2-14-2018

Effects of Data Quality Vetoes on a Search for Compact Binary Coalescences in Advanced LIGO's First Observing Run

B. P. Abbott  
*California Institute of Technology*

K. Gill  
*Embry-Riddle Aeronautical University*

B. Hughey  
*Embry-Riddle Aeronautical University*

J. Pratt  
*Embry-Riddle Aeronautical University*

M. J. Szczepańczyk  
*Embry-Riddle Aeronautical University*

*See next page for additional authors*

Follow this and additional works at: https://commons.erau.edu/publication

Part of the Astrophysics and Astronomy Commons

Scholarly Commons Citation
Authors

This article is available at Scholarly Commons: https://commons.erau.edu/publication/785
Effects of Data Quality Vetoes on a Search for Compact Binary Coalescences in Advanced LIGO’s First Observing Run

B P Abbott¹, R Abbott¹, T D Abbott², M R Abernathy³, F Acernese⁴,⁵, K Ackley⁶, C Adams⁷, T Adams⁸, P Addesso⁹, R X Adhikari¹, V B Adya¹⁰, C Affeldt¹⁰, M Agathos¹¹, K Agatsuma¹¹, N Aggarwal¹², O D Aguiar¹³, L Aiello¹⁴,¹⁵, A Ain¹⁶, B Allen¹⁰,¹⁷,¹⁸, A Allocca¹⁹,²⁰, P A Altin²¹, S B Anderson¹, W G Anderson¹⁷, K Araï¹, M C Araya¹, C C Arceneaux²², J S Areeda²³, N Arnaud²⁴, K G Arun²⁵, S Ascenzi²⁶,¹⁵, G Ashton²⁷, M Ast²⁸, S M Aston²⁹, P Astone²⁹, P Aufmuth³⁰, C Aulbert³⁰, S Babak³⁰, P Bacon³¹, M K M Bader¹¹, P T Baker³², F Baldacchini³³,³⁴, G Ballardin³⁵, S W Ballmer³⁶, J C Barayoga¹, S E Barclay³⁷, B C Barish¹, D Barker³⁸, F Barone³⁹, B Barr⁴⁰, L Barsotti¹², M Barsuglia³¹, D Bart³⁹, J Bartlett³⁸, I Bartos⁴⁰, R Bassiri³¹, A Basti¹⁹,²⁰, J C Batch³⁸, C Baune³⁰, V Bavigadda³⁵, M Bazzan³²,³³, M Bejger³⁴, A S Bell³⁷, B K Berger¹, G Bergmann³⁰, C P L Berry⁴⁵, D Bersanetti⁴⁶,⁴⁷, A Bertolini¹¹, J Betzwieser⁷, S Bhagwat³⁶, R Bhandare⁴⁸, I A Bilenko⁴⁹, G Billingsley¹, J Birch⁷, R Birney⁵⁰, S Biscans¹², A Bisht¹⁰,¹⁸, M Bitossi³⁵, C Biwer³⁶, M A Bizouard²⁴, J K Blackburn¹, C D Blair³¹, D G Blair³¹, R M Blair³⁸, S Bloemen⁵², O Bock¹⁰, M Boer⁵³, G Bogaert⁵³, C Bogan¹⁰, A Bohe⁴⁰, C Bond⁴³, F Bondu⁵⁴, R Bonnand⁵⁵, B A Boom¹¹, R Bork¹, V Boschi¹⁹,²⁰, S Bosso³⁵,¹⁶, Y Bouffanais³¹, A Bozzi³⁵, C Bradaschi²⁰, P R Brady¹⁷, V B Braginsky⁴⁹, M Branchesi⁵⁶,⁵⁷, J E Braun⁵⁸, T Briant⁵⁹, A Brillet³³, M Brinkmann¹⁰, V Brisson²⁴, P Brockill¹⁷, J E Broda⁶⁰, A F Brooks¹, D A Brown³⁶, D D Brown⁴⁵, N M Brown¹², S Brunett¹, C C Buchanan², A Buikema¹², T Bulik⁶¹, H J Bulen⁶²,¹¹, A Buonanno³⁰,⁶³, D Buskulic⁸, C Buy³¹, R L Byer¹¹, M Cabero¹⁰, L Cadonati⁶⁴, G Cagnoli⁶⁵,⁶⁶, C Caihill¹, J Calderón Bustillo⁶⁴, T Callister¹, E Calloni³⁷,⁵, J B Camp⁶⁸, K C Cannon⁶⁹, J Cao⁷⁰, C D Capano¹⁰, E Capocasa³¹, F Carbognani³⁵, S Caride⁷¹, J Casanueva Diaz²⁴, C Casentini²⁶,¹⁵, S Caudill¹⁷, M Cavaglià²², F Cavalier²⁴, R Cavalieri²⁵, G Cella²⁰, C B Cepeda¹, L Cerboni Baiardi⁵⁶,⁵⁷, G Cerretani¹⁹,²⁰, E Cesarini²⁶,¹⁵, S J Chamberlin⁷², M Chan⁷⁷, S Chao⁷³, P Charlton⁷⁴, E Chassande-Mottin³¹, B D Cheeseboro⁷⁵,
Data Quality for CBC Searches in O1

Data Quality for CBC Searches in O1


(LIGO Scientific Collaboration and Virgo Collaboration)


