is then to strengthen our confidence in a possible detection by submitting the candidate to the detection checklist that a gravitational-wave signal should pass succesfully. Such an outlier candidate should stand well above the estimated background, lying inside the empty dotted box shown in Fig. 1).

For the S4 Binary Neutron Star search (see Fig. 1), the candidates (triangles) were found consistent with the estimated background (crosses with one standard-deviation range). The candidate's follow-up confirmed that no detection was found.

4. Detailed examples of the checklist

This section will highlight a few examples of tests used for the review of candidate-events. In order to illustrate the expected results for an inspiral gravitational-wave signal we will refer to a simulated inspiral signal. This simulation was performed by acting on the arm mirrors of the interferometers to generate a differential motion of the interferometer arm cavities equivalent to the expected effect of a gravitational wave. This simulated gravitational-wave signal was detected by the CBC analysis and stands as an outlier above the estimated background. We will also illustrate the behaviour of the detection checklist when it is applied to an example of accidental coincident trigger (called a background trigger). In the following subsections, we will refer to the simulated gravitational wave as *Candidate G* and to the background trigger as *Candidate B*.

4.1. Status of the interferometers

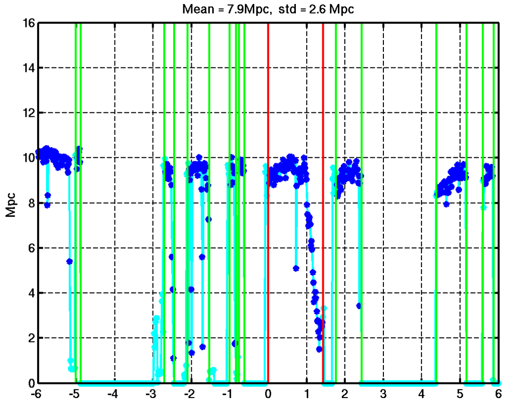

Figure 2. Inspiral range as a function of time at Livingston. The time along the x axis is expressed in units of hour, and the inspiral range along the y axis is expressed in units of Mpc.

A part of the detection checklist consists in verifying the status of the interferometers and the data quality in the segment containing the candidate. This includes examining the figures of merit (detectors state, sensitivity, seismic trends) posted in the detectors log, and scanning the database which contains the list of data quality flags. The goal of this study is to check for a possible unusual behavior of the detectors or an unsual excess of noise which could translate into a higher rate of accidental coincident triggers and thus reduce our condifence in the candidate-event. For instance we check how the detectors sensitivity varies in time and how it might affect the performances of our searches.

Figure 2 shows an example of figure of merit displaying the minute trends of the inspiral search horizon distance at the Livingston's site, so-called *inspiral range*, for a period of twelve hours. The *inspiral range* is defined as the horizon of the search for 1.4-1.4 $M_\odot$ systems (binary neutron stars systems), averaged over all possible sky locations and source polarizations, assuming a detection with *Signal-To-Noise Ratio* equal to 8. In figure 2 the vertical lines demarcate the edges of the science segments, during which the mean value of the *inspiral range* is lying between 8.5 and 10 Mpc for this twelve-hour period. However one can notice that the segment between *hour=0* and *hour 1.5* in figure 2 (near the center of the plot) terminates with a dropping *inspiral range* for about twenty five minutes. The inspiral trigger associated with *Candidate B* (background trigger) in the Livingston data was found inside this segment while the *inspiral range* was about 2.5 Mpc, that is to say well below the averaged sensitivity reached by the interferometer during this day. This indicates that the *Candidate B* was detected while the Livingston interferometer was exceptionally noisy. Such an observation reduces our confidence in *Candidate B*, although this check does not prove that the candidate itself is due to detector noise. A confirmation of the nature of the candidate is brought by its *Signal-To-Noise Ratio* or χ^2 time series as discussed in section 4.3.

4.2. Environmental and instrumental causes

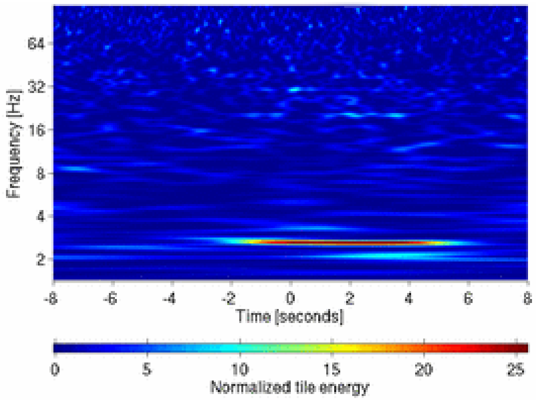

Figure 3. *Q spectrogram* of a transient in a seismometer located near the end mirror test mass of the Hanford 2 km interferometer. This transient is coincident with the inspiral trigger associated to the *Candidate G* (simulated gravitational-wave signal).

