

Syracuse University

SURFACE

Dissertations - ALL

SURFACE

August 2019

IMPROVING THE SENSITIVITY OF ADVANCED LIGO THROUGH DETECTOR CHARACTERIZATION

Derek Davis
Syracuse University

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Davis, Derek, "IMPROVING THE SENSITIVITY OF ADVANCED LIGO THROUGH DETECTOR CHARACTERIZATION" (2019). *Dissertations - ALL*. 1066.

<https://surface.syr.edu/etd/1066>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

This dissertation focuses on the impact of detector characterization work on searches for gravitational waves from compact binary coalescences (CBCs) in Advanced LIGO's second observing run (O2). This observing run started on November 30, 2016, and lasted until August 25, 2017, and resulted in the identification of 8 unambiguous gravitational-wave signals, including the first observation of a binary neutron star merger. The role of detector characterization is to leverage knowledge of both the interferometers and the data in order to improve aLIGO's ability to observe gravitational-waves.

I focus on the construction of the O2 noise subtracted data set that was searched as a part of the LIGO-Virgo Collaboration's first gravitational-wave catalog, GWTC-1. This data set was processed with a noise subtraction pipeline to remove the excess noise identified at each interferometer that resulted in a 30% improvement in the sensitive volume that aLIGO was able to probe. Equally important to the finalized data set is the inclusion of data quality vetoes that indicate periods of instrumental artifacts.

I also examine how these instrumental artifacts can mimic gravitational-wave waveforms and reduce the sensitivity of searches for CBC signals, with particular emphasis on the PyCBC pipeline. Understanding this connection is one of the key ways that gravitational waves are differentiated from instrumental artifacts. Finally, I detail the final results presented in the GWTC-1 catalog from a detector characterization perspective, and discuss how the efforts highlighted in this dissertation allowed for the detection of new gravitational-wave events and improved analyses of previously identified events.

**IMPROVING THE SENSITIVITY OF
ADVANCED LIGO THROUGH DETECTOR
CHARACTERIZATION**

By

Derek Davis

B.A. Physics, University of Chicago
B.S. Mathematics, University of Chicago

DISSERTATION
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN PHYSICS

Syracuse University
August 2019

Copyright © 2019 Derek Davis
All rights reserved.

ACKNOWLEDGEMENTS

As I finish up my time in graduate in school, it's an important point to acknowledge that no one gets this far all on their own. There's always people along that way lending a helping hand, and I feel so grateful to have so many of those people in my life.

I first want to thank my advisor, Peter Saulson. I definitely couldn't have had a better mentor looking out for me. Your excitement and experience about gravitational waves helped me keep going through to the finish, and you were always available to support me and push me forward. I am certainly grateful for all your help throughout my entire graduate career.

To professors Stefan Ballmer and Duncan Brown, for all of the guidance and knowledge that you've shared with me as a part of the gravity group here in Syracuse. The breadth of knowledge in the group made it a stimulating environment that always forced me to learn new things. The melding of all of these perspectives is one thing I will truly miss.

I also want to thank Ryan Fisher and Jax Sanders, for dealing with my many questions and helping me through the finer points in a fair share of projects. Working through ideas in your offices helped me get to the next step.

To all of the other grad students who were in this with me at Syracuse: Daniel Vander-Hyde, Ari Pedersen, Steven Reyes, Lindsay DeMarchi, Lorena Magaña Zertuche, Varun Srivastava, Daniel Finstad, Soumi De, Chaitanya Afle, Erik Muniz, and Nick Didio. You made coming into the office each day exciting. It was a pleasure to have this experience together, and for those that are still here, I know that you'll all make to the finish line.

To Jess McIver, TJ Massinger, Laura Nuttall, Marissa Walker, Andy Lundgren, Duncan MacLeod, Josh Smith, and the rest of Team Detchar for helping me grow as a scientist and as a person. You were always there to help me solve problems big and small, as well as give words of encouragement or share a laugh. I've always enjoyed being a part of a collaboration, and you all have been an important part of making that experience everything that it was.

Thank you to my committee members, Peter Wilcoxon, Stefan Ballmer, Walter Freeman, Joseph Paulsen, and Mitch Soderberg, for taking the time and effort to review my dissertation.

Thank you to all of my amazing peers and mentors in LIGO, who made me feel like a part of something bigger. And to all those at Syracuse University, who made me feel at home.

And finally, to my parents and sisters, for always believing in me and for all your support on my journey.

Contents

List of Tables	ix
List of Figures	xii
Preface	xiv
1 Gravitational Waves: Theory, Sources, and Detectors	1
1.1 Gravitational Waves in General Relativity	3
1.2 Gravitational Wave Sources	5
1.2.1 Compact Binary Coalescence	7
1.2.2 Gravitational Wave Transients	13
1.3 Gravitational Wave Interferometers	14
1.3.1 Interferometer Design	14
1.3.2 Fundamental Sources of Noise	24
2 Searches for Gravitational Waves	32
2.1 Searches for Gravitational Wave Bursts	33
2.2 Matched Filter Searches	34
2.3 PyCBC Pipeline	36
2.3.1 PyCBC Detection Statistic	37
2.3.2 Background Estimation	43
2.4 Limiting Sources of Transient Noise	47
2.4.1 Classes of Noise Artifacts	47
2.4.2 Detector Characterization Tools	56

3	O2 Data Quality Efforts	57
3.1	Data Quality Products	58
3.1.1	Data Quality Flags	58
3.1.2	Gating	62
3.2	Instrumental Artifact Investigations	64
3.2.1	Power Mains Transients	66
3.2.2	Chiller Switch Transients	66
3.2.3	Optical Lever Transients	68
3.3	Effect of Data Quality Work on the Searches	71
3.3.1	Data Quality flags	73
3.3.2	Gating	76
4	Impact of Gravity Spy Glitches on PyCBC	78
4.1	PyCBC Trigger Set	79
4.2	Gravity Spy Classification	80
4.3	How Different Kinds of Glitches Mimic Traits of Signals	82
4.4	Probability of Creating Triggers	85
4.5	A Method to Utilize Glitch Classification in Significance Estimates	90
4.6	Future Applications	95
5	Linear Noise Subtraction	97
5.1	Subtraction Pipeline Overview	98
5.1.1	Measurement of Transfer Functions	98
5.1.2	Calculation of Coupled Noise	100
5.1.3	Workflow Implementation	101
5.2	Noise sources	103
5.2.1	Jitter Noise	103
5.2.2	Line Artifacts	104
5.3	Diagnostics	105
5.3.1	Sensor Safety	105
5.3.2	Recovery of Simulated Compact Binary Coalescence Signals	105
5.3.3	Simulated Noise Tests	106
5.3.4	Effect on Calibration	108
5.3.5	Impact of Nonstationary Data	108

5.4	Results	109
5.4.1	The O2 Data Set	109
5.4.2	Effects on the Noise Curve	110
5.4.3	Effect on Astrophysical Analyses	110
5.5	Application in future observing runs	114
6	O2 Catalog of Gravitational Wave Signals	116
6.1	PyCBC Search of noise-subtracted data	117
6.1.1	Classification of Significant Triggers	117
6.1.2	Identified Gravitational Wave Signals	119
6.1.3	Identified Marginal Triggers from PyCBC	119
6.1.4	Identified Marginal Triggers from GstLAL	122
6.2	Impact of Noise Subtraction on Search Result	122
6.3	Impact of Noise Subtraction on Parameter Estimation	128
6.4	Impact of Noise Transients on Search Result	135
7	Future Prospects	144
	Bibliography	146

List of Tables

1	Glitch class overview	49
2	Total deadtime from flags during O2	61
3	Properties of the marginal triggers identified from the two matched-filter CBC searches	123
4	GWTC-1 Triggers before and after subtraction	130
5	Recovered properties of GW170104 before and after subtraction	133
6	Marginal triggers from the two matched-filter CBC searches along with a data quality statement for each trigger	136

List of Figures

1	Effect of the different polarizations of gravitational waves on test particles	6
2	Graphic of the observed stellar remnants as of December 2018	9
3	Observed gravitational wave signal from GW150914	11
4	Basic Michelson Interferometer design	15
5	The basic structure of cavities inside of Advanced LIGO	19
6	A diagram of a Fabry-Perot cavity	20
7	Detailed diagram of aLIGO optics	23
8	Fundamental noise sources for Advanced LIGO	25
9	Noise budget for LIGO-Hanford at the start of O2	30
10	Spectrogram of GW150914	34
11	Example of cross correlation	35
12	Overlay of GW170814 template and data	36
13	Template bank used by PyCBC during O2	38
14	Example of background fitting	42
15	A visualization of how timeslides are created	44
16	Background estimation for PyCBC Live	46
17	Daily variation in the trigger rate	48
18	Variation of timeslide background distribution	49
19	Omega scans of common glitch classes present in Advanced LIGO data	50
20	ASD of tomte glitch	52
21	Overlay of a CBC trigger with a blip and whistle glitch	54
22	Diagram of how scattered light occurs	55
23	Omicron Triggers for a typical day	59
24	Loud glitching producing Omicron and PyCBC Live Triggers	60
25	Stability of the data during a category 1 flag	63

26	Diagram of how a gate is applied	65
27	PEM array during O2	67
28	Power mains spikes	68
29	Comparison of magnetometer noise with glitching	69
30	Spectrogram of optical lever	70
31	Omicron triggers in coincidence with oplev glitching	71
32	Omega scan of an IMBH candidate and an oplev glitch	72
33	Rates of single detector triggers	74
34	The effect of data quality on the VT of PyCBC	75
35	Recovered newSNR for a simulated signal near a gate	77
36	Omega scans of four problematic Gravity Spy glitch classes	81
37	An overlay of a timeseries of a representative CBC template and de- tector data around a blip glitch	83
38	Histograms of the maximum sine-Gaussian SNR recorded for triggers found in different parts of the PyCBC template bank parameter space during time periods corrupted by each glitch class	84
39	Probability of producing a trigger above sine-Gaussian SNR of 7.0 in specific regions of the template bank parameter space for each glitch class	87
40	Probability of producing a trigger above a reduced SNR of 7.0 for blip and koi fish classes	88
41	A comparison of the rate of triggers based on all times in the analysis and only times coincident within 2 seconds of blip	92
42	Recovered IFAR for a set of injections compared against the back- ground from all time during an analysis versus the background during only blip times	94
43	Coherence between witness sensors and gravitational wave strain for three types of instrumental noise subtracted in O2 at LIGO-Hanford: beam jitter, power mains, and calibration lines	100
44	Visualization of how transfer function measurements are tiled in time	102
45	Recovery of a hardware injection before and after cleaning	105
46	Recovered network ranking statistic for simulated gravitational wave signals before and after applying noise subtraction	107

47	Amplitude spectral density of interferometer gravitational wave strain data before and after noise subtraction from a representative day of data during O2	111
48	Inspirational range of the LIGO-Hanford detector over the course of O2 before and after noise subtraction.	112
49	The ratio of volume-time the PyCBC search was sensitive to during O2 before and after noise subtraction binned by chirp mass	113
50	Amplitude spectral density of interferometer gravitational wave strain data before and after noise subtraction, from a representative day of data during an engineering run	114
51	Range of LIGO-Livingston, LIGO-Hanford, and Virgo during O2	118
52	Waveform reconstructions of the 10 BBH signals	120
53	PyCBC search result with signal and background model overlaid	121
54	Spectrograms of GW170818 with the inspiral track overlaid	127
55	Spectrograms of trigger 170616 with the inspiral track overlaid	129
56	SNR Timeseries of GW170818 at LIGO-Hanford.	130
57	SNR Timeseries of trigger 170616 at LIGO-Hanford.	131
58	SNR Timeseries of GW170817 at LIGO-Hanford.	131
59	SNR Timeseries of O2 BBH signals at LIGO-Hanford.	132
60	Comparison of measured component masses for GW170104 before and after subtraction	134
61	Spectrograms of triggers impacted by scattered light	137
62	Spectrogram of trigger 170208	139
63	Spectrograms of triggers impacted by scratchy glitches	140
64	Probability of PyCBC triggers in O2 during scratchy glitches, with catalog triggers marked	141
65	Spectrograms of triggers impacted by short-duration, high amplitude glitches	143

Science is fun. Science is curiosity. We all have natural curiosity. Science is a process of investigating. It's posing questions and coming up with a method. It's delving in.

—Sally Ride

Preface

The work presented in this thesis stems from my participation in the LIGO Scientific Collaboration (LSC). This work does not reflect the scientific opinion of the LSC and it was not reviewed by the collaboration.

Chapter 1

Gravitational Waves: Theory, Sources, and Detectors

In 1915, Albert Einstein published the first paper on general relativity [1], revolutionizing our understanding of gravity. In addition to explaining previously unsolved problems such as the procession of the perihelion of Mercury, the theory made a wide variety of predictions that had yet to be observed at the time of writing. One of the first such predictions to be tested was the deflection of light by gravity.

One of the central tenets of general relativity was that masses deformed spacetime, which then in turn influenced other masses. In the words of John Wheeler, “Space tells matter how to move, matter tells space how to curve” [2]. However, spacetime curvature also affects the motion of particles with no mass such as photons. In order for this effect to be large enough to be observed at the time of Einstein’s prediction, photons needed to pass very close to an extremely massive object. The total solar eclipse of May 29, 1919 presented the perfect opportunity for both conditions to be met. With the light from the sun blocked out, detailed measurements were taken by Arthur Eddington of the deflection angle of photons from starlight passing very near the sun, which agreed with the expected value [3]. For the first time, Einstein’s theory of general relativity was able to make a precise prediction that can explain a new phenomenon in our universe.

In the one hundred years since this initial success, Einstein’s theory of general relativity has been subjected to more rigorous scrutiny; the theory has correctly predicted effects such as the size of black holes and their related photon rings [4], frame

dragging [5], gravitational redshift [6], and the strong equivalence principle [7]. One of the most contentious predictions of general relativity was gravitational waves. After the initial wave solution published by Einstein in 1916 [8], the existence of gravitational waves was deeply debated. It was not until the mid 1950s that the scientific community converged on support of the concept. In 1974, the discovery of a binary system consisting of two neutron stars by Hulse and Taylor [9] provided the first opportunity for this theory to be tested. As gravitational waves are emitted by a system, they carry away energy that causes the orbit to slowly decay. Careful measurements of the orbital phase of the system showed that the period was decreasing exactly at the rate predicted due to gravitational wave emission [10]. This provided the first indirect evidence for gravitational waves.

What remained elusive was the direct detection of gravitational waves. The predicted low amplitude of these waves initially led Einstein to assert that direct detection may never be possible. However, advancements in precision measurements culminating in the development of the Laser Interferometer Gravitational Wave Observatory (LIGO) finally made such a discovery within the realm of possibility [11]. On September 14, 2015, the twin LIGO detectors made coincident observations of gravitational waves from the merger of two compact objects, each about thirty times the mass of the sun [12]. This signal, GW150914, was not only the first gravitational-wave detection, but also the first evidence for a population of stellar mass black holes above $20 M_{\odot}$ [12].

Since this initial discovery, ten additional detections of gravitational waves [13] have been made by LIGO along with the Virgo interferometer [14]. Of the eleven total, all but one were from binary black hole mergers. The remaining detection, GW170817, was from the merger of binary neutron stars; this led to one of the largest astronomical follow up campaigns in history, marking the beginning of multi-messenger astronomy with gravitational and electromagnetic waves [15, 16].

In this dissertation, I will outline the theory and techniques that are critical for gravitational wave detections with interferometry. I will then discuss sources of broadband and transient instrumental noise that limit the sensitivity of Advanced LIGO and searches for signals from compact binary coalescences (CBCs). I will evaluate mitigation strategies used in O2 to increase the sensitivity of aLIGO to gravitational waves, including identification of time periods corrupted by short duration transient

noise, subtraction of correlated noise over long durations, and studies to differentiate transient noise from genuine astrophysical signals. Finally I will present the most recent results from LIGO-Virgo observations, and how these mitigation methods have allowed for the detection of new gravitational-wave events and improved analyses of previously identified events. The focus will be on the second observing run (O2) of Advanced LIGO, which ran from November 2016 to August 2017.

The remainder of Chapter 1 details how gravitational waves arise in general relativity, the astrophysical sources of such waves, and the basic design of the interferometers used to detect them. Chapter 2 outlines methods used to identify gravitational waves in interferometer data, along with an overview of the limiting sources of noise for these analyses. Chapter 3 explains how sources of transient noise are identified in the data and how instrumental improvements are tested to alleviate such problems. Chapter 4 further explores the relationship between common sources of transient noise and the algorithms used in gravitational wave detection, explicitly discussing problems of both false negatives and false positives in the analysis. Chapter 5 discusses how the overall sensitivity of the detector was improved with by the subtraction of broadband noise sources. Finally, Chapter 6 details the final search results for the second observing run, and how the mitigation methods discussed in this dissertation affect the astrophysical results from the O2 dataset.

1.1 Gravitational Waves in General Relativity

Gravitational waves are one of the central features of Einstein’s theory of general relativity. Gravitational waves can be constructed in general relativity by starting with a spacetime metric in a vacuum and adding a metric perturbation that obeys a few basic assumptions. In this section, I will work through the assumptions that are made to generate gravitational waves, and how gravitational waves appear to a stationary observer¹.

To see how gravitational waves can be constructed, we want to develop our spacetime metric for gravitational waves, $g_{\mu\nu}$. To produce this spacetime metric, we begin with the standard Minkowski metric for flat space, $\eta_{\mu\nu}$, that describes spacetime in the absence of matter:

¹This section loosely follows a derivation found in Flanagan [17]

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \quad (1.1)$$

We can now add a perturbation to this metric given by $h_{\mu\nu}$ such that $\|h_{\mu\nu}\| \ll 1$. This means our chosen metric is

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} . \quad (1.2)$$

We will now focus on this perturbation term, $h_{\mu\nu}$. We will make three additional assumptions about how this perturbation is constructed:

$$\begin{aligned} \text{Spatial:} \quad & h_{tt} = h_{ti} = 0 \\ \text{Traceless:} \quad & h = h^i_i = 0 . \\ \text{Transverse:} \quad & \partial_i h_{ij} = 0 \end{aligned} \quad (1.3)$$

These conditions are referred to as the ‘‘Transverse-traceless gauge’’. In order to fully satisfy these assumptions, this perturbation needs to obey the differential equation

$$\square h_{\mu\nu} = \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) h_{\mu\nu} = 0 . \quad (1.4)$$

It is in this way that we have shown that with these gauge choices, the metric perturbation must follow a classical wave equation. Furthermore, this metric is described by only two values,

$$\begin{aligned} h_+ \equiv h_{xx} &= -h_{yy} \\ h_\times \equiv h_{xy} &= h_{yx} \end{aligned} . \quad (1.5)$$

This means that the full metric perturbation is

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \quad (1.6)$$

These two variables, h_+ and h_\times , define the two polarizations of gravitational waves, called the plus and cross polarizations, respectively.

In order to consider how this perturbation would impact test particles, we consider a gravitational wave with only one polarization, h_+ , for simplicity. In this scenario, if the gravitational wave was propagating in the \hat{z} direction, the metric perturbation is

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & 0 & 0 \\ 0 & 0 & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (1.7)$$

Therefore, this perturbation would produce an equal but opposite effect in the \hat{x} and \hat{y} dimensions. Namely, there would occur a stretching of spacetime in the \hat{x} direction, with compression in the \hat{y} direction. As the wave passes, oscillations would occur, with each dimension successively being stretched and compressed. An example of this behavior on a ring of test masses can be seen in Figure 1.

The amplitude of each term, $||h_{\mu\nu}||$, is referred to as the *strain* produced by the gravitational wave. Strain is the fractional change in distance in a particular direction due to the stretching of space caused by gravitational waves. If we consider our ring of test particles from Figure 1, let the initial diameter, when no gravitational waves are present, to be L . Then when a gravitational wave with strain amplitude h passes through, the ring will be compressed and stretched by a factor of h , so that the ring width and height will now be $L(1 - h)$ and $L(1 + h)$, respectively.

1.2 Gravitational Wave Sources

With the knowledge that gravitational waves are a well described component of general relativity, the next step is to understand the known emission sources for these waves, along with the strain amplitude that can be observed on Earth. In general, the optimal scenario for gravitational wave emission involves highly massive object moving quickly in an asymmetric manner. Specifically, a variation in the quadrupole moment of a mass distribution leads to gravitational wave emission. Thankfully, a number of sources that fit this description have been identified.

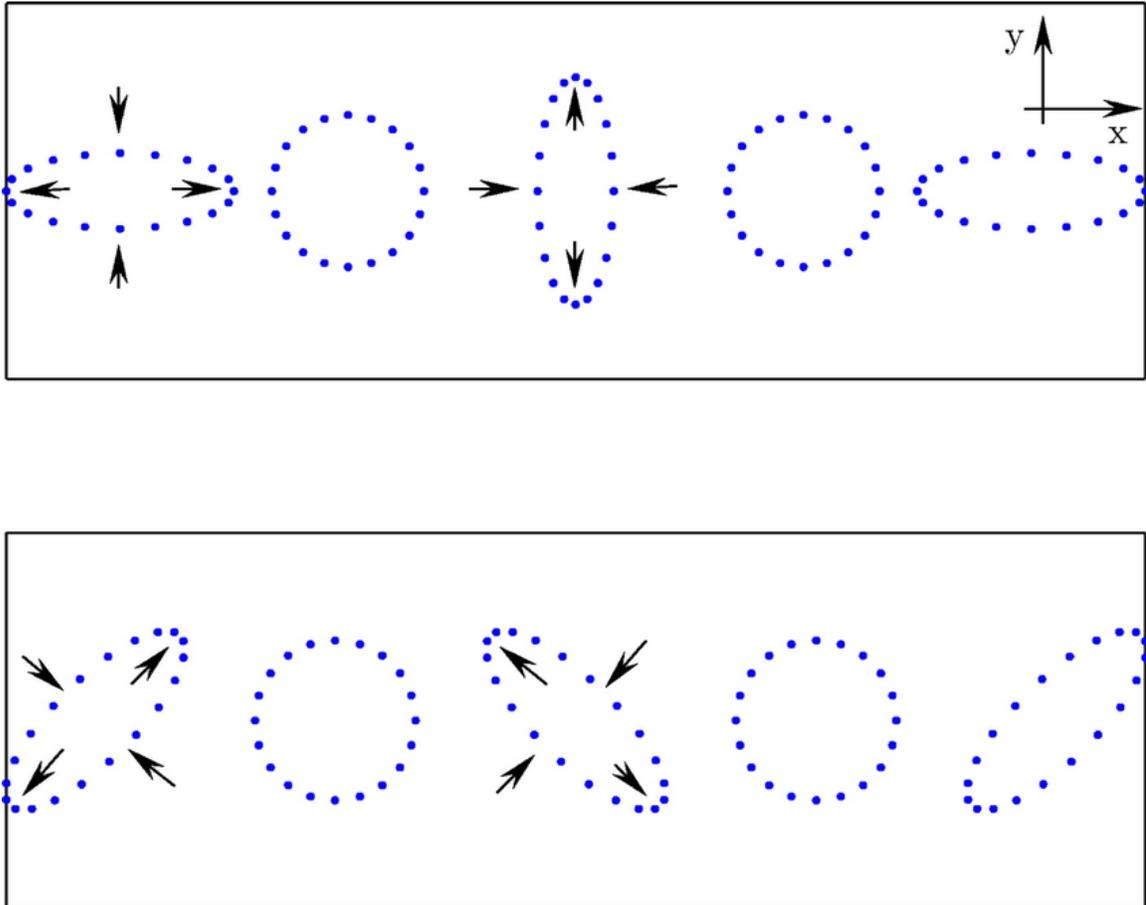


Figure 1: Top: The response of a ring of test particles to plus polarized gravitational waves. Note the stretching and compression of orthogonal axes. Bottom: The response of a ring of test particles to cross polarized gravitational waves. Reproduced from [18].

1.2.1 Compact Binary Coalescence

The most straightforward way to vary the quadrupole moment of a system is to have two objects orbiting around each other². The expression for the observable strain amplitude from gravitational waves is

$$h_{\mu\nu} = \frac{2G}{Rc^4} \ddot{I}_{\mu\nu} . \quad (1.8)$$

To see how strong gravitational waves can be, we derive $\ddot{I}_{\mu\nu}$ for a system that contains two objects of mass m orbiting at a radius of r and frequency f_{orb} .

To derive $\ddot{I}_{\mu\nu}$, we start with the expression for a generic quadrupole in general relativity:

$$I_{\mu\nu} = \int dU (x_\mu x_\nu - \frac{1}{3} \delta_{\mu\nu} r^2) \rho(r) . \quad (1.9)$$

If we assume that the gravitational wave is polarized in the plus direction, we can focus on a single term in this tensor, I_{xx} . We write this term down and then differentiate twice with respect to time:

$$\begin{aligned} I_{xx} &= 2Mr_0^2 (\cos^2 2\pi f_{orb} t - \frac{1}{3}) \\ \dot{I}_{xx} &= 4Mr_0^2 (2\pi f_{orb}) (\cos 2\pi f_{orb} t \sin 2\pi f_{orb} t) . \\ \ddot{I}_{xx} &= 16\pi^2 Mr_0^2 f_{orb}^2 (\cos 4\pi f_{orb} t) \end{aligned} \quad (1.10)$$

Therefore the total strain in the \hat{x} direction is

$$h_{xx} = \frac{32\pi^2 G}{Rc^4} Mr_0^2 f_{orb}^2 (\cos 4\pi f_{orb} t) . \quad (1.11)$$

Solving for h_{yy} will give a result with identical amplitude. Therefore we find that the expected strain is

$$|h| = \frac{32G}{Rc^4} 2Mr_0^2 f_{orb}^2 . \quad (1.12)$$

While we do expect some deviations from Keplerian orbits due to general relativity, this approximation will hold to first order. With that in mind, we can calculate the expected orbital frequency from the masses and radius of the orbit, given by

²The derivations in this section follows from Saulson [19], which in turn is sourced from Misner, Thorne, and Wheeler [2].

$$f_{orb}^2 = \frac{2Gm}{32\pi^2 r^3} . \quad (1.13)$$

A particularly useful quantity when discussing gravitational waves from compact binaries is chirp mass, given by [20]

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} . \quad (1.14)$$

The chirp mass is one of the best measured parameters from a gravitational wave signal [20] because it gives, to leading order, the rate of evolution of the orbital frequency of the binary.

Another important value is the minimum radius of an object in general relativity, known as the Schwarzschild radius [21],

$$R_s = \frac{2GM}{c^2} . \quad (1.15)$$

The Schwarzschild radius is the location of the event horizon as predicted by general relativity for a black hole with mass of M . This radius sets the minimum value for the compactness of an astrophysical object, and therefore the minimum separation distance of orbiting objects that are capable of emitting gravitational waves.

In order to achieve the highest measurable strain amplitude for a binary system, we need extremely dense objects moving quickly. The best identified targets for gravitational waves are stellar remnants such as black holes, neutron stars, and (for gravitational waves at lower frequencies) white dwarfs. While it is possible to have a system with different classes of object (for example a system with one neutron star and one black hole) we will assume in this section for simplicity that objects are paired with like objects. The spread of masses for the population of measured black holes and neutron stars from both gravitational waves and electromagnetic observations can be seen in Figure 2. A system of two binary objects will have the most rapid acceleration of each component object just as the components are about to collide and merge. Therefore, the largest amplitude gravitational waves are related to compact binary coalescence (CBC) signals.

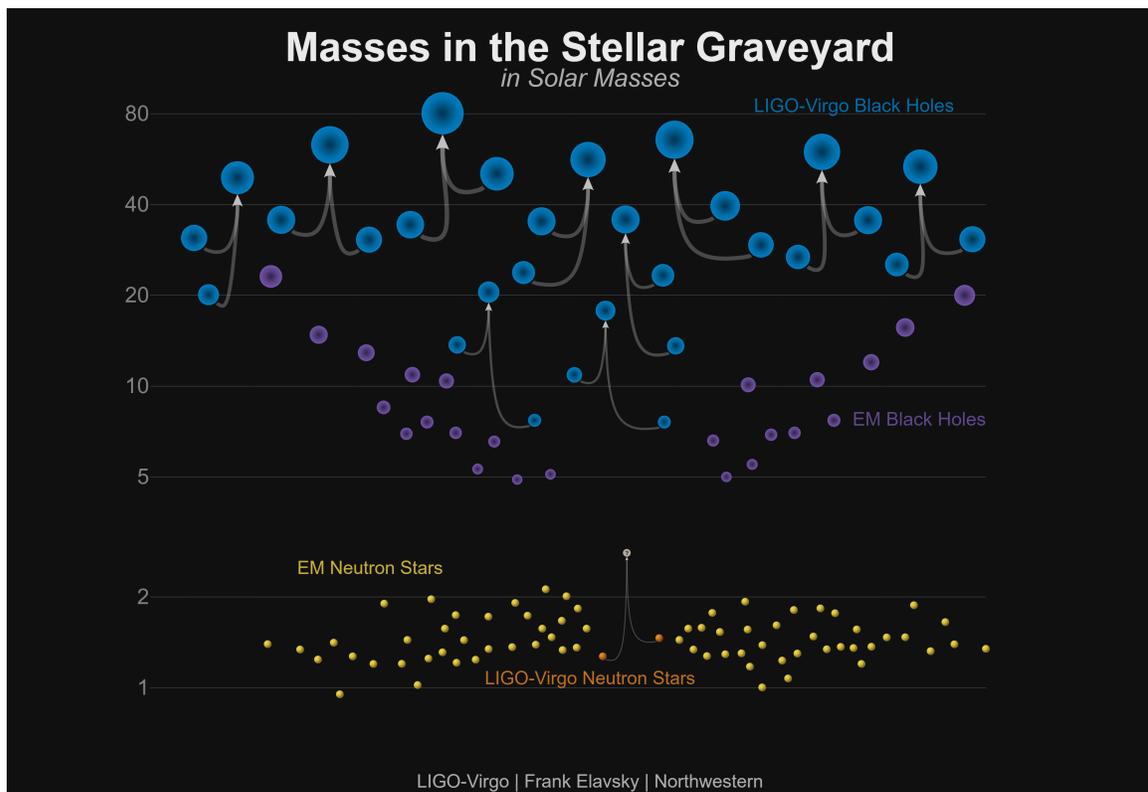


Figure 2: Graphic representation of the catalog of gravitational wave event observed in Advanced LIGO's first and second observing runs, alongside black holes and neutron stars with measured masses from electromagnetic observations. Arrows connecting pairs of black holes represent binary mergers observed by LIGO and Virgo. Reproduced from [22].

Binary Black Holes

In order to have the largest spacetime distortion, it is reasonable to first consider the most dense objects possible moving at the fastest speeds. Black holes represent the most compact objects in the universe, and are the expected result of a main sequence star over $25 M_{\odot}$ [23]. While binary black holes are not visible electromagnetically, they are the most common type of system observed via gravitational waves due to their small radii and large masses. We can calculate the expected strain amplitude of a binary black hole system using the parameters of the first BBH observed by LIGO, GW150914 [12]. We therefore set each black hole to have a mass of $35M_{\odot}$ and the separation distance of 350 km, (a few multiples of the black hole's Schwarzschild radius) we find that the orbital frequency would be 75 Hz, and hence a gravitational wave frequency of 150 Hz. This approximates one of the final orbits before merger, when the black holes would be at their closest point. We place our binary system at 400 Mpc, again matching the measured parameters of GW150914. At this range, we find using Equation 1.15 that the maximal observable strain amplitude would be

$$|h| \approx 2 \times 10^{-21} . \quad (1.16)$$

This simple approximation actually estimates the overall amplitude of gravitational waves quite well! Comparing this expected strain to the observed value for GW150914, $|h| \approx 1 \times 10^{-21}$, this is within a factor of 2. The observed strain for this event, showing the increase in frequency and amplitude as the binary system inspirals closer and closer before merger, can be seen in Figure 3.

Binary Neutron Stars

The next class of object we can consider is neutron stars, aptly named as they can be approximated as consisting only of neutrons. These objects are also stellar remnants, and are held up primarily by neutron degeneracy pressure. A neutron star is the likely final result for a main sequence star that began with the mass of 10 - 25 M_{\odot} [23]. One clear advantage of a system with binary neutron stars, unlike binary black holes, had been previously observed with electromagnetic waves before their discovery with gravitational-waves, the most notable of which is the Hulse-Taylor system. From observations of the double neutron star population in the Milky Way, merger rates

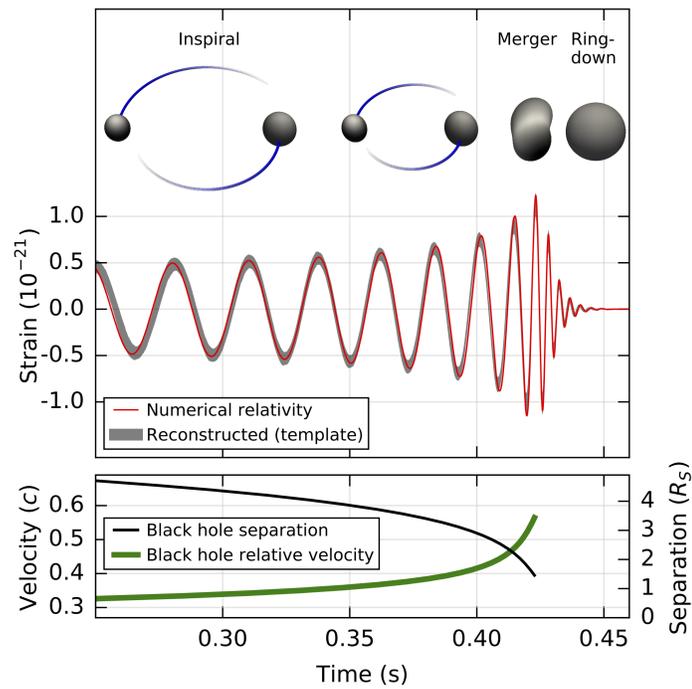


Figure 3: The observed gravitational wave signal from GW150914, along with the separation distance and orbital velocity as the binary system approaches merger. Reproduced from [12]

have been estimated to be ≈ 40 per million years in our galaxy [24].

Since these objects have been observed via electromagnetic methods, their parameters have been measured. The expected mass of a neutron star based on Milky Way observations, as well as predictions of the point at which a star would collapse into a neutron star [25], is $1.4M_{\odot}$ [26]. The radius of a neutron star is less well measured, and is dependent on the equation of state for matter at densities above those found in the nucleus of an atom. However, a reasonable assumption about their radius is 10 km.³

Taking these two values for radius and mass, we can then calculate the orbital frequency for when the two neutron stars touch. We find that f_{orb} is 1085 Hz, meaning f_{GW} is 2170 Hz. Using these values, and the measured distance to GW170817, ≈ 40 Mpc [29], we find that the maximum strain amplitude for a this binary neutron star merger is

$$|h| \approx 1 \times 10^{-21} . \quad (1.17)$$

An additional feature of neutron star mergers that makes them an interesting target is the fact that since neutron stars (unlike black holes) are comprised of matter, it is expected that electromagnetic emission will occur alongside the merger. The first neutron star merger was observed in coincidence with a prompt gamma ray burst [30] as well as emission in other wavelengths that lasted for weeks after the event [16].

White Dwarfs

The final stellar remnant we will consider is white dwarfs. These are the remnants for main sequence stars less than $10M_{\odot}$, including the Sun [23]. A white dwarf is what remains after the star has shed its hydrogen and helium shell, leaving behind the core of the star. Primarily held up by electron degeneracy pressure, white dwarfs are compact as compared to a main sequence star, but much larger than neutron stars or black holes. As the primary pressure comes from electron degeneracy, white dwarfs decrease in radius as their mass increases, until a critical point is reached at $1.4M_{\odot}$, called the Chandrasekhar mass. At this point the object implodes, causing a Type 1b supernova, and creating a neutron star.

³This value of the radius is consistent with the measured values from GW170817. [27, 28]

Using the parameters of the nearest white dwarf with a measured mass, Sirius B [31], we set each white dwarf in our binary to have a mass of $1.0 M_{\odot}$, which leads to a radius of approximately 7000 km (around the size of the Earth). Following the procedure for neutron stars, we estimate the frequency of orbit for when the white dwarfs touch. We find an orbital frequency of 3.5×10^{-5} Hz, and therefore a gravitational wave frequency of 7.0×10^{-5} Hz, much lower than for black holes or neutron stars.