1LIGO, California Institute of Technology, Pasadena, CA 91125, USA
2Louisiana State University, Baton Rouge, LA 70803, USA
3American University, Washington, D.C. 20016, USA
4Università di Salerno, Fisciano, I-84084 Salerno, Italy
5INFN, Sezione di Napoli, Complesso Universitario di Monte S.Angelo, I-80126 Napoli, Italy
6University of Florida, Gainesville, FL 32611, USA
7LIGO Livingston Observatory, Livingston, LA 70754, USA
8Laboratoire d’Annecy-le-Vieux de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy-le-Vieux, France
9University of Sannio at Benevento, I-82100 Benevento, Italy and INFN, Sezione di Napoli, I-80100 Napoli, Italy
10Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany
11Nikhef, Science Park, 1098 XG Amsterdam, The Netherlands
12LIGO, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
13Instituto Nacional de Pesquisas Espaciais, 12227-010 São José dos Campos, São Paulo, Brazil
14INFN, Gran Sasso Science Institute, I-67100 L’Aquila, Italy
15INFN, Sezione di Roma Tor Vergata, I-00133 Roma, Italy
16Inter-University Centre for Astronomy and Astrophysics, Pune 411007, India
17University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
18Leibniz Universität Hannover, D-30167 Hannover, Germany
19Università di Pisa, I-56127 Pisa, Italy
20INFN, Sezione di Pisa, I-56127 Pisa, Italy
21Australian National University, Canberra, Australian Capital Territory 0200, Australia
22The University of Mississippi, University, MS 38677, USA
23California State University Fullerton, Fullerton, CA 92831, USA
24LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
25Chennai Mathematical Institute, Chennai 603103, India
26Università di Roma Tor Vergata, I-00133 Roma, Italy
27University of Southampton, Southampton SO17 1BJ, United Kingdom
28Universität Hamburg, D-22761 Hamburg, Germany
29INFN, Sezione di Roma, I-00185 Roma, Italy
30Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-14476 Potsdam-Golm, Germany
Data Quality for CBC Searches in O1

31 APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, F-75205 Paris Cedex 13, France
32 Montana State University, Bozeman, MT 59717, USA
33 Università di Perugia, I-06123 Perugia, Italy
34 INFN, Sezione di Perugia, I-06123 Perugia, Italy
35 European Gravitational Observatory (EGO), I-56021 Cascina, Pisa, Italy
36 Syracuse University, Syracuse, NY 13244, USA
37 SUPA, University of Glasgow, Glasgow G12 8QQ, United Kingdom
38 LIGO Hanford Observatory, Richland, WA 99352, USA
39 Wigner RCP, RMKI, H-1121 Budapest, Konkoly Thege Miklós út 29-33, Hungary
40 Columbia University, New York, NY 10027, USA
41 Stanford University, Stanford, CA 94305, USA
42 Università di Padova, Dipartimento di Fisica e Astronomia, I-35131 Padova, Italy
43 INFN, Sezione di Padova, I-35131 Padova, Italy
44 CAMK-PAN, 00-716 Warsaw, Poland
45 University of Birmingham, Birmingham B15 2TT, United Kingdom
46 Università degli Studi di Genova, I-16146 Genova, Italy
47 INFN, Sezione di Genova, I-16146 Genova, Italy
48 RRCAT, Indore MP 452013, India
49 Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia
50 SUPA, University of the West of Scotland, Paisley PA1 2BE, United Kingdom
51 University of Western Australia, Crawley, Western Australia 6009, Australia
52 Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
53 Artemis, Université Côte d’Azur, CNRS, Observatoire Côte d’Azur, CS 34229, Nice cedex 4, France
54 Institut de Physique de Rennes, CNRS, Université de Rennes 1, F-35042 Rennes, France
55 Washington State University, Pullman, WA 99164, USA
56 Università degli Studi di Urbino “Carlo Bo,” I-61029 Urbino, Italy
57 INFN, Sezione di Firenze, I-50019 Sesto Fiorentino, Firenze, Italy
58 University of Oregon, Eugene, OR 97403, USA
59 Laboratoire Kastler Brossel, UPMC-Sorbonne Universités, CNRS, ENS-PSL Research University, Collège de France, F-75005 Paris, France
60 Carleton College, Northfield, MN 55057, USA
61 Astronomical Observatory Warsaw University, 00-478 Warsaw, Poland
62 VU University Amsterdam, 1081 HV Amsterdam, The Netherlands
63 University of Maryland, College Park, MD 20742, USA
64 Center for Relativistic Astrophysics and School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, USA
65 Laboratoire des Matériaux Avancés (LMA), CNRS/IN2P3, F-69622 Villeurbanne, France
66 Université Claude Bernard Lyon 1, F-69622 Villeurbanne, France
67 Università di Napoli “Federico II,” Complesso Universitario di Monte S. Angelo, I-80126 Napoli, Italy
68 NASA/Goddard Space Flight Center, Greenbelt, MD 20771, USA
69 RESCEU, University of Tokyo, Tokyo, 113-0033, Japan.
70 Tsinghua University, Beijing 100084, China
71 Texas Tech University, Lubbock, TX 79409, USA
72 The Pennsylvania State University, University Park, PA 16802, USA
73 National Tsing Hua University, Hsinchu City, 30013 Taiwan, Republic of China
74 Charles Sturt University, Wagga Wagga, New South Wales 2678, Australia
75 West Virginia University, Morgantown, WV 26506, USA
76 University of Chicago, Chicago, IL 60637, USA
77 Caltech CaRT, Pasadena, CA 91125, USA
Abstract.

The first observing run of Advanced LIGO spanned 4 months, from September 12, 2015 to January 19, 2016, during which gravitational waves were directly detected from two binary black hole systems, namely GW150914 and GW151226. Confident detection of gravitational waves requires an understanding of instrumental transients and artifacts that can reduce the sensitivity of a search. Studies of the quality of the detector data yield insights into the cause of instrumental artifacts and data quality vetoes specific to a search are produced to mitigate the effects of problematic data. In this paper, the systematic removal of noisy data from analysis time is shown to improve the sensitivity of searches for compact binary coalescences. The output of the PyCBC pipeline, which is a python-based code package used to search for gravitational wave signals from compact binary coalescences, is used as a metric for improvement. GW150914 was a loud enough signal that removing noisy data did not improve its significance. However, the removal of data with excess noise decreased the false alarm rate of GW151226 by more than two orders of magnitude, from 1 in 770 years to less than 1 in 186000 years.
1. Introduction

The Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) is comprised of two dual-recycled Michelson interferometers [1] located in Livingston, LA (L1) and Hanford, WA (H1). A gravitational wave passing through a LIGO interferometer will induce a strain on spacetime, stretching and squeezing the 4 km arms and generating an interferometric signal at the antisymmetric port of the beamsplitter.