In order to check for possible instrumental artifacts that could be responsible for background triggers, we examine the auxiliary channels of the detectors in a few seconds long window around the candidate. For this purpose time-frequency maps of auxiliary channels are being analysed, using an event visualization tool called *QScan*, which is based on a Q-transform. More details on this tool can be found in [5]. *Qscan* produce time-series and *Q spectrograms* of the auxiliary channels in which transients are detected. The *Q spectrograms* correspond to time-frequency decompositions using sinusoidal Gaussians characterized by a central time, central frequency, and a quality

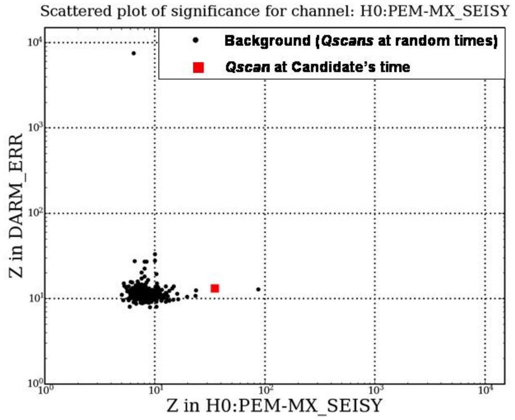

Figure 4. Scattered plot of *QScan significance* in a seismometer channel (*H0:PEM-MX_SEISY*) and in the error channel of the differential mode control loop of the H2 interferometer (*H2:LSC-DARM_ERR*). The dots refer to the *QScan significance* measured at times randomly distributed over the first calendar year of the S5 run, and the square refers to the *QScan significance* measured at the time of the *Candidate G*.

factor Q. An example of *Q spectrogram* is provided in Figure 3 showing a transient in a Hanford seismometer which is found in coincidence with the *Candidate G* (simulated gravitational-wave signal). When a candidate-event is simultaneous with a transient in an auxiliary channel, further investigations are performedas explained below.

Statistical significance of instrumental transients

The statistical relevancy of the transient found in the auxiliary channel is calculated by comparing its strength (which is characterized by a parameter called *Z significance*, equivalent to a *Signal-To-Noise Ratio* for the Q-transform) to an estimated background. For each auxiliary channel the background is estimated by running *QScan* at times which are randomly distributed over large epochs of the data-taking. This provides a distribution of *Z significances* corresponding to the background of the auxiliary channel. Figure 4 shows a scattered plot of the *Z significance* measured in the seismometer versus the *Z significance* measured in the error channel of the differential mode control loop of the H2 interferometer (which is the main port sensitive to gravitational waves). The dots refer to the estimated background while the square shows the significances measured in both channels at the time when the *Candidate G* was detected. A seismic transient whose *Z significance* would be comparable with the median of the background distribution could be ignored as irrelevant. On the contrary the seismic transient shown in Figure 3 has a higher significance than the background, which makes it statistically relevant. If we faint to ignore the nature of the *Candidate G*, the next question that needs to be addressed is to identify whether or not this environmental transient couples to the interferometer output port.

Coupling of environmental disturbances into the interferometer output port

The coupling of an environmental distrubance into the interferometer output port can be proven by comparing the *Q spectrogram* of the auxiliary channel to the *Q spectrogram* of the interferometer output port and by looking for possible correlations between this two channels. A high frequency disturbance might couple linearly into the gravitational-wave bandwidth of the output port. When a measured transfer function from the auxiliary channel to the output port is available one can then compare it to the amplitude ratio measured in the *Q spectrograms* of the two channels. If such a coupling is proven, this leads to the rejection of the candidate as a possible detection.

In the case of a low frequency seismic transient, noise upconversion mechanisms might induce a false-alarm trigger in the gravitational-wave bandwidth. One can notice in Figure 4 that the *Z significance* measured in the error signal of the interferometer differential mode at the time of the *Candidate G* is comparable to the significances obtained at random times. Therefore, despite the statistical relevancy of the transient in the seismometer, Figure 4 does not argue in favor of a possible coupling into the interferometer output port. Knowing that the *Candidate G* is a simulated gravitational-wave signal, the presence of an inspiral trigger is indeed not related to the seismic transient.

4.3. Candidate appearance

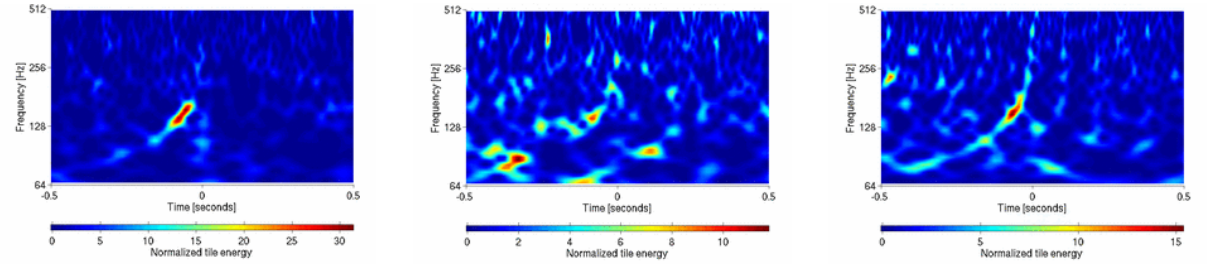

Figure 5. *Q spectrograms* of the data containing the *Candidate G* (simulated gravitational-wave signal) in each detector (from left to right): Hanford 4 km, Hanford 2 km, Livingston 4 km. A "chirp" waveform is visible in the data of both Hanford 4 km and Livingston 4km.