Furthermore, when we use these inputs for even an maximally close merger at the center of the Milky Way, only 7000 pc, we find a maximum strain amplitude of

$$|h| \approx 1 \times 10^{-27} . \quad (1.18)$$

As the expected amplitude of gravitational wave strain observable from these systems in the most optimistic scenario is much lower than for neutron star or black hole binaries, in addition to being at an extremely different frequency band, white dwarf mergers are not a promising candidate for detection by ground-based gravitational-wave observatories. These systems may be observable by space-based observatories, such as LISA [32], that will be sensitive in this frequency band and able to observe white dwarf systems for multiple years before merger in order to better separate the signal from background noise.

1.2.2 Gravitational Wave Transients

In addition to CBC signals, there are other proposed sources of short lived gravitational waves. One of the most understood sources is supernova [33, 34, 35]. However there are other sources such as bursts from magnetars [36] or bremsstrahlung radiation from black holes in dense clusters passing close to each other [37], that may be possible.

Supernovae, particularly Type II supernova may be good candidates for a gravitational wave transient. During the collapse of the star, it is possible that a proto-neutron star will be formed with high spin and high eccentricity that is capable of emitting measurable gravitational waves. Furthermore, matter shockwaves from the supernova are expected to not be isotropic, which creates an additional mechanism for gravitational-wave emission. While the amplitude of such an event is highly model

dependent, an estimate for the strain amplitude at 100 Hz from a supernova at 10 kpc is [35]

$$|h| \approx 1 \times 10^{-22} . \quad (1.19)$$

While this amplitude is comparable to the expected strain from a compact binary merger, galactic supernova are extremely rare, with an expectation of only ≈ 3 per century in the Milky Way [38].

Additional sources of gravitational waves, such continuous waves from pulsars [39] or a gravitational-wave stochastic background [40], are predicted to exist, but are detected with highly orthogonal data analysis methods to gravitational wave transients.

1.3 Gravitational Wave Interferometers

With our knowledge of the loudest potential sources of gravitational waves, the next reasonable step to examine the methods that allow the detection of a gravitational-wave strain as small as 10^{-21} .

The first design for a gravitational wave detector was a resonant bar detector designed by Joseph Weber [41]. While it did not meet the sensitivity requirements that we are aiming for (and hence did not detect gravitational waves) it was the first attempt to directly measure these waves. In 1972, Rainer Weiss proposed [42] that a detector based on the design of a Michelson interferometer [43] would be highly effective.

The target sensitivity was finally reached with the construction of the Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) [11] in 2015. The rest of this section will discuss the overall design of interferometers and the specific design components of aLIGO that allowed it to reach its current sensitivity, along with the limiting sources of background noise.

1.3.1 Interferometer Design

A Michelson interferometer is designed to be highly sensitive to changes in the path length of two arms that are orthogonal to each other. A diagram of this basic setup can be seen in Figure 4. A light source is sent into a beam splitter, which then sends

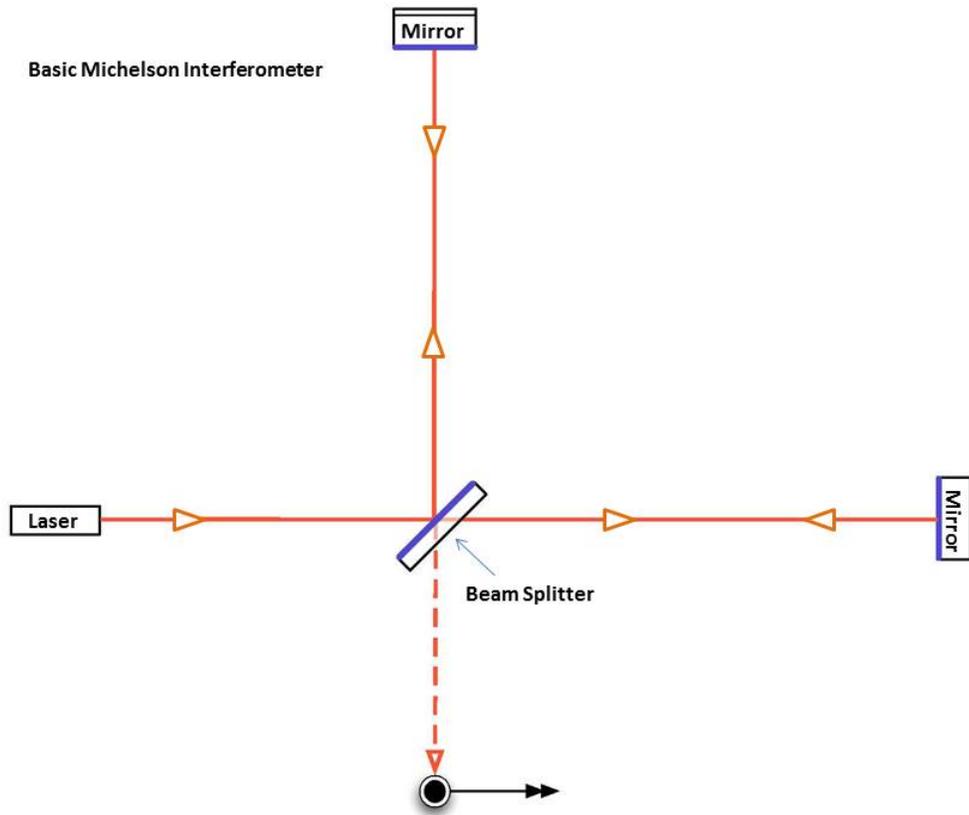


Figure 4: Basic Michelson Interferometer design showing how light is split into two perpendicular arms. Reproduced from [44]

the light half down one arm and half the other. At the end of each arm, there is a mirror that reflects the light back to the beam splitter. At the beam splitter, the light from each arm is split again, such that half of the light is sent back towards to input, while the other half is sent to the output.

When the light is recombined at the output port, either constructive or destructive interference occurs based on the path distance along each arm. If the path length differs by an integer number of wavelengths, constructive interference occurs, while a difference of a half wavelength results in destructive interference. It is in this way that the interferometer can sense slight variations in relative length. In fact, we can model the behavior of the interferometer from first principles using the equation of a electromagnetic wave.

We first assume the input to the interferometer is⁴

$$E_{in} = E_0 e^{i(2\pi ft - kx)} . \quad (1.20)$$

This input field then encounters a 50/50 beamsplitter with reflection and transmission amplitude coefficients of

$$\begin{aligned} r &= 1/\sqrt{2} \\ t &= i/\sqrt{2} . \end{aligned} \quad (1.21)$$

Therefore the light transmitted down the \hat{x} and \hat{y} axes is described as

$$\begin{aligned} E_x &= \frac{i}{\sqrt{2}} E_0 e^{i(2\pi ft - kx)} \\ E_y &= \frac{1}{\sqrt{2}} E_0 e^{i(2\pi ft - ky)} . \end{aligned} \quad (1.22)$$

After travelling down the length of each arm, and reflecting off of the end mirrors, each field picks up a factor of -1 . Reflecting back to the beam splitter and recombining, the final output field is

$$\begin{aligned} E_{out} &= \frac{i}{\sqrt{2}} E_0 e^{i(2\pi ft - kL_x)} + \frac{i}{\sqrt{2}} E_0 e^{i(2\pi ft - kL_y)} \\ &= i E_0 e^{i(2\pi ft - k[L_x + L_y])} \cos k(L_x - L_y) . \end{aligned} \quad (1.23)$$

However, the measurable quantity of relevance in this case is not the field amplitude, but the power. Since this is simply the field amplitude squared, we find the the total power output is

$$\begin{aligned} P_{out} &= E_0^2 \cos^2 k(L_x - L_y) \\ &= P_{in} \cos^2 k(L_x - L_y) , \\ &= P_{in}/2 (1 + \cos 2k(\Delta L)) \end{aligned} \quad (1.24)$$

where ΔL is the difference between the two arm lengths. It is the $\cos 2k(\Delta L)$ term that ultimately is sensitive to the total relative arm distance. As light travels at a fixed speed down the arms, this difference in length causes the light to return to the beamsplitter with a small phase difference in our final result.

⁴This derivation follows the treatment from Saulson [19], which is in turn based on Haus [45].

The next step is to now examine how a passing gravitational wave will effect the interferometer. In the Transverse-Traceless gauge, the spacetime coordinates of the mirrors are not affected by the gravitational-wave perturbation, allowing us to focus solely on how light in the interferometer is affected by a passing gravitational wave. If we consider the world line that the light takes through the arms, it must be $ds^2 = 0$. When a passing wave perturbs the spacetime metric, this means that

$$ds^2 = (\eta_{\mu\nu} + h_{\mu\nu}) dx^\mu dx^\nu . \quad (1.25)$$

To simplify the calculation, we will assume that the period of the gravitational wave is much larger than the round trip time in the interferometer such that we can assume the metric perturbation is constant. With this assumption, we can construct the world line of light in the \hat{x} arm as follows:

$$ds^2 = -c^2 dt^2 + (1 + h_{11}) dx^2 . \quad (1.26)$$

To find the full light travel time, we can integrate:

$$\begin{aligned} \int_0^{\tau_{out}} dt &= \frac{1}{c} \int_0^L \sqrt{1 + h_{11}} dx \\ &\approx \frac{L}{c} + \frac{L}{2c} h_{11} . \end{aligned} \quad (1.27)$$

As the return trip takes an equal amount of time, the total round trip duration is

$$\tau_{rt,x} = \frac{2L}{c} + \frac{L}{c} h_{11} . \quad (1.28)$$

Similarly, the round trip duration for light in the y arm is

$$\tau_{rt,y} = \frac{2L}{c} + \frac{L}{c} h_{22} . \quad (1.29)$$

This means that the difference in the duration is

$$\Delta\tau = \frac{L}{c} (h_{11} - h_{22}) . \quad (1.30)$$

If we assume that the wave is only plus polarized, then $h = h_{11} = -h_{22}$. Therefore for a particular strain amplitude, h , the light travel time difference is

$$\Delta\tau = \frac{2L}{c} h . \quad (1.31)$$

This time difference will result in a phase difference at the output port, and create an observed power as was shown for a difference in length between the arms. In fact, if we consider that $\Delta L = Lh$, the coupling between the strain amplitude of gravitational wave and the effect on the interferometer becomes clear. Armed with the knowledge that our interferometer is sensitive to these waves, our next step is to design an optimal detector for the target sources. An example of the basic aLIGO design, including all of the relevant components that will be discussed in this chapter can be seen in Figure 5.

Fabry-Perot Cavity

In order to maximize the sensitivity of our instrument to gravitational waves, it is important to first note the difference between the quantity we have been using to define gravitational wave amplitude, strain, versus the value our basic interferometer measures, the difference in path length. Strain is defined as $\Delta L/L$ for some length L , while the total path length difference is $2\Delta L$. Therefore if we can increase L , we can increase ΔL for a fixed strain amplitude.

In a basic interferometer, this quantity L is simply the arm length. Therefore the most straightforward approach to increase sensitivity is to increase the arm length. However, there is often a limitation to this approach, not simply in terms of physics but of construction. We can, however, look back to an innovation by Michelson on his original design for inspiration. In the 1887 interferometer Michelson constructed he chose to increase the path length further as compared to his original design by folding the arms [43]. This involved adding additional mirrors at both ends of each arm such that the light was reflected back and forth multiple times, leading to the thus constructing a longer path length for the light in each arm.

This has been proposed in gravitational wave interferometer design (e.g. the Herriott Delay Line [47]), but was not utilized in either Initial LIGO or Advanced LIGO due to the complexity of designing the large reflective surface needed to complete the necessary reflections. Instead, aLIGO uses a different set up to increase the effective path length beyond the 4 km length of the arms, namely Fabry-Perot cavities in each of the arms.

In a Fabry-Perot cavity, the arm length is increased by multiple reflections, but instead of using distinct paths, each successive reflection overlaps. If the arm length

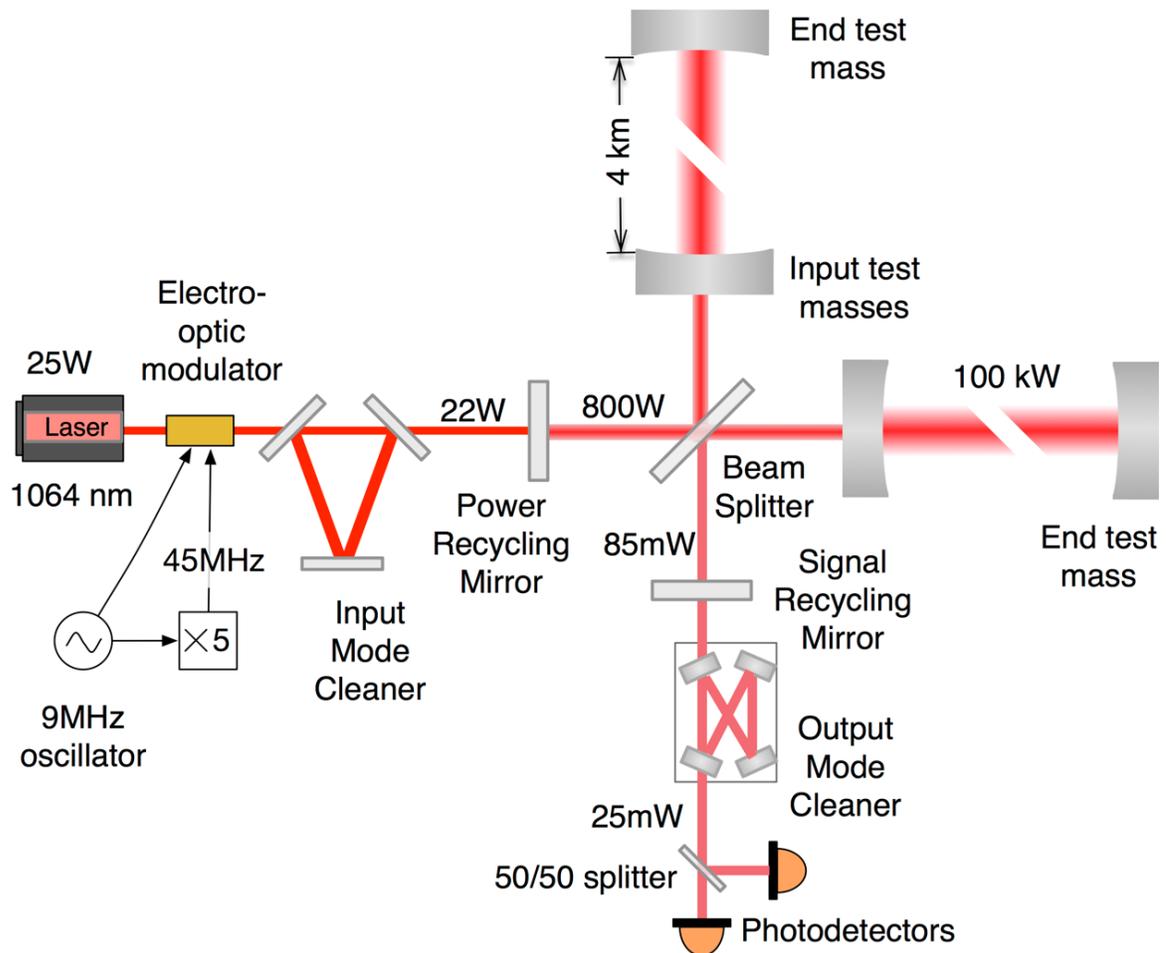


Figure 5: The basic structure of cavities inside of Advanced LIGO. Note the presence of Fabry-Perot cavities in each arm, the input mode cleaner, the power and signal recycling mirrors, and the output mode cleaner. Arm length is not to scale. Reproduced from [46]

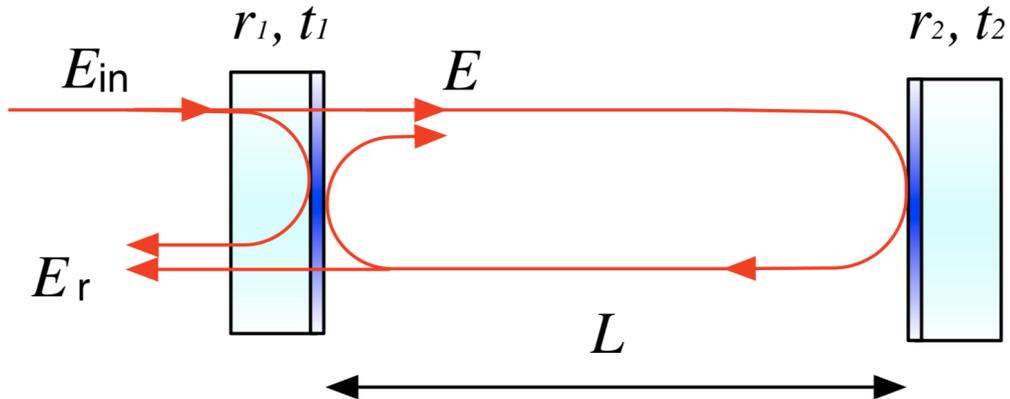


Figure 6: A diagram of a Fabry-Perot cavity showing both reflected and transmitted light. This cavity is constructed of two mirrors with reflective and transmission coefficients of $r_{1,2}$ and $t_{1,2}$. Inside of the cavity there will be a circulating field, given by E . Similar to an arm of a generic Michelson interferometer, length variations in the length of the cavity will add a phase shift to the reflected field, E_r . Reproduced from [48].

is precisely controlled, the light will resonate inside the cavity with minimal losses. A diagram of how light resonates in a Fabry-Perot cavity can be seen in Figure 6. If one solves the field equations for this system, one can see that variations in the length of the cavity will add a large phase shift to the light resonant in the cavity (given by E in Figure 6), as well as the reflected light (given by E_r in Figure 6). Therefore a Fabry-Perot interferometer can measure arm length changes in the same way that a generic Michelson interferometer can.

One method of measuring the equivalent number of round trips light takes in a cavity is through the finesse. This value is based on the reflection coefficient of the two mirrors in the cavity, r_1 and r_2 , and is [19]

$$\mathcal{F} = \frac{\pi\sqrt{r_1 r_2}}{1 - r_1 r_2} . \quad (1.32)$$

Each of the aLIGO arm cavities has a finesse \mathcal{F} of 450 [11], which, combined in an arm length of 4 km, results in a light storage time of $\tau_s \approx 2$ ms. This is equivalent to an ‘effective arm length’ of $\tau_s c/2 \approx 2800$ km.

One additional comment about the arm length is that there is a physical reasoning

that prevents the arm length from being extended indefinitely. In order for many of the assumptions made in this chapter to be valid, the light storage time must be less than the period of the gravitational wave. If this is not true, the consecutive maxima and minima from the same gravitational wave will cancel out the change in phase of the light in the arm. Hence a cavity is optimally sensitive to gravitational waves with periods of twice the light storage time of the cavity, and has reduced sensitivity to those with shorter periods. For aLIGO, this means the interferometer is most sensitive for wave with a period of 4 ms (a frequency of 250 Hz).

In addition to cavities present in the arms, aLIGO heavily utilizes cavities throughout the interferometer. A more detailed diagram of the interferometer showing some of these components can be found in Figure 7.

Power Recycling

In the conditions that no gravitational wave is present, almost no light is sent to the output (anti-symmetric) port, while almost all of the light is sent to the input (symmetric) port. In our basic interferometer design, this means that only a minimal amount of light actually reaches the photodetector making our measurement.

In order to get an additional benefit from the light that leaves the symmetric port, aLIGO utilizes a power recycling mirror placed at the symmetric port. This mirror then forms an additional cavity, resonant at the frequency of the light [49]. This reflects the light not sent to the output photodetector back into the interferometer, allowing the interferometer to increase the intensity of the light in arms.

Signal Recycling

Similar to the power recycling cavity, the signal recycling cavity is formed with an additional mirror at the anti-symmetric port that makes a cavity with the arms [49]. The arms themselves are such that the frequency of the laser, f_l , is resonant in the arm cavity. When a gravitational wave passes through the detector, the light in the arms is modulated by the gravitational wave frequency, creating side bands at $f_l \pm f_{sig}$. As the detector is by default in a configuration where minimal light reaches the anti-symmetric port, the path length change by a passing gravitational wave will cause constructive interference at the signal recycling mirror. Hence we can consider these

sidebands to be directed towards the signal recycling mirror.

Contrary to the name of this component, the signal recycling mirror in aLIGO is configured such that the sidebands related to a gravitational wave signal are transmitted through the mirror to the photodetector. This configuration is known as resonant sideband extraction [49]. By preferentially transmitting the gravitational wave signal, this effective finesse of the arm cavities to light modulated by gravitational waves is reduced. Recalling that the sensitive bandwidth of the interferometer is increased by decreasing the finesse of the arm cavities, this configuration allows for aLIGO to have a larger bandwidth than Initial LIGO, which did not include a signal recycling mirror [11]. Furthermore, as the frequency of the laser, f_l , is not transmitted through the signal recycling mirror, there is no change in the finesse of the cavity for this frequency [49]. This allows the arm cavities to maintain high finesse, and hence high power, while maximizing the bandwidth of gravitational waves aLIGO is sensitive to.

Mode Cleaners

In order to maximize the sensitivity of the interferometer to gravitational waves, the input laser beam must have a precise frequency and angular direction. To help stabilize both of these degrees of freedom in the beam, aLIGO uses an input mode cleaner, a triangular cavity that is just after the laser beam is sent into the interferometer. This cavity is used to stabilize the input beam and reduce noise due to angular and frequency fluctuations [50]. An additional concern is the presence of higher order modes of the laser beam in the interferometer. The input mode cleaner also functions to remove any higher order modes in the input beam [50].

An additional feature of aLIGO is the presence of radio frequency sidebands added to the main laser that are utilized by sensors to control the interferometer. While important, these sidebands are removed before the differential arm motion is measured. This is done with the output mode cleaner, placed just before the output photodetector. This cavity is only resonant for the zeroth order mode of the main beam, preventing both the added sidebands and higher order modes of the beam (arising from misalignment of the interferometer optics) from exiting the interferometer [51]. If these were not removed from the main beam, there would be excess power incident on the photodetector that was not related to the differential arm length.

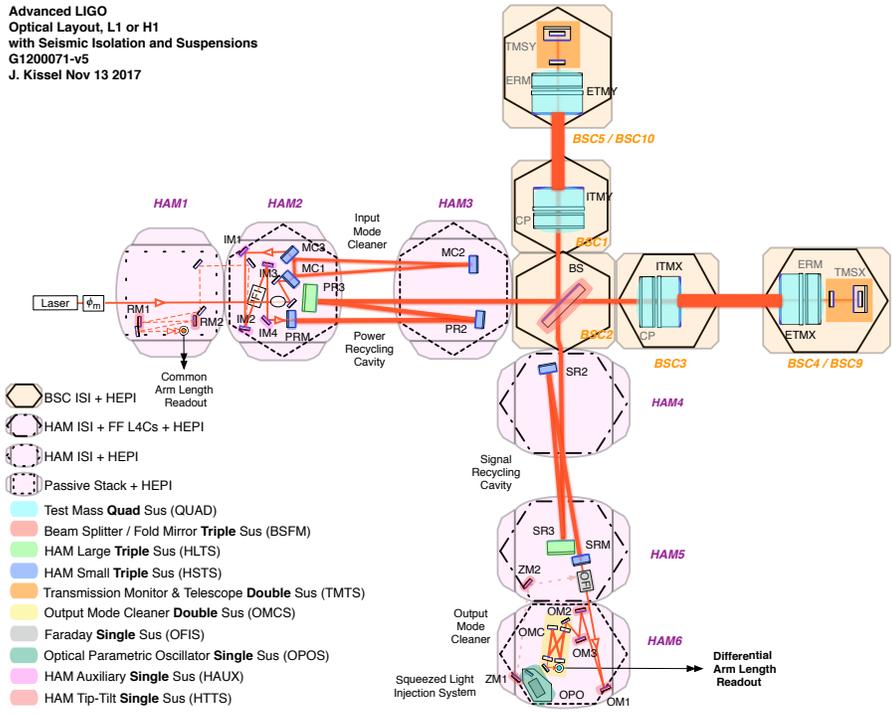


Figure 7: Detailed diagram of aLIGO optics showing the horizontal access modules (HAMs) that each optical component is placed in. Reproduced From LIGO Document number G1200017-v5.

1.3.2 Fundamental Sources of Noise

With the aLIGO design discussed, there are a number of sources of broadband noise in the detector that limit the sensitivity of the interferometer to astrophysical sources. Many of these sources are stationary, Gaussian processes, meaning that they do not vary in time and can be approximated as a random variable with a Gaussian distribution. However, many sources that are neither stationary or Gaussian can still limit the sensitivity. The overall level of these noise sources set the minimum strength of a detectable signal at a given frequency for aLIGO. A comparison of the dominant sources of broadband fundamental noise in aLIGO can be seen in Figure 8.

Additional sources of noise that occur on short time scales that can affect our ability to isolate astrophysical signals are also highly problematic, but do not affect the overall noise curve. An overview of these transient noise sources can be found in Chapter 2.

Quantum Noise

Currently the most limiting fundamental noise source at frequencies above 60 Hz is from quantum noise. This class of noise can be separated into two competing categories of noise that are dependent on the power of the laser used in the interferometer.

The first class of quantum noise is shot noise. Roughly speaking, this is the counting error due to the particle nature of light. Therefore, if the number of photons, and hence the overall power is increased, the fractional error in the power measurement will be lowered and the relative shot noise will decrease. More precisely, when converted into the error in the length measurement, this is [52, 53]

$$h_{shot}(f) = \frac{1}{NL} \sqrt{\frac{\hbar c \lambda}{2\pi P_{in}}} \quad (1.33)$$

with L the arm length, N the number of round trips a photon makes on average through the arm cavity, λ the wavelength of light, and P_{in} the input power into the interferometer.

Competing against shot noise is radiation pressure noise. As the power increases, size of pressure fluctuations on the test masses increases due to the number of photon collisions. This leads to an additional motion of the test masses that limits the

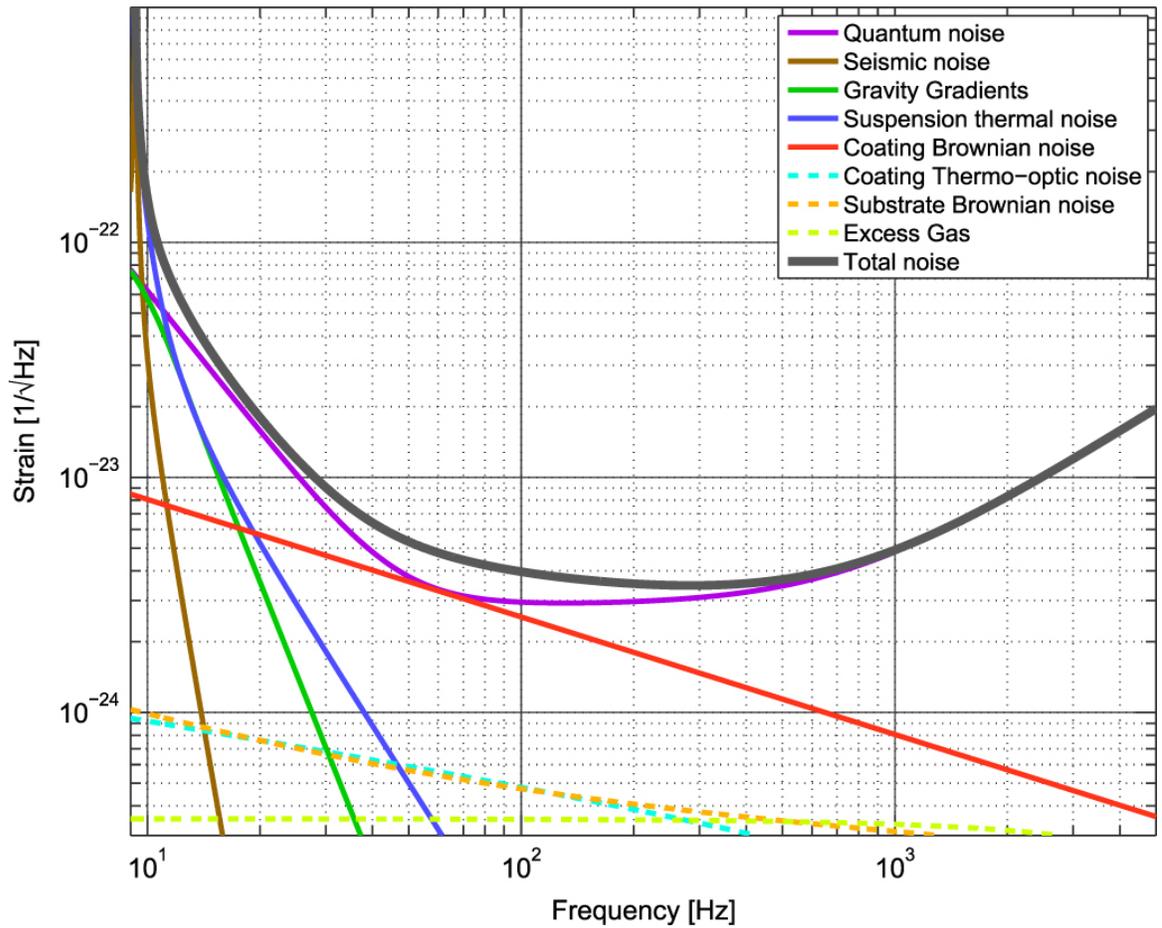


Figure 8: Fundamental noise sources for Advanced LIGO interferometers at design sensitivity. The total noise represents the sum of the fundamental noise sources shown in this figure. Additional noise from technical sources not shown in this plot is expected to also contribute to the final sensitivity curve. Reproduced from [11]

sensitivity to gravitational waves. Similar to shot noise, the relevant equation for how this limits the strain sensitivity is [52, 53]

$$h_{rad}(f) = \frac{N}{mf^2L} \sqrt{\frac{2\hbar P_{in}}{\pi^3 c \lambda}} \quad (1.34)$$

with a new factor m , the mass of the test masses.

One important difference between these two types of quantum noise is their functional dependence on frequency, f . Namely, $h_{rad} \propto f^{-2}$, while h_{shot} has no frequency dependence. Examining how their overall amplitude is impacted by the power in the detector, we see that shot noise is proportional to $\sqrt{1/P_{in}}$, radiation pressure is proportional to $\sqrt{P_{in}}$. This means that in order to maximize our sensitivity, we need to choose a input power that balances both of these competing factors. For a given frequency, our sensitivity is maximized when $h_{shot}(f) = h_{rad}(f)$. Solving for P_{in} , we get

$$P_{in} = \pi c \lambda m f^2 / 2N^2 . \quad (1.35)$$

Since we add these two noise sources in quadrature, the total noise is

$$h_{quan}(f) = \sqrt{h_{rad}^2(f) + h_{shot}^2(f)} . \quad (1.36)$$

Plugging in the optimal power we find that the quantum limit is

$$h_{quan}(f) = \frac{1}{\pi f L} \sqrt{\frac{2\hbar}{m}} . \quad (1.37)$$

If we choose to target the loudest source previously discussed, we would like to find the minimal noise relevant for a BBH merger. Since our ideal BBH reaches a maximum strain at 150 Hz, we use this as the frequency of interest. We use the aLIGO test mass value of 40 kg and an arm length of 4 km [11]. We also assume 75 round trips in the cavity, based on the light storage time. At this frequency, the optimal power from Equation 1.35 is ≈ 70 kW. These values set the minimum noise limit to

$$h_{quan}(150 \text{ Hz}) \approx 2 \times 10^{-24} / \sqrt{\text{Hz}} . \quad (1.38)$$

We can then take this value and calculate the *root-mean-square* (RMS) signal for a short duration pulse around this frequency. If we approximate the BBH signal as a pulse of duration 0.01 seconds, we have an effective bandwidth of 100 Hz. Hence the RMS noise for this bandwidth, can be found via

$$h_{rms} \approx \sqrt{h^2(f)\Delta f} . \quad (1.39)$$

for our computed quantum limit over this bandwidth, we find a strain rms of $h_{rms} \approx 2 \times 10^{-23}$. This is 50 times lower than the strain amplitude of GW150914!

Even if one was to run at a lower circulating power of 850 W, (the power that aLIGO was operating at during the first observing run) the approximate strain rms from quantum noise would be $h_{rms} \approx 1 \times 10^{-22}$, still well below our expected astrophysical strain amplitude. When one considers that power recycling can be used to increase the circulating power in the interferometer, this value is even more achievable. The aLIGO configuration uses a power recycling mirror which increases circulating power in the interferometer ≈ 42.7 times, allowing for this sensitivity to be achieved with a modest input power of 20W [11].

Thermal Noise

At present, the next most limiting source of noise from 60 Hz to 300 Hz is thermal noise. As the test masses are physical objects at finite temperature, their motion is limited by the laws of thermodynamics. This residual thermodynamic motion of the individual atoms in the surface of the test mass is an analog of Brownian motion, where the test mass is in thermal equilibrium with an external heat bath. Mechanical coupling with the heat bath comes through mechanical friction, and one of the primary sources of friction in an interferometer is the coatings on the mirrors. Coating thermal noise produces motion of the mirror surface with a noise spectrum given by [54, 55, 56]

$$h_{therm}(f) = \frac{1}{L} \sqrt{\frac{2k_b T}{\pi^{3/2} f} \frac{1 - \sigma^2}{wY}} \phi \quad (1.40)$$

where σ and Y are the Poisson ratio and Young's modulus of the material, w is the radius of the beam, and ϕ is the loss angle of the material. Note that this is nearly white noise, with a shallow $f^{-1/2}$ dependence in the spectrum.

At currently measured levels, thermal noise from coatings is below that of quantum noise [56]. However, as the laser power increases, (so that quantum noise decreases) this will soon be the limiting noise source. As well motivated models for predicting the amplitude are not currently available, in order to develop improved coatings, the loss angle, ϕ , (and hence the thermal noise levels) for each material must be experimentally measured [57].

An additional source of thermal noise in the interferometer is from suspension thermal noise [58, 59, 60]. Suspension thermal noise is from vibrations of the fibers holding the mirrors. For frequencies above resonance (and in the frequency band targeted by aLIGO), the amplitude of the noise spectrum will rapidly drop, so that suspension thermal noise will be much lower than coating thermal noise. However, suspension thermal noise is significant at lower frequencies and is one of the contributing factors that set the minimum frequency that aLIGO can probe.

Seismic Noise

At low frequencies, the current limiting noise source is due to broadband noise from ground motion. In order to reduce the impact of this ground motion on the test masses, they are suspended with wires. Any ground motion is sent through the response function of a pendulum, given by f^{-2} above the resonance frequency. Additional stages further increases the isolation of the test masses from ground motion. Adding additional isolation stages to the test mass suspensions was one of the key improvements for aLIGO heavily which reduced the overall impact of seismic noise as compared to Initial LIGO by [11, 61]. With four stages in aLIGO, the total response of the suspension system is hence

$$h_{sus}(f) \propto \frac{1}{f^8} . \quad (1.41)$$

Despite this mitigation, when ground motion is still limits aLIGO sensitivity at frequencies below 20 Hz. The overall amplitude of this seismic noise is variable. High seismic motion is generally caused by earthquakes and anthropogenic sources [62, 63].

Newtonian Noise

Due to the location of the detector on the surface of the Earth there will be local gravitational gradients from uneven distributions of matter around the detector that create motion of the mirrors. As this noise source is due to the Newtonian gravitational attraction between the optics and surrounding matter, it is often referred to as ‘Newtonian noise’.

Newtonian noise is dominated by the movement of the buildings housing the detectors and from density fluctuations in the ground under the detectors. To counteract forces from building movement, it is possible to adjust the tilt of the optics to reduce the relative movement [64]. For ground density fluctuations, no mitigation strategies have been developed at present. As most fluctuations in density are due to passing seismic waves, this noise source is nonstationary, and cannot be counteracted by reducing the relative motion. However, the frequencies of density fluctuations are quite low, below 10 Hz, and hence are not one of the dominant sources of broadband noise at present.