Advanced LIGO's first observing run (O1) lasted from September 12, 2015 to January 19, 2016. A primary goal of this observing run was the detection of gravitational waves from compact binary coalescences (CBC) [2]. This goal was achieved with the detections of GW150914 and GW151226, both signals from binary black hole systems, which mark the first direct detections of gravitational waves [3, 4]. These detections were part of a broader search for CBC signals carried out by multiple search pipelines during O1 [5, 6, 7, 8, 9, 10] and searches for unmodeled transients [11, 12, 13, 14].

Searching for gravitational waves requires an understanding of instrumental features and artifacts that can adversely affect the output of a gravitational wave search pipeline. Throughout the observing run, noisy data were identified in the form of data quality (DQ) vetoes to ensure that the analysis pipelines did not analyze data known to be contaminated with excess noise [15]. These vetoes are discussed further in Section 4. This study measures the effects of removing data with excess noise on the output of PyCBC [9, 5, 10], a python-based pipeline used to search for CBC signals. Section 3 contains a brief description of the PyCBC search pipeline and its internal DQ features.

Section 2 outlines the data selection and noise characterization processes. The DQ vetoes that are generated in the noise characterization process are described in Section 4. The methodology of this study is discussed in Section 5. Section 8 describes the limiting noise sources for the PyCBC search. This paper focuses on two specific subsets of the O1 data set. The first data set, from September 12 - October 20, 2015, was used for background estimation for GW150914. This data set is discussed in Section 6. The second data set, from December 3, 2015 - January 19, 2016, was used for background estimation for GW151226. This data set is discussed in Section 7.

2. Data Selection

Data were marked as suitable to be used in a gravitational wave search based on a set of conditions applied to each detector. The first condition indicates that the detector is in its nominal configuration or observation state according to software monitors used to control the instrument. The second condition indicates that no excitations or test signals are being applied to the instrument and that the instrument is undisturbed. This condition is set by the on-duty instrument operator on site who is continuously monitoring the detector performance.

The gravitational wave strain data measured at the output of the detectors are typically non-stationary and non-Gaussian and contain transient noise artifacts of varying durations. The longer duration non-stationary data can affect the overall sensitivity of the search, but they do not result in loud background events as they occur on a time-scale that is longer than any CBC waveform. The transient noise artifacts, however, can reduce the sensitivity of CBC searches by producing loud background
events.

Data quality studies must be performed to search for causes of transients in the data that generate loud events in a gravitational wave search. If the source of noise is identified, a veto is generated to flag times when transient noise makes the data unsuitable for analysis. Section 4 describes DQ vetoes that are used to indicate when the detector data are known to have excess noise [16, 17, 15, 18]. The exception to this process is gating, which is a feature internal to the CBC searches. Gating removes large transients from the data regardless of their source and is discussed in Section 3.1.

3. The PyCBC search pipeline

The PyCBC pipeline is designed to search for gravitational wave transients from CBCs [5]. It employs a matched filter algorithm, which correlates expected CBC waveforms with detector data and outputs a ranking statistic, the signal-to-noise ratio (SNR). If the ranking statistic exceeds a specified threshold, an event, or “trigger”, is generated. The SNR of each trigger is weighted based on a signal consistency test [19], resulting in a refined ranking statistic called re-weighted SNR. Section 3.2 discusses this signal consistency test further.

To perform this search, the matched filter algorithm needs to know what to search for. A collection of model CBC waveforms is generated before the analysis [20, 21]. Each of these waveforms is called a template and the full collection of waveforms is referred to as the template bank. This template bank is constructed to span the astrophysical parameter space included in the search [22]. Each waveform is defined by the mass and spin of each compact object in the binary system. It is often convenient to combine the effects of each object’s spin into one parameter called effective spin \( \chi_{\text{eff}} \), which is the mass-weighted spin of the system [7]:

\[
\chi_{\text{eff}} = \frac{\chi_1 m_1 + \chi_2 m_2}{m_1 + m_2},
\]

where \( \chi_i \) is the component of the dimensionless spin parameter [23] that is aligned with the orbital angular momentum, and \( m_i \) is the detector frame mass for each compact object in the binary system. The component masses are also used to calculate the chirp mass, which is used to parameterize gravitational wave signals in general relativity. Chirp mass is defined as [24]

\[
\mathcal{M} = \left(\frac{m_1 m_2}{m_1 + m_2}\right)^{3/5},
\]

where the \( m_i \) are the detector frame component masses of the compact objects in the binary system.

The search algorithm is run separately at each detector and a set of single detector triggers is generated. The two sets of single detector triggers are then compared to search for any events that were recorded within a 15 ms coincidence window, which reflects the travel time of a gravitational wave between the detectors and allows for uncertainty in the arrival time of a signal [5]. Any triggers that are found in coincidence with the same source parameters in both detectors represent potential gravitational wave signals and are referred to as foreground events. Some of these foreground events will be chance coincidences between noise in each detector, which is expected given the number of events in each data set.
To calculate the statistical significance of foreground events, a background distribution is generated. To generate the background, all coincident triggers are removed from the set of triggers generated for each detector, effectively removing all potential gravitational wave signals from the data set. The remaining triggers are then a realization of the background noise in each detector. These two sets of triggers, one from each detector, are then time shifted by a duration longer than the light travel time between the detectors. This time shift ensures that the two sets of triggers are astrophysically uncorrelated and do not contain any gravitational wave signals. The coincidence test is then performed again with the time shifted triggers, resulting in a coincident trigger set which represents background noise alone.