In this section two examples of qualitative checks of the candidate's appearance are illustrated: a check of the candidate's time-frequency map, and a check of the output of the match-filtering algorithm used to search for inspiral gravitational-wave signals.

Q spectrograms of the candidate

A *QScan* of the data in which a candidate-event as been detected is examined in order to perform the following checks:

Figure 6. *Signal-To-Noise Ratio* (left) and χ^2 (right) time series obtained after match-filtering the Livingston data containning the *Candidate G* (simulated inspiral gravitational-wave signal). The time origin on the x axis coincides with the time of the inspiral trigger.

Figure 7. *Signal-To-Noise Ratio* (left) and χ^2 (right) time series obtained after match-filtering the Livingtson data containning the *Candidate B* (background trigger). The time origin on the x axis coincides with the time of the inspiral trigger.

- The presence of a possible known signal waveform that might confirm the detection is verified. However, a low *Signal-To-Noise Ratio* inspiral signal is not expected to be visible in a *Q spectrogram*. Therefore the absence of visible known waveform in the spectrogram does not rule out a possible detection.
- The presence of an obvious excess of noise in the data is also checked.

Figure 5 shows the *Q spectrograms* of the *Candidate G* in each interferometer where this simulated gravitational-wave signal was injected.The transient visible in the H1 and L1 data corresponds to the typical "chirp" pattern that is characterestic of an inspiral signal. The simulated gravitational-wave signal is thus visible in these two spectrograms. The reason why it is not clearly visible in the H2 data is due to the lower *Signal-To-Noise Ratio* in this interferometer.

Output of the match-filtering algorithm

Another example of check for the candidate's appearance that is used by the CBC group consists in examining the time-series of the *Signal-To-Noise Ratio* obtained after match-filtering the data with inspiral waveforms, as well as the time-series of a χ^2 which aims to test the consistency between the triggered waveform and the signal present in the data.

An example of the expected time-series for a simulated gravitational-wave signal is shown in Figure 6. On the left plot, the *Signal-To-Noise Ratio* time series shows a short central peak corresponding to the time of the trigger associated with the simulated inspiral signal. On the right plot, the χ^2 time-series presents a very characteristic shape for a few tens of milli-seconds around the inspiral trigger, which corresponds to the expectations for gravitational-wave signal in stationary gaussian noise.
Figure 7 shows the *Signal-To-Noise Ratio* and χ^2 time-series around the time of the *Candidate B* at Livingston. Multiple peaks of *Signal-To-Noise Ratio* are visible, which indicates highly non-stationary data. Moreover the χ^2 time series shows large values for the whole two seconds window surrounding the candidate, which again indicates a very noisy stretch of data. Accordingly the *Candidate B* can be ruled out as a possible detection, which confirms the first suspicions born from the analysis of the inspiral horizon in section4.1.

5. Conclusions and perspectives

The Burst and CBC groups are pursuing the refinement of the detection checklist for candidate-event validation. Part of these tests are still under development. Efforts of the groups are currently aiming to automate this detection checklist in order to build a candidate follow-up pipeline which will improve the swiftness of the analysis. The detection checklist will be a crucial tool to analyse the candidate-events obtained by the searches running on the LIGO S5 science run.

6. Acknowledgments

The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory and the Particle Physics and Astronomy Research Council of the United Kingdom, the Max-Planck-Society and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the Natural Sciences and Engineering Research Council of Canada, the Council of Scientific and Industrial Research of India, the Department of Science and Technology of India, the Spanish Ministerio de Educacion y Ciencia, the John Simon Guggenheim Foundation, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation.

7. References

[1] Abbott B *et al* (LIGO Scientific Collaboration), “LIGO: The Laser Interferometer Gravitational-Wave Observator” (2007), *Preprint* arXiv:0711.3041

[2] Blackburn L, Cadonati L, Caride S, Caudill S, Chatterji S *et al*, "The LSC Glitch Group: Monitoring Noise Transients During the 5th LIGO Science Run", To be published in Class. Quantum Grav.

[3] Abbott B *et al* (LIGO Scientific Collaboration), Phys. Rev. D **77**, 062002 (2008)

[4] Klimenko S., Mohanty S., Rakhmanov M., Mitselmakher G., Phys. Rev. D **72**, 122002 (2005)

[5] Chatterji S, Ph.D thesis, MIT (2005)