In future detectors, Newtonian noise will become one of the limiting noise sources. To attempt to mitigate this, methods have been proposed to measure local density fluctuations with arrays of sensors and subtract the noise from the data [64, 65, 66]. Such methods are similar to the noise subtraction techniques discussed in Chapter 5.

Additional Noise Sources

Beyond the specific broadband noise sources discussed in this chapter, aLIGO sensitivity is limited by a variety of additional known and unknown noise sources [46]. The relative amplitudes of each measured noise source, along with a comparison between the sum of all known noise sources and the total noise, can be seen in Figure 9. As expected, the sum of estimated noises (black) matches well with the measured noise level (red) above 100 Hz. In this region, the overall noise level is dominated by shot noise. However, below this point the estimated and measured noise levels differ significantly. This indicates that additional noise contributions are from sources not measured or understood.

A number of known noise sources not previously discussed in this chapter are shown, with the most significant being length control noise, alignment control noise,

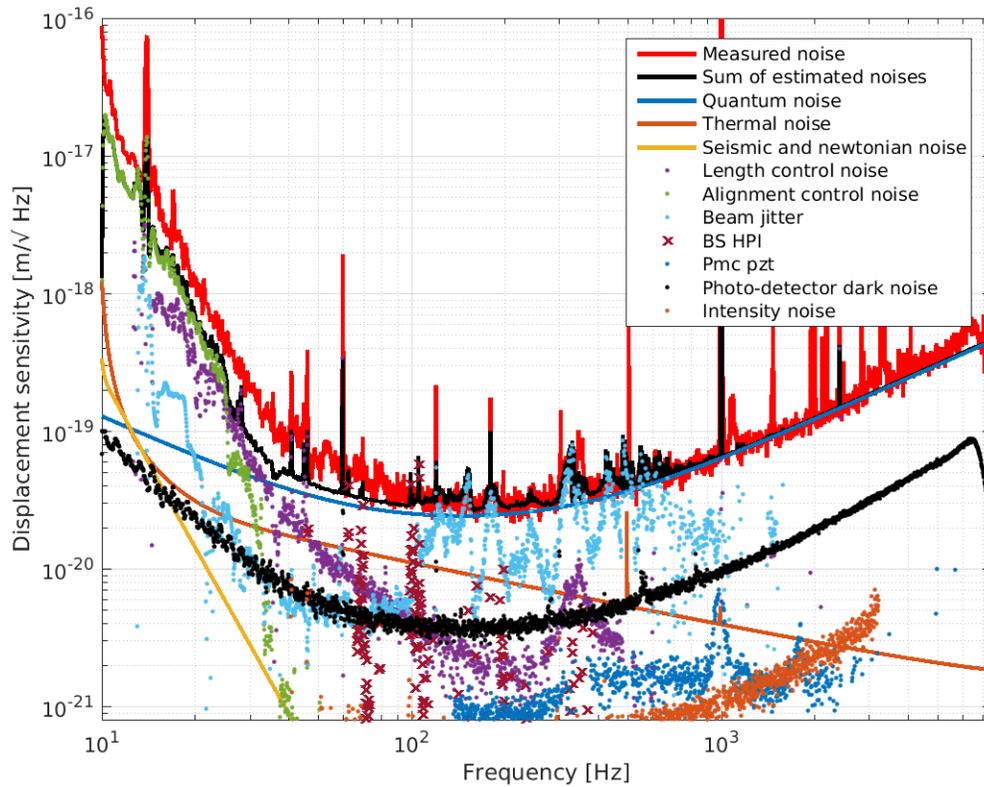


Figure 9: Noise budget for LIGO-Hanford at the start of O2. The sum of estimated noises (black) matches the measured noise level (red) well above 100 Hz, but underestimates the total noise level below this frequency. Reproduced from [67].

and beam jitter. Length and alignment control noise is a result of noise in the control loops that keep the detector stabilized. Beam jitter, on the other hand, is from fluctuations in the amplitude and angle of the input beam. This noise source was especially problematic at LIGO-Hanford in O2, with amplitudes higher than quantum noise in the 100-500 Hz region. Further discussion of beam jitter, along with mitigation methods, is found in Chapter [5.2](#).

Chapter 2

Searches for Gravitational Waves

While the design of aLIGO makes it sensitive enough to detect gravitational waves from astrophysical sources, the rate at which such events occur in the nearby universe is low. In aLIGO’s first and second observing runs, the observed rate of detections was about once every 15 days of coincident detector data [13]. If we considered a standard binary black hole merger signal to last one second, this means that only 0.0001% of data contains measurable gravitational waves! In order to identify these rare signals in interferometer data, computing algorithms known as ‘searches’ process the data with the goal of identifying time periods that may contain gravitational waves.

In the rest of this section, I will give an overview of the main methods used by the LIGO-Virgo collaboration to search for transient gravitational wave signals. Emphasis will be given primarily to the PyCBC pipeline, a specific search algorithm optimized to identify gravitational waves from compact binary coalescences (CBCs) as predicted by general relativity. I will also discuss how these searches are limited by the presence of short instrumental artifacts known as ‘glitches’, along with efforts to characterize and mitigate these noise sources.

Portions of Section 2.4 of this Chapter is adapted from an appendix of [13] titled, “Characterization of transient noise relevant to catalog triggers.” I was the lead editor of this appendix, which presented an overview of what was understood about the glitch classes at the time. This section also features original work that was not previously featured.

2.1 Searches for Gravitational Wave Bursts

The most general way to search for gravitational waves is with no assumptions about the expected waveform of the signal. Instead, one can look for cases when a short, loud signal passes through multiple detectors at almost the same time. These searches are referred to as looking for ‘gravitational wave bursts’.

In order to understand how a gravitational wave signal can be isolated in the data, it is useful to think about the data as not only a wave with a specific amplitude at a given time, but also with a specific frequency profile. Therefore if we are able to identify what specific frequencies we have excess power at, we can reduce the rate of false positives by requiring that both detectors observe similar signals.

In order to represent the frequency structure of the data, a number of techniques are used. The first is known as the short Fourier transform, where the signal is decomposed into a basis of exponentials. For a timeseries $x(t)$, the Fourier transform $X(f)$ is

$$X(f) = \sum_{n=0}^{N-1} x(n\Delta t) \exp(2i\pi f(n\Delta t)/N) . \quad (2.42)$$

The variable f represents a given frequency, for a timeseries with N samples and a sample rate of $1/\Delta t$. In this way the signal is broken down to specific frequency components. If N is chosen to break the data stream into short segments, and if the Fourier transform is repeated for every small segment of the data, the changes in amplitude at each frequency can be tracked. This is often described as the ‘time-frequency’ representation of the data, or a ‘spectrogram’.

In practice, burst searches tend to use more complicated bases to decompose the timeseries. These methods allow the relevant frequencies to be better isolated in both frequency space and time. Some of the more common bases are the Wilson-Daubechies-Meyer transform [68] or the Q transform [69]. An example of spectrograms around the time of GW150914 produced with the Q transform can be seen in Figure 10.

Once the data is broken down into this representation, burst searches look for excess power at the same frequencies at the same time in both detectors. By not making any assumptions about the signal, there is a risk of confusing excess power from instrumental artifacts and excess power from astrophysical sources, thereby reducing

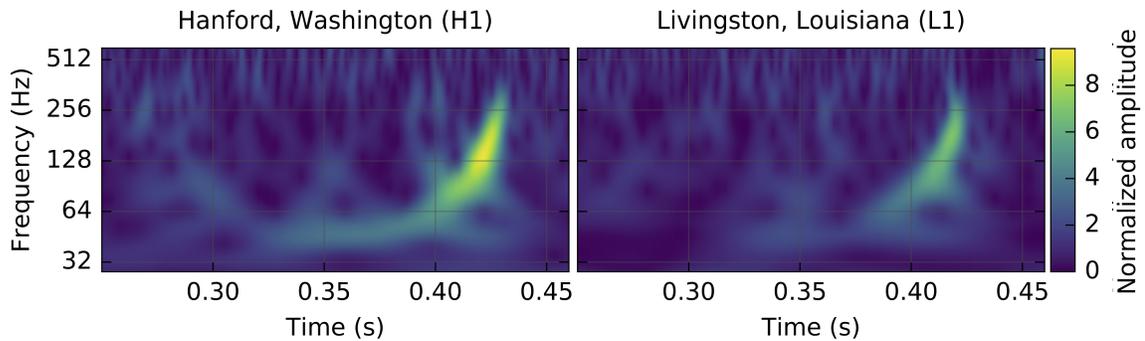


Figure 10: Spectrograms of the data around GW150914 from each detector, produced with the Q transform, adapted from [12]. The excess power from the gravitational-wave signal is clearly visible.

the sensitivity of the search. However, this agnostic approach also allows for the possibility of unexpected signals to be found by the search, such as gravitational waves not fully predicted by general relativity or from a source that is not well modeled, such as the signal from a supernova.

While useful for novel sources, this method should also be sensitive to well understood sources, such as gravitational waves from binary mergers. This search approach was in fact how GW150914 was first identified in the data [70]. During O1 and O2, a number of burst searches have been utilized, including cWB [71] and oLIB [72].

In cases when the expected signal is well modeled, however, this search method is not optimal. Even making simple assumptions about the signal, such as a frequency that increases with time, increases the sensitivity [13, 71]. For a waveform that is known precisely, the optimal techniques instead rely upon a method known as matched filtering.

2.2 Matched Filter Searches

Before discussing matched filtering, it is worthwhile to discuss the more general approach of using time-domain cross-correlation. Cross-correlation is used to measure the similarity between two time-series. For two timeseries, f and g , the cross-correlation is defined as

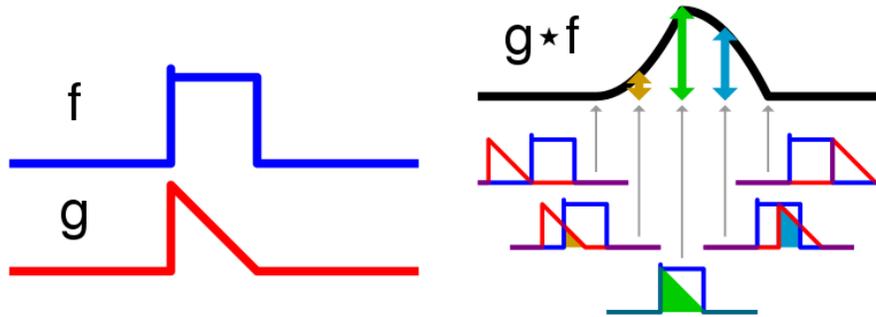


Figure 11: Example of cross correlation, showing how cross-correlation can be thought of as the overlap between two functions. Adapted from [73].

$$(f \star g)(\tau) = \int_{-\text{inf}}^{+\text{inf}} f^*(t)g(t + \tau)dt \quad (2.43)$$

where f^* indicates the complex conjugate. This can be imagined as sliding the two timeseries against each other, and recording how well the two overlap. This process is shown in Figure 11.

The matched filter is an application of the cross-correlation, using it to measure the overlap between a known template timeseries, and recorded timeseries data. For the matched filter, the overlap is measured using the quantity *signal-to-noise ratio* (SNR). The SNR for a matched filter with a specific waveform template h is [74]

$$\rho^2(t) \equiv \frac{\| \langle s|h \rangle \|^2}{\langle h|h \rangle} \quad (2.44)$$

where the inner product is defined as

$$\langle a|b \rangle(t) = 4\text{Re} \int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} e^{2\pi itf} df \quad (2.45)$$

with s the strain data, h the template, and $S_n(f)$ the estimated power spectral density for the time in question. This is equivalent to cross-correlation in the frequency domain.

In the regime where the waveform is precisely known, this method is well-known to be optimal for identifying periods likely to contain a known signal. In practice, CBC waveforms are known only up to approximations and often do not contain all effects that are predicted to exist from general relativity. However, recent comparisons of

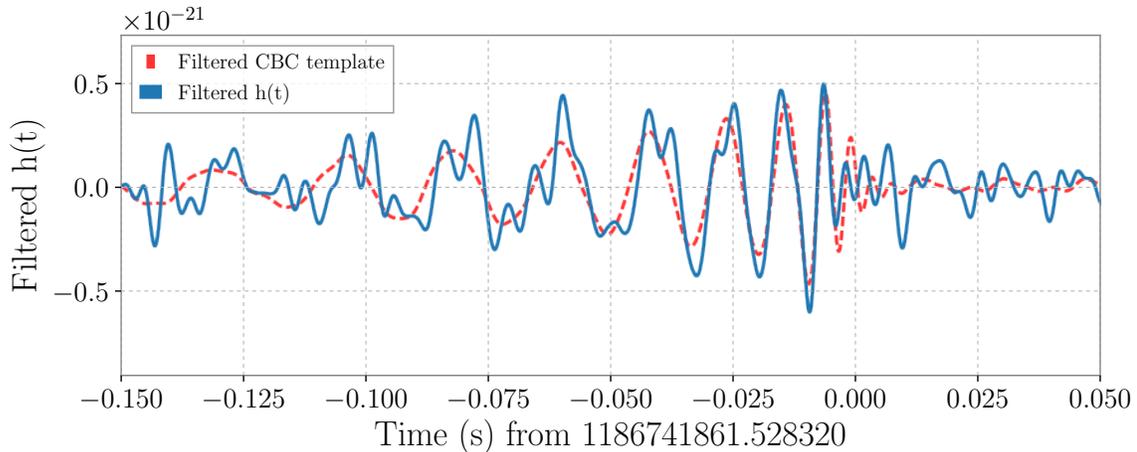


Figure 12: Filtered data around the time of GW170814 [76] with the template from the PyCBC search pipeline overlaid. Both the data and the template are filtered with a bandpass filter and a series of notch filters to suppress line features. Note the high level of match between the waveform and the data.

these waveforms with gravitational wave signals observed by LIGO have found no measurable differences [75]. An additional practical constraint that is encountered is that it is not possible to consider every possible parameter value for a signal. Instead, a finite number of waveforms are used that approximate the full parameter space with finite resolution.

A variety of pipelines based on matched filtering have been developed to search LIGO data. PyCBC [77], GstLAL [78], MBTA [79], and SPIIR [80] were all developed initially by members of the LIGO Scientific Collaboration (LSC) and are still used internally by the collaboration. With the advent of open data Venumadhav et al. [81] have also developed a matched filter pipeline independent to the LSC. The remainder of this section will be devoted to detailing the PyCBC pipeline, which will be the pipeline this dissertation primarily focuses on.

2.3 PyCBC Pipeline

The PyCBC pipeline is one of the matched filter searches used to identify gravitational waves from CBC events in the first and second observing runs of Advanced LIGO [13]. During the second observing run, the overall search space spanned, in terms of

individual masses, from $1M_{\odot}$ up to $500M_{\odot}$, including BNS, NSBH, and BBH systems [82]. In order to effectively search this large parameter space, over 4×10^5 templates are used. Templates in the bank are chosen so that for any arbitrary set of parameters for a signal in the search space, there will be a template in the bank that recovers at least 96.5% of the SNR of the signal. This minimum match is higher for some regions of the bank. This requirement produces a bank that is extremely dense for long duration templates (such as low-mass BNS templates), and comparatively sparse for short duration templates (such as high-mass BBH templates). The full search space, in terms of component masses and effective spin, is shown in Figure 13.

The search process begins by conducting a matched filter search of every single template in the bank. Preliminary identification of candidates is performed by looking for peaks in the SNR timeseries. If a peak is above a specific threshold (SNR of 5.5), this time is identified as a trigger. Once a trigger is noted, additional signal consistency tests based on the details of the trigger, the parameters of the template that produced said trigger, and the quality of the data around that time are used to further vet interesting candidates. This process results in a single *detection statistic* which is then further used to evaluate the significance of such a trigger.

2.3.1 PyCBC Detection Statistic

If aLIGO noise was perfectly Gaussian, the matched filter SNR alone would be an optimal detection statistic. However, since the data is neither Gaussian nor stationary over long time periods [63, 83, 84], numerous additional signal consistency tests are required to discriminate between common instrumental artifacts and astrophysical signals. An ideal signal consistency test would result in true astrophysical signals being unaffected while noise triggers fail this test. In PyCBC, this is accomplished by dividing the value of the SNR detection statistic of triggers by a factor that quantifies the degree of mismatch between the data and what is expected for a true signal.

Chi-Squared discriminator

To provide this discriminatory power, one of the most useful tests for gravitational-wave signals is the chi-squared discriminator [85]. The test is constructed by dividing the frequency space spanned by the waveform template into bins of equal power,

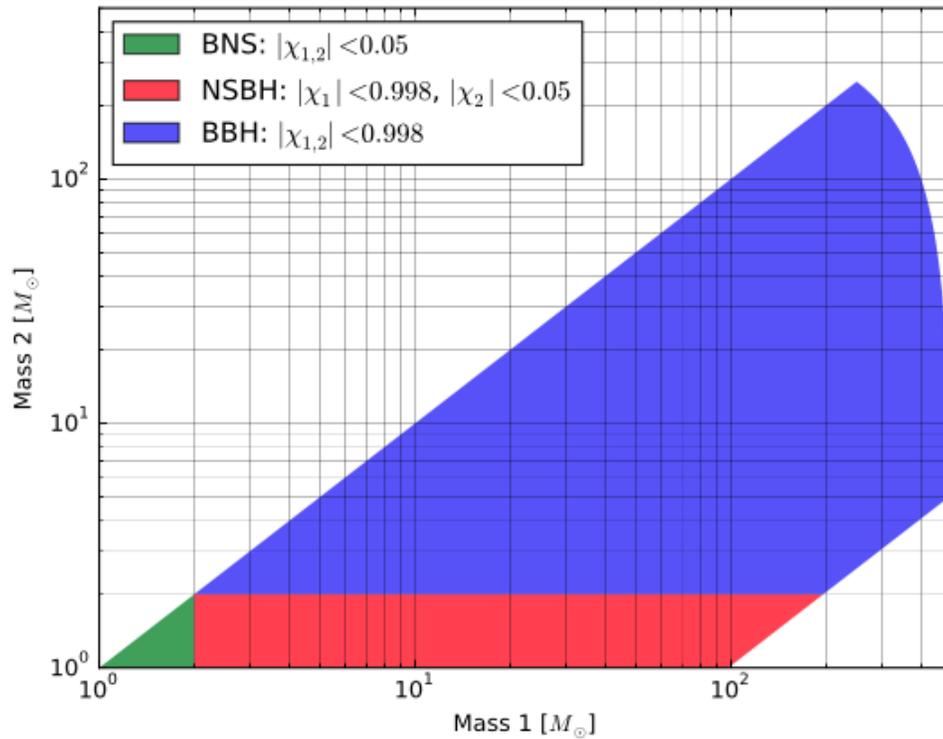


Figure 13: Overview of the component mass parameter space spanned by the template bank used by PyCBC during O2 [82]. In addition to showing component masses, the colored regions correspond to different bounds on the effective spins of the templates, motivated by astrophysical observations of related objects.

and checking if each bin contributes the expected amount of power. This has the effect of measuring how well a candidate trigger matches the signal morphology of the template. Specifically, the chi-squared discriminator for a trigger is

$$\chi_r^2 = \frac{1}{2p-2} \sum_{i=1}^p \|\langle s|h_i \rangle - \langle h_i|h_i \rangle\|^2 . \quad (2.46)$$

This value should follow a reduced χ^2 distribution with $2p - 2$ degrees of freedom. The choice of p is dynamically scaled based on the duration of each template, so that a sufficient number of bins with measurable power are used. For example, the GW170814 template has a duration of 0.44 seconds and uses 29 bins, while the GW170817 template has a duration of 74.43 seconds and uses 247 bins. If the value of the chi-squared test is greater than unity, the detection statistic for the related trigger is reduced to produce a “re-weighted SNR”, $\tilde{\rho}$. This is

$$\tilde{\rho} = \begin{cases} \rho & \text{for } \chi_r^2 \leq 1 \\ \rho \left[\frac{1}{2} \left(1 + (\chi_r^2)^3 \right) \right]^{-1/6} & \text{for } \chi_r^2 > 1 \end{cases} . \quad (2.47)$$

The effectiveness of the chi-squared discriminator had been shown to be dependent on the duration of the signal and the number of bins used in the test [83]. For long duration signals, the test provides excellent rejection of many classes kinds of glitches. For short duration signals, this test has reduced efficiency. To help address this effect for short duration templates, an additional signal consistency test, the sine-Gaussian discriminator, is utilized.

Sine-Gaussian discriminator

The sine-Gaussian discriminator is designed to downrank triggers with excess power at frequencies above the expected maximal frequency of the signal at merger [86]. If excess power is detected above this frequency, it indicates that the excess power is not likely to have come from a CBC signal. To quantify the excess power present at high frequencies, a number of sine-Gaussian wavelets with frequencies above this maximum are matched filtered against the data. These wavelets are parameterized by their frequency, f_0 , central time t_0 , and quality factor Q . In the time domain, each wavelet can be written as

$$g(t) = \exp\left(-4\pi f_0^2 \frac{(t-t_0)^2}{Q^2}\right) \cos(2\pi f_0 t + \phi_0) . \quad (2.48)$$

A new signal discriminator can be written down as the sum of the measured matched filter SNR squared of each individual sine-Gaussian tile. In the case of N different tiles, this is

$$\chi_{r,sg}^2 \equiv \frac{1}{2N} \sum_{i=1}^N \rho_i^2 = \frac{1}{2N} \sum_{i=1}^N \langle s | \tilde{g}_i(f, f_0, t_0, Q) \rangle^2 . \quad (2.49)$$

Similar to the chi-squared discriminator, this statistic should follow a reduced χ^2 distribution with $2N$ degrees of freedom for astrophysical signals. The result of this test is then used to compute a new detection statistic, $\tilde{\rho}_{sg}$, defined as

$$\tilde{\rho}_{sg} = \begin{cases} \tilde{\rho} & \text{for } \chi_{r,sg}^2 \leq 4 \\ \tilde{\rho} (\chi_{r,sg}^2/4)^{-1/2} & \text{for } \chi_{r,sg}^2 > 4 \end{cases} . \quad (2.50)$$

The value of 4 (as opposed to 1) is chosen as the threshold to account for the expected variability of $\tilde{\rho}_{sg}$ in Gaussian noise. Values above 4 are indicative of non-Gaussian features in the data.

This test is most useful in cases of broadband noise combined with a high mass waveform, which has a lower termination frequency. One of the most common cases of a glitch creating a trigger is with a short duration instrumental artifact that rings off a short duration template. The shortest templates in the PyCBC bank are only 0.15 seconds, and only are capable of supporting a few chi-squared bins [82]. The chi-squared discriminator has reduced power to rule out glitches that match these templates. In this scenario, the frequency at merger of the template is quite low (below 100 Hz) and hence the sine-Gaussian discriminator is extremely sensitive to the observed excess power, as a broadband glitch is likely to have significant excess power above 100 Hz. Therefore, the sine-Gaussian discriminator is able to easily rule out a class of glitches that are not able to be down-ranked by the chi-squared discriminator.

Background dependent reweighting

If we assume that the rate of triggers is given by a Poisson process, we would expect that the probability of a trigger in the detector is given by a Gaussian probability distribution

$$P(\rho^2) \propto \frac{1}{\alpha} \exp\left(-\frac{\rho^2}{\alpha^2}\right). \quad (2.51)$$

However, since a large parameter space is spanned by the template bank used in the PyCBC search, it is unlikely that the rate and distribution is the same for vastly different parameters. This can be accounted for by allowing the standard deviation variable, α , to be dependent on the template parameters, $\bar{\theta}$, and by including an additional parameter-dependent prefactor, μ , that accounts for the total number of triggers. With these additions, the expected number of triggers for a specific template is

$$N(\rho^2, \bar{\theta}) = \mu(\bar{\theta}) P(\rho^2, \bar{\theta}) \propto \frac{\mu(\bar{\theta})}{\alpha(\bar{\theta})} \exp\left(-\frac{\rho^2}{\alpha^2(\bar{\theta})}\right). \quad (2.52)$$

In order to directly compare the significance of templates in the bank that have vastly different parameters, one component of the PyCBC search is to directly measure the values of $\alpha(\bar{\theta})$ and $\mu(\bar{\theta})$ based on the observed distribution of single detector triggers. Templates that have large values of $\alpha(\bar{\theta})$ or $\mu(\bar{\theta})$ are down-ranked to account for the increased chance of producing a trigger at a given ρ value. This process is known as *template-dependent background reweighting* [87].

In O2, this reweighting process was done based on the duration of the template, so that $\bar{\theta}$ was equivalent to the template duration. However, including additional parameters, such as the chirp mass and effective spin, is possible, and has been implemented in other searches [88].

Identifying coincident triggers

Once triggers have been assigned a significance in each detector, the triggers are compared between the two detectors to test if any occur in coincidence with each other. In this context, a coincident trigger is one where the same template produced a trigger above threshold with a peak time difference less than the light travel time

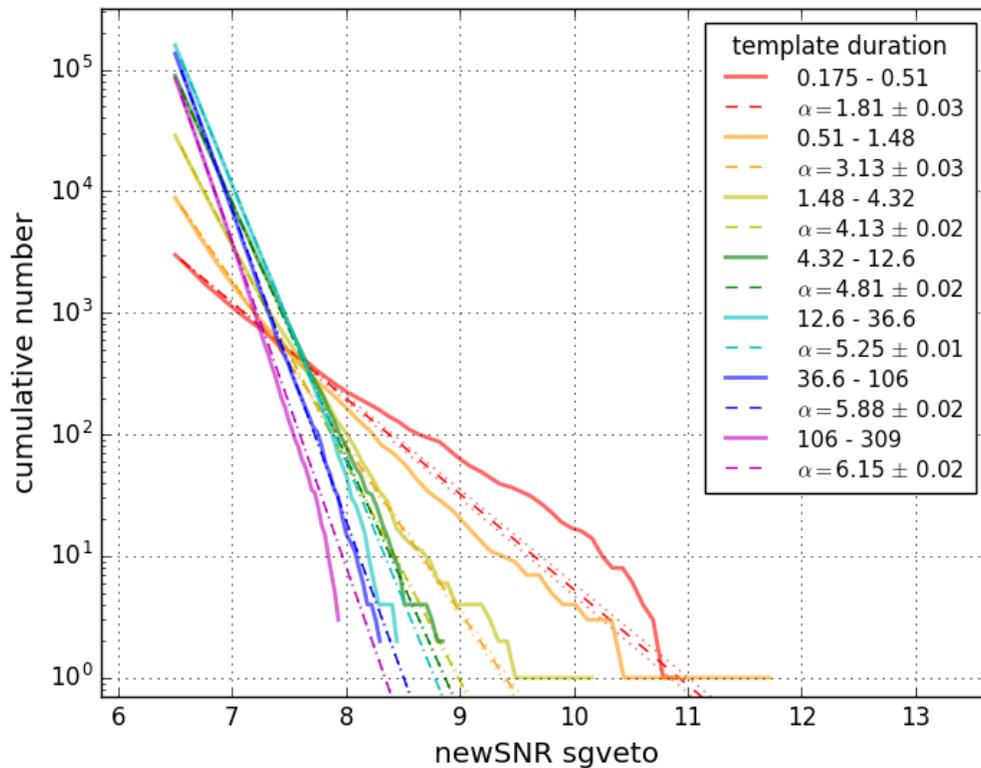


Figure 14: An example of fits generated from template dependent background reweighting for a representative period during O2. Templates are binned based on template duration and the cumulative number of triggers is then plotted (solid line) along with an exponential fit (dotted line). The value of $\alpha(\bar{\theta})$ for each template bin fit is labeled in the legend (Referred to here as α). Note that templates with shorter durations have lower α values and hence are more likely to produce high-SNR triggers.

between the detectors. For the LIGO-Livingston and LIGO-Hanford network, this is 7 ms. To account for uncertainty in the peak time, this window is further lengthened to 10 ms [77].

Since the search looks for hundreds of thousands of different waveform templates in the same data, there is the potential for a given time period to have triggers from multiple different templates. It is common (and expected) that an astrophysical signal will produce multiple triggers at the same time with a variety of templates. However, only the trigger with the highest statistic is used to estimate the significance of the trigger.

Time-phase consistency test

If a coincident trigger is found, an additional step of computing the final detection statistic is testing whether the observed trigger parameters in each detector are likely to occur from an astrophysical trigger. Due to the known sensitivity, antenna pattern, and orientation of each detector, there exist certain combinations of the arrival time, phase offset, and amplitude for the coincident trigger that are more likely than others. For triggers due to noise, no correlation between these values is expected. To include this information in the detection statistic, a series of injections that modeled the sky location and distance of an astrophysical population of signals are performed to empirically measure the relative likelihood of each combination.

This information is then used to down rank coincident triggers that are recovered with unlikely combinations. Specifically, the new detection statistic is [87]

$$\tilde{\rho}^2 = \hat{\rho}_c^2 + 2 \log \left(\frac{p^S(\bar{\theta})}{p_{\max}^S} \right). \quad (2.53)$$

where $\bar{\theta}$ is parameterized by the relative arrival time, phase offset, and SNR of the coincident trigger in each detector. The variable p_{\max}^S is the probability of the most likely set of parameters, and $p^S(\bar{\theta})$ is the probability of the observed parameters.

2.3.2 Background Estimation

Once a detection statistic is assigned to a coincident trigger, the significance of this value must be computed. One possible method is by estimating the rate of chance coincidences by using a specific model of the noise background in the detectors, and

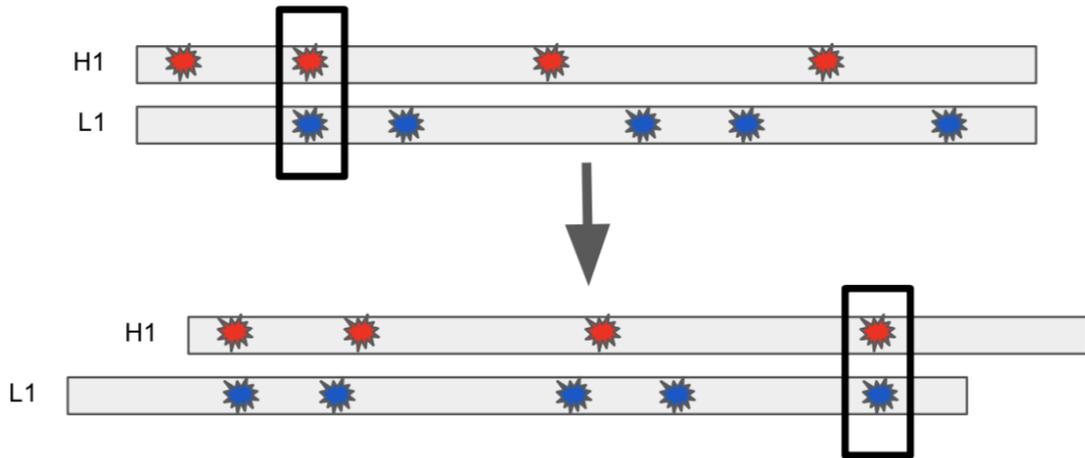


Figure 15: A visualization of how timeslides are created. The upper panel shows a series of triggers in each detector timeseries, with one coincident trigger marked with a box. The lower panel shows the timeseries shifted so that different triggers are recovered in coincidence. This shifting procedure is repeated until the requisite background duration is achieved.

how often different types of noise occur that are similar to the signals being searched for [78]. However, in PyCBC the significance is empirically measured using *timeslides* [77].

If one was only to use the coincident data available for each detector to estimate this false alarm rate, then the lowest rate of coincidences that could be empirically measured is one per the analysis time duration. In addition to the low threshold this sets, this method would also be incapable of differentiating a chance coincidence with one that is due to the signal being astrophysical. In a typical experiment, it is possible to calibrate the experiment by shielding the apparatus from what is being measured, and measuring the properties of the noise without any signal present. However, as there is no way to prevent gravitational-wave events from occurring, or from preventing the interferometer from detecting them, the rate of false alarms must be evaluated using data that has both noise and real signals therein. With timeslides, one is capable of generating large amounts of simulated data that has a low contamination by astrophysical coincidences.

To produce timeslides, the data from one detector is shifted by much more than the light travel time between the detectors (7 msec)¹. Triggers in each detector are tested again for new coincidences at this time shift. A visualization of this process can be seen in Figure 15. Because the time shifts are larger than the largest time separation possible for a real signal, any coincidences in the new shifted dataset can not be due to the same astrophysical signal. The number of timeslides for a given analysis period duration T_{init} and a time shift of t_{shift} is

$$N_{\text{slides}} = (T_{\text{init}})/t_{\text{shift}} , \quad (2.54)$$

meaning that the total amount of simulated background time that is constructed via this method is

$$T_{\text{total}} = T_{\text{init}} \times N_{\text{slides}} = (T_{\text{init}})^2/t_{\text{shift}} . \quad (2.55)$$

For a typical analysis period of 5 days and a timeshift of 0.1 seconds between time slides, this method allows $\approx 100,000$ years of simulated data to be produced, allowing the rate of random coincidences to be measured much more precisely. An example of the measured background via timeslides over a variety of situations can be seen in Figure 16.

To further prevent astrophysical signals from biasing estimation of the background, any coincidences found in the original data can be removed. However, this choice has the potential to remove chance coincidences from the data stream that are not astrophysical in origin, and result in underestimating the rate of random coincidences. Both of these choices may slightly bias the measured rate of random coincidences, but in cases when astrophysical events are rare (less than a few per analysis period) these biases are not significant [90].

One additional source of bias in this procedure is the assumption that any given time is equally likely to produce a trigger. If this is true, then the rate of triggers in the coincident data is well modeled by the background. Tests of this assumption have noted small variations in the distribution of triggers, but no clear presence for or against coincident background data producing a higher rate of triggers. An example of the variation of trigger rates over a two week period in O2 can be seen in Figures

¹In practice, the data is typically shifted by > 100 ms.

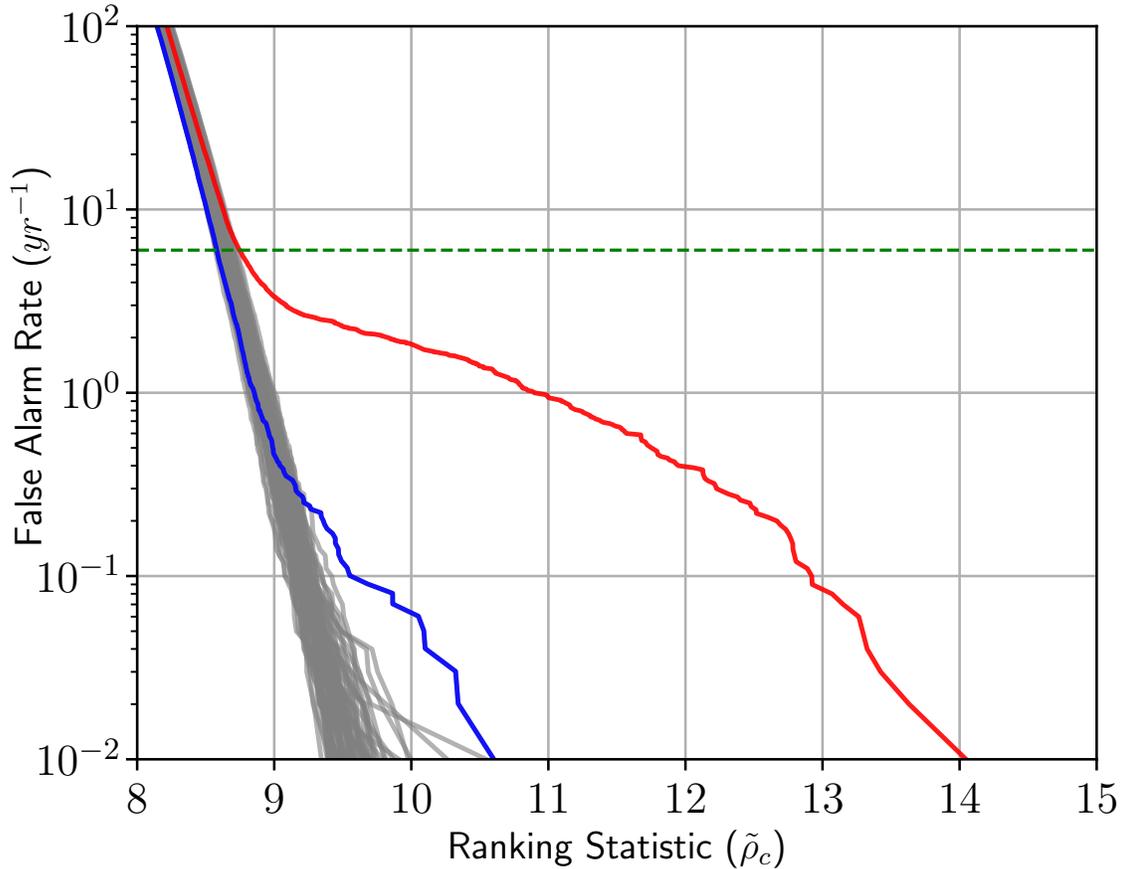


Figure 16: Background estimation for PyCBC Live [89] showing the variability of the relationship between the detection statistic and the false alarm rate (FAR) for independent analysis periods during O1. A green line is placed at the threshold for astronomical alerts (1 per 2 months). Background estimates were computed over 5 hour analysis periods. The mapping between the detection statistic and the FAR is stable with the notable exception of classes of astrophysical (blue) and instrumental (red) outliers. In this case, the blue curve is from a time containing GW150914, which produces a tail in the distribution at low FAR. The red curve is from a time period that contained an extremely high rate of instrumental artifacts. This time period was removed from later analyses due to the amount of corruption from artifacts. Reproduced from [89].