The statistical significance of any candidate gravitational wave is evaluated by calculating the rate of background events from detector noise that are at least as loud as the candidate event. This statistic is called the false alarm rate (FAR). Any loud triggers that appear as the result of instrumental transients will extend the background distribution and the influence the measured false alarm rate. The purpose of the DQ effort as a whole is thus two-fold: to ensure that the search is using representative detector data in the background noise estimation and to suppress the rate of loud events that will pollute both the background and the foreground distributions. Two additional stages of DQ that are internal to the PyCBC pipeline, gating and the $\chi^2$ signal consistency test, are discussed below.

3.1. Gating

The PyCBC search includes a data conditioning stage that applies preventative cuts on the input data stream. This gating [5] uses a window function to remove times containing large transients from the input data stream. This window function smoothly sets the value of the data to zero, excising a large transient. The time domain input data are Fourier transformed into the frequency domain and whitened using the measured amplitude spectral density. The data are then inverse Fourier transformed back into the time domain and compared to a threshold value. If the whitened time domain data have excursions that exceed this threshold, a gating window is constructed to remove these data from the input to the search.

3.2. $\chi^2$ signal consistency test

A further layer of effective DQ that is internal to the PyCBC pipeline is the application of the $\chi^2$ signal consistency test [19]. The SNR produced by the matched filter in PyCBC is an integral in the frequency domain. The $\chi^2$ test divides each CBC waveform into frequency bins of equal power, checking that the SNR is distributed as a function of frequency as expected from an actual CBC signal. Each trigger that comes out of the matched filter search is down-weighted based on the results of the $\chi^2$ test. This is folded into a new ranking statistic for CBC triggers, which is called re-weighted SNR and is denoted by $\hat{\rho}$. The ranking statistic for coincident events in the PyCBC search is the network re-weighted SNR, $\hat{\rho}_c$, which is the quadrature sum of the re-weighted SNR from each detector. Since a real signal has a power distribution that matches the template waveform, it will not be down-weighted by the $\chi^2$ test; the SNR and the re-weighted SNR will be the same.

This test is extremely powerful, as shown in Figure 1, which shows the distribution of single detector PyCBC triggers generated from September 12 to October 20, 2015.
Figure 1a shows the distribution of triggers in SNR. The extensive tail of triggers with high SNR, which extends beyond SNR 100, is down-weighted in the re-weighted SNR distribution, leaving behind a tail that extends to $\hat{\rho} \approx 10.5$ as seen in Figure 1b. This re-weighted SNR tail represents the loudest single detector background triggers in the CBC search. Investigating this set of loudest background triggers guides DQ efforts in defining the current limiting noise sources to the CBC search.

4. Data quality vetoes

As seen in Figure 1, the $\chi^2$ test is a powerful tool, but there is still a considerable tail in the single detector trigger distribution that will limit the attainable false alarm rate of the PyCBC search. This tail is often caused by transient instrumental noise. If these noise sources can be linked to a systematic instrumental cause or a period of highly irregular instrumental performance, they can be flagged and removed from the analysis in the form of a DQ veto.

DQ vetoes are produced for all analysis time based on systematic instrumental conditions without any regard for the presence of gravitational wave signals. All data are treated equally; the removal of data with excess noise has the ability to remove real gravitational wave signals as well as background events. There are two types of vetoes implemented in the PyCBC search: category 1 and category 2.

4.1. Category 1 and 2 vetoes

Category 1 vetoes are intended to mark times when significant instrumental issues are present and the data should not be used in any analysis. Category 1 vetoes often indicate time when the character of the data has drastically changed and should not be combined with noise estimations from times of nominal performance. An example of this from O1 is an electronics failure that dramatically changes the character of the background noise and creates transient noise artifacts. As such, category 1 vetoes remove data before any analysis pipelines are run. This ensures that severely problematic data are not used for background noise estimations and that no triggers will be generated at these times.

Category 2 vetoes are intended to mark short, noisy times that should not be treated as clean data. Category 2 flags are used to flag transients that could potentially generate loud triggers, but do not corrupt the surrounding data badly enough that they need to be excluded at the input to the pipeline. An example of this from O1 is a transient electronics saturation that only impacts the output data for 1 second. Data designated as category 2 will still be used to compute background noise estimations for the matched filter search, but any triggers generated during category 2 vetoed times will be excluded before background trigger distributions are calculated.

Further details on the application of DQ vetoes in the first observing run are available in a paper detailing the transient noise in the detectors at the time of GW150914 [15].

5. Measuring the Effects of Data Quality Vetoes

To test the effects of DQ vetoes, the PyCBC search pipeline was run with and without applying vetoes. The only vetoes that were used in all runs are those that indicate that the data were not properly calibrated, that a data dropout occurred, or that there
Figure 1: Histograms of single detector PyCBC triggers from the Livingston (L1) detector. These triggers were generated using data from September 12 to October 20, 2015. These histograms contain triggers from the entire template bank, but exclude any triggers found in coincidence between the two detectors. (1a) A histogram of single detector triggers in SNR. The tail of this distribution extends beyond SNR = 100. (1b) A histogram of single detector triggers in re-weighted SNR. The chi-squared test down-weights the long tail of SNR triggers in the re-weighted SNR distribution. The triggers found using only the Hanford detector have a similar distribution.
were test signals being injected into the detectors. Gating is internal to the search pipeline and was applied in all of the analyses. Two methods were used to understand the effects of applying vetoes. The first, described in Section 5.1, considers the average sensitivity of the search pipeline to gravitational wave signals. The second, described in Section 5.2, compares the measured search backgrounds and the false alarm rates of recovered gravitational wave signals.

5.1. Measuring search sensitivity

The metric used to measure the sensitivity of the search pipeline is sensitive volume. Sensitive volume is measured by injecting simulated gravitational wave signals into the data and attempting to recover them using the search [5]. The ability of the pipeline to recover signals at a given false alarm rate is then measured by analyzing the number of missed and recovered injections.

In addition to the sensitive volume, the amount of time used in the analysis must be considered when removing noisy data. If a search is rejecting too much data, it will miss the opportunity to detect signals. To address this, the sensitive volume of the search is multiplied by the amount of analysis time to create a new metric called VT. If time is removed from an analysis, the sensitive volume of the search must increase to make up for the shorter analyzed time.