17 and 18.

2.4 Limiting Sources of Transient Noise

Despite the numerous signal consistency tests included as part of the PyCBC detection statistic, there still remain a wide variety of known instrumentally-generated transients that limit sensitivity. The fundamental, broadband sources of noise discussed in the previous section limit the recovered SNR of signals, and hence prevent the pipeline from identifying the event. However, the pipeline sensitivity is also bounded by the rate of loud instrumental artifacts with short time durations.

These artifacts, often referred to as instrumental transients or *glitches*, can cause the matched filter to respond as it would to a real signal, and hence can produce significant triggers in the search. The presence of these transients limits the sensitivity by adding a long tail to the distribution of noise triggers (As is shown by the red curve in Figure 16), forcing the analysis to set a higher threshold than would otherwise be required to achieve a target false-alarm rate. An additional concern is when such glitches overlap a gravitational wave signal, causing it to fail signal consistency checks [91].

2.4.1 Classes of Noise Artifacts

A large variety of different classes have been found to produce noise artifacts that impact the search. Of these, some classes have well known causes that have yet to be mitigated, while others are yet to be understood. In this section, I will detail some of the most common and problematic classes of glitches in the data for PyCBC, along with an overview on what is known about the causes. I will also discuss possible mitigation strategies that may exist for each glitch class. An outline of each glitch class is detailed in Table 1. Spectrograms of each of the glitch classes discussed in this section can be seen in Figure 19. Further details on how specific glitches can mimic gravitational wave signals with different parameters can be found in Chapter 4.

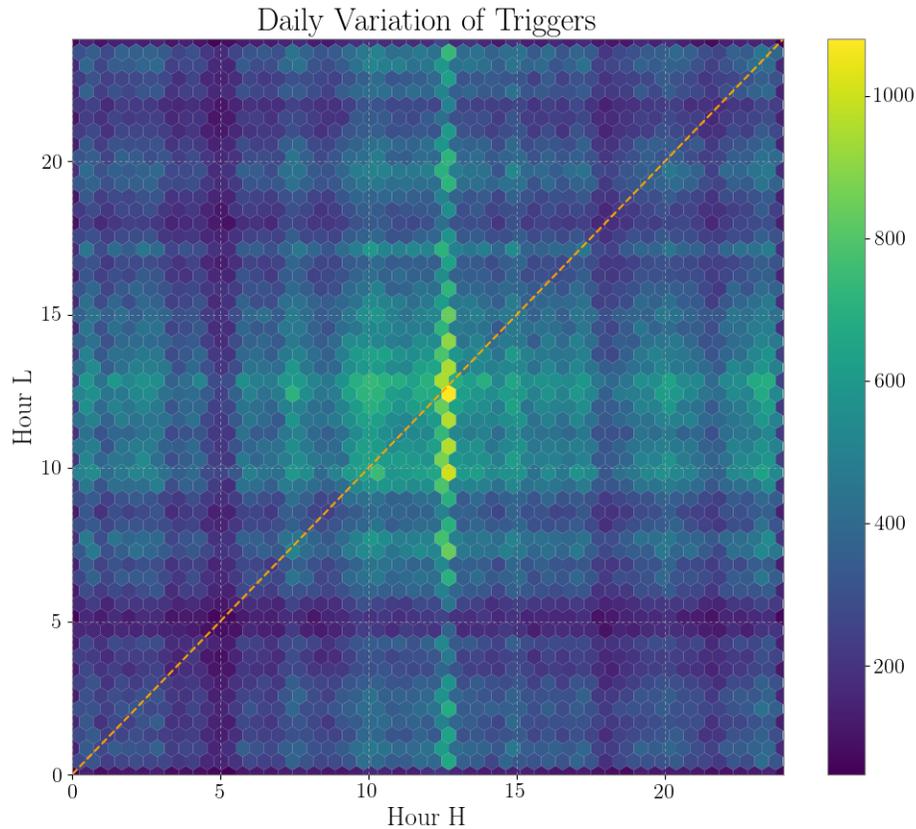


Figure 17: Daily variation in the rate of triggers at each site over a two week period in O2. Coincident trigger times generated from time slides are plotted against the local time at each site, given by Hour H (Hanford) and Hour L (Livingston). The color refers to the density of the triggers at each combination of times. A dotted line is also plotted indicating the location of zero lag triggers. The increased density near 13 UTC in Hanford is related to triggers from GW170817. Overall, there is indeed some variation in the rate of triggers for different time offsets (excluding increased trigger density due to astrophysical events).

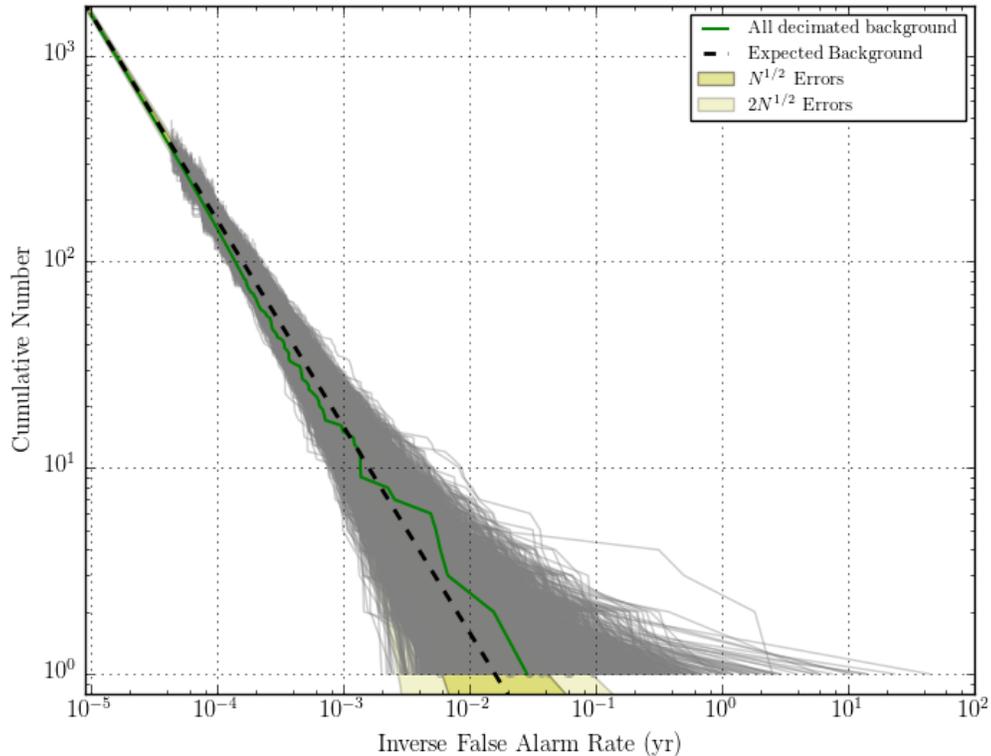


Figure 18: Variation of timeslide background distribution over a five day analysis period. Each curve represents the background distribution from one 5-day timeslide. The zero-lag background distribution is shown in green. Note that the green curve is consistent with the median of the timeslide background curves.

Glitch	SNR	Frequency Range	Duration	Shape
Blip	≈ 10	60 - 500 Hz	$\ll 1$ sec	Pulse
Koi Fish	> 100	10 - 1000 Hz	$\ll 1$ sec	Pulse
Tomte	≈ 10	10 - 100 Hz	$\ll 1$ sec	Pulse
Whistle	< 10	> 100 Hz	$\ll 1$ sec	V-shape
Scattering	≈ 10	60 - 200 Hz	2 - 10 sec	Arch
Scratchy	< 10	< 50 Hz	10 sec - 2 min	Drifting frequencies

Table 1: An overview of each glitch class listing the typical SNR, frequency range, duration, and shape of the class. See Figure 19 for spectrograms of each glitch class.

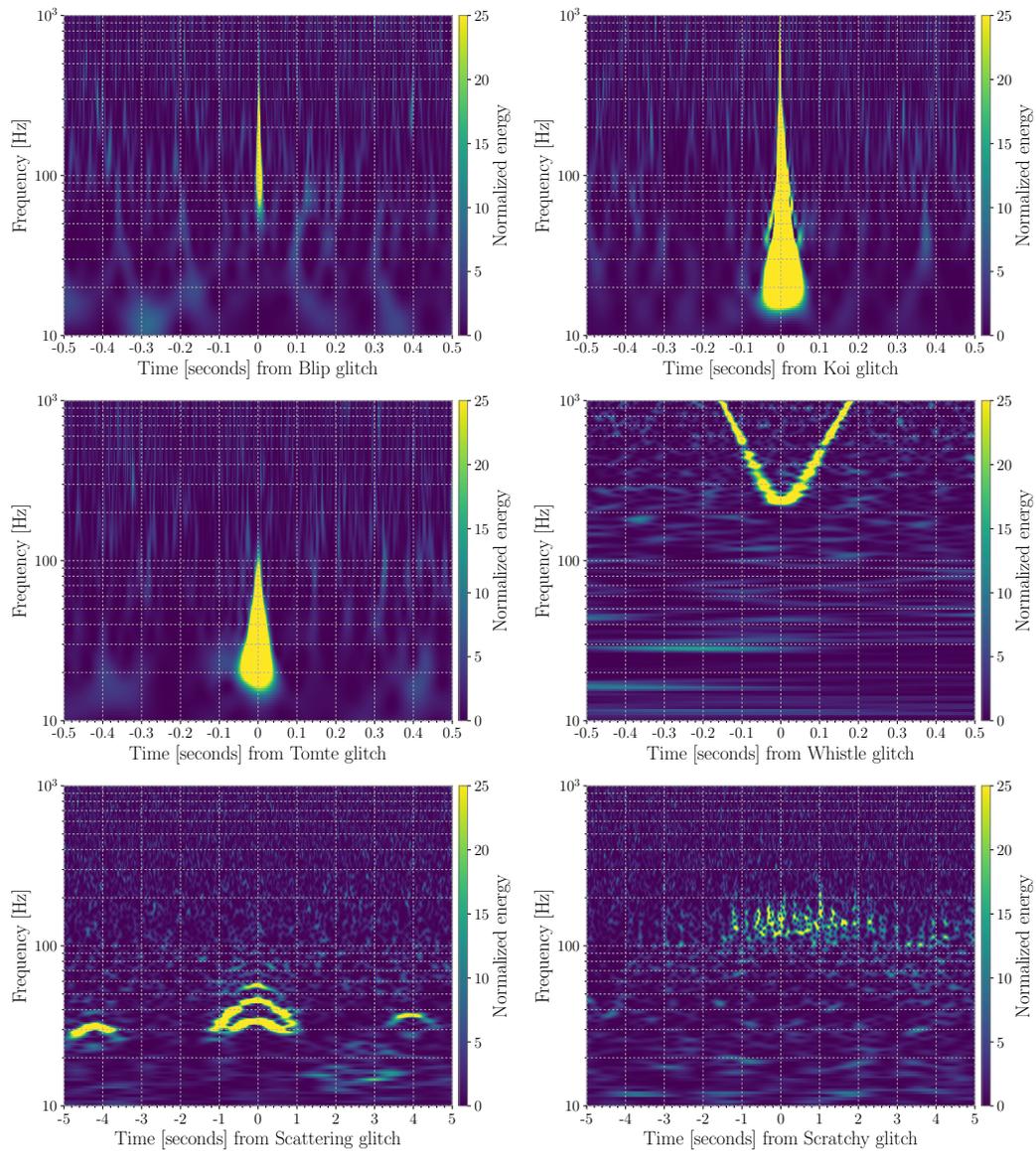


Figure 19: Omega Scans [69] of the common glitch classes present in Advanced LIGO data that are discussed in this chapter. From top-left to bottom-right, these shown classes are blip, koi fish, tomte, whistle, scattering, and scratchy.

Blips

Blip transients [63] are short, band-limited transients that occur in both LIGO detectors at a rate of roughly once per hour. Although these glitches occur in both detectors, they do not occur in coincidence at each site. Because of their sub-second duration and large bandwidth, these transients often have significant overlap with the shortest templates used in matched-filter searches. Blip transients are particularly problematic as they typically are not seen in any of the witness sensors used to monitor the detector. This makes it difficult to systematically remove them from the analyses. As such, these transients were the limiting source of loud transients in PyCBC for high mass compact binary coalescences in O1 and O2 [83, 63, 86, 70]. An example of how a blip can mimic a short duration template can be seen in Figure 21.

Investigations into blips have identified multiple causes [92], but the vast majority of blips remain unexplained. For the majority of blips, they follow the rate expectations of a Poisson process, in both time and amplitude. A subset of these blips occur in large storms, with hundreds of glitches in a few hours. Although these transients cannot be removed from the analysis entirely, the sine-Gaussian discriminator is able to somewhat mitigate their effect on the the PyCBC search [86].

Koi Fish and Tomte Glitches

Koi fish and tomte glitches are classes that appear similar to blip glitches, but differ in both SNR and frequency space. Koi fish are distinguished by their high SNR values, often above SNR 100 (while blips tend to be $\text{SNR} \approx 10$). Tomte, on the other hand, are often lower SNR than blips, and only impact frequencies below 100 Hz. Work investigating tomte glitches has found some correlation with discharges in the mirror that lead to a step response of the electro-static drive that controls the mirror [93]. This response appears in the frequency space as power given by $\propto f^{-3}$. This can be seen for a representative tomte glitch in Figure 20.

Whistles

Whistles are due to the beating of different radio frequencies that are present at each site [63]. Radio frequency (RF) modulators are used throughout the interferometer to sense and control a number of length and angular degrees of freedom. These

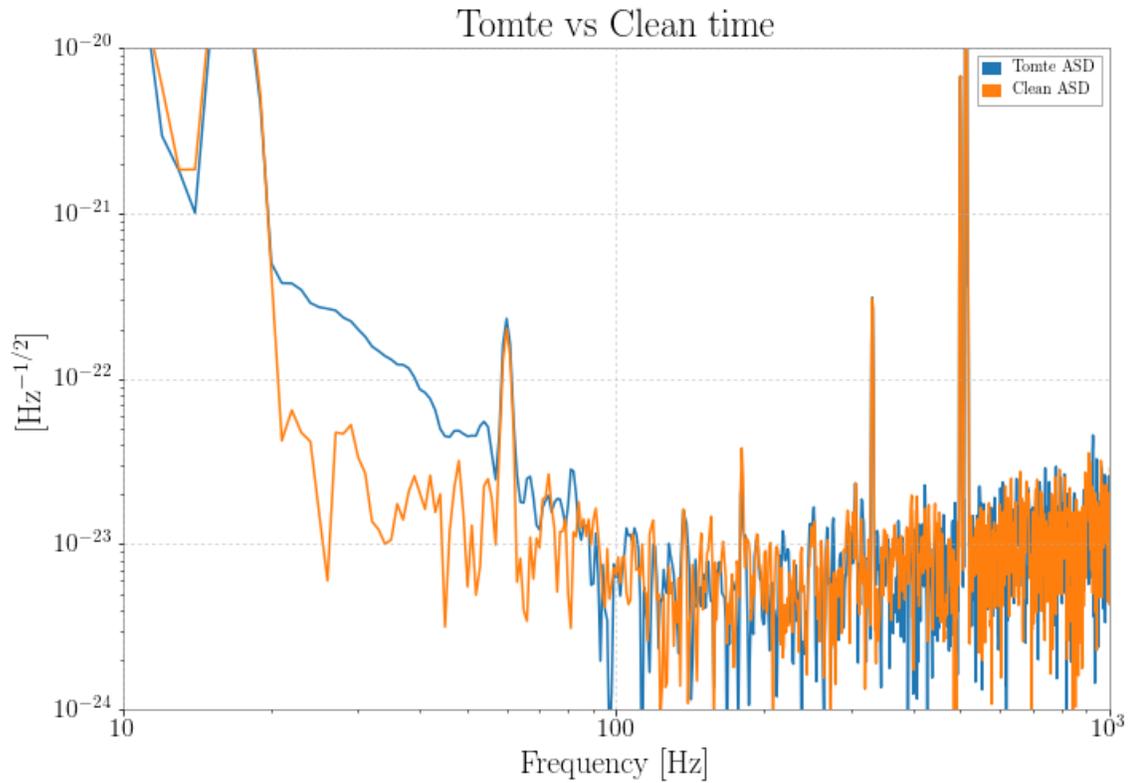


Figure 20: ASD for a short time period around a tomte glitch as compared to a nearby time containing no transients (referred to as a ‘clean time’). Each ASD is produced by using a Kaiser window with a high central peak that emphasizes the transient. Note the excess amplitude at lower frequencies proportional to f^{-3} .

modulators are referred to as voltage controlled oscillators (VCOs). When two VCO frequencies cross, the beat structure between the two sources couples into the channels that the VCO is used by. For example, one particularly problematic VCO is a part of the control loop to reduce the variability of the input laser. When this VCO crosses particular frequencies, the beat notes that occur pollute the main gravitational-wave channel [94]. This behavior allows instances of whistle glitches to be predicted based on observed radio frequencies. Whistles heavily pollute auxiliary channels, yielding them useless for their intended purpose of detecting correlated noise around the site.

The high frequency that whistles are found at, combined with the increasing frequency profile means that whistles can mimic long duration signals that terminate at comparatively high frequencies. An example of a whistle glitch with a long duration template overlaid can be seen in Figure 21.

In the past, solutions to prevent whistles have included shielding channels from radio waves, or carefully controlling the radio frequencies present at the site as to prevent any beatnotes from developing [95]. In O2, a combination of these strategies allowed whistles to have only minimal impact on the gravitational wave strain channel.

Scattering

Scattered light glitches are caused by stray light that is reflected off of one of the main interferometer mirrors that comes back into the main interferometer path, as shown in Figure 22. This adds a spurious phase shift at the interferometer output that can mimic a signal [63, 96, 97]. Motion of a reflective surface, such as optic mounts, phase-shift the reflected light. Large motions result in arch-like shapes in the time-frequency spectrograms. It is these larger motions that impact transient searches.

The frequency of the scattering arch, $f_{fringe}(t)$, can be predicted from the optic motion by the equation [96]

$$f_{fringe}(t) = \left| 2 \frac{v_{sc}(t)}{\lambda} \right| \quad (2.56)$$

where $v_{sc}(t)$ is the velocity of the optic motion and λ is the wavelength of the main laser. The strain noise amplitude of scattered light instrumental artifacts correlates with the velocity of the corresponding ground motion. Additionally, higher velocities

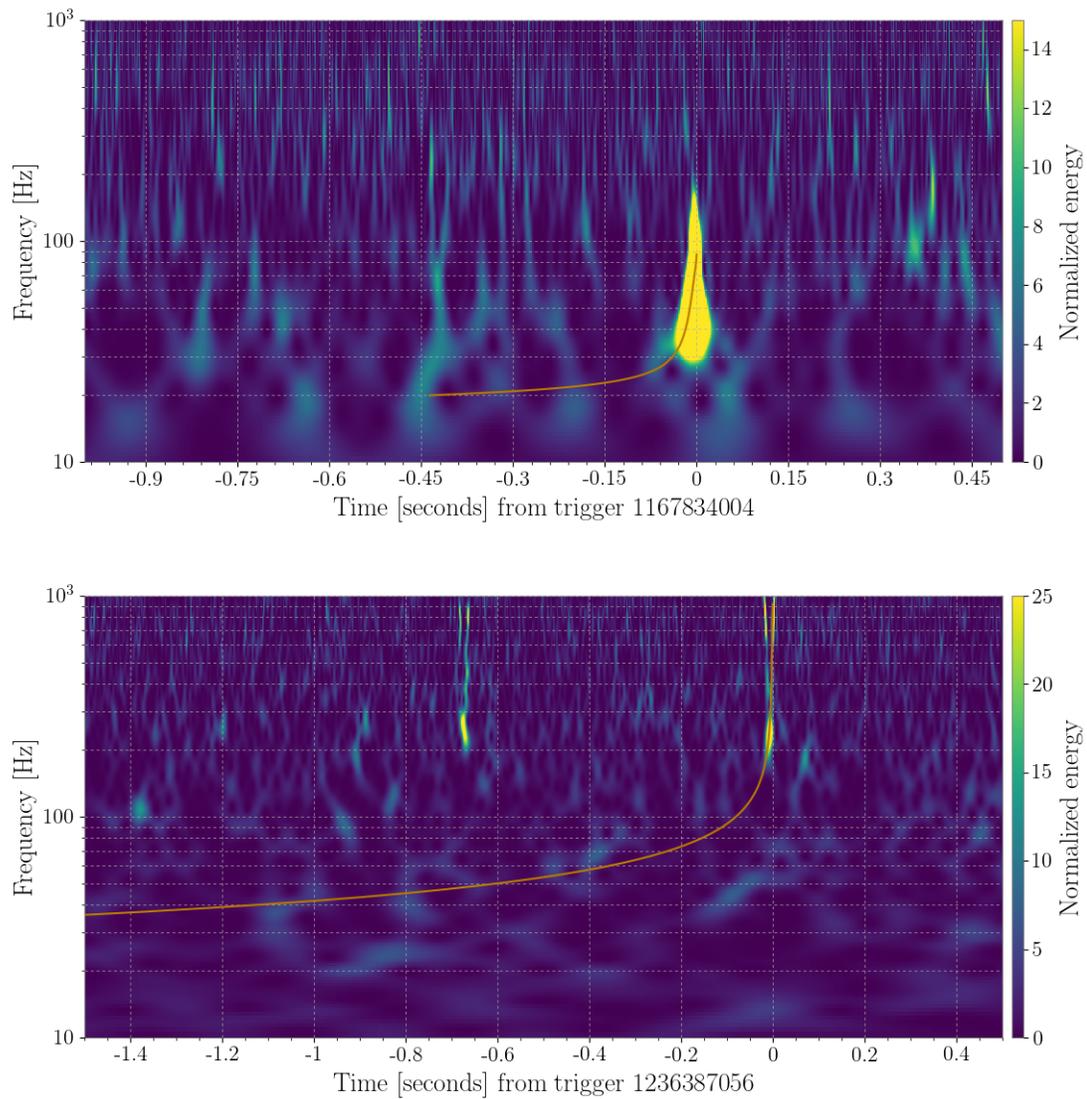


Figure 21: Overlay of a CBC trigger template (orange) with glitches. Top: For a short duration template, a blip glitch can easily mimic the final pulse of the waveform before merger. At this duration, the chi-squared discriminator is not able to completely mitigate the effect of these glitches. Bottom: The end of the inspiral matches well with the structure of a whistle glitch, producing a large number of significant triggers when whistles are present. Note that the duration of the plotted template is multiple seconds, underscoring how whistles are problematic for long duration, low mass templates.

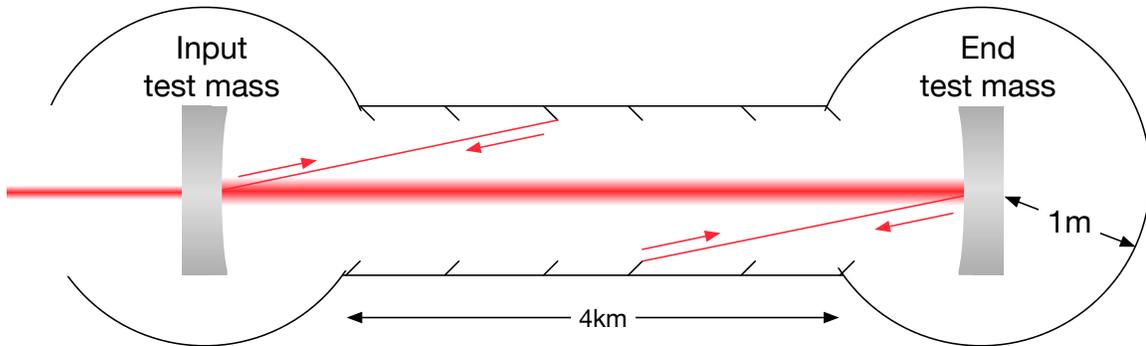


Figure 22: Diagram of how scattered light occurs, with, showing the process in which light is scattered and reflected back into the beam path. Reproduced from [46].

of optic motion lead to higher frequency content in the scattered light [96].

This motion is often observed by auxiliary sensors and is used to help identify periods of scattering [98]. Scattered light can be present in the data for stretches of multiple hours during periods of elevated ground motion, making it difficult to effectively veto without removing large amounts of data [63]. Because of this, scattered light is one of the most common sources of background triggers in PyCBC [83, 99, 100].

Since scattered light typically affects low frequencies, it is often possible to mitigate the impact by using only data from frequencies above the affected region. Because it is often related to optics where the motions is measured, subtraction of the artifacts based on a non-linear relationship with this optic motion detailed above may be possible.

Scratchy

A 60-200 Hz nonstationarity, or “scratchy” glitching, appears in time-frequency spectrograms as excess power with slowly varying frequencies in clusters of multiple minutes that generally consist of individual glitches lasting 10-100 seconds [83]. The structure of the excess power appears similar to scattered light, but at higher frequencies than predicted by available measurements of optic motion. This type of nonstationarity occurs in both LIGO-Hanford and LIGO-Livingston at a rate of 1-2 clusters per day. While correlations between excess anthropogenic seismic noise and the nonstationarity exist, no clear witnesses have been identified. The structure of

this class of artifact changed after mitigation of the motion of baffles used to block stray light [101, 102]. This change suggests that baffles (sheets installed to reduce stray light in the interferometer) may be involved in the production of the 60 - 200 Hz nonstationarity.

In previous observing runs, the rate of these artifacts caused multiple hours of data to be vetoed. In time that is not vetoed, these transients can create significant triggers in the background of PyCBC [83]. The long duration, variable frequencies, and lack of a clear witness make this nonstationarity a difficult target for noise subtraction.

2.4.2 Detector Characterization Tools

One of the key science goals of the Detector Characterization Group in the LIGO Scientific Collaboration is to study the main sources of transients and quantify their impact on the astrophysical analyses. To support this effort, a wide variety of tools are utilized to identify glitches [103], find statistical correlations [104, 105], and evaluate the long term behavior of the detector [84, 106, 107].

Another important tool to understand the impact of noise is the investigation of the loudest triggers from PyCBC Live [89], a version of PyCBC designed to analyze the data in low latency that uses a simplified version of the detection statistic introduced in this chapter. Daily result pages are produced for each detector, and allow investigators to easily understand what sources of noise are appearing as the loudest triggers in the background.

In addition to providing follow up of noise issues, visual inspection of spectrograms based on significant triggers from PyCBC Live as a part of standard data quality investigations has allowed gravitational-wave events to be identified even in cases when both detectors were not originally in an operational state or when the SNR in one of the detectors was below the threshold of the search. This was the case for GW170104 [108], GW170608 [109], and GW170818 [13].

Chapter 3

O2 Data Quality Efforts

As discussed in the previous chapter, one of the primary limitations on the sensitivity of a search pipeline to detect gravitational waves is the presence of instrumental transients in the data. These transients are relatively common, happening, on average, once every 100 seconds. If such transients can be identified and removed from the data stream, the overall sensitivity of the search will improve. This chapter discusses methods to identify and mitigate such transients, and their overall effect on the sensitivity of the PyCBC search during O2.

The main goal of these mitigation methods is to cut out loud recognizable transients with known causes. We expect this to primarily help searches for short signals, where signal consistency tests have less discriminating power. For long duration signals, we expect only marginal benefit as the long duration, combined with the chi-squared discriminator, allows triggers produced by glitches to be rejected. In addition to removing glitches, an additional goal of these methods is to ensure that they do not have the potential to reduce sensitivity to genuine signals.

As a part of the detector characterization group, I generated the data quality products and took a part in the instrumental investigations discussed in this chapter. I also led the group's efforts to follow up individual noise triggers and quantify the impact of this work on CBC searches. Further discussion of this work, along with how it fits into the broader goals of the LIGO detector characterization group, can be found in [110].

3.1 Data Quality Products

The primary methods to remove transients from the data for CBC matched filter analyses are a combination of data quality flags and gates. The choice of method used in each case is dependent on the amplitude and duration of the identified noise source.

3.1.1 Data Quality Flags

One of the key components of gravitational wave searches is understanding when data is available and ready to be analyzed. Initially, searches only utilize time for which a detector was in ‘observing mode’, as marked by operators at each site. This state means that the interferometer was locked, at an acceptable sensitivity as measured by the power spectral density, and free of any human interference. Despite this standard for observing, there are numerous cases when additional noise (both steady and transient) impacts the data quality. Data quality flags are designed to indicate time periods corrupted by instrumental noise. While the time periods that the detector is running are recorded at the time of data production (along with a small set of other data quality flags), the vast majority of flags must be generated post-facto.

Problematic noise at each interferometer is first identified in a number of ways. Monitors that are included as part of the LIGO web-based summary pages [106] such as the scattering prediction monitor [96] or overflow monitors can easily note the presence of instrumental artifacts. Other methods include visual inspection of plots of omicron triggers to identify any consistently loud noise sources or daily spectrograms. An example of a daily omicron plot showing a band of triggers can be seen in Figure 23. Once abnormal glitching is observed with these methods, additional plots of the behavior of various subsystems in the interferometer allows the cause of the glitching to be traced. The general strategy for tracking glitch behaviour is to attempt to identify a specific witness in the interferometer that follows the same pattern as the glitching. This generally relies upon the timescale and frequency structure of the glitching, such as repeating glitches at a particular frequency or glitching isolated to a unique time period but across a wide range of frequencies.

Problematic noise sources are also identified using the search methods themselves. As part of the PyCBC production analysis, lists of the triggers found in only a single

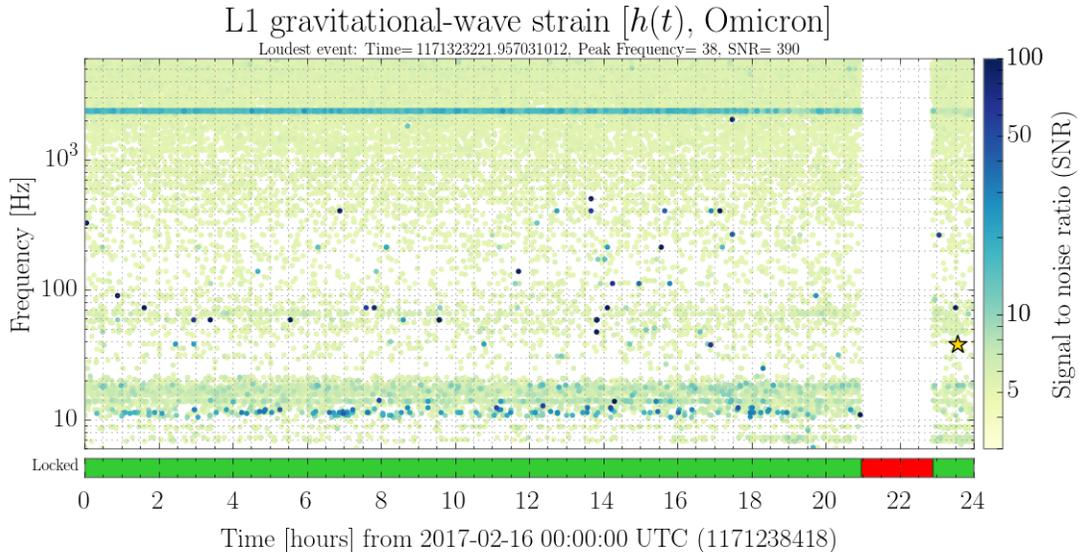


Figure 23: Omicron triggers produced for a typical day. Note the presence of specific frequencies with a large number of glitches, time variations in the glitch rate, and the presence of glitches with SNR values above 100.

detector (as opposed to in coincidence) are vetted by analysts. As noted in the previous chapter, the results of low latency searches, such as PyCBC Live [89], are examined for single detector triggers that may be indicative of problematic noise. An example of a daily output for PyCBC, showing the pipeline producing triggers near glitches can be seen in Figure 24. Using triggers from the searches helps further focus data quality efforts beyond just identifying excess power in the detector. Often the most problematic glitches are not the loudest glitches, but quieter ones that can better mimic the structure of a CBC signal. Triggers identified by these methods are considered due to instrumental causes as opposed to an astrophysical ones because they are not found in both detectors at the same time.

In order to produce a data quality flag for a problematic time period, the source of the noise needs to be witnessed by an auxiliary sensor, as opposed to only being seen in the gravitational-wave strain data. This standard is used to prevent a signal from accidentally being removed from the data stream. This means that many classes of glitch that are known, but do not have understood causes or any available witnesses, (e.g. blips) are not able to be removed via these methods.

The central goal of data quality flags is that they increase the overall sensitivity

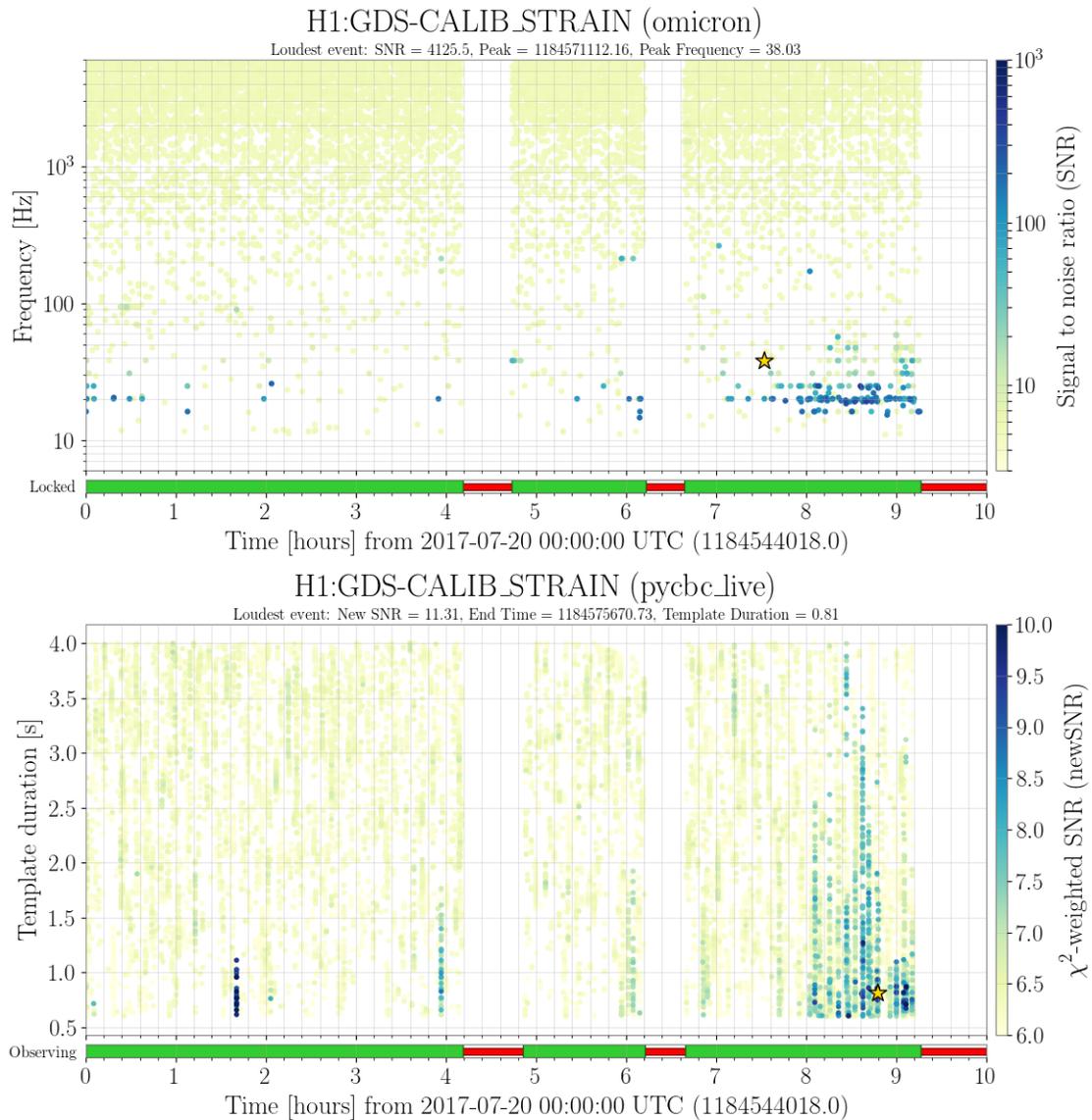


Figure 24: Loud glitching producing Omicron and PyCBC Live triggers. Each vertical “band” in the PyCBC graph corresponds to a single time that produced a trigger for a large range of templates, symptomatic of a glitch. In the period of 8-9 UTC, there are extremely loud glitches witnessed by Omicron which correlated with significant triggers from PyCBC Live. These types of plots serve as tools to investigate sources of noise producing triggers in the search and pinpoint times to focus on.