The sensitivity of a search varies as a function of how significant candidate gravitational wave events are. The VT ratios are therefore calculated at both the 1 per 100 year and the 1 per 1000 year levels. These significance levels are expressed as inverse false alarm rates (IFAR).

5.2. Comparing search backgrounds

In the first observing run, the bank of CBC waveform templates used in the PyCBC search was divided into three bins [22]. The significance of any candidate gravitational wave found in coincidence between the two detectors is calculated relative to the background in its bin. Waveforms with different parameters will respond to instrumental transients in different ways. This binning is performed so that any foreground triggers are compared to a background generated from similar waveforms. As such, the effects of removing data from the PyCBC search are variable depending on which bin is considered. The actual gravitational wave signals discovered in the PyCBC search, GW150914 and GW151226, were part of a full search that was broken into 3 bins but reported as a single table of results. Because of this, their reported false alarm rates include a trials factor of 3. The background distributions shown in Sections 6 and 7 were measured on a bin-by-bin basis, so the cumulative trigger rates have not been divided by 3.

The first bin is called the binary neutron star (BNS) bin and contains all waveforms with $M < 1.74$. The second bin is the edge bin, which is defined based on the peak frequency $f_{\text{peak}}$ of each CBC waveform. These waveforms are typically shorter in duration than binary neutron star waveforms and are comprised of both binary black hole (BBH) and neutron star-black hole (NSBH) binary waveforms. Waveforms in the edge bin typically have high masses and negative $\chi_{\text{eff}}$. In this analysis, the edge bin contained waveforms with $f_{\text{peak}} < 100$ Hz. The third bin is the bulk bin, which contains all remaining waveforms needed to span the parameter space of the search. This contains BBH and NSBH waveforms with a variety of mass ratios
Data Quality for CBC Searches in O1

6. Analysis containing GW150914

This analysis lasted from September 12 - October 20, 2015 and contained a total of 18.2 days of coincident detector data. After category 1 vetoes were applied, 16.9 days of coincident data remained. After category 2 vetoes were applied, 16.8 days of coincident data were used in the final analysis. There were two interesting events that occurred in this analysis period. The first is GW150914, a gravitational wave signal from a binary black hole merger that marked the first direct detection of gravitational waves \[3\]. The second is a marginal candidate gravitational wave event, LVT151012, which stands out from the background distribution but does not have enough statistical significance to be quoted as a confident detection \[22, 25\].

6.1. Search sensitivity

To measure the effects of DQ vetoes on the sensitivity of the search, the analysis containing GW150914 was performed with and without applying data quality vetoes. The resulting measurements of VT were divided to calculate a VT ratio.

Figure 2 shows the change in VT when vetoes are applied for two values of IFAR and several chirp mass bins. The lowest chirp mass bin, which contains BNS signals, does not show any improvement in sensitivity when DQ vetoes are applied. For marginally significant signals at IFAR = 100, the measured value of VT increases by 3-32% in higher chirp mass bins. For highly significant signals at IFAR = 1000, the measured value of VT increases by 34-62% in higher chirp mass bins.

Figure 2: The change in search sensitivity when DQ vetoes are applied for the analysis containing GW150914. The error bars show the 1 \(\sigma\) error from each VT calculation combined in quadrature. The lowest chirp mass bin, which contains BNS signals, does not show any improvement in sensitivity. For marginally significant signals at IFAR = 100, the measured value of VT increases by 3-32% in higher chirp mass bins. For highly significant signals at IFAR = 1000, the measured value of VT increases by 34-62% in higher chirp mass bins.
6.2. BNS bin

Binary neutron star systems have the longest waveforms in the template bank, often spanning up to 60 seconds in duration. With such long waveforms, the $\chi^2$ test is effective at reducing the impact of transients on the BNS search. Typical instrumental transients have a small number of cycles and a duration of less than 1 second. As such, the overlap between a transient and a BNS signal is a small fraction of the total duration of the BNS waveform and is easily distinguished as noise in the re-weighted SNR calculation. This is demonstrated in Figure 3, which shows the distribution of single detector triggers in SNR and re-weighted SNR. The tail of high SNR triggers is down-weighted, resulting in a re-weighted SNR distribution that extends to $\hat{\rho} \approx 8.3$.

![Figure 3: Histograms of Livingston (L1) single detector triggers found in the BNS bin.](image)

Since the $\chi^2$ test is so effective in this bin, it is rare to see strong outliers in the re-weighted SNR distribution. Figure 4 shows the background distribution of the BNS bin in the PyCBC search for the analysis containing GW150914. The cumulative rate of background events in a given bin indicates the rate of false alarms expected in that bin for a given re-weighted SNR. In this bin, there is no substantial improvement for any value of $\hat{\rho}_c$. See Sections 6.3 and 6.4 for a contrary case.
6.3. Bulk bin

Figure 5 shows the background distribution in the bulk bin for the analysis containing GW150914. The first noticeable change is that the loudest background event is at $\hat{\rho}_c = 14$ in the presence of noisy data compared to 12 when all DQ vetoes are applied. This new loudest event does not show up as a small outlier; there is a significant shoulder in the distribution that persists up to $\hat{\rho}_c = 12$ before falling off. Considering the two distributions as a whole, there is a separation between the two curves beginning at $\hat{\rho}_c = 9$, which reaches an order of magnitude discrepancy at $\hat{\rho}_c \approx 11$ and continues to diverge at higher values of $\hat{\rho}_c$.

6.3.1. LVT151012

The second most significant trigger in the analysis containing GW150914 was LVT151012, recorded on October 12, 2015. This trigger was recovered in the bulk bin with $\hat{\rho}_c = 9.75$ with a false alarm rate of 0.33 yr$^{-1}$. This is not significant enough to be claimed as a confident detection but is nevertheless interesting. The false alarm rate decreases by a factor of 2.1 when DQ vetoes are applied, as shown in Table 1.

<table>
<thead>
<tr>
<th>Analysis configuration</th>
<th>False alarm rate (yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All vetoes applied</td>
<td>0.33</td>
</tr>
<tr>
<td>No vetoes applied</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Table 1: Table of bulk bin false alarm rates for LVT151012.
6.4. Edge bin

Figure 6 shows the background distribution in the edge bin before and after data with excess noise have been removed from the analysis. If noisy data are not removed from the analysis, there is a noticeable extension of the tail of loudest events. The loudest background event with no data removed from the analysis is at $\hat{\rho}_c = 15.5$ compared to $\hat{\rho}_c = 13.3$ when all vetoes are applied. There is a visible separation between the two curves that increases for larger values of $\hat{\rho}_c$, indicating that the ability of the search pipeline to make confident detections is diminished.

6.4.1. GW150914  The gravitational wave signal GW150914, produced from the inspiral and merger of a binary black hole system, was detected on September 14, 2015 and was recovered by the PyCBC search with $\hat{\rho}_c = 23.6$ [3]. GW150914 was louder than any background event in the analysis regardless of what data were considered. This being the case, DQ vetoes do not improve the false alarm rate for GW150914. Since GW150914 is a loud enough event that sits well above the search background, it is not the type of event that is expected to benefit from vetoes. This is quantified in Table 2.

7. Analysis containing GW151226

The extended analysis containing GW151226 lasted from December 3, 2015 - January 19, 2016 and contained a total of 16.7 days of coincident detector data. After category 1 vetoes were removed, 15.9 days of coincident data remained. After category 2 vetoes
Figure 6: The background distribution in the edge bin before and after applying DQ vetoes. (6a) The cumulative rate of background triggers in the edge bin as a function of re-weighted SNR. (6b) A histogram of background triggers in the edge bin. The red traces indicate the distribution of background triggers without noisy data removed from the analysis and the cyan traces indicate the distribution of background triggers with all data quality vetoes applied.

Table 2: Table of bulk bin false alarm rates for GW150914. GW150914 is loud enough that its false alarm rate is not strongly affected by removing noisy data from the analysis. The slight change in false alarm rate is due to each analysis configuration containing different amounts of analysis time.

<table>
<thead>
<tr>
<th>Analysis configuration</th>
<th>False alarm rate (yr(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>All vetoes applied</td>
<td>&lt; 5.17 × 10(^{-6})</td>
</tr>
<tr>
<td>No vetoes applied</td>
<td>&lt; 4.43 × 10(^{-6})</td>
</tr>
</tbody>
</table>

were removed, 15.6 days of coincident data were used in the final analysis. This analysis time provided an extended background estimation for the binary black hole merger GW151226 [4], which was detected by the aLIGO detectors on December 26, 2015.

7.1. Search sensitivity

Figure 7 shows the change in VT when DQ vetoes are applied to the analysis containing GW151226. For this analysis, the lowest chirp mass bin, which contains BNS signals, shows a slight improvement when vetoes are applied. This improvement is discussed further in section 7.2. Similar to the analysis containing GW150914, the higher chirp mass bins show an improvement in search sensitivity for both values of IFAR.

7.2. BNS bin

The BNS bin shows a slight improvement when DQ vetoes are applied. Figure 8 shows the background distributions in the BNS bin before and after removing data with vetoes. Although the \(\hat{\rho}_c\) of the loudest background event does not change considerably,
there is a noticeable gap between the two background distributions that is visible at \( \hat{\rho}_c > 9.7 \) and widens to an order of magnitude difference in FAR at \( \hat{\rho}_c \approx 10.5 \). The removal of noisy data does marginally improve the background in the BNS bin, but the two distributions are still similar and would not be limiting to a CBC search.

Figure 7: The change in search sensitivity when DQ vetoes are applied for the analysis containing GW151226. The error bars show the 1 \( \sigma \) error from each VT calculation combined in quadrature. The lowest chirp mass bin, which contains BNS signals, shows a small improvement in sensitivity when vetoes are applied, though the error bars are consistent with a VT ratio of 1. For marginally significant signals at IFAR = 100, the measured value of VT increases by 27-62% in higher chirp mass bins. For highly significant signals at IFAR = 1000, the measured value of VT increases by 45-90% in higher chirp mass bins.

Figure 8: The background distribution in the BNS bin before and after applying DQ vetoes. (8a) The cumulative rate of background triggers in the BNS bin as a function of re-weighted SNR. (8b) A histogram of background triggers in the BNS bin. The red traces indicate the distribution of background triggers without noisy data removed and the cyan traces indicate the distribution of background triggers with all vetoes applied.
7.3. Bulk bin

The bulk bin benefited from the application of DQ vetoes. Figure 9 shows the bulk bin background distribution before and after DQ vetoes were applied. The first notable effect is that if DQ vetoes are not applied, then the loudest background event is at $\hat{\rho}_c = 14.3$ rather than $\hat{\rho}_c = 12.4$. This effect limits the values of $\hat{\rho}_c$ for which a significant detection could be claimed. The second effect is the visible separation between the two curves, indicating an increase in false alarm rate for any trigger with $\hat{\rho}_c > 9$.

![Figure 9: The background distribution in the bulk bin before and after applying DQ vetoes. (9a) The cumulative rate of background triggers in the bulk bin as a function of re-weighted SNR. (9b) A histogram of background triggers in the bulk bin. The red traces indicate the distribution of background triggers without noisy data removed, the cyan traces indicate the distribution of background triggers with all DQ vetoes applied. The dotted line indicates GW151226, which was recovered with $\hat{\rho}_c = 12.7$. If no vetoes are applied, the tail of loud background triggers extends to $\hat{\rho}_c = 14.3$ instead of $\hat{\rho}_c = 12.4$. The impact of this change is apparent when considering GW151226, which is no longer the loudest event in this bin (see Section 7.3.1).](image)

7.3.1. GW151226  The binary black hole system GW151226 was recovered by the PyCBC pipeline in the bulk bin with $\hat{\rho}_c = 12.7$ [4]. The significance of GW151226 was improved by the application of DQ vetoes. These changes in significance are quantified in Table 3. If noisy data are not removed from the analysis, GW151226 is not the loudest event in the bulk bin and its false alarm rate is 1 in 770 years. When all vetoes are applied to the analysis, GW151226 is the loudest event in the bulk bin and has a false alarm rate of less than 1 per 186000 years. The application of DQ vetoes decreases the false alarm rate by over two orders of magnitude and elevates GW151226 from a detection candidate to a clear detection.