	LIGO-Hanford	LIGO-Livingston
Total Science Time	158.4 days	155.5 days
Time lost to Category 1	3.2 days (2.00%)	0.8 days (0.51%)
Time lost to Category 2	0.6 days (0.34%)	0.3 days (0.18%)

Table 2: Total deadtime from flags during O2. Only a small amount of the total available time is vetoed.

of the search by removing a large number glitches without significantly reducing the amount of time available to analyze. This is assessed during flag production by the metrics of efficiency and deadtime. The efficiency of a flag is the fraction of omicron triggers above a fixed SNR that are removed by this flag, while the deadtime is the fraction of time removed. If the ratio of efficiency to deadtime is high, this is considered a effective veto. Generally, ratios above 10 are required before a flag is used. However, many flags have ratios of over 100. If this ratio is low, the flag may still be removing the desired transient noise, but the time lost due to the flag is too high. Due to this issue, many noise sources can not be flagged despite the existence of a useful witness; this is the case for scattering, for example.

Time segments matching each flag definition during O2 are available from [111, 112] as part of the open data release. These flags are broken down for CBC searches into two categories, referred to as category 1 and category 2 flags. A breakdown of the total time lost to each category at each site in O2 can be seen in Table 2.

Category 1

Category 1 flags are designed to indicate times when the data recorded is not usable, despite the initial assessment that the detector had been in observing mode. The main figure of merit used to flag such a time is based on the non-stationary of the strain strain channel PSD. If the detector is not stationary enough for PSD estimation to be valid over long periods, then any resulting result from such a time would be invalid. This is often easily checked by examining the stability of the *inspiral range*, which is the average distance at which a detector could observe a BNS system (1.4 - 1.4 M_{\odot}) at a SNR of 8.

Category 1 flags mark extreme data quality issues; thus, these flags are utilized

by all analysis groups, including those undertaking long-duration analyses that are not typically sensitive to transients. A time period when the detector was sufficiently non-stationary to yield a category 1 flag, along with the impact of this data on the search, can be seen in Figure 25. Examples of category 1 flags include times where the calibration model is tuned incorrectly, when there are failures of the computing systems at the site, excessive anthropogenic noise such as fire alarms, and generically high rates of glitching. In some of these cases, such as calibration issues, the data may be recoverable, but usually a category 1 flag indicates that this time will not be analyzed.

Category 2

Category 2 flags are used to indicate short time periods when transient instrumental artifacts with a known coupling are present in the data. While such short transients do not impact PSD estimation, they still can create significant triggers in CBC searches. By indicating time periods when a known source of noise is present, the rate of false alarms due to noise in the background of CBC searches is reduced. This has the overall effect of increasing the significance of genuine astrophysical sources.

Since the goal of category 2 flags is to increase the sensitivity of searches, they are designed for maximum efficiency as opposed to capturing every case of noise. For these flags, the efficiency and deadtime statistics are of the most concern. An additional difference between category 1 and 2 flags is that category 2 flags are often search-specific. As different searches have different limiting sources of noise, a flag that may increase the sensitivity of one search would result in only time lost in another.

In the PyCBC search, category 2 flags are applied by removing triggers in the final processing stages of the pipeline. This allows continuous segments of data to be analyzed without known transients polluting the final results.

3.1.2 Gating

The final class of data quality product used in O2 is *gates*. These differ from flags in that the data during these times is altered to mitigate the related instrumental artifact as opposed to the time merely being flagged. These are used to indicate exceptionally loud transients that would impact the matched filter or whitening process. A common

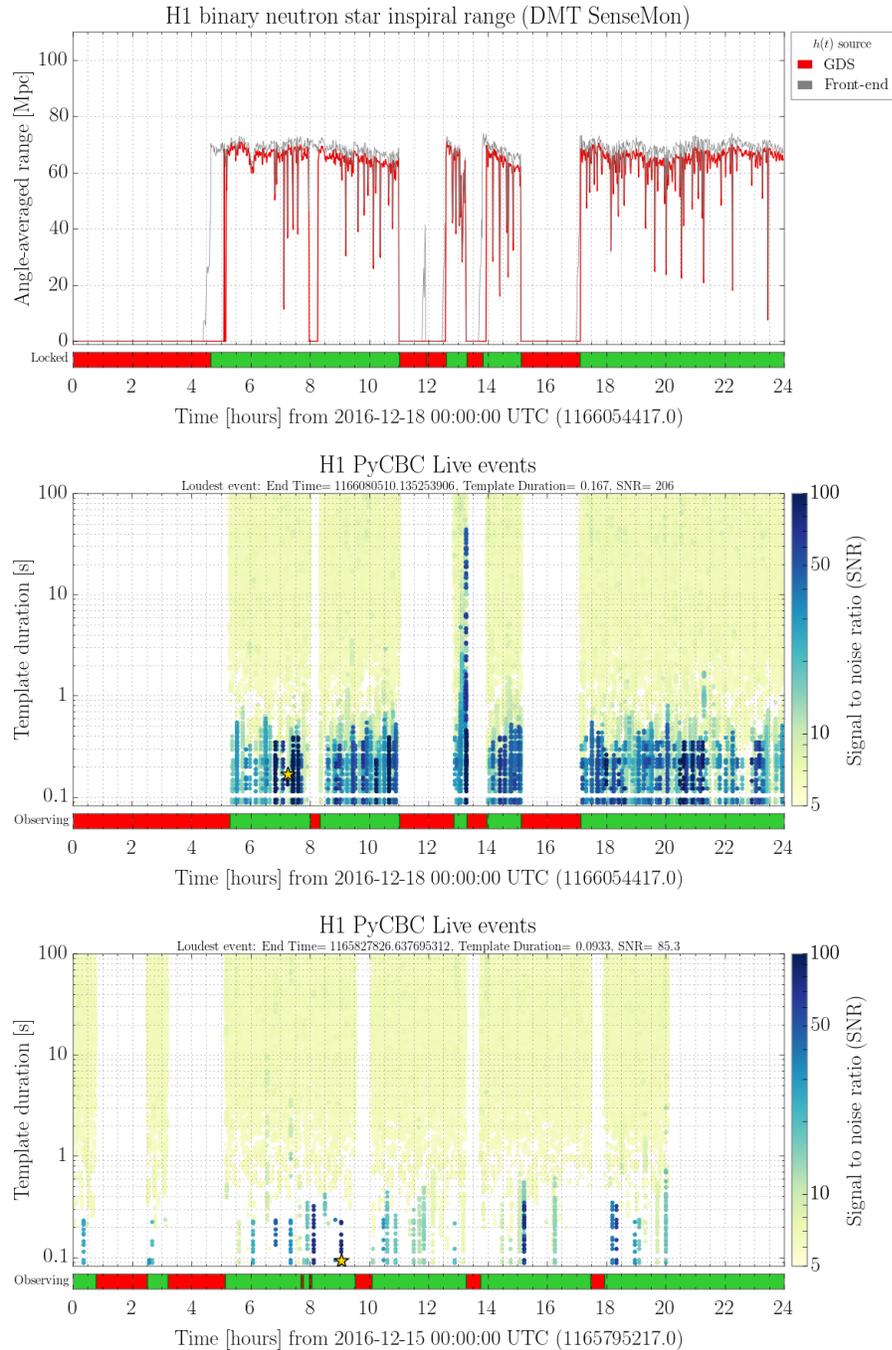


Figure 25: The stability of data during a category 1 flag due to electronics glitching at LIGO-Hanford. Top: the range during the day, showing a lack of stability, indicated by frequent drops and slow declines during locks. Middle: PyCBC Live triggers during the same day, showing that the glitching produced a high rate of high-SNR triggers for short duration templates. Bottom: PyCBC Live triggers from another day that had a stable range, showing a low rate of high-SNR triggers.

class of noise requiring gating is glitches that cause overflows of the control loop that provides active damping to the test masses. This overflow causes the servo to rail against its maximal movement, resulting in a large amount of additional broadband noise. In PyCBC, the gating is applied as described in [77], with a Tukey window. An example of a gate applied during a glitch coincident with GW170817 can be seen in Figure 26.

If a gate was not applied in this instance, the excess power would impact either the whitening of the data or the matched filter response. This in turn would lead to a large number of unphysical triggers in the search. With gating, care must be taken to ensure that the period zeroed is sufficiently short that PSD estimation is not impacted. If the gate is too long, the time window used in PSD estimation may be mostly zeroes. To account for this, if a designed gate would exceed 3 seconds in duration, the time period is instead flagged as a category 1 veto.

3.2 Instrumental Artifact Investigations

In order to provide beneficial data quality products to the analysis groups in a timely manner, sources of problematic glitches need to be quickly identified and mitigated. To identify problematic time periods, a wide variety of monitors are used; for example, statistical measures of correlations between the gravitational wave strain channel and auxiliary sensors using the tool called HVeto [104] and inspections of the loudest triggers in the search after initial analyses are completed are both used to identify instrumental artifacts. Once a problematic time window is recognized, additional investigations take advantage of the large array of sensors throughout the sites. A small subset of the available sensors is the physical environmental monitoring (PEM) array. In order to record how the physical environment of the site changes, this array consists of a large number of different types of sensors, including magnetometers, seismometers, microphones, thermometers, and cosmic ray detectors. The location of each sensor in this array, along with the specific aspect of the environment being monitored can be seen in Figure 27. Each of these routes was utilized to produce flags and improve the overall sensitivity and confidence in the results during O2. In this section I will focus on three specific instrumental investigations during O2 that led to the development of data quality flags and/or a fix at the site that resolved the

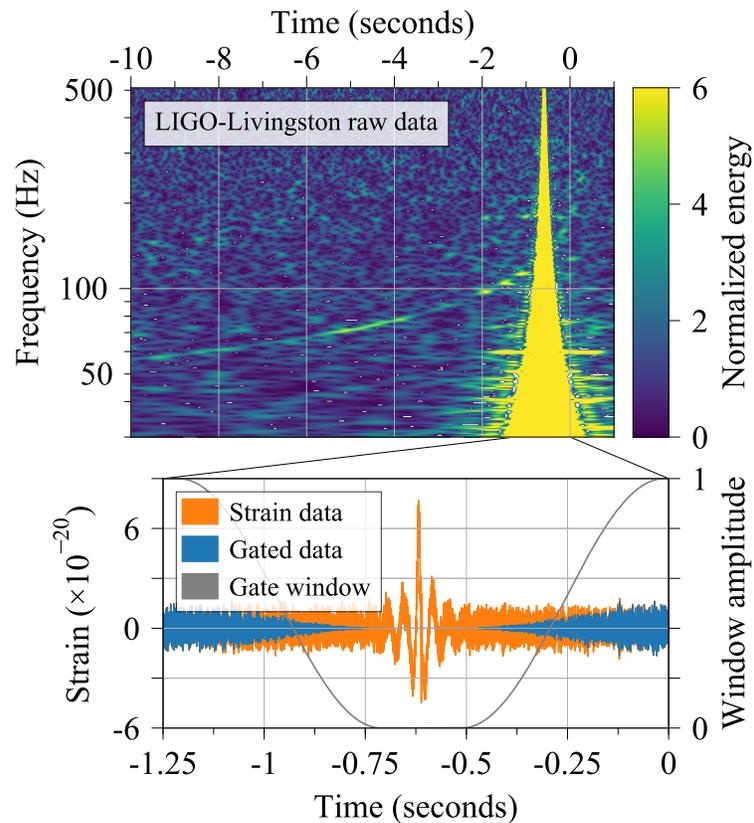


Figure 26: Diagram describing how the the large glitch associated with GW170817 was gated and modeled. This same gating procedure is used in the searches to address similar glitching. Specifically, a Tukey window is applied width the duration of the glitch zeroed, and then a 0.5 second roll off from unity before and after the glitch. After applying the gating window (gray), the gated data (blue) is zeroed at the time of the glitch. Adapted from [15].

identified instrumental issue.

3.2.1 Power Mains Transients

Power mains transients are an example of glitch class that was first identified from investigations into the background of the PyCBC search. Once the glitch was noted, spectrograms were produced of all auxiliary channels with excess power at the time of the glitch to manually identify any glitching correlated with this time period. Through inspection of these spectrograms, sensors monitoring the power mains at the site were shown to have the strongest correlation.

Focusing in on the power mains channels, studies were completed to construct an efficient veto that captured the identified glitches. An example of a time period linked to power mains glitching, with the threshold used in the veto can be seen in Figure 28.

While these glitches occurred only a few times per analysis period, they generally were found in the list of loudest background triggers. Similar glitching that is rare, but still problematic for a search, is difficult to identify by other means. Generally, singular glitches garner little attention if the glitch is not identified as problematic by a search pipeline.

3.2.2 Chiller Switch Transients

This class of glitches was initially identified with HVeto [104], an algorithm that looks for statistic correlations between glitches in the gravitational-wave strain and glitches in witness sensors, as a band of periodic glitches near 60 Hz. The winning channel from HVeto was related to the PEM subsystem, specifically a magnetometer. The large amplitude of the associated magnetometer transient allowed for a highly efficient veto to be developed. This veto was particularly important for burst searches, as this glitching was one of the loudest components of the background at the time.

Further follow up of this glitch class found that the glitching occurred in coincidence with temperature fluctuations of water used to cool lasers in the end stations. This correlation allowed for the source of the magnetometer noise to be isolated to a chiller in the EX station that periodically switched on and off. When this occurs, a loud glitch appeared in the gravitational wave strain channel. A comparison of the

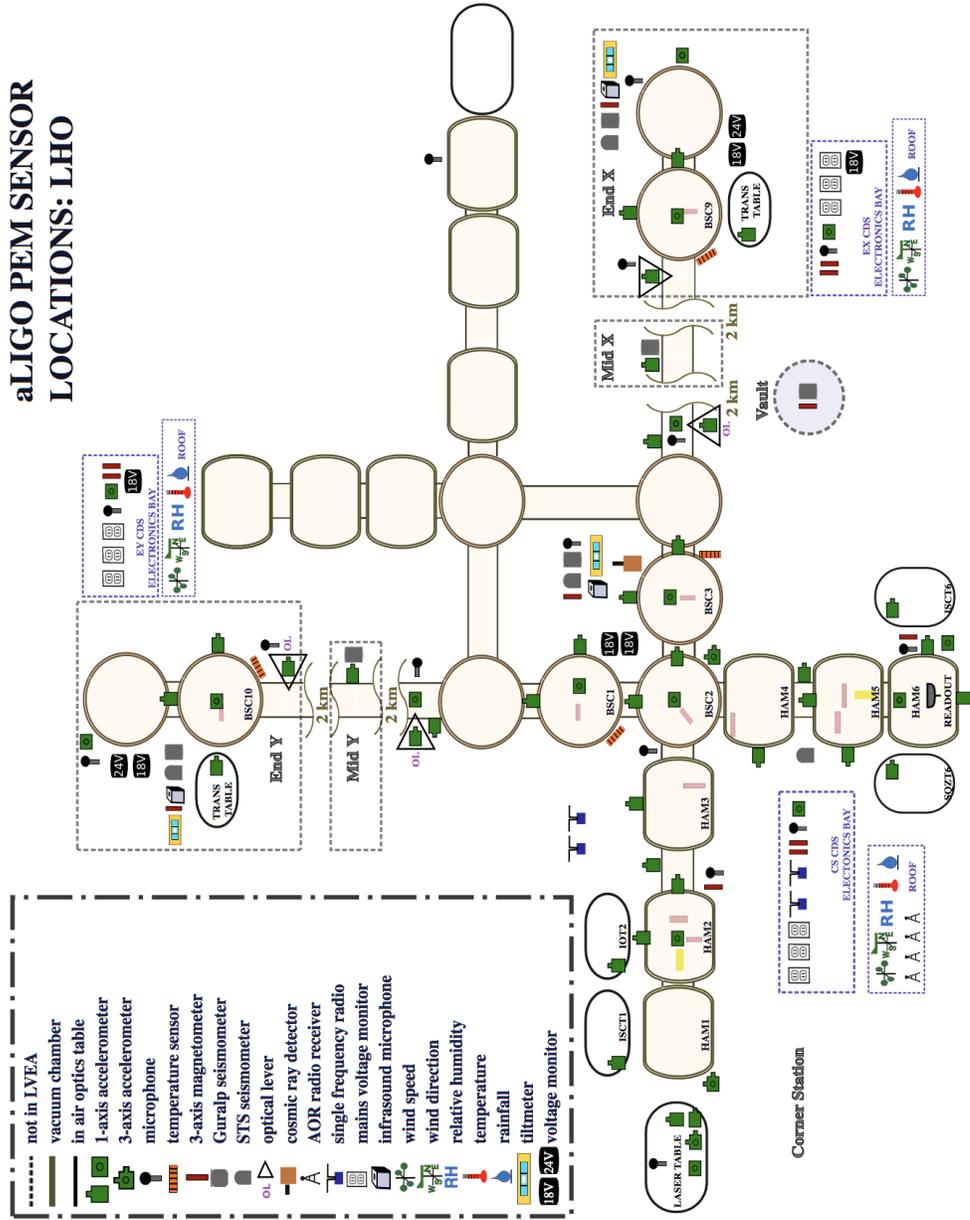


Figure 27: A diagram showing the type and location of the sensors that made up the physical environmental monitoring (PEM) subsystem during O2 at LIGO-Hanford. These sensors allowed for detailed followup and testing of how environmental noise couples into the interferometer. Reproduced from [113].

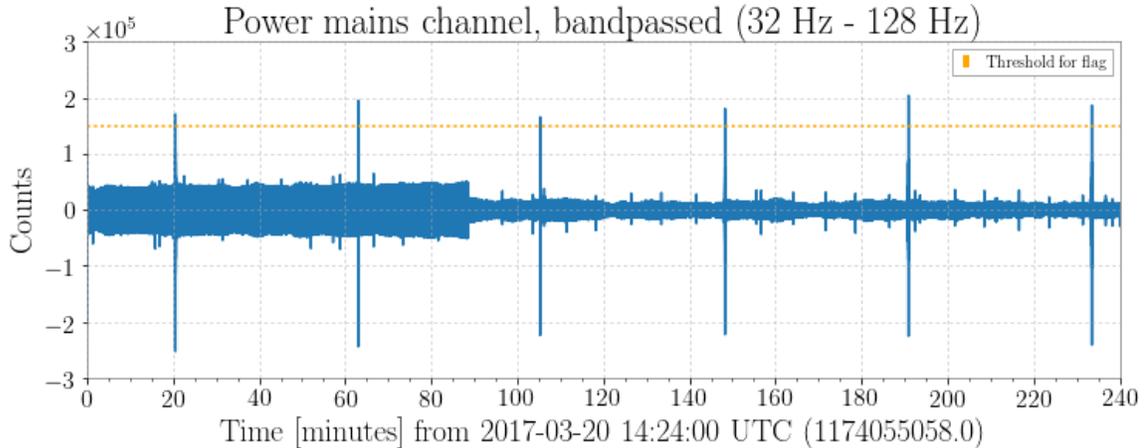


Figure 28: A plot of the timeseries of the power mains monitor channel at LIGO-Livingston used to produce a flag during O2. The threshold used to generate flag segments is indicated with a dotted line. Note the large spikes corresponding to periods of glitching.

magnetometer noise, the chiller behavior, and glitching in the strain channel can be seen in Figure 29.

Once the source of this noise was identified, work was completed at the site to change the grounding configuration of the electronics at the chiller. After tests of multiple configurations and checks to see if glitching remained, a successful fix was found that prevented further coupling between chiller power surges and the gravitational wave strain channel [114].

3.2.3 Optical Lever Transients

Optical levers (oplevs) are used in the control loop at each site to assist in alignment of the optics [115]. A laser beam reflecting off of the mirror is used in a servo loop to both damp angular motion of an optic, as well as record any residual movement. During O2, the lasers began to experience frequent bursts in the overall power, disturbing the oplev servos. The resulting angular disturbance of the mirrors couples into the gravitational wave strain channel. At the time, this class of glitch accounted for the vast majority of the loudest triggers in the PyCBC background [99]. Since there was a well understood witness of this glitching, an efficient veto was developed based each glitching oplev photodetector. At the worst point, oplev glitching occurred in spurts

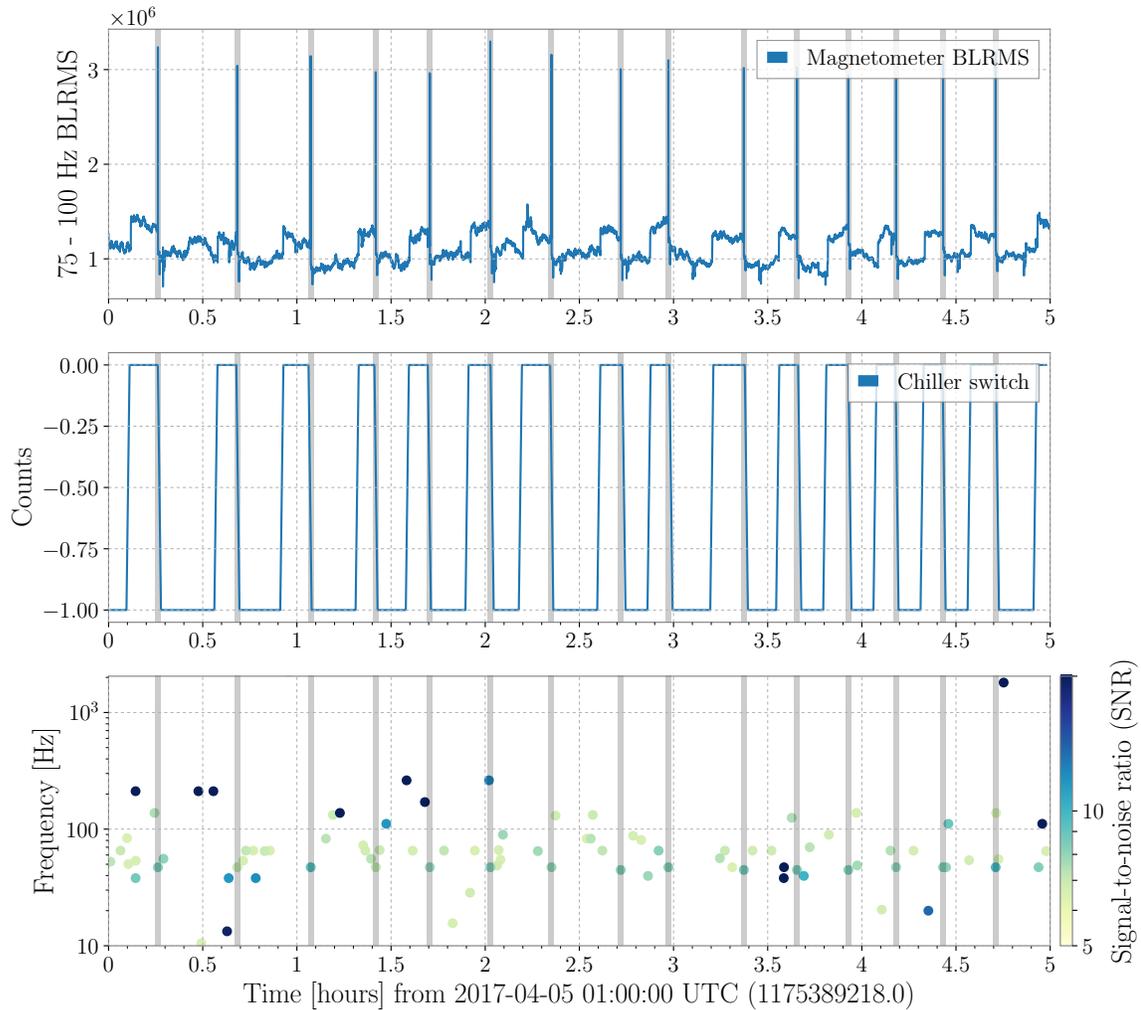


Figure 29: A comparison of how chiller glitches appeared in different monitors. The time periods corresponding to chiller glitches is shaded in each plot. Top: the band-limited RMS of a magnetometer channel in the end station that witnessed the glitching, and showing large spikes during each glitch period. Middle: the recorded value of the relevant chiller switch, showing a drop from 0 to -1 at the moment of glitching. Bottom: a plot of omicron triggers in the gravitational-wave strain channel, showing the presence of glitching near 60 Hz related to the chiller behavior.

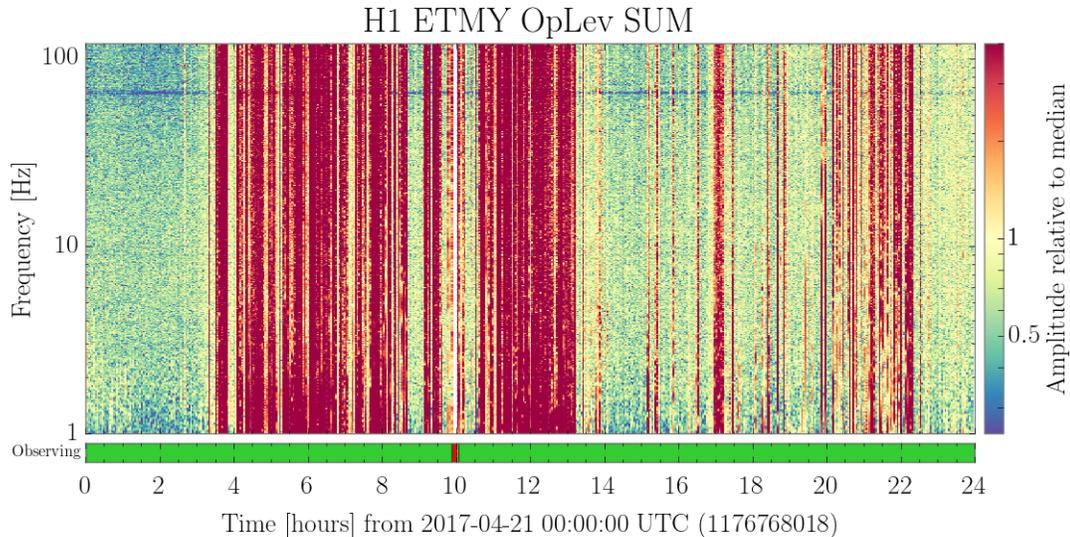


Figure 30: A spectrogram of a channel recording the output of an optical lever at LIGO-Hanford over the course of a day. The red bands correspond to periods of extremely loud, broadband noise.

once every 10 minutes throughout the entire day. A spectrogram of the witness sensor observing these power bursts can be seen in Figure 30.

Glitching was noted for multiple optical levers at LIGO-Hanford, which all required follow-up by commissioners at the site. After adjustments, further investigations were completed to test if any correlation between the gravitational wave strain and the oplev remained. An example of observed oplev glitching before and after successful mitigation can be seen in Figure 31. Here omicron triggers in coincidence with oplev glitches are colored red. After mitigation, the prevalence of excess power in the gravitational wave strain channel related to this particular oplev stopped completely.

In analyses completed after the end of O2, the most significant outlier from the cWB search for high mass BBH signals¹ was noted as similar in morphology to oplev glitches [116]. Furthermore, the time of the candidate trigger was during the calendar period impacted by oplev glitches. A comparison of spectrograms of the candidate trigger and a representative oplev glitch can be found in Figure 32.

Investigations into this time period noted that the candidate trigger did occur in

¹The high mass cWB search is a burst algorithm that looks for excess power in the data that increases in frequency with time.

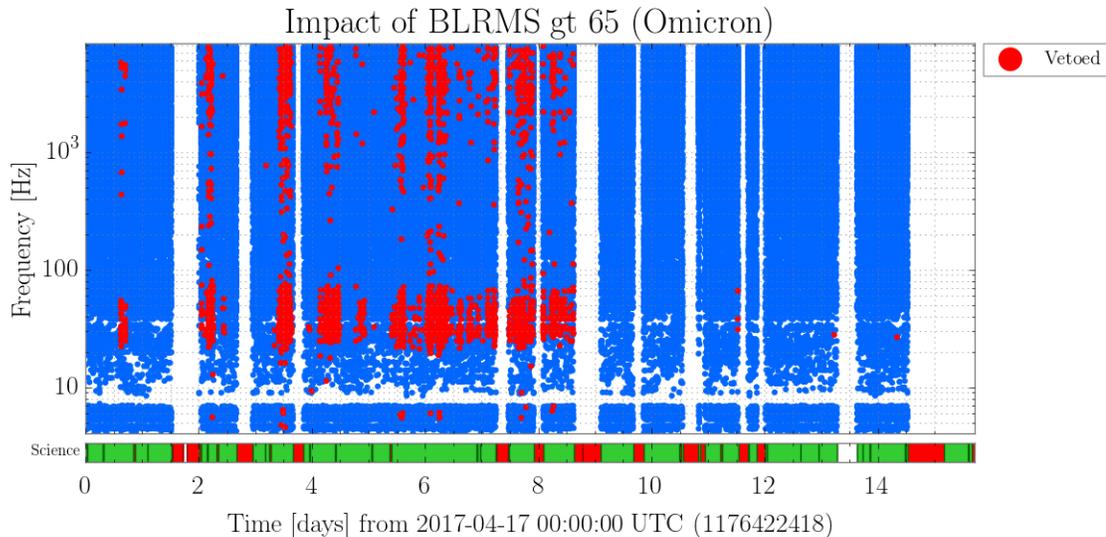


Figure 31: A plot of omicron triggers found in coincidence with the oplev flag. Note the high number of omicron triggers produced and the abrupt stop in the red region, showing when the problem was resolved.

coincidence with a glitch in an oplev at LIGO-Hanford. The sensor witnessing this glitching was the same sensor used to develop the original oplev flag. Although the glitching was not loud enough to meet the original flag threshold, when the power observed in the oplev and in the gravitational wave strain channel is compared, the ratio matches expectations from time periods vetoed. While the observed power was lower than a typical oplev glitch that produced loud triggers in search backgrounds, it was still loud enough for this chance coincidence to produce a marginal trigger.

This episode illustrates an additional use of data quality products, to validate a trigger to ensure that it wasn't caused by a known source of noise. Validation of candidate triggers is further discussed in Chapter 6.

3.3 Effect of Data Quality Work on the Searches

While each data quality product is designed with a different goal in mind, the end result should ideally be an increase in observing time, during which the detector is sensitive to gravitational wave events from the largest possible volume of space. To evaluate the effect of these products, I will focus primarily on their impact on the

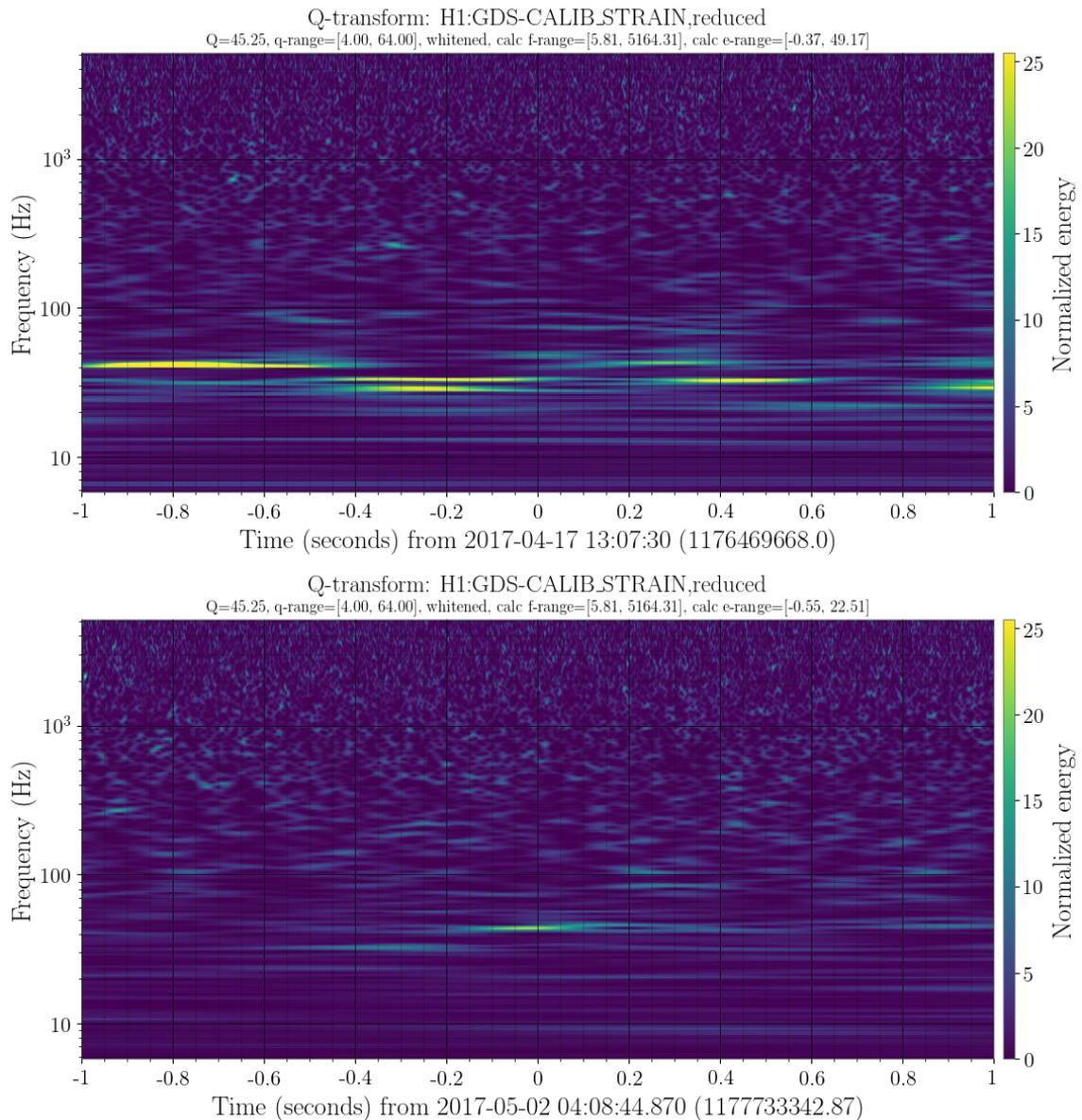


Figure 32: A comparison of the most significant IMBH trigger identified in O2 and a known time period impacted by glitching from optical levers. Top: an omega scan of a representative oplev glitch in May at LIGO-Hanford. Bottom: an omega scan of the loudest foreground trigger in the IMBH search. Note the similar morphologies and frequencies.