7.4. Edge bin

The background distribution in the edge bin looks dramatically different if DQ vetoes are not applied. This is not surprising, given that templates in the edge bin will have a short duration and will be susceptible to instrumental transients. Figure 10 shows
Data Quality for CBC Searches in O1

<table>
<thead>
<tr>
<th>Analysis configuration</th>
<th>False alarm rate (yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All vetoes applied</td>
<td>$&lt; 5.39 \times 10^{-6}$</td>
</tr>
<tr>
<td>No vetoes applied</td>
<td>$1.30 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 3: Table of bulk bin false alarm rates of GW151226. The false alarm rate of GW151226 increases from less than 1 in 186000 years to 1 in 770 years if data with excess noise is not removed from the analysis.

The background distribution in the edge bin before and after vetoes have been applied. The loudest event in this bin with all vetoes applied was at $\hat{\rho}_c = 15$, which was already inconveniently loud for a search hoping to recover a signal in this bin. When noisy data are not removed from the analysis, the loudest background event is at $\hat{\rho}_c = 17.5$, which further restricts the region where a confident detection could be made. Further, there is a notable separation between the two background distributions at all values of $\hat{\rho}_c$.

8. Limiting noise sources

After applying DQ vetoes, there are still noticeable tails in the bulk and edge bin background distributions that limit the sensitivity of the search. This section aims to identify the types of instrumental features that are causing triggers with a high re-weighted SNR and acting as limiting noise sources. This section studies the analysis containing GW150914, which was discussed in Section 6.
8.1. Loud transients

A reasonable hypothesis is that the search is limited by loud transients with an SNR below the gating threshold. To study the impact of loud transients, a cut was applied to the CBC triggers to exclude all triggers with SNR $> 20$. The histograms of Livingston re-weighted SNR triggers in Figure 11 show the results of this test. The green histogram in the foreground of the plot has had all single detector triggers with SNR $> 20$ removed. The yellow histogram plotted in the background contains all single detector triggers from the analysis. The cut does remove a small number of triggers with $\hat{\rho} > 8$, but the overall structure of the tail is not significantly affected. None of the triggers with $\hat{\rho} > 10$ are removed by this cut. Most of the high SNR triggers are down-weighted below $\hat{\rho} = 6$ and are not visible on this histogram.

![Figure 11: A histogram of single detector re-weighted SNR triggers for the Livingston (L1) detector. The green bins indicate triggers with an SNR $< 20$. The yellow bins indicate all triggers in the data set. The triggers removed by the SNR cut do not significantly impact the loudest events which form a tail in the re-weighted SNR distribution. A small number of triggers at $\hat{\rho} > 9$ are removed by the SNR cut, but the population is not fully removed. The majority of the distribution is unchanged.](image)

8.2. Blip transients

The transients that are able to pass the $\chi^2$ test and populate the tail in the re-weighted SNR distribution are in fact those with a specific morphology which resembles that of short duration CBC waveforms. The most common and problematic source of transient noise that causes high re-weighted SNR triggers are called “blip transients” [15]. These transients are often the source of the highest re-weighted SNR triggers at
both the Livingston and Hanford detectors. Although blip transients are seen in both detectors, they are not found as coincident triggers and do not represent gravitational wave signals.

Blip transients show up as short duration, band-limited impulses that have power in the \(\sim 30-300\) Hz frequency range (see Figure 12). They do not couple into any monitors of detector performance and are not loud enough to exceed the gating threshold applied in the PyCBC search.

![Figure 12: A time-frequency representation [26] of the Livingston strain channel at the time of a blip transient. This visualization of a blip transient demonstrates their typical features: band-limited, short duration, and little visible frequency structure.](image)

A time-domain analysis reveals why blip transients are so damaging to the CBC searches. Figure 13 shows a filtered time-domain representation of a blip transient in the Livingston strain channel. The data have been filtered with a bandpass filter with notch filters to attenuate strong lines in the strain spectrum, double-passed to be zero-phase. Overlaid on top of the strain data is a CBC waveform that reported a high re-weighted SNR value at the time of the blip transient under study. The two curves show significant overlap in the few cycles where the template has appreciable amplitude.

The CBC template that reported a high re-weighted SNR when filtered against the blip transient in Figure 13 represents a neutron star-black hole binary system with a total mass \(M_{\text{total}}\) of \(98.34\) \(M_\odot\) and a highly anti-aligned effective spin of \(-0.97\), resulting in a short template duration. The waveform spends less than 0.1 seconds at the frequencies that aLIGO is sensitive to, which, as shown in Figure 13, is the approximate time scale of some instrumental transients. This time scale is in stark contrast to that of a binary neutron star waveform, which can have a duration on the order of 1 minute and contain ample signal for use in the \(\chi^2\) test.

Although blip transients are capable of creating high re-weighted SNR triggers, their effects are constrained to a fairly small region of the CBC parameter space.
Data Quality for CBC Searches in O1

Figure 13: A filtered time-domain representation of the Livingston strain channel $h(t)$ at the time of a blip transient. Overlaid on the strain plot is a filtered CBC waveform that reported a high re-weighted SNR value at the time of the blip transient. Both sets of data have been zero-phase bandpass filtered to isolate the frequency range that aLIGO is sensitive to. The two curves show significant overlap in the few cycles where the template has appreciable amplitude. The similarity between these two curves causes the $\chi^2$ test to be ineffective at down-weighting these transients.