PyCBC search. PyCBC is one of the searches used to tune data quality products, and is hence is likely to achieve the maximal effect of this work. However, the improvements discussed in this section should apply broadly to other searches of LIGO data for gravitational-wave transients, such as those focused on identifying gravitational-wave bursts.

3.3.1 Data Quality flags

The most straightforward way to evaluate the impact of data quality flags is the number of triggers present in the search before and after the inclusion of data quality flags. By design, the total number of triggers should be reduced after the inclusion of flags. We can gain further insight by investigating the effect on templates of different duration. The background distribution for different template duration is known to drastically differ [83]. Specifically, the number of loud short duration triggers is much higher than long duration triggers.

Looking at an analysis of 5 days of coincident data during O2, we plot the total number of triggers for various template durations. A comparison for cases with and without data quality flags can be seen in Figure 33. As expected, the bin with the shortest duration has the most triggers at high values of the PyCBC ranking statistic. This bin has the largest improvement due to the inclusion of data quality flags.

Examining the number of triggers binned by template duration is ideal in the context of background reweighting, discussed in Section 2.3.1. With the inclusion of this feature in the detection statistic, an excess of triggers with short template durations does not affect the sensitivity of the search to long-duration signals. Improvements due to data quality flags were concentrated in the part of the template bank most in need of improvement.

Perhaps the best measure of the benefit of data quality flags is their effect on the overall “reach” of the search, as quantified in the volume-time product (VT). While flags decrease the overall time that data is available, removing periods corrupted by noise increases the range to which the search can identify a gravitational-wave signal. This leads to an overall increase in VT when flags are applied.

We assess the VT to which the search is sensitive using “software injections” of a simulated astrophysical population. Each individual injection is analyzed by the search and assigned a significance. We can then calculate the maximum distance the

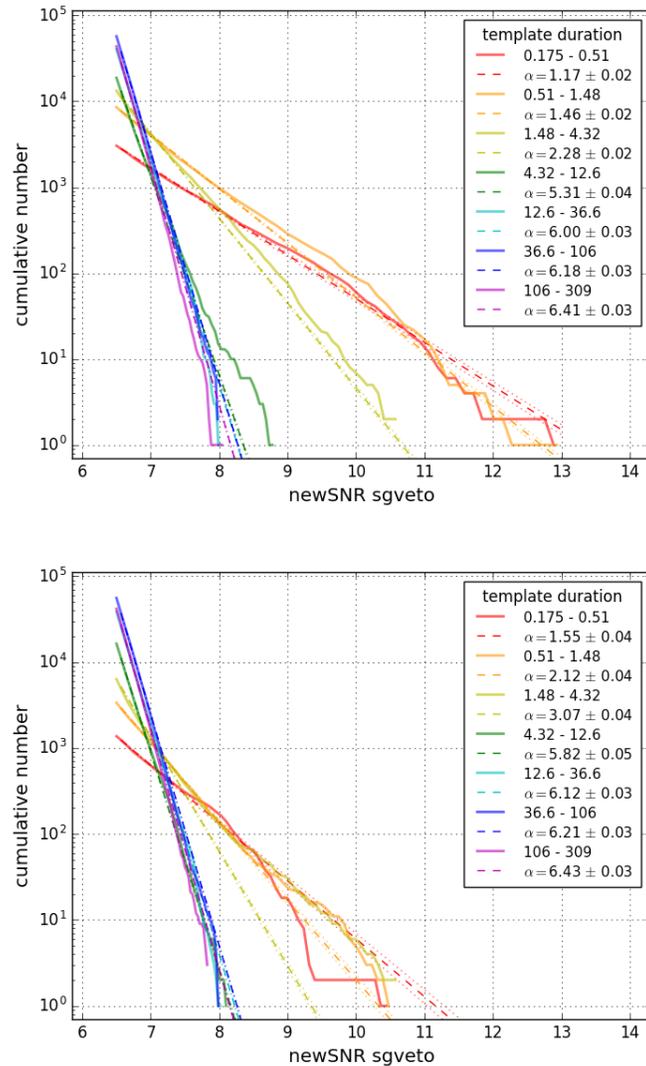


Figure 33: Plots of the number of single detector triggers before and after the inclusion of data quality flags. Triggers are binned based on the duration of the template and plotted separately. For each duration bin, the solid line corresponds to the cumulative number of triggers, while the dotted line is an exponential fit of the data. Left: single detector triggers produced by an analysis without data quality flags. Right: single detector triggers produced by an analysis with data quality flags. Note the large drop in total number of short-duration triggers after the inclusion of data quality flags.

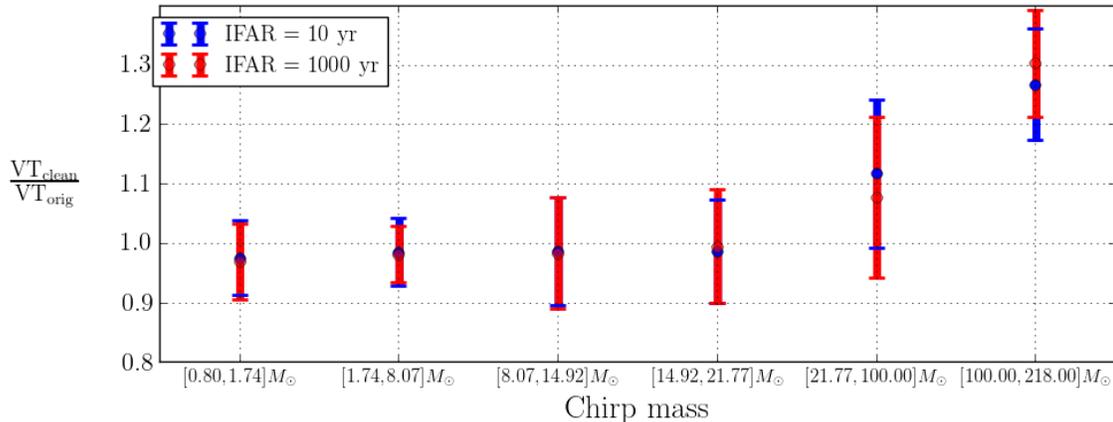


Figure 34: A plot of the ratio of volume-time (VT) that the PyCBC search was sensitive to before and after the inclusion of data quality flags. These flags had negligible impact for low mass templates, but up to a 30% increase in VT for the highest mass templates.

search is able to recover a signal below a chosen FAR value. This limit, combined with the total time analyzed by the search, is used to compute the total VT of the search at a chosen FAR.

We analyze periods of data during O2 with the PyCBC pipeline, with and without the category 1 and 2 flags. After calculating the VT of each analysis, we plot the ratio of the two values in Figure 34. Error bars are the maximal error if the measurements were fully independent, and hence overestimate the actual error by a significant amount.

Similar to the result seen in Figure 33, the largest impact is for the largest chirp mass templates (which have the shortest duration). At low chirp mass, there is a minimal loss in VT, indicative of the case where data quality flags have only marginal gains in the distance that the search can observe a signal. This, combined with a loss in analyzable time due to data quality flags, leads to a reduction in VT. As the effect of data quality vetoes differs so drastically across the bank, it may be possible to increase the sensitivity of the search by only applying vetoes to parts of the PyCBC template bank where these vetoes lead to improvements.

Comparing this result with [83], there is a reduced benefit from data quality flag

during O2 than during O1. This difference is related to numerous search improvements since O1. Specifically, the inclusion of template-duration-dependent background reweighting in O2 has made it so that data quality issues only affect short-duration parts of the bank, while coherence tests, along with the sine-Gaussian discriminator, reduced the likelihood of a glitch producing a significant trigger. If the ranking statistic from O1 were still being used, there would have been a substantially greater benefit to applying data quality products to the search.

3.3.2 Gating

When a long-duration signal occurs, there is a non-trivial probability that a glitch will overlap with the signal. In this scenario, the detection statistic of the trigger will be reduced; a short-duration glitch will add excess power to one of the chi-squared bins, even though the signal itself may match the template quite well. While a typical ($\text{SNR} \approx 8$) glitch does not heavily bias this signal consistency test, this that could happen for a glitch with high enough SNR; we apply gates to remove a small amount of data around the high SNR glitch, allowing the long signal to be recovered cleanly.

In contrast to data quality flags, the main goal of gating is not to increase the VT of the search. The rarity of glitches that requiring gating means that the number of triggers removed by gates does not have a significant impact on the integrated volume of the sensitivity of the search. Furthermore, the overall effect of ex post facto gates developed in O2 was reduced in the PyCBC search by the inclusion of features in PyCBC [77] that result in a similar set of times being automatically gated in the data conditioning step of the search. Instead, the main benefits of the gating products discussed in this chapter is to ensure that loud glitches do not prevent the recovery of long-duration signals, and to ensure that no unphysical triggers are produced due to these glitches.

An example of the re-weighted SNR recovered for an injection overlapping an overflow (that was not gated) can be seen in Figure 35. Without gating the glitch, the injection would not be recoverable. This is true no matter where in the signal the glitch occurs. While an exceptionally loud event may still be recoverable ($\text{SNR} > 20$), a signal at a more likely SNR value ($\text{SNR} 10\text{-}15$) would not be. The inclusion of gates was especially important for the estimation of the significance of GW170817. The presence of an extremely loud glitch that occurred within one second of merger

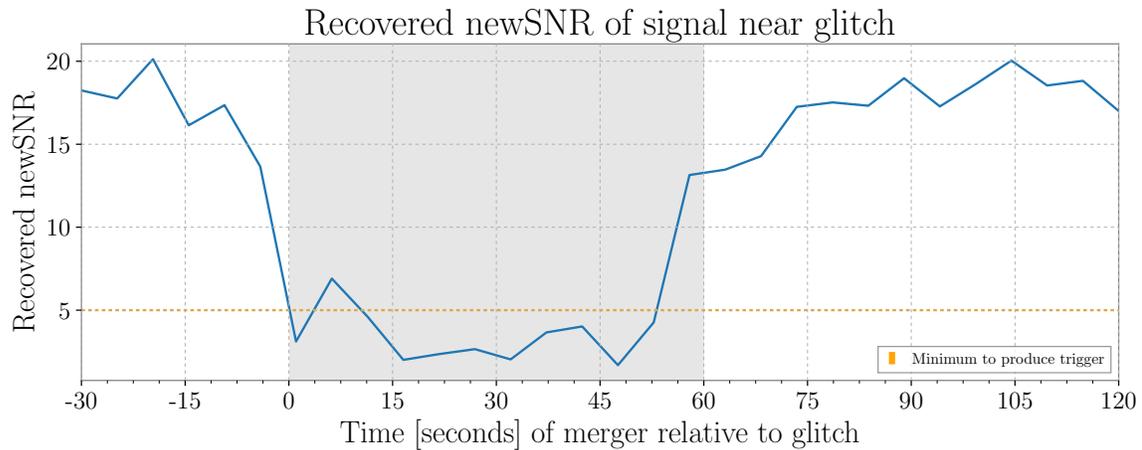


Figure 35: The change in recovered newSNR due to the presence of an overflow without gating. A binary neutron star merger template is injected at various time around the glitch and then recovered by PyCBC. In all cases, the optimal SNR was 20. The region where the template directly overlaps the glitch is shaded. The reduced recovery values in this region are due to signal consistency tests down-weighting the trigger. This shows the importance of gating such glitches; note that the effect continues through the entire length of the template.

for this signal, as shown in Figure 26, prevented this event from being identified as a coincident event by the search. Instead, GW170817 was first identified as an extremely significant signal in only the LIGO-Hanford detector [15]. Manual inspection of the LIGO-Livingston data revealed the presence of both the signal and the loud glitch that prevented recovery. Searches of the data later used the gates discussed in this chapter, which allowed GW170817 to be identified as a highly significant, coincident event.

Chapter 4

Impact of Gravity Spy Glitches on PyCBC

While it has already been discussed how glitches in the data can impact the sensitivity of the PyCBC search, this has only been in the context of removing time periods corrupted by glitching from the search. In this chapter, I will instead focus on how we can utilize knowledge of individual glitch classes to study periods impacted by such transients. As these periods are known to produce triggers in the search, this data is only usable if we can effectively differentiate a trigger caused by a glitch from a trigger related to an astrophysical signal.

The importance of this approach is highlighted by the case of GW170817. The detection of GW170817 was initially complicated by the presence of a loud instrumental noise transient that had to be removed before analysis could be completed [15, 117]. When a candidate signal is identified by the search pipeline, rigorous studies can be undertaken regarding the time in question to understand if the trigger is related to instrumental causes [13, 99, 118]. Tests to quickly evaluate the data quality around a candidate signal were also routinely completed as part of the O2 EM follow up process [119]. For GW170817, the presence of the glitch was found to not be correlated to the gravitational-wave signal due to the vastly different time-scales involved (the signal was clearly visible in spectrograms for over a minute, while the glitch lasted less than a second), as well as due to the high SNR of the gravitational-wave signal. However, the differentiation between glitch and signal is not as clear in most cases.

This chapter demonstrates how the most problematic glitch classes affect matched

filter searches for gravitational waves from CBC sources in order to improve our ability to better differentiate glitches from gravitational waves. In order to identify periods that are corrupted by known classes of glitches, we take advantage of Gravity Spy, a machine-learning based image classifier [102]. I examine the response to these glitches by the PyCBC search pipeline, one of the pipelines used to find CBC signals with aLIGO [77, 87, 120]. I show how these glitches can mimic waveforms from astrophysical sources of gravitational waves in Section 4.3 and quantify the likelihood of a given glitch from each glitch class to create a highly significant trigger in PyCBC in Section 4.4. I then propose a method in Section 4.5 that takes advantage of Gravity Spy classifications to more accurately measure the impact of these glitch classes on the estimation of false alarms in the background of PyCBC and to evaluate if candidate signals are consistent with the expected response of the search to a population of glitches. These studies will be detailed in a paper currently in preparation, of which I am the lead author.

4.1 PyCBC Trigger Set

In order to directly examine how the PyCBC search couples with the detector data, in this chapter we will be taking a close look at the full set of triggers produced in the search. In this chapter we utilize triggers generated from the the PyCBC searches associated the GWTC-1 catalog [13]. A more detailed look at the most significant triggers from this set is in Chapter 6.

Throughout this chapter, we will be looking at both single-detector and coincident triggers. In this situation, these terms are not exclusive, such that the term ‘single-detector’ refers to all triggers from a single detector, including those that may be later identified as coincident. In each case, however, the detection statistic used differs, as is discussed in Chapter 2.

Single-detector triggers are ranked using only SNR, the Chi-squared discriminator, and the sine-Gaussian discriminator. These signal consistency tests are all based on the quality of the data in a single detector, and provide the most direct look at how the data can mimic a CBC template.

Later in this chapter, coincident triggers are also considered. In this situation, the impact of background dependent reweighting and the time-phase consistency tests are

included. Background dependent reweighting is especially of interest in this chapter, as the ability of a specific glitch class to mimic a CBC templates is strongly dependent on the parameters of the template.

While Chapter 3 detailed how problematic noise sources can be identified and vetoed, that process is parameter-agnostic, meaning that the pipeline is not capable of applying different data quality products for different template parameters. In this chapter, however, we will carefully examine how different glitch classes impact different parts of the bank.

4.2 Gravity Spy Classification

In order to classify glitches in this chapter, we use Gravity Spy[102, 121], a machine learning based classification tool that utilizes citizen science efforts. Gravity Spy has been used by the detector characterization group in efforts to quantify glitch rates and identify large sets of similar glitches [122, 123, 124, 125, 126, 127, 128, 129, 130, 131], as it can quickly and accurately identify common classes of instrumental artifacts in the detector. These studies have generally been aimed at understanding detector performance and the sources of these glitch classes. However, Gravity Spy is also useful for studies such as those discussed in this chapter that are focused on searches for gravitational waves, as it provides a method to develop an initial dataset of glitches to investigate that is independent of the ways the data is analyzed by the searches. In this section I will detail the relevant data selection process for the Gravity Spy pipeline including possible selection effects relevant to this study. Full details on the classification methods for Gravity Spy can be found in [102].

At its core, Gravity Spy is an image classifier based on convolutional neural network methods [132]. Before this is possible, time periods containing glitches must be identified and the relevant detector data translated into an image format that the neural network can process. To identify a glitch, the Gravity Spy pipeline takes advantage of the Omicron pipeline [103]. Omicron uses a set of sine-Gaussian wavelets to identify excess power events in detector data. Any excess power event with a SNR above 7.5 is reported to the pipeline. Once a time is selected, the time series data is transformed into a spectrogram using an Omega Scan [69]. This representation provides the input that both the machine learning classifier and citizen scientists will

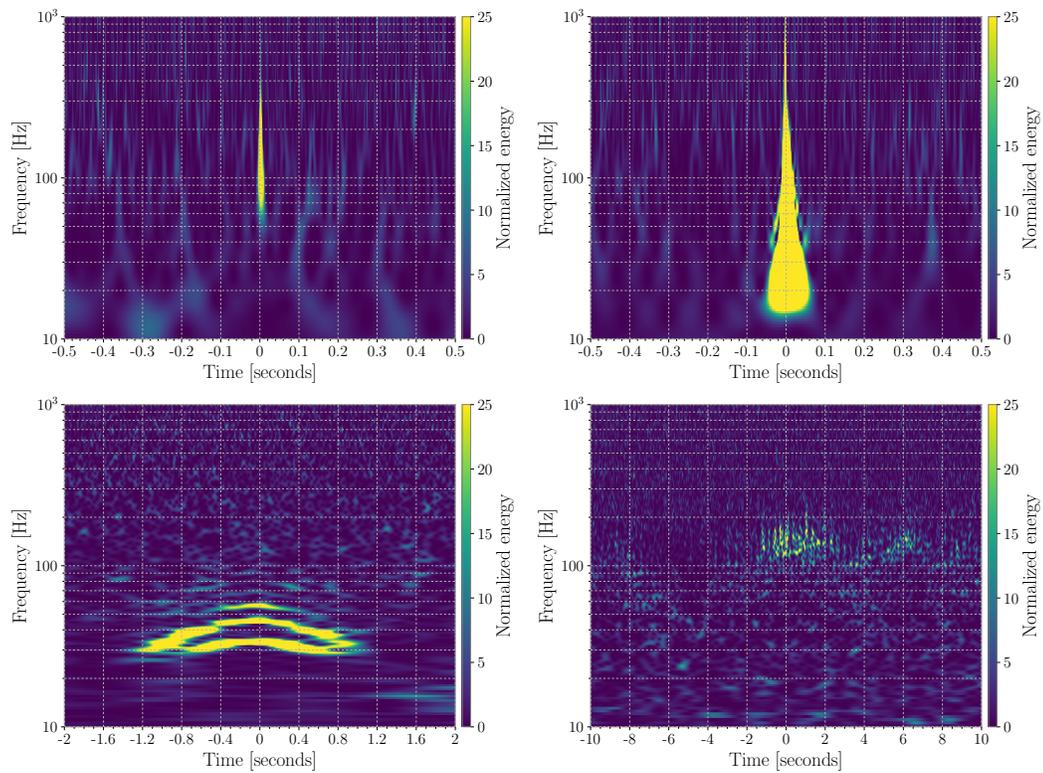


Figure 36: Omega Scans [69] of the four problematic Gravity Spy glitch classes discussed in this chapter. The glitch classes blips (top-right), koi fish (top-left), scattering (bottom-left) and scratchy (bottom-right) are highlighted here due to their known impact on the PyCBC search. Note the diversity in duration and morphology for each glitch class.

use in their classification efforts. This glitch image is fed into the classifier and a confidence score ranging from 0.0 to 1.0 is given for each category. The total sum over all categories is 1.0, with the highest numeric value representing the most likely classification.

The SNR threshold from Omicron is used to ensure that a clearly defined glitch will be visible in the Omega Scan representation. In the context of using these classifications for understanding the effect of glitches on search pipelines, this does provide some bias, as noise sources problematic to the searches may not meet this threshold. This consideration is especially important for long duration signals that are not expected to be identifiable in this representation.

Another important consideration is the set of possible classifications that Gravity Spy can provide. While there is a “None of the Above” and “No Glitch” class that the pipeline can utilize, the classification is mostly limited to predetermined classes from a training set [121]. Therefore, if a glitch unknown to the pipeline is classified, the result has a much higher chance of being incorrect, and hence contaminating the glitch set. To counteract this issue, we set a minimum confidence of 0.95 for all glitch classes to reduce the risk of contamination.

Of the classes Gravity Spy has in its training set, we will focus on four in this chapter: “blip”, “koi fish”, “scattered light”, and “scratchy”. Omega Scans of representative examples of each of these glitches can be seen in Figure 36. These four are chosen as they have been previously identified as problematic for searches for gravitational waves from compact binaries [13, 83]. These classes are also some of the most common glitches in the LIGO detectors, allowing for a broad statistical study. Finally, each of these classes has yet to be completely mitigated via instrumental means in the LIGO detectors. Due to this, it is likely these glitch classes will be present in future observing runs and continue to limit the sensitivity of searches. Additional details on the sources of these glitch classes can be found in Chapter 2.

4.3 How Different Kinds of Glitches Mimic Traits of Signals

PyCBC signal consistency tests have been shown to discriminate between glitches and astrophysical signals [83, 86, 77, 87, 74]. However, the wide range of template parameters included in the search [82], combined with the wide variety of instrumental

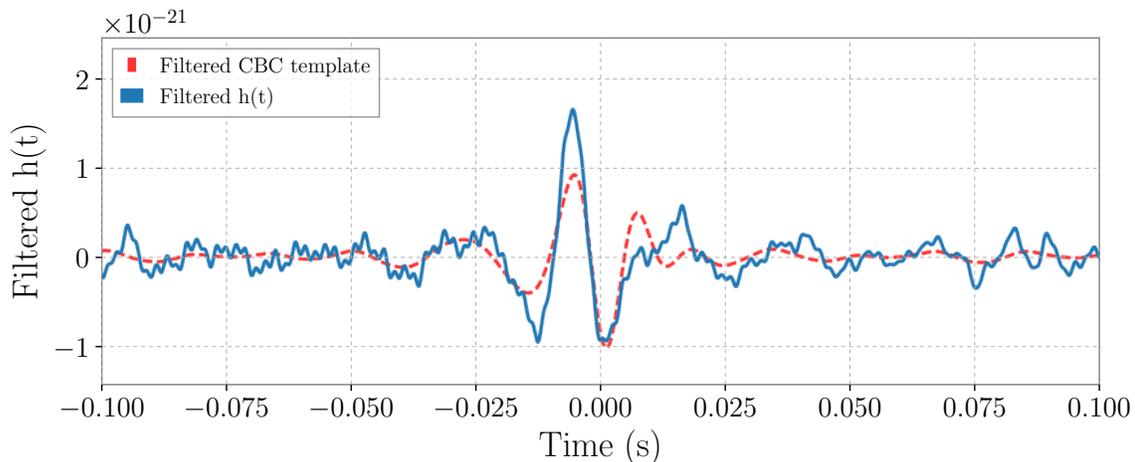


Figure 37: An overlay of a timeseries of a representative CBC template and detector data around a blip glitch. Both of the timeseries have been filtered with bandpass filters to isolate the most sensitive region of 35-500 Hz. This visualization serves to show the similarity between a blip glitch and a CBC template after the response of the detector is considered.

artifacts, means that this discriminating power is not uniformly effective across the entirety of the search parameter space [83].

One of the glitch classes where this concern is easily demonstrated is the blip class. When plotted against a timeseries of the data around a representative blip glitch, the match between a blip and a particular template is apparent, as shown in Figure 37. As visible in Figure 37, the characteristics of blip waveform are a few short, loud cycles, similar to a sine-Gaussian pulse [63, 86, 92]. As the detector is less sensitive at lower frequencies, a template that reaches merger by 100 Hz will be in the observable band for only a few gravitational wave cycles and qualitatively match the model for a representative blip. Such templates correspond to some of the most massive systems in the PyCBC template bank, and have been noted as one of the limiting sources of noise for searches for gravitational waves from high mass binary black holes [133, 70].

In order to understand how the PyCBC search responds to the classes of glitches we are focusing on, we analyzed short segments of data around glitches noted by Gravity Spy as from a specific glitch class and recorded all PyCBC triggers that met a minimum detection statistic value. Multiple seconds of data after the glitch were included to ensure that triggers intersecting the inspiral component of the waveform

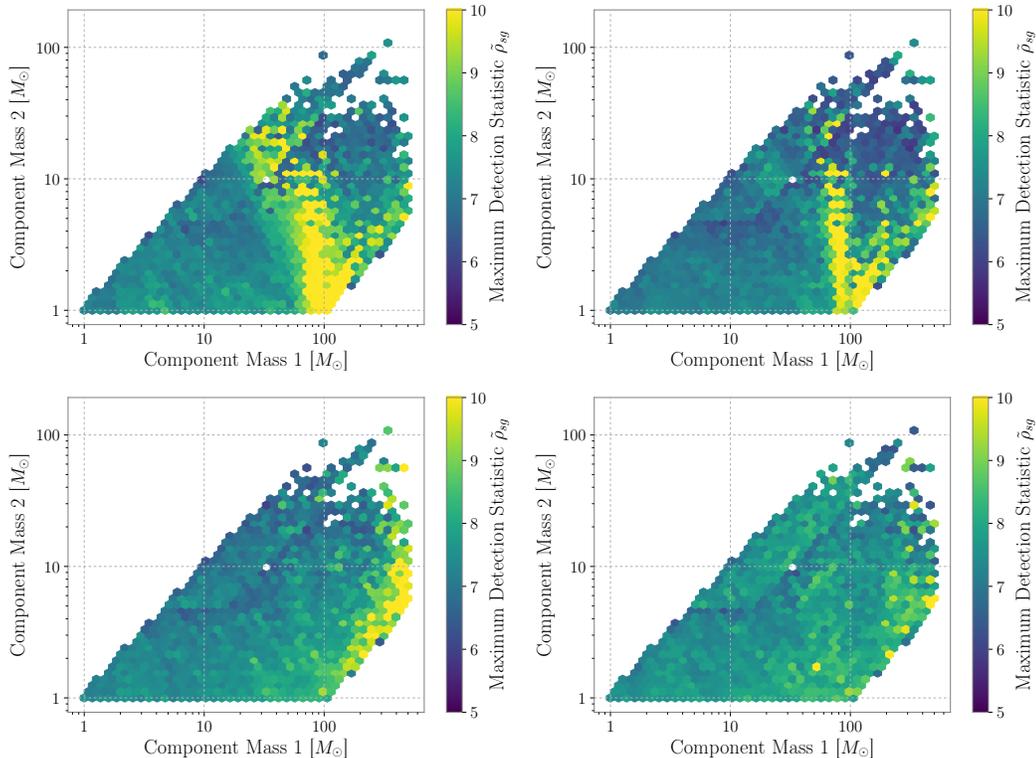


Figure 38: Histograms of the maximum sine-Gaussian SNR, $\tilde{\rho}_{sg}$ recorded for triggers found in different parts of the PyCBC template bank parameter space during time periods corrupted by each glitch class. Triggers are plotted based on the component masses of the relevant templates. The glitch classes of blips (top-right), koi fish (top-left), scattering (bottom-left) and scratchy (bottom-right) show different maximum detection statistics overall as well as varying number of triggers.

were correctly recorded. We then plot these triggers based on their component masses to understand which parameters are more likely to record loud triggers due to the glitching. This process is completed for each of the four classes of glitch focused on in this chapter. Results of this visualization for single detector triggers from the LIGO-Hanford detector can be seen in Figure 38. Results for triggers from the LIGO-Livingston detector are qualitatively similar.

One of the more interesting features that appears in the projection onto the template bank is the presence of a tower of loud triggers at $m_1 = 100M_\odot$ and a high mass ratio seen in the blip example. This particularly stands out as the loudest triggers are not found by the templates with the largest component masses. Investigating

this further, we noted that this feature corresponded to the shortest templates in the template bank with maximally anti-aligned spins. In the O2 PyCBC template bank, a minimum duration of .15 seconds was set for all templates [82]. When the compact objects have spins that are maximally anti-aligned, this minimum duration corresponds to a maximum total mass of $M_{tot} \approx 100M_{\odot}$. Furthermore, this configuration leads to a more rapid frequency evolution in comparison to aligned templates with the minimal template duration. We also note that the loudest single detector triggers from the search were related to blip glitches. Triggers related to koi fish glitches (Figure 38 top-right) show qualitatively similar behavior to blip triggers, with the largest probabilities found at high masses and spins.

Examining the scattered light triggers shown in Figure 38 (bottom-left), we observe the behaviour that was initially expected for blips. Namely that the loudest triggers correspond to the templates with the largest component masses. These triggers also correspond to the minimal template duration for the bank, but with maximally aligned spin instead of anti-aligned. Triggers related to scratchy glitches (Figure 38 (bottom-right)) show no clear clustering of loud triggers in this representation, indicative of the broad range of parameters for triggers during scratchy periods.

While this visualization method does demonstrate which glitch classes contribute to the loudest background triggers observed by PyCBC, this contribution could be from a single outlier in a given glitch class. In order to assess how pervasive the impact is of each glitch class, the behavior of the entire glitch population needs to be assessed. If a glitch class is capable of producing significant triggers during every recorded instance, this is much more likely to produce a coincident trigger in the search than an individual loud outlier from a single detector.

4.4 Probability of Creating Triggers

In order to present the likelihood of a common glitch producing a significant trigger in a way that can be used in both automated and human based follow up of candidate triggers, we next attempt to assess the likelihood of a single glitch producing a trigger. In this section we calculate the probability of a given glitch in a specific class creating at least one trigger of a fixed detection statistic. We then further plot this information against total mass and effective spin for a given system. This method prevents a single

glitch from dominating our statistics, and helps us understand the population more fully.

For each class of glitch in our data set, we test to see the likelihood that a single glitch from this class produces a trigger above a detection statistic threshold (in this case $\tilde{\rho}_{sg} > 7.0$ in each bin) in each part of the template bank parameter space. This value was chosen because a trigger of this detection statistic value combined with the minimal value possible in the other detector ($\tilde{\rho}_{sg} \approx 5.5$) would result in a candidate with a combined network detection statistic of $\tilde{\rho}_{sg,net} > 9.0$, which is sufficient to separate itself from the background. We choose to bin the parameter space spanned by the O2 PyCBC template bank uniformly in both χ_{eff} (the effective spin of the system) and $\log(M_{total})$ to help account for the reduced density of templates at high M_{total} .

In each template bank bin, we count the fraction of glitches in a given class that produced at least one trigger above the chosen threshold that have parameters consistent with the bin in question. This value is used as the probability of creating a trigger in that template bank bin. Since each bin probability is calculated independently, the total probability summed over all bins is not bounded by 1. In fact, it is possible for a given glitch instance to create triggers recorded in multiple bins. This allows the full extent of the glitch overlap with the template bank to be recorded. The results of this study can be seen in Figure 39 for the four glitch classes considered in this chapter and for all of the template bank bins.

For blip glitches, (Figure 39 top-left) we see clustering of the most likely part of the parameter space near maximal total mass and anti-aligned spin, but with a low probability of producing a trigger. Outside of this cluster the probability is negligible. This contrasts the results shown in Figure 38 (top-left) that suggested a more pervasive overlap. Combined, these results show that while singular blip glitches can produce significant triggers in the search, the overall population is far less susceptible. It is important to note that blips are one of the most common classes of instrumental transient found in the detectors, so the low probability shown here may still impact the sensitivity of the search.

As shown in Figure 39 (top-right), koi fish glitches have a similar profile to blip glitches, producing very few triggers and only at high masses.

Scattered light glitches impact short duration highly aligned spin templates, as

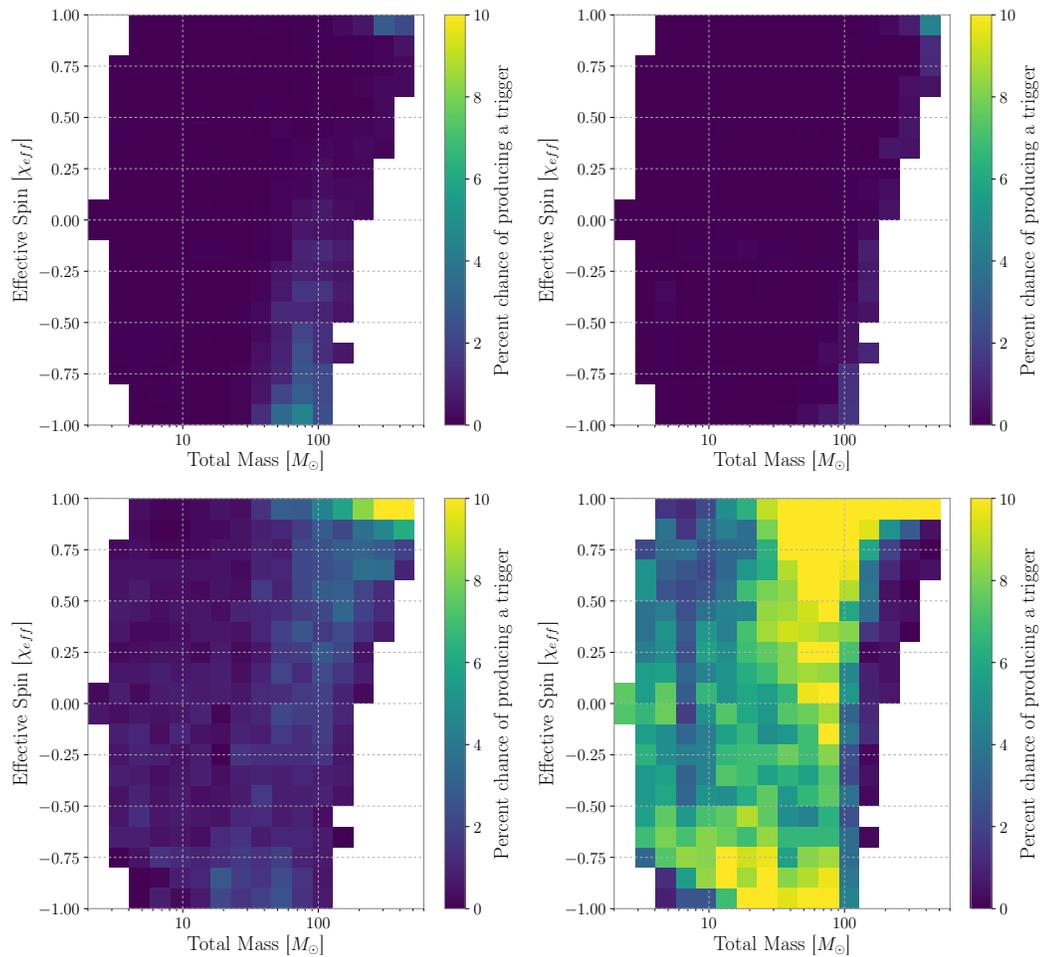


Figure 39: Probability of producing a trigger above sine-Gaussian SNR of 7.0 in specific regions of the template bank parameter space for each glitch class. blips (top-right), koi fish (top-left), scattering (bottom-left) and scratchy (bottom-right) each show maximum probabilities in different parts of the parameter space.

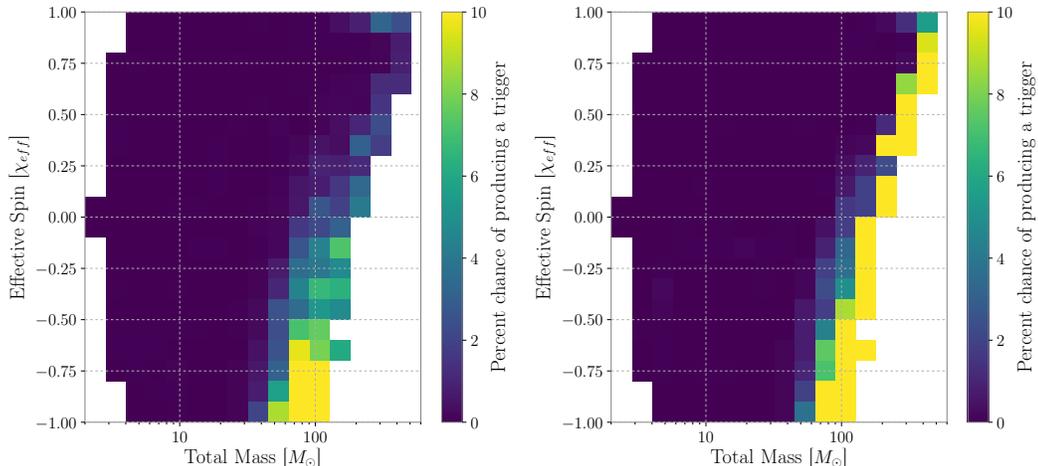


Figure 40: Probability of producing a trigger above a reduced SNR ($\tilde{\rho}$) of 7.0 for blip (left) and koi fish (right) classes. Note the large difference with the probability plots which use the sine-Gaussian discriminator as a component of the detection statistic. The scattered light and scratchy classes (not shown) saw only minimal changes.

shown in Figure 39 (bottom-left). These templates experience a ‘hang-up’ effect [134], resulting in a template that can match the arch-like morphology of a representative scattered light glitch.