Figure 14 shows single detector triggers from Livingston binned by total mass and effective spin. The bottom right corner of the plot, bounded by $M_{\text{total}} > 80M_\odot$ and $\chi_{\text{eff}} < -0.5$, contains all of the shortest duration templates and the highest re-weighted SNR triggers. This represents a small fraction of the CBC parameter space, containing only 65 waveform templates out of 249077 total. The loudest triggers in the plot are even further constrained, corresponding to waveform parameters similar to those in Figure 13.

Further investigation reinforces that the loudest triggers correspond to the templates with the shortest duration. Figure 15 shows single detector triggers from Livingston as a function of template duration and peak frequency of the CBC template. There is a systematic clustering of loud triggers below a template duration of 0.1 seconds, which is the timescale of typical instrumental transients. Constraining the loudest triggers using the peak frequency of the waveform template is not as successful. While the region corresponding to $f_{\text{peak}} < 100$ Hz does include the templates that are most susceptible to instrumental transients, it also includes numerous templates with a duration between 0.1 - 1.0 seconds that do not report any triggers with a high re-weighted SNR.

8.3. 60-200 Hz noise

Another limiting noise source for the CBC search is present only at Livingston and has commonly been referred to as the “60-200 Hz” noise. This noise occurs in clusters that can last multiple minutes and are typically comprised of a series of individual flares of noise that seem to last about 10-100 seconds each. These storms of noise correlate visibly with dips in the inspiral range, a figure of merit for the CBC searches that estimates the effective range at which detection of a binary neutron star inspiral
Data Quality for CBC Searches in O1

Figure 14: A plot of single detector triggers from the Livingston detector binned by total mass and effective spin. The color of each bin indicates the highest re-weighted SNR trigger found in that bin. The highest re-weighted SNR triggers are constrained to the bottom corner of the plot, bounded by $M_{\text{total}} > 80$ and $\chi_{\text{eff}} < -0.5$. This corner contains the shortest duration templates and is susceptible to instrumental transients such as blip transients.

is possible based on the shape of the noise curve. This noise contributes to the tail of loudest background triggers in the PyCBC search and is responsible for the cluster of loud triggers with a template duration of 4.4 s in Figure 15. Figure 16 shows the time-frequency representation of this noise on a 20 minute timescale.

A more focused look at these noisy periods reveals a structure that is reminiscent of scattered light [27], appearing as arc-like traces in the time-frequency plane as seen in Figure 17. However, the frequency of this noise is higher than is typically expected from scattered light and investigations have not been able to find an associated source of scattered light during these noisy periods. This noise was a common source of high re-weighted SNR triggers in the Livingston data throughout the first observing run, second only to blip transients.

9. Conclusions

Data quality vetoes improved the sensitivity of the PyCBC search in Advanced LIGO’s first observing run. Although the sensitivity of the search to BNS signals was not
Figure 15: A plot of single detector triggers from the Livingston detector binned by template duration and waveform template peak frequency. The loudest triggers in re-weighted SNR are constrained to the area of the parameter space with template durations < 0.1s, which is the timescale of typical instrumental transients, most notably blip transients. The small cluster of loud triggers with a template duration of roughly 4.4 s corresponds to the 60-200 Hz noise discussed in Section 8.3.

dramatically affected, VT improved significantly for higher mass sources when DQ vetoes are applied.

The gravitational wave signal GW150914 was strong enough that it was louder than all background events regardless of what data were removed from the search. As such, DQ vetoes did not improve its significance. The false alarm rate of LVT151012, which occurred during the same analysis period, was improved from 0.69 yr$^{-1}$ to 0.33 yr$^{-1}$ when vetoes were applied. The false alarm rate of the second gravitational wave signal discovered in O1, GW151226, was decreased by over two orders of magnitude when DQ vetoes were applied, which resulted in a clear detection. The production and application of DQ vetoes was critical for increasing overall sensitivity in Advanced LIGO’s first observing run and similar methods were employed during the second observing run.
Figure 16: A time-frequency spectrogram of the 60-200 Hz noise. This noise appears in storms that often last for many minutes. This time scale and frequency range is damaging to CBC searches and has often been found responsible for loud background events.

Figure 17: A zoomed in time-frequency spectrogram of the 60-200 Hz noise. This period of noise caused a loud trigger in the PyCBC background. The arc-like shape of the noise is reminiscent of noise due to scattered light, but the frequency of the noise is higher than expected.

10. Acknowledgements

The authors gratefully acknowledge the support of the United States National Science Foundation (NSF) for the construction and operation of the LIGO Laboratory and Advanced LIGO as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State
of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, the Department of Science and Technology, India, the Science & Engineering Research Board (SERB), India, the Ministry of Human Resource Development, India, the Spanish Agencia Estatal de Investigación, the Vicepresidència i Conselleria d’Innovació, Recerca i Turisme and the Conselleria d’Educació i Universitat del Govern de les Illes Balears, the Conselleria d’Educació, Investigació, Cultura i Esport de la Generalitat Valenciana, the National Science Centre of Poland, the Swiss National Science Foundation (SNSF), the Russian Foundation for Basic Research, the Russian Science Foundation, the European Commission, the European Regional Development Funds (ERDF), the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the Lyon Institute of Origins (LIO), the National Research, Development and Innovation Office Hungary (NKFI), the National Research Foundation of Korea, Industry Canada and the Province of Ontario through the Ministry of Economic Development and Innovation, the Natural Science and Engineering Research Council Canada, the Canadian Institute for Advanced Research, the Brazilian Ministry of Science, Technology, Innovations, and Communications, the International Center for Theoretical Physics South American Institute for Fundamental Research (ICTP-SAIFR), the Research Grants Council of Hong Kong, the National Natural Science Foundation of China (NSFC), the Leverhulme Trust, the Research Corporation, the Ministry of Science and Technology (MOST), Taiwan and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, MPS, INFN, CNRS and the State of Niedersachsen/Germany for provision of computational resources.

11. References