Scratchy glitches (Figure 39 bottom-right) overlap best with templates that have total masses of 10-100 M_{\odot} and are highly spinning. Further investigation into this parameter space shows that this region also corresponds to templates with high mass ratios, such as those from neutron star - black hole (NSBH) systems. While scratchy glitches are rarer than the other classes identified in this chapter, the likelihood of a single glitch producing a significant trigger is high for a wide range of parameters, a contrasting situation to blips and koi fish.

One important note is that the total duration of scratchy glitches is quite significant, up to multiple minutes, as opposed to blip and koi fish glitches that last tenths of seconds. While this does increase the chance of a trigger being due to chance, the excess shown in Figure 39 (bottom-right) is higher than we would expect from colored Gaussian noise alone.

We also can examine how specific aspects of the detection statistic impact the likelihood of a trigger being produced. Instead of using the sine-Gaussian SNR statistic, we switch to using reduced SNR, the detection statistic used for initial analyses of

aLIGO’s first observing run [77, 12]. For the blip and koi fish classes, this substantially changed the likelihood of producing a trigger. The results of this can be seen in Figure 40.

For blip glitches, we see clustering of the most likely part of the parameter space near maximal total mass and anti-aligned spin. In this region, particular parts of the space cause significant triggers up to 10% of the time. The rest of the parameter space shows extremely low likelihood of any blip glitch causing a significant trigger. The confinement of the probability to a small part of the parameter space suggests that there is a high match between these template parameters and the glitches.

Without the benefit of the sine-Gaussian discriminator, koi fish glitches can mimic a wide variety of high mass templates, indiscriminate of their effective spin. This corresponds to the entire region of the template bank where the templates are at the minimum duration. This suggests that the largest contributing factor to whether a koi fish glitch can produce a trigger with specific template parameters is the template duration as opposed to the similarity between the template and the glitch. At this short template duration, PyCBC consistency tests have less discriminating power, leading to an increased chance of a template being recovered with a significant detection statistic.

The high level of mitigation seen between Figures 39 and 40 underscores the importance of signal consistency tests in increasing the sensitivity of searches. As the sine-Gaussian discriminator was originally developed to combat both blips and koi fish [86], it is unsurprising that there is such a significant change. However, the scratchy and scattering classes only saw minimal improvement. If similar glitch-targeted tests for these classes can be developed, they are likely to lead to significant improvements in the sensitivity of the search.

For each glitch class, there does appear to be a range of parameters for which a candidate trigger in time coincidence would likely be caused by the presence of the glitch. For example, a high mass, anti-aligned BBH trigger noted during a time flagged by Gravity Spy is consistent with the expected triggers produced by the observed blip population. Conversely, the mass and spin parameters of previously observed BBH and BNS signals [13] are not consistent with blips, and hence there is minimal risk of falsely using these results to cast doubt on a similar signal. At lower masses, the most problematic region is the correlation between templates with

high mass ratios and high spins (such as from an NSBH waveform) and scratchy glitches. Such correlations can be used to quickly follow up candidates after they are identified by the search. If no overlap is found, the trigger is unlikely to be related to the identified instrumental artifact. Alternatively, if the template parameters are consistent with those expected from the relevant glitch class, additional investigations to understand causality between the artifact and candidate trigger are warranted.

4.5 A Method to Utilize Glitch Classification in Significance Estimates

As discussed in the previous section, some of the most likely regions in the PyCBC parameter space for a glitch to produce a trigger are those with high spins and high ratios of component masses. Combining current upper limits on the rate of mergers from highly spinning BBH and NSBH systems with the rate of problematic glitches in aLIGO data suggests that any given candidate trigger in time coincidence with a glitch is most likely due to the presence of the instrumental artifact. However, this does not preclude the possibility of a real astrophysical signal occurring in coincidence with a glitch. In this scenario, it may be possible to include the additional information we have about the expected overlap with the candidate trigger parameters and the glitch population to re-evaluate the significance of the signal. In this section we outline a procedure to calculate the significance of a trigger found in time coincidence with a time flagged by Gravity Spy as belonging to a specific glitch class. We focus on blip glitches as a test case since they are one of the most common glitches in both detectors and have very defined regions of the parameter space where overlap between signals and glitches occur.

To understand the the trigger rate changes during glitching in different parts of the parameter space, we compare trigger rates from generic times versus only times that are known to contain glitches. We then sort triggers from each period into a BNS category ($M_{chirp} < 2.0M_{\odot}$) and a BBH category ($M_{chirp} > 5.0M_{\odot}$). A comparison of the rate of triggers versus network ranking statistic $\tilde{\rho}_{sg,net}$ for the entire analysis period and around blip glitches is shown in Figure 41. If we examine the trigger rate during the entire analysis period versus during short time periods around blip glitches, we see that there is indeed an increased rate of triggers at fixed ranking

statistic for times around blip glitches. Furthermore, this increase is only apparent for the BBH category, with only minimal increases for the BNS category. This agrees with our expectation that blip glitches are responsible for a high rate of triggers.

As was shown with GW170817, it may still be possible to detect a gravitational-wave signal during a time period corrupted by glitching. A critical component of this detection was establishing that the observed glitch could not have accounted for the BNS signal in the data [15, 117]. In order to facilitate significance estimates of additional candidates, Gravity Spy classifications, combined with our knowledge of the overlaps between template parameters, can be used to evaluate if the candidate trigger could plausibly be related to the observed transient.

Blip glitches present a clear use case for this follow up. This class is known to impact only an isolated part of the parameter space. Specifically, we would expect that low mass BNS triggers would be unrelated, while high mass BBH triggers may be due to the presence of a glitch. In order to account for the expected variation in the background distribution across the template bank, we include parameter dependent background reweighing that measures the rate of triggers with respect to template duration and downranks templates that are shown to occur more frequently [87]. As BNS and BBH signals have vastly different template durations, we would expect both classes of signals to be affected differently by the inclusion of this term to the ranking statistic when compared against our expected triggers from blips.

To demonstrate how the significance of each signal model is affected by a correlation with blip glitches, we perform a series of astrophysical software injections of both BNS and BBH signals into real aLIGO data. We first calculate the inverse false alarm rate (IFAR) of each injection using the background distribution measured from a 5 day analysis period as a control. We empirically measure the distribution of single detector background triggers during times period flagged by Gravity Spy as blips with 0.95 confidence, and use this as the input for the background re-weighing procedure. We then reevaluate the IFAR of each injection with this new background distribution. As these signals were not injected to directly overlap periods corrupted by glitching, they represent an ideal scenario where there is a glitch has no impact on the recovery of an unrelated astrophysical signal. A comparison of the recovered IFAR for each injection with the standard and blip-focused background distribution can be seen in Figure 42, along with a 1-1 line indicating where the recovered IFAR is consistent

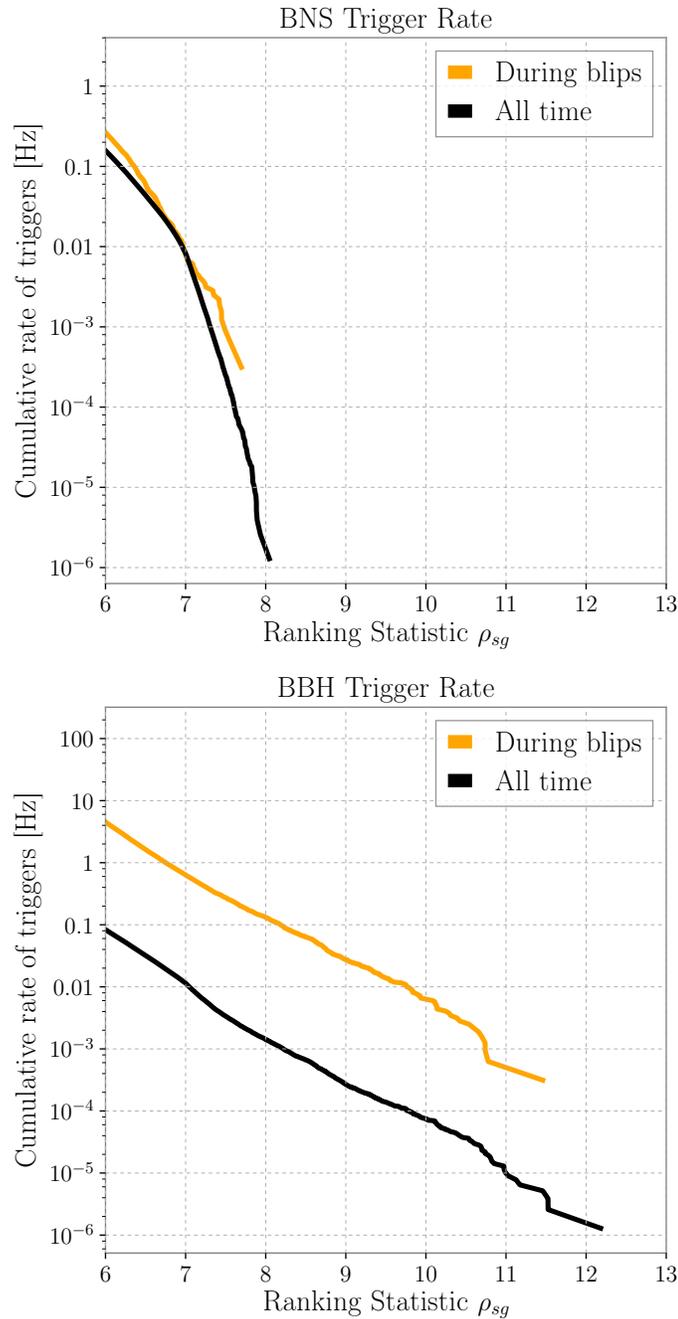


Figure 41: A comparison of the false alarm rate (FAR) based on all times in the analysis and only times coincident within 2 seconds of blip. The trigger rate is much higher during blip periods for BBH triggers, supporting the previous conclusion that there is an increased chance of producing a trigger coincident with this source of noise. Top: Difference in trigger rate for BBH triggers. Bottom: Difference in trigger rate for BNS triggers.

between the two cases. Note that Figure 42 (top) shows only BNS injections and Figure 42 (bottom) shows only BBH injections.

Comparing the two injection sets, there is a clear difference in the distribution with respect to the 1-1 line. BNS injections were recovered at approximately the same IFAR in both cases, showing that times corrupted by blips are equally likely to produce a BNS trigger as an average time. The BBH injections, on the other hand, have a clear separation from the 1-1 line. Specifically, the IFAR of injections recovered is lower when a background based on blips is used versus a standard time. Since blip times are more likely to produce triggers with BBH parameters, these injections are naturally downranked. However, loud injections are still recovered at IFAR of greater than 1 year, which would be sufficiently significant to separate itself from the background and to be identified as a candidate of interest.

This method provides a natural way to evaluate candidate triggers that occur during periods of known noise. In the above example, a BNS signal in coincidence with a blip would not be down ranked, while a BBH signal at this time would be have a reduced significance due to this time correlation with the glitch. While only blips were focused on in this section, this procedure can be repeated for all glitch classes that are sufficiently common enough for an expected background distribution to be measured. As a high fraction of significant outliers from matched filter searches have been shown to occur during such periods of transient noise [13], this method can provide further quantitative evidence that a candidate trigger is unlikely to be related to the instrumental artifact that it occurs in time coincidence with. In general, the presence of a glitch does not preclude the possibility that a candidate trigger is astrophysical, but only reduces the likelihood of astrophysical origin as compared to a candidate trigger that does not overlap a known instrumental artifact.

As this method provides a method to rerank triggers based on classification information from Gravity Spy, such a procedure can be implemented in the ranking statistic. One possible method to include this information is to separately evaluate triggers during times categorized by Gravity Spy as glitches and times where no glitch is identified. This would allow the search to benefit from an increased sensitivity during time periods where no glitching occurs, and to more accurately rank candidates related to glitching. Down ranking triggers during glitches instead of removing them from the analysis has the benefit of not removing the possibility of detecting a signal

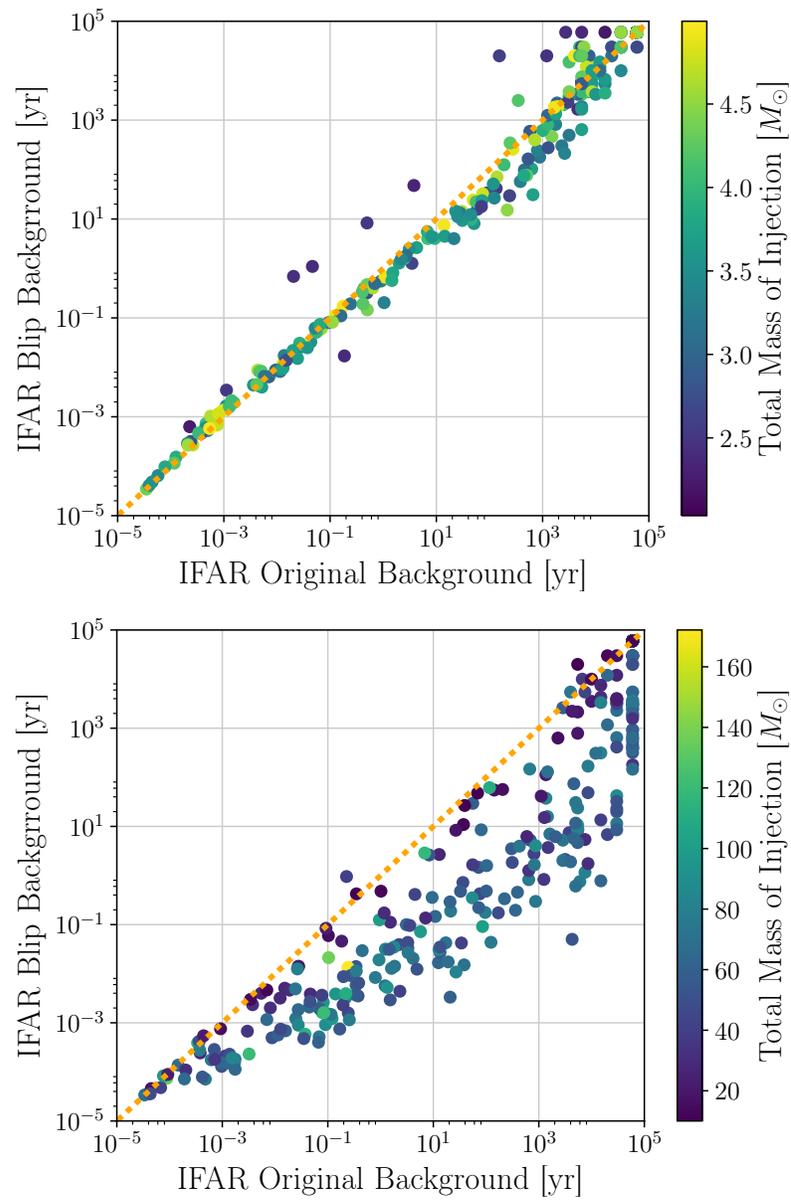


Figure 42: Recovered IFAR for a set of injections compared against the background from all time during an analysis (x-axis) versus the background during only blip times (y-axis). In each plot a 1-1 line is included for reference. For BNS injections (top), there is no affect on the recovered IFAR. For BBH injections (bottom), the high rate of triggers with similar template duration during blip times reduces the significance of the injections. This provides a natural way to evaluate a candidate trigger that is coincident with a known source of noise in the detector.

during a glitch. If an astrophysical signal was classified as a glitch due to a highly similar morphology with a glitch class, this method would not prevent detection. Further studies evaluating the safety of this method, and the utilization of Gravity Spy classifications to rank candidate triggers, is warranted.

One limitation of this method is based on which parameters are used as part of the background reweighting. As has been shown in this chapter, one single variable will not neatly delineate templates into groups impacted and not impacted by a glitch class. Therefore, this method will only be effective in the situation that the background is modeled by parameters that effectively describe the overall distribution of triggers. Detailed investigations into what parameters would be best for each glitch class may be resolved in future investigations.

4.6 Future Applications

This chapter emphasizes that the non-stationary features of aLIGO data currently are one of the limiting factors for sensitivity of the analyses identify gravitational waves. This is especially problematic for novel sources including mergers of intermediate mass black holes [133, 135] and neutron star - black hole [136] systems. As each of these regions of the current template bank are impacted differently by each of the glitch classes, there is unlikely to be a single method to efficiently differentiate these novel sources from common instrumental artifacts. Focused work to design consistency tests that account for known problematic glitch morphologies is needed. Alternatively, developing robust mitigation techniques for each of these common glitch classes will have tangible effects on the overall sensitivity of the searches.

At the present, the quantifiable metrics developed in this chapter can also be used to guide event validation of candidate triggers. When evaluating whether to initiate a search for an electromagnetic counterpart, being able to predict whether a significant trigger is likely due to the presence of a common glitch will allow more informed responses. As many EM counterparts to a gravitational wave signal occur within minutes of merger [137], quick follow up is critical. Gravity Spy classifications are planned to be included in automated follow up in LIGO-Virgo's third observing run and the results of this study can be used to translate these classifications into easily used metrics to determine the likelihood of the candidate being related to a common

glitch. Since the regions where a glitch is most likely to produce a trigger correspond with regions where a lower event rate is expected, understanding if a candidate of interest is a rare astrophysical signal or a common glitch is especially important to guide astronomical observations.

As aLIGO reaches design sensitivity and the rate of detections increases, signals found near noise transients in the data will become a much more common situation. Already in the recent results from O2, a large number of marginal triggers have been identified in time coincidence with noise transients [13]¹. Future progress will be facilitated both by continual instrumental work to reduce the rate of glitches as well as by further studies of how to distinguish glitches from genuine signals.

¹See Chapter 6 for additional discussion of the impact of these noise transients.

Chapter 5

Linear Noise Subtraction

It has been previously shown that it is possible to increase the sensitivity of the aLIGO detectors by subtracting instrumental noise from the gravitational-wave strain data [65, 138, 139, 140, 66]. For source parameter estimation [141] of previously published gravitational-wave signals from O2, a MATLAB-based noise subtraction algorithm was used to subtract instrumental noise using an associated witness sensor for 4096 seconds around identified events [142, 143]. This was the first instance of noise subtraction being used in the analysis of gravitational-wave events. However, this process was not designed with the intention of subtracting noise from the entire O2 data set.

Since this initial analysis, a Python-based implementation of noise subtraction was developed that prioritizes parallel processing and computational efficiency with the goal of subtracting instrumental noise over the entirety of the second observing run. Considering that each individual interferometer recorded over 150 days of data, one of the key considerations was the size of the data set that this noise subtraction pipeline needed to process. The methods used in this pipeline are general enough to allow any linearly coupled noise source with a clear witness to be subtracted out efficiently.

This chapter describes the method used to subtract noise due to beam jitter, detector calibration lines, and mains power lines in O2 and reports the improvement to search sensitivity gained by applying this method. Section 5.1 outlines the workflow used to process the data set in parallel. Section 5.2 characterizes the instrumental noise sources that were subtracted from the O2 data set. Section 5.3 describes the

tests that were done to ensure that the subtraction process was not capable of removing genuine astrophysical signals. Section 5.4 presents the effects of noise subtraction on the aLIGO noise spectrum and on the sensitivity to simulated astrophysical signals.

I was the lead developer for the noise subtraction pipeline and produced the finalized noise-subtracted dataset discussed in this chapter. This chapter is adapted from [144], of which I am the lead author. The complete noise-subtracted dataset is available as a part of bulk data release for aLIGO’s second observing run, and can be accessed through the Gravitational-Wave Open Science Center [112, 111].

5.1 Subtraction Pipeline Overview

5.1.1 Measurement of Transfer Functions

The assumption of a linear transfer function is motivated by the high coherence between witness sensor signals and gravitational-wave strain data. Figure 43 shows the coherence between witness sensors and gravitational-wave strain for three types of instrumental noise subtracted in O2. These noise sources are further detailed in Section 5.2.

For a given noise source, we assume that our measured gravitational wave strain data, $h(t)$, contains a noise component that can be modeled as the convolution of an unknown transfer function $c'(t)$ and the output of a witness sensor $a(t)$,

$$h(t) = h'(t) + a(t) * c'(t). \quad (5.57)$$

This noise component can be removed from the strain data by filtering the witness sensor data with this transfer function and subtracting its contribution to the measured strain, resulting in a residual strain denoted $h'(t)$. This transfer function can be conveniently calculated in the frequency domain, so that the subtraction takes the form

$$\tilde{h}(f) = \tilde{h}'(f) + \tilde{a}(f) \cdot \tilde{c}'(f). \quad (5.58)$$

Adapting the methodology and notation from [145], we begin by considering our data as time series that are sampled at time interval Δt over a time period T . This results in $M = T/\Delta t$ samples, denoted by $Y(j)$ for $j = 0, \dots, M - 1$. We denote the

Discrete Fourier Transforms of each data stream as $\tilde{Y}(k)$ for $k = -M/2, \dots, M/2$, so that the k 'th bin corresponds to a frequency $f = k/T$. We then split the frequency space into bands of width F given by

$$f \in [f_b, f_{b+1}) \text{ with } f_b = \frac{bF}{T} \quad (5.59)$$

for $b = 0, \dots, M/2F$. The transfer function is measured independently over each of these frequency bands and is constructed using frequency domain inner products between the relevant data sets. For two data sets Y_1 and Y_2 the inner product over a specific frequency band b is calculated as the cross-power spectrum summed over that frequency band:

$$\tilde{c}_{12}(f_b) = \sum_{f=f_b}^{f_{(b+1)}} \tilde{Y}_1(f) \tilde{Y}_2^*(f). \quad (5.60)$$

A measurement of the transfer function for uncorrelated noise, for which each frequency bin has a random phase, should find no significant coupling as multiple uncorrelated data points are averaged over to calculate the transfer function. To help reduce the risk of spurious correlations being measured, we set a minimum threshold on the value that the transfer function can take as a fraction of the maximum value and set the value of the transfer function to zero in any band whose value is below that threshold. We found that a uniform fractional threshold of 2.5×10^{-9} was sufficient for the noise sources considered in this work, but in practice this value can be tuned for different use cases.

In the case when multiple sensors witness the same noise, there will be a measurable correlation between each of the sensors, resulting in oversubtraction if not accounted for. For N witness sensors Y_1, \dots, Y_N and a target data stream to subtract noise from, Y_0 , the set of frequency domain transfer functions $\tilde{c}'_{01}, \dots, \tilde{c}'_{0N}$ that contain independent noise is the solution to the matrix equation [146]

$$\begin{bmatrix} \tilde{c}'_{01}(f_b) \\ \tilde{c}'_{02}(f_b) \\ \vdots \\ \tilde{c}'_{0N}(f_b) \end{bmatrix} = \begin{bmatrix} \tilde{c}_{11}(f_b) & \dots & \tilde{c}_{N3}(f_b) \\ \tilde{c}_{12}(f_b) & \dots & \tilde{c}_{N2}(f_b) \\ \vdots & \ddots & \vdots \\ \tilde{c}_{1N}(f_b) & \dots & \tilde{c}_{NN}(f_b) \end{bmatrix}^{-1} \begin{bmatrix} \tilde{c}_{01}(f_b) \\ \tilde{c}_{02}(f_b) \\ \vdots \\ \tilde{c}_{0N}(f_b) \end{bmatrix} \quad (5.61)$$

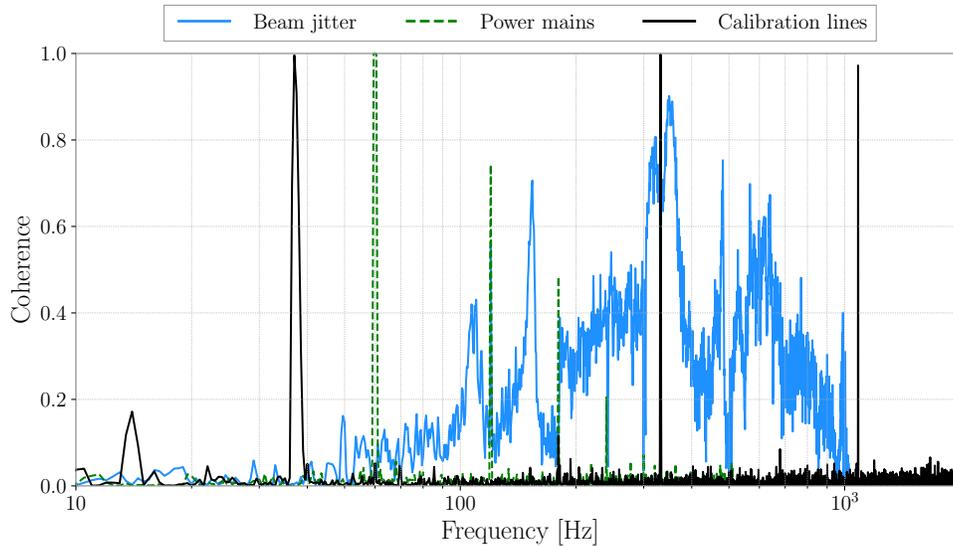


Figure 43: Coherence between witness sensors and gravitational wave strain for three types of instrumental noise subtracted in O2 at LIGO-Hanford: beam jitter, power mains, and calibration lines. The measured coherence demonstrates significant linear coupling between these witness sensors and the strain data, motivating the use of linear subtraction methods. Reproduced from [144].

These independent transfer functions can then be used for noise subtraction as described in Equation 5.58. This process allows additional sensors that may witness different features of the same noise source to be added to the noise subtraction process without risking oversubtraction.

5.1.2 Calculation of Coupled Noise

Advanced LIGO data is not stationary on the time scale of hours [83, 84], meaning the transfer functions used to subtract each noise source will vary over the time period that the noise subtraction is applied. This necessitated the development of methods to understand the stability and accuracy of the transfer function estimation on long timescales.

High amplitude non-Gaussian instrumental artifacts that can impact the measurement of transfer functions are removed from the strain data before transfer functions are calculated. This is done by applying an inverse Tukey window that zeroes the

data containing each instrumental transient. Instrumental transients are identified for removal by marking any times where the whitened time series exceeds a value of 100. This process is identical to the windowing done in [15] and described in Chapter 3. A continuous measurement for long stretches of data is approximated by calculating transfer functions with overlapping finite measurement windows, called “sections.” A visualization of this process is shown in Figure 44. After calculating the transfer functions and projected noise contributions for each individual window, each section of projected noise is multiplied by a Hann window and smoothly added together with 50% overlapping sections.

In the case that the transfer function is truly constant, this method is identical to applying a single transfer function over the entire period. The transfer function for each witness sensor is constructed to be uncorrelated with the transfer functions from other witness sensors, resulting in noise projection time series that are also independent. This allows each noise time series to be subtracted from the strain data independently. Once all targeted noise contributions are subtracted, we refer to the data as “cleaned”.

5.1.3 Workflow Implementation

One of the key features of this implementation is the throughput at which the subtraction can be done over long stretches of data. The pipeline takes advantage of the Pegasus workflow methods implemented in the PyCBC software package [147, 77, 120], which allows for parallelized calculation of transfer functions. Since the transfer functions for different noise sources can be calculated independently, the workflow was able to measure transfer functions for each noise source and generate projected strain data in parallel. In addition, the data set was broken up into distinct sections of continuous detector operation that were processed in parallel. The limiting factor in the subtraction process is the availability of computing nodes. Applying this method using available resources with 14 witness sensors allowed for two weeks of data from one detector, approximately 65 gigabytes, to be processed in only a few hours.

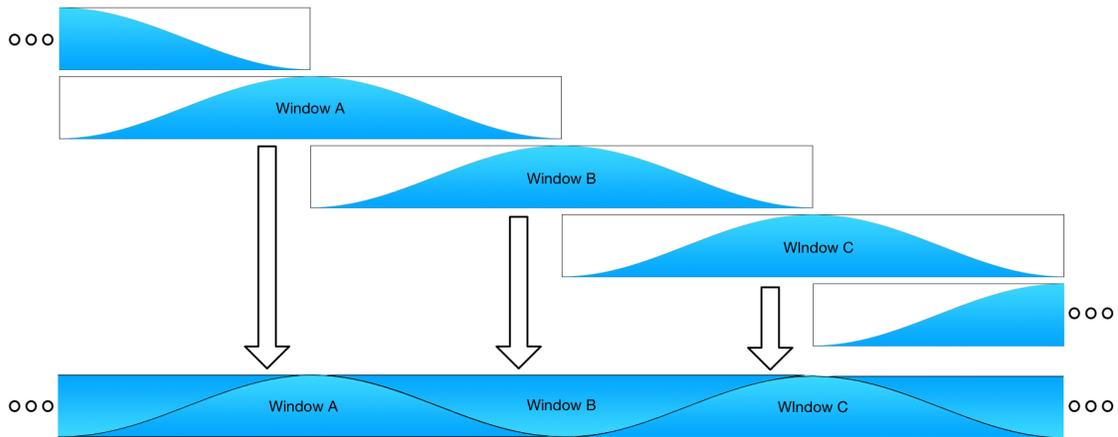


Figure 44: Visualization of how transfer function measurements are tiled in time. Transfer functions are measured in time windows (typically 1024 seconds) with 50% overlap. For a given time, the transfer function between the witness sensor and $h(t)$ is measured and the witness data are filtered to generate their projected contributions to $h(t)$. A Hann window is applied to each section of projected data before adding them together, resulting in a single projected $h(t)$ time series that has incorporated the time dependence of the transfer functions. Reproduced from [144].

5.2 Noise sources

During O2, multiple sources of linearly coupled noise were identified. These fell into two main categories: beam jitter noise that led to broadband noise contributions and narrow line artifacts from power mains and calibration lines. Both of these noise classes were identified and subtracted for analyses on previously published events, as described in [142]. This section describes each noise source and the witness sensors used in the subtraction process.

5.2.1 Jitter Noise

The main source of linearly coupled noise identified during O2 was related to jitter of the pre-stabilized laser (PSL) beam in angle and size [143, 148, 149]. The PSL is responsible for generating the frequency- and intensity-stabilized input laser beam that is injected into the interferometer. Upgrades to this subsystem undertaken in preparation for O2 led to different configurations of the PSL between LIGO-Hanford and LIGO-Livingston.

The configuration of the PSL at LIGO-Hanford during O2 included the addition of a high powered oscillator (HPO) that was designed to increase the laser power injected into the interferometer up to 200 W [148, 11]. The optical components used in the HPO required continuous heat dissipation via water cooling. Vibrations from water flow coupled to the table that supports the optical components used to control the beam angle, introducing jitter in beam angle and size [143, 149].

Fluctuations in beam angle are measured using quadrant photodiodes that sense the light reflected from the input mode cleaner (IMC) [150], which is used to filter higher order optical modes from the input beam. In February 2017, an additional sensor sensitive to radial beam distortions was installed [151]. In total, 7 readouts of beam angle and size (4 derived from quadrant photodiodes and 3 derived from the bullseye photodiode) were used to measure and subtract noise due to beam jitter.

During O2, the coupling of beam jitter into the output of the detector was further complicated by the presence of an axially asymmetric point absorber that was present on one of the test masses at LIGO-Hanford [152]. Thermal deformations are generally corrected with the use of the the Thermal Compensation System (TCS), which heats and deforms the mirrors [153], but this system is not capable of compensating for a

pointlike deformation. This deformation may have caused beam size and angle fluctuations to more strongly couple into the gravitational-wave strain data. Mitigating beam jitter noise required replacement of the HPO stage in the PSL and the test mass with the point absorber. Due to the invasive nature of this work, mitigation was not possible until after the end of the observing run.

Jitter noise related to beam size and beam angle fluctuations was present at LIGO-Hanford throughout all of O2, with increased coupling towards the end of the run. Variations in the beam angle led to broadband noise contributions, while variation in beam angle was coupled strongly at mechanical resonances of optic mounts between 100 and 700 Hz. The sensors used to witness these noise sources were digitally sampled at 2048 Hz, which sets the maximum frequency at which this jitter noise can be subtracted at 1024 Hz. The broadband coupling may have introduced noise above this frequency, but is not addressed in this work.

At LIGO-Livingston, the HPO was not included in the O2 configuration, and no asymmetries in the test masses were noted, leading to no noticeable jitter coupling in the gravitational-wave strain data. For this reason no jitter subtraction was done with the LIGO-Livingston data, which accounts for the lack of broadband noise subtraction seen in the spectrum shown in Figure 47.

5.2.2 Line Artifacts

The gravitational-wave strain data demonstrates several noise features that are narrowband, appearing as sharp lines in the frequency domain that can affect long duration searches and parameter estimation. The strain data contains excess noise at 60 Hz and its harmonic frequencies at both sites due to coupling of the power mains. These lines can be subtracted out using 3 witness sensors that directly measure the 3-phase voltage provided by the mains power grid at each observatory. In addition, the calibration lines discussed in Section 5.3.4 are applied using two methods. One set of calibration lines are digitally injected into actuation signals that control the position of the optics. Each digital excitation signal can be subtracted using the recorded excitation at the injection point. A second set of calibration lines are applied to the test masses via radiation pressure using the photon calibrator [154] and can be measured and subtracted using a single photodetector that monitors the power of the photon calibrator beam.

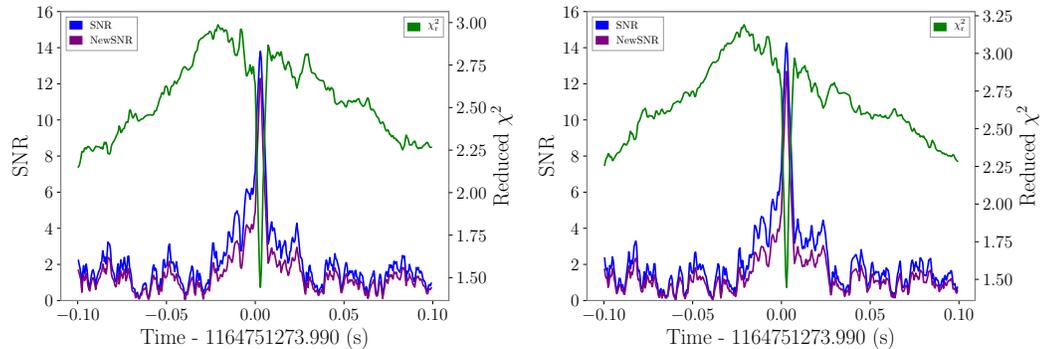


Figure 45: Recovery of a hardware injection before and after cleaning. Left: SNR timeseries around the injection before noise subtraction. Right: SNR timeseries around the injection after noise subtraction. After subtraction, the signal is recovered at an increased SNR without impacting the timing of the signal, demonstrating that the subtraction process does not negatively impact recovery of signals in the data.

5.3 Diagnostics

5.3.1 Sensor Safety

Before using the witness sensors described in Section 5.2 to subtract correlated noise, each sensor’s sensitivity to gravitational waves, or “safety”, was estimated. To establish safety, a series of sine-Gaussian waveforms were injected into the detector to excite the degree of freedom that is sensitive to gravitational waves [155]. If an excitation of this degree of freedom coupled into the readout of any witness sensors in a statistically significant way [104], those sensors were considered capable of accidentally subtracting away real gravitational-wave signals and were marked as unsafe. All of the witness sensors used for noise subtraction were determined to be incapable of witnessing and subtracting away gravitational-wave signals.

5.3.2 Recovery of Simulated Compact Binary Coalescence Signals

To ensure the noise subtraction process would not corrupt an astrophysical signal, a set of simulated compact binary coalescence (CBC) waveforms was digitally inserted over five days worth of aLIGO data from both detectors. The data set containing these simulated signals was processed by the PyCBC astrophysical search algorithm

[77, 120], used to search for signals from stellar-mass neutron star and black hole binaries, in order to compare the recovery of the simulated signals before and after noise subtraction. For each recovered signal, a coincident ranking statistic that represents the significance of an event found in multiple detectors in the detector network is calculated. Figure 46 shows the recovered coincident ranking statistic of each simulated signal before and after subtracting noise from the data set. After noise subtraction, all of the simulated signals were recovered with a ranking statistic that is consistent with or better than the ranking statistic before subtraction. In addition, there is a population of simulated signals that were not recovered in the original analysis but were found as coincident events after noise subtraction. As a final test, hardware injected CBC signals [155] were successfully recovered after performing noise subtraction. These hardware injections were recovered with increases in ranking statistic consistent with changes seen in software injections, as is shown in Figure 45.

5.3.3 Simulated Noise Tests

To verify that the noise subtraction process is effective for generic noise sources, artificial noise was added to aLIGO strain data and processed using the same method. The first test attempted to subtract artificial correlated noise. This noise was constructed by generating Gaussian noise, passing it through a transfer function that had similar features to the jitter transfer function, and summing it into the strain data. When provided with the strain data and the Gaussian noise, the noise subtraction algorithm was able to reconstruct the transfer function used to project the Gaussian noise into the strain data and subtract out the excess noise. The amplitude spectral density of the resulting data was consistent with the original data to within $\pm 3\%$ at all frequencies.

The second test was to subtract out random, uncorrelated noise which had *not* been added to the strain data. When provided with the strain data and the uncorrelated Gaussian noise, the algorithm subtracted a minimal amount of random noise. Similarly, the amplitude spectral density of the resulting data was consistent with the original data to within $\pm 2\%$ at all frequencies.

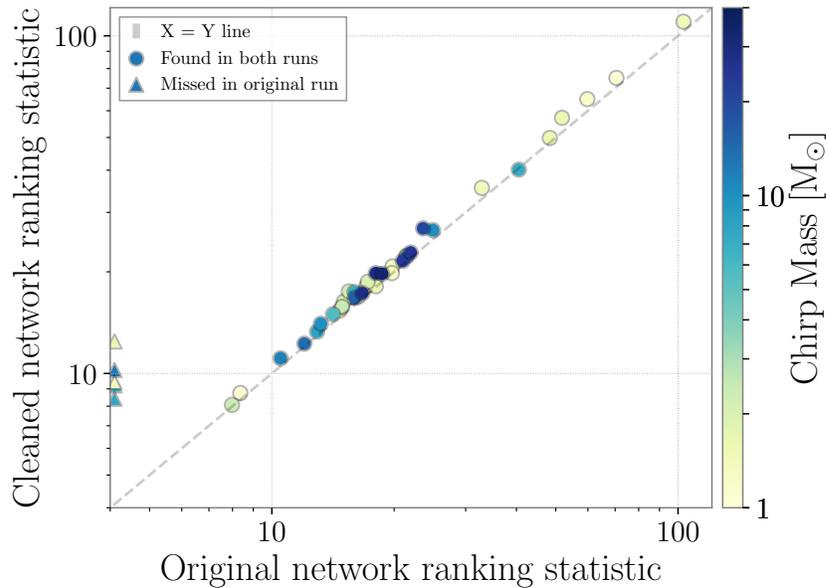


Figure 46: Recovered network ranking statistic for simulated gravitational wave signals before and after applying noise subtraction. The colorbar indicates the chirp mass [20] of each event, which spans a large astrophysical parameter space including binary neutron star, neutron star - black hole, and binary black hole signals. After noise subtraction, the simulated signals are recovered with a network ranking statistic that is greater than or equal to the ranking statistic without noise subtraction. In addition, several quiet simulated signals that were below the threshold of the search pipeline were recovered after noise subtraction due to being below the minimum threshold of signal to noise ratio of 5.5 in both detectors. These are indicated with triangles. Reproduced from [144].

5.3.4 Effect on Calibration

One important feature of aLIGO data is the presence of continuous, narrowband sinusoidal injections, or “calibration lines”, which are used to calibrate the data [156]. This calibration is performed on data that does not have noise subtracted, therefore tests were conducted to ensure that the calibration of the data was still valid after cleaning. A set of noise-subtracted data was produced using data from LIGO-Hanford that did not subtract away calibration lines in order to measure the impact of broadband noise subtraction on the data calibration process. Both the cleaned and uncleaned strain data were demodulated at the calibration line frequencies and the amplitude and phase were averaged in 300 second bins. The amplitude ratio and phase offset of each resulting measurement were calculated and are used as metrics for consistency.

To accumulate a statistically significant measurement of the calibration line consistency, 6.65 days of data were analyzed and the 1σ errors on the distribution of amplitude ratios and phase offsets are reported. For the 36.7 Hz and 1083.7 Hz calibration lines at LIGO-Hanford, the amplitude ratio was consistent with 1 to within $\pm 0.014\%$ and the phase offset was consistent with 0° to within $\pm 0.0078^\circ$. The 331 Hz calibration line (Located at a frequency where a non-negligible amount of power is expected to be subtracted off due to beam jitter) has an amplitude ratio that is consistent with 1 to within $\pm 0.15\%$ and a phase offset that is consistent with 0° to within $\pm 0.087^\circ$. As typical calibration uncertainties are $\pm 4\%$ [157], these measurements confirm that the noise subtraction process did not significantly impact the overall calibration of the strain data.

5.3.5 Impact of Nonstationary Data

While generally stable, the witness sensors used for transfer function estimation sometimes contain transient noise. In cases where the witness sensor has transient excess power that is not linearly correlated to the gravitational-wave strain, the transfer function is overestimated and the noise subtraction algorithm removes too much projected noise from the strain data. However, when the transient noise is linearly correlated to the gravitational-wave strain, transient noise can be subtracted from the gravitational-wave strain data. This linear subtraction of transient noise was

commonly found during periods of transient noise in the power mains.

The most impactful cases of oversubtraction due to excess power that is not linearly correlated with the gravitational-wave strain were noticed during review of the cleaned data set, and occur when there is transient noise in the photon calibrator used to inject calibration lines into the detector. To avoid this overestimation, the noise subtraction process is halted for 3 seconds around these transient noise artifacts. As these noise artifacts last less than one second, this veto period was chosen to ensure that the effect of the transient on transfer function measurement was completely mitigated. Once times where witness sensors contain transient noise are removed, the nearby noise-subtracted data shows no evidence of oversubtraction as compared to time periods disjoint from the excess noise.

Additional oversubtraction may occur if a feature of a witness sensor is spuriously correlated with the gravitational wave data. While such features are not observed on the timescales that the subtraction process is computed over, narrowband noise features from beat notes in the photon calibrator system may appear when signals are averaged on the timescale of multiple hours. The total bandwidth affected by these spurious correlations is less than 0.1 Hz and can be removed from long timescale analyses with the use of notch filters [158].

5.4 Results

5.4.1 The O2 Data Set

The noise subtraction algorithm was used to clean the entire data set from Advanced LIGO's second observing run, which spanned 9 months. The final version of calibrated data [156, 157] was used as the input to the noise subtraction pipeline. For computational efficiency, data which were considered unfit for astrophysical analysis [83, 94, 99, 118] were not processed. Additional time losses were due to removal of time periods corrupted by bandpass filters applied in the subtraction process and the excision of data where witness sensors were unsuitable for reliable transfer function measurement, as noted in Section 5.3.5. In all, only 0.05% of strain data was discarded as a result of the noise subtraction process. The final cleaned data set contains 118 days of coincident data. This value is greater than the coincident livetime reported in [15] due to the inclusion of additional time with updated calibration [157, 156].

5.4.2 Effects on the Noise Curve

The noise subtraction process is capable of removing both narrowband and broadband spectral features. Figure 47 shows the amplitude spectral density of the strain data from the Hanford and LIGO-Livingston detectors before and after noise subtraction. The narrow lines removed at 33, 60, 120, 180, 331, and 1083 Hz, detailed in Section 5.2.2, are related to detector calibration lines and power mains harmonics. Broadband subtraction in LIGO-Hanford data is a result of removing noise related to beam jitter. Due to the 2048 Hz sampling rate of the witness sensors and a low pass filter applied to reduce corruption near the Nyquist frequency, broadband noise is only subtracted up to 1024 Hz. A high pass filter applied at 13 Hz set the minimum frequency at which broadband noise was subtracted.

The same procedure was used to address noise sources present in the LIGO-Livingston detector. Line artifacts due to calibration lines and harmonics of the power mains were removed. As previously noted, beam jitter noise did not contribute significantly to the LIGO-Livingston data and was not subtracted.

We can characterize the benefit of the noise subtraction process with the “inspiral range”, which is the average distance at which a detector could observe a BNS system (1.4 - 1.4 M_{\odot}) at a signal to noise ratio (SNR) of 8. The inspiral range during O2 at LIGO-Hanford before and after noise subtraction is shown in Figure 48. After noise subtraction, the inspiral range at LIGO-Hanford increased by $\sim 20\%$ when averaged over all of O2 with a peak increase of $\sim 50\%$ towards the end of the observing run. The change in inspiral range at LIGO-Livingston was negligible over the course of O2 due to the lack of broadband noise subtraction. The effect of noise subtraction on the overall network sensitivity of the detectors is discussed in Section 5.4.3.

5.4.3 Effect on Astrophysical Analyses

The figure of merit used for quantifying sensitivity of a search compact binary coalescences is the sensitive volume of the search multiplied by the time duration of analyzed data, which is known as volume-time (V-T). While this volume can be approximated using the inspiral range as a measure of sensitive distance, that method does not fully account for the effects of data containing non-Gaussian noise artifacts on astrophysical search sensitivity, as well as the sensitivity of the entire interferometer network.

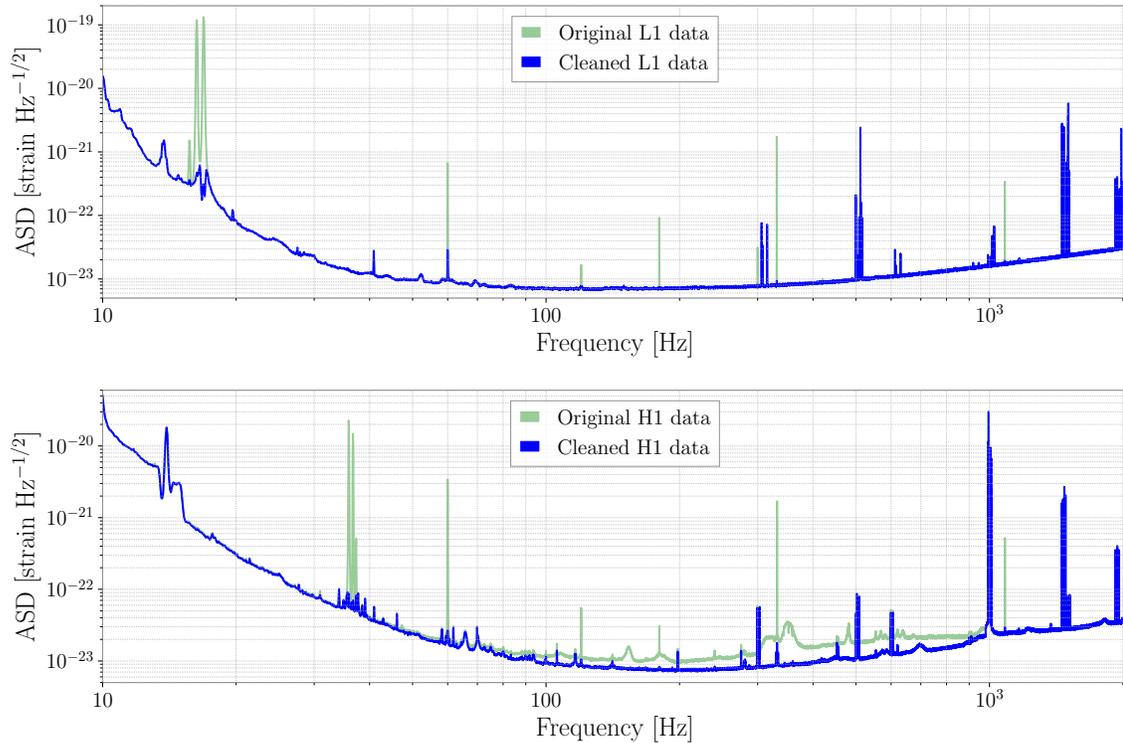


Figure 47: Top: Amplitude spectral density (ASD) of LIGO-Livingston (L1) gravitational wave strain data before (green) and after (blue) noise subtraction from a representative day of data during O2. Narrowband features from calibration lines and power mains were subtracted from the L1 strain data. There were no broadband noise sources with an appropriate witness sensor that could be subtracted from the L1 strain data. Bottom: Amplitude spectral density of LIGO-Hanford (H1) gravitational wave strain data before and after noise subtraction from a representative day of data during O2. In addition to narrowband features from calibration lines and power mains, broadband noise was subtracted between 80 - 1000 Hz using beam jitter witness sensors. Reproduced from [144].

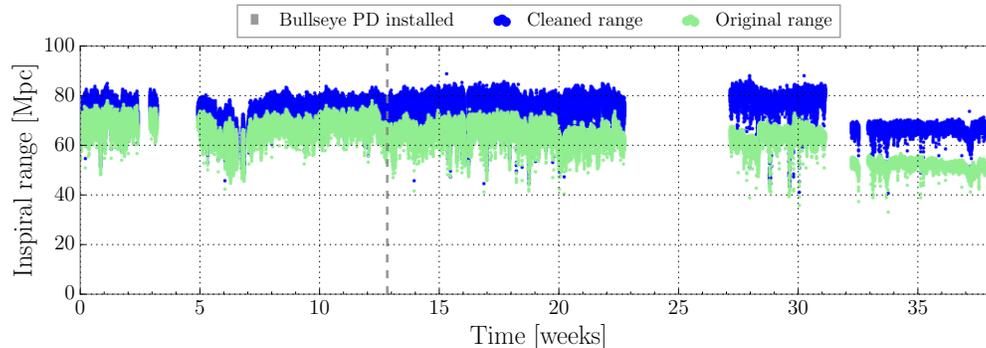


Figure 48: Inspiral range of the LIGO-Hanford (H1) detector over the course of O2 before (green) and after (blue) noise subtraction. The dashed line indicates the installation of the bullseye photodiode, a witness sensor used for subtraction of noise due to beam jitter. The large decrease in range after week 32 for both the Original and Cleaned range was due to the impact of an earthquake near the site [159]. LIGO-Livingston (L1) had no broadband noise subtraction, the increase in inspiral range was negligible and is not shown. Reproduced from [144].

V-T can be measured by injecting a population of simulated gravitational-wave signals into the data and attempting to recover them with a search pipeline [77]. For each search pipeline, a background distribution is generated that excludes coincident events in order to estimate the effects of detector noise on the search algorithm. Each recovered signal is then compared to this background and assigned an inverse false alarm rate (IFAR) that quantifies how likely it is that such a signal was caused by coincident instrumental artifacts rather than an astrophysical source. To estimate the increase in sensitivity due to noise subtraction, V-T of the PyCBC search was measured before and after noise subtraction using identical injection sets. As multiple values can be used as a cutoff to determine if a signal is recovered, we examined the V-T for IFAR values of both 100 years and 1000 years. This is the same process as was used to estimate the change in sensitivity before and after the inclusion of data quality products in Chapter 3. The ratio of V-T before and after noise subtraction binned by chirp mass is shown in Figure 49.

The large increase to the sensitive volume of the detector network, combined with a negligible reduction in available coincident time, led to a significant increase in V-T over the course of O2. Averaging over all mass bins, a 30% increase in V-T

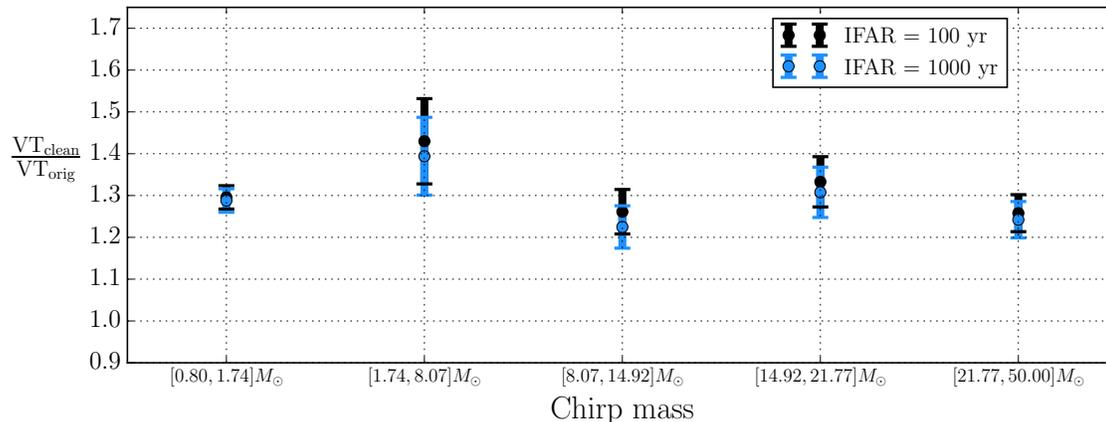


Figure 49: The ratio of volume-time (V-T) the PyCBC search was sensitive to during O2 before (original) and after noise subtraction (clean) binned by chirp mass. Black represents the volume-time that the search was sensitive to for signals with an inverse false alarm rate (IFAR) of 100 years, while Blue corresponds to signals with an IFAR of 1000 years. Error bars show 1 sigma error. On average, a 30% increase in V-T was measured over the course of O2. Reproduced from [144].

was measured over the course of O2. Particularly, the largest gains in sensitivity were realized for chirp mass between $1.74 M_{\odot}$ and $8.07 M_{\odot}$. This is a parameter space that aLIGO has not previously detected signals in, and hence has a largely unconstrained rate, in addition to being the location of the observed NS-BH mass gap [160, 161, 162].

One observed effect that led to a difference in measured V-T versus V-T extrapolated from estimating sensitive volume as a sphere with radius equal to the inspiral range was the impact of the noise subtraction process on instrumental artifacts. While the noise subtraction process reduced the broadband noise in the detector, it did not affect the amplitude of noise artifacts unrelated to the noise sources addressed by the noise subtraction pipeline. With a lower noise floor and no change in their absolute amplitude, artifacts already present in the data were found to increase in SNR. As one of the primary limitations of an astrophysical search's ability to recover signals is the rate of loud noise artifacts [83], these SNR increases limit the increase in V-T due to noise subtraction.

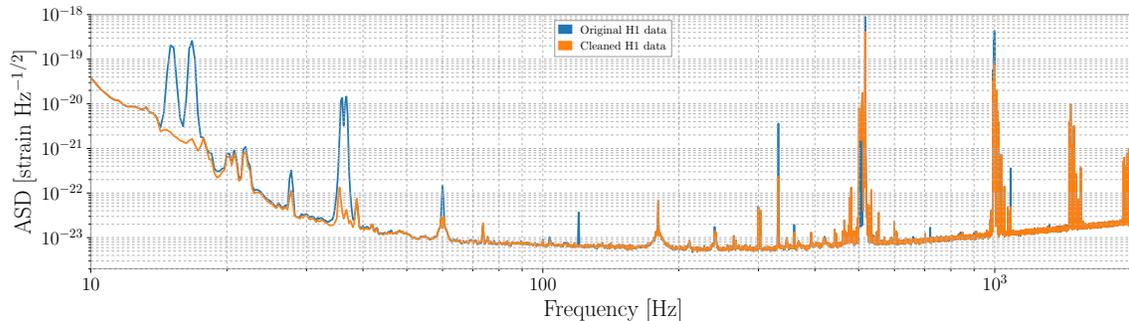


Figure 50: Amplitude spectral density of LIGO-Hanford (H1) gravitational wave strain data before and after noise subtraction from a representative day of data during an engineering run preceding O3. In addition to narrowband features from calibration lines and power mains, a small amount of broadband noise was subtracted below 40 Hz related to alignment sensing control (ASC) noise.

5.5 Application in future observing runs

The versatility of the noise subtraction method discussed in this chapter to remove known sources of noise makes it a useful tool to address a variety of noise sources in future observing runs. While beam jitter and line artifacts are the only noise sources subtracted from this data set, noise from feedback loops used to sense and control the length and alignment of optical cavities in the aLIGO detectors have been shown to contribute noise at lower frequencies [142] and are potential candidates for subtraction in future observing runs. Before aLIGO’s third observing run, the test mass with the point absorber at the LIGO-Hanford detector was replaced and the PSL configuration updated, which has reduced noise contributions due to beam jitter. Preliminary tests of noise subtraction with this pipeline for noise sources identified in the engineering run preceding aLIGO’s third observing run have found only minimal improvements in the range [163]. A comparison of the noise curve for LIGO-Hanford during this time period can be seen in Figure 50.

The sensitivity gains demonstrated in this paper show that a robust offline noise subtraction pipeline is an integral aspect of achieving maximum sensitivity in gravitational-wave detectors. The 30% increase in sensitivity of aLIGO to compact binary coalescences after noise subtraction allowed for an increased volume of spacetime to be searched for gravitational waves. In the following chapter, the results of the CBC

analysis based on this dataset is discussed. Although not quantified in this chapter, the noise subtraction process will also lead to general increases in the sensitivity of searches for gravitational waves using aLIGO data, such as those for continuous waves [39], stochastic [40], and unmodeled burst sources [72, 71, 164].

Chapter 6

O2 Catalog of Gravitational Wave Signals

This chapter discusses the final astrophysical search results of aLIGO's second observing run, and the impact on the searches of the data quality products that were discussed in preceding chapters. Two matched filter pipelines, GstLAL [78] and PyCBC [77], analyzed the noise-subtracted data. In this chapter, the PyCBC results will be highlighted and discussed, although triggers that were only identified by GstLAL will also be mentioned in the context of understanding how noise subtraction and data quality issues affected the analysis of these candidates. However, as PyCBC was used in quantifying the impact on sensitivity of both the noise subtraction and data quality products in this dissertation, it is this pipeline that provides the best opportunity to test if the final results met the expectations discussed in the previous chapters.

This chapter prominently features results that were a part of the LVC collaboration's GWTC-1 catalog [13]. I was involved in this work as the lead for the PyCBC reanalysis of noise-subtracted data, both running the pipeline and interpreting the performance in this new dataset. I also served as the event validation lead for the LIGO Detector Characterization Group, coordinating follow up investigations of the individual triggers that made up this catalog; I was also the technical liaison from the group for the catalog. Many of the investigations discussed in this chapter were also performed as part of the final review process for the noise-subtracted dataset.

6.1 PyCBC Search of noise-subtracted data

Initial results from the searches, first in low latency (e.g. PyCBC Live [89]) and then over five day periods were performed with a preliminary calibration and without noise subtraction¹. With finalized calibration and with the significant increase in sensitivity due to noise subtraction, the data was searched again with the hope of identifying new events. The sensitivity of each detector over this entire period is shown in Figure 51.

Outside of the improved gravitational-wave strain channel, this analysis was able to use the finalized version of all data quality products for the observing run; the most important addition was the use of gates based on auxiliary channel information for the entirety of O2. No changes were applied to the detection statistic calculation or background estimation methods between the analysis of GW170817 and this search.

As a part of this analysis, individual run periods were analyzed to identify if the noise subtraction had resulted in unanticipated effects on the search, such as creating non-physical triggers or preventing injections from being recovered. No problematic artifacts were identified through this check. However, it was noted that a large number of short-duration glitches produced higher SNR triggers after noise subtraction. This was most likely due to the lower noise levels after subtraction that elevated the prominence of the glitches. This effect led to interesting connections between the PyCBC detection statistic and the search results that will be discussed later in this chapter.

6.1.1 Classification of Significant Triggers

The catalog considered triggers that were found with a false alarm rate below 1 per month. These triggers were labeled as ‘GW’ if they had a high enough probability that the trigger was astrophysical. This probability is based on a comparison between the probability that the trigger is drawn from the background distribution of noise versus the probability that it was drawn from the distribution of astrophysical signals. If signals are distributed homogeneously in the universe, the distribution of signal SNRs

¹While noise subtraction was available for analyses of short time periods around some events, noise subtraction was not applied to the data that was searched.

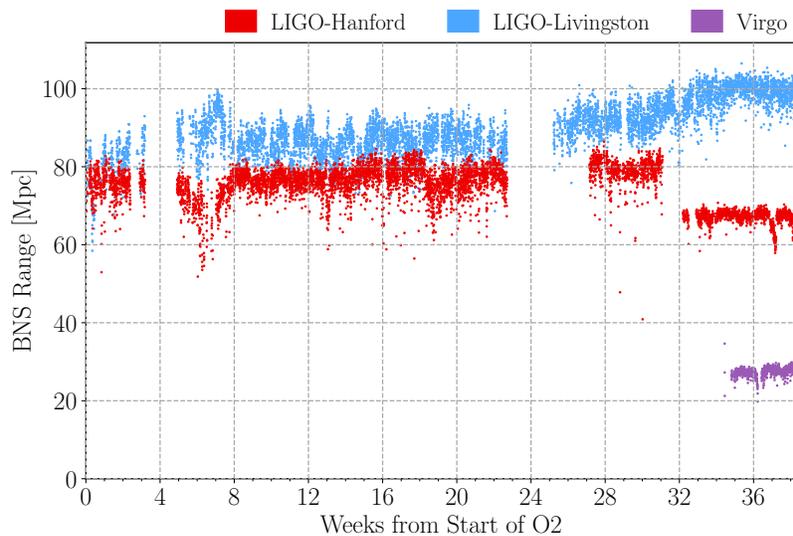


Figure 51: Range of LIGO-Livingston, LIGO-Hanford, and Virgo during O2. Note that this is after noise subtraction is applied for both LIGO sites. Reproduced from GWTC-1 [13].

(ρ), is such that

$$s(\rho) = R\rho^{-4} \quad (6.62)$$

where the normalizing constant R is based on the overall rate of signals. If the measured background distribution, $b(\rho)$, is also known, then the astrophysical probability is [165]

$$p_{astro}(\rho) = s(\rho)/[s(\rho) + b(\rho)] . \quad (6.63)$$

As this model relies upon the rate of signals to accurately calculate the astrophysical probability, it is possible to measure this probability with a significant number of observations of the relevant type of signal, but difficult when this isn't the case. This probability can be further refined by considering the probability distribution of multiple different signal populations [166].

For GWTC-1, candidates that had an astrophysical probability of above 0.5 were then labeled 'GW' [13]. All other triggers are referred to only by their date.

6.1.2 Identified Gravitational Wave Signals

Over the course of O2, seven gravitational-wave candidates with the GW label were identified with the PyCBC pipeline. Of these, four had been announced prior to the release of the catalog: GW170104 [108], GW170608 [109], GW170814 [76], and GW170817 [15]. Three unreleased events were found in the noise subtracted analysis: GW170729, GW170809, and GW170823. One additional gravitational wave event, GW170818, was also identified by (only) the GstLAL analysis of the noise-subtracted data. Spectrograms of these BBH events, along with reconstructions of the underlying waveform using a variety of methods, can be seen in Figure 52. Each event shows the typical chirp structure, with varying durations and intensities.

The gravitational-wave event GW170809 was identified by both PyCBC and GstLAL due to the use of noise-subtracted data. The event GW170818 was also identified by GstLAL for the same reasons.

Among the eight identified O2 events, seven are binary black hole mergers, and one is a binary neutron star merger (GW170817). A cumulative histogram of the triggers identified by PyCBC in O1 and O2, with GW events labeled, can be seen in Figure 53. Models for the noise background and the astrophysical signal foreground are shown, along with their sum. The cumulative number of triggers above a fixed detection statistic is shown to match the plotted signal model, which assumes that events are distributed homogeneously in the universe. Particularly of note is the steep slope in the noise model contrasted with the shallow slope of the signal model. This clearly shows that above the value where the two models predict an equal number of triggers, almost all triggers must be astrophysical, while the converse is true below this value.

6.1.3 Identified Marginal Triggers from PyCBC

The reanalysis identified only 1 trigger in O2 with false alarm rate below 1/month that was not labeled a ‘GW’. This trigger, labeled by date, is 170616. This trigger was of particular interest due to its false alarm and the recovered masses from the search. A spectrogram of the time around the trigger with the inspiral track overlaid can be seen in Figure 55. Additional details about 170616 can be found later in this section.

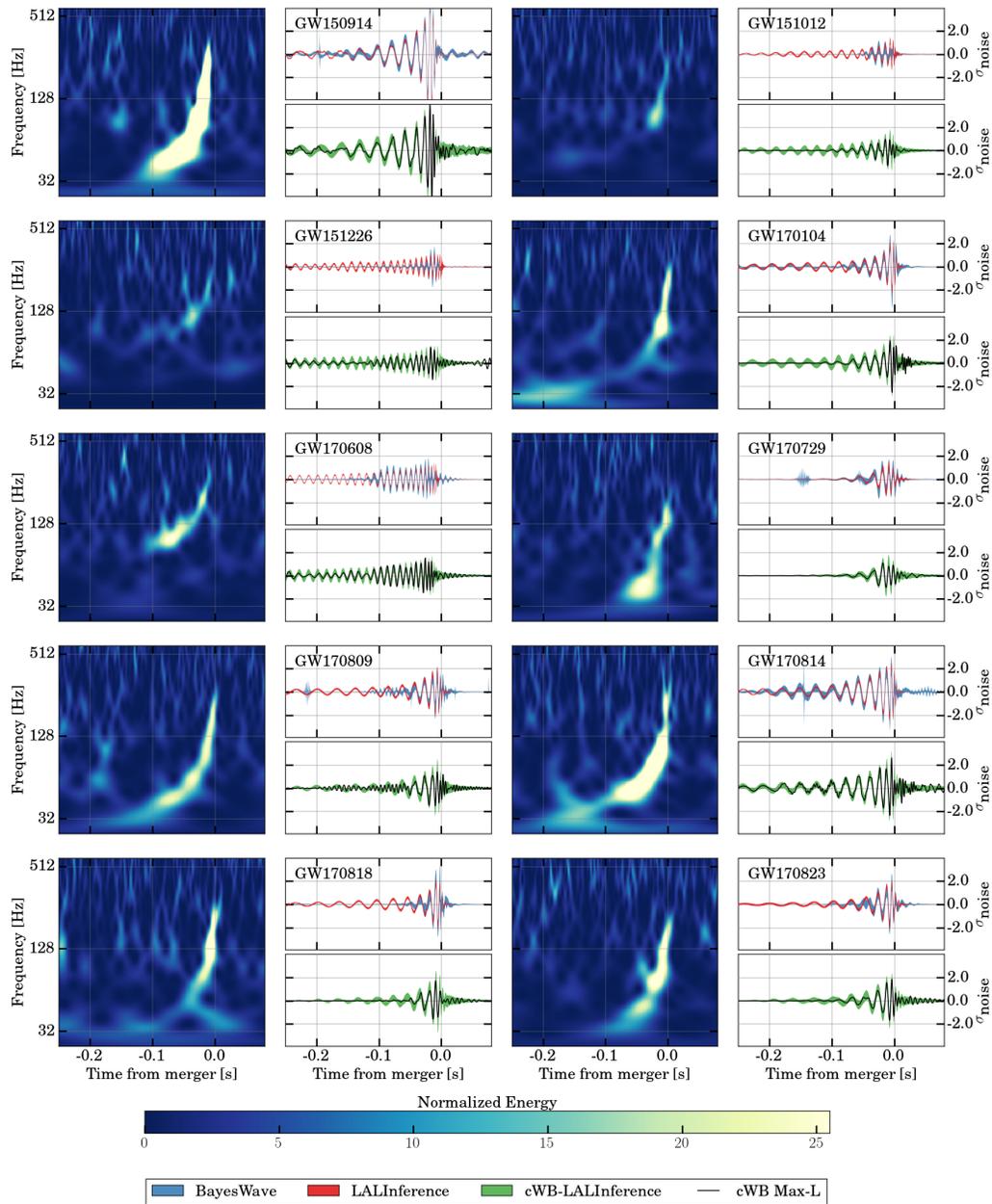


Figure 52: Waveform reconstructions of the 10 observed BBH signals observed in O1 and O2. For each signal, the left panel shows a spectrogram of the signal at LIGO-Livingston, while the right shows the reconstructed waveform based on BayesWave [164], LALInference [141], and cWB [167]. Reproduced from GWTC-1 [13].

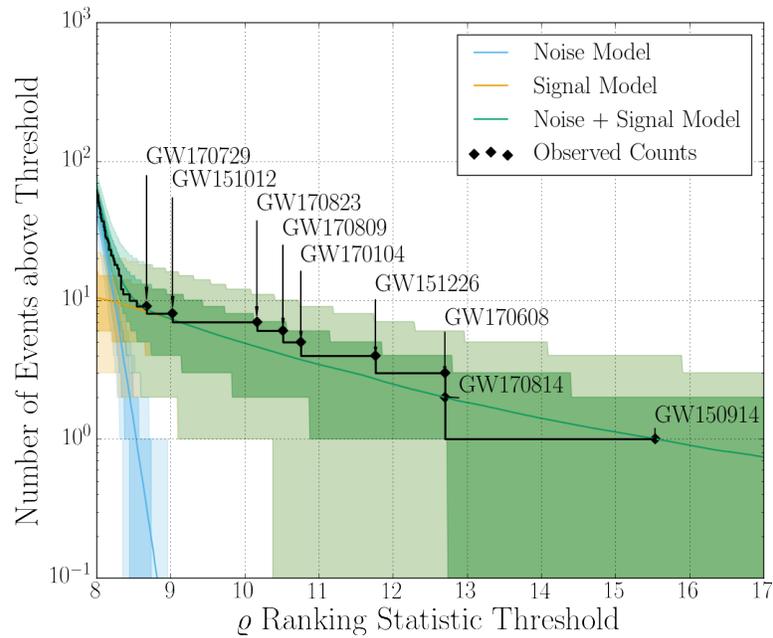


Figure 53: PyCBC cumulative histogram of identified BBH triggers during O1 and O2 with signal and background model overlaid. The observed triggers are shown in black, with significant individual triggers labeled. The overall search result is in agreement with a simple model of the distribution of triggers from the combination of an astrophysical signal and background noise model. Reproduced from GWTC-1 [13].

Based purely on the false alarm threshold of 1 per month and the analysis time of ≈ 4 months, we would expect to identify ≈ 4 triggers passing this threshold due to Gaussian noise alone. While recovering only one is below the median value, it is not out of line with expectations. Based on the expected counts from a Poisson distribution, recovering this many events or fewer will occur $\approx 10\%$ of the time. Furthermore, this number is highly sensitive to the threshold chosen, as the total number of events at this significance due to noise is low.

6.1.4 Identified Marginal Triggers from GstLAL

The analysis of the noise-subtracted data by GstLAL identified an additional 10 triggers that were below a false alarm rate of 1 per month but were not labeled GW. These 10 were distinct from the single such trigger identified by PyCBC. This difference in trigger sets is likely due to configuration differences between the pipelines, in both the data that makes up the analyses and the analyses themselves. These differences result in each pipeline being impacted by various noise sources differently, and hence would recover a different set of triggers that are due to noise.

The total number of marginal triggers identified by GstLAL (10) is much higher than the expected number based on a threshold FAR of 1/month. This large number is highly unlikely to be due to chance if the data obeys a Poisson distribution. However, at FAR thresholds close to the amount of time analyzed, it is difficult to precisely calculate a false alarm rate due to the large variety of different noise features that are present throughout the observing run.

These triggers, along with the trigger identified by PyCBC, are listed in Table 3. The triggers are listed with their UTC time, the recovered FAR, network SNR, and chirp mass of the most significant trigger from the search.

6.2 Impact of Noise Subtraction on Search Result

The noise-subtracted dataset that was prepared for both LIGO-Hanford and LIGO-Livingston was the first time that noise-subtracted data was used in the search. This section will discuss how the final result compared to expectations and how noise subtraction impacted the recovery of individual events.

Date	UTC	Search	FAR [y^{-1}]	SNR	$\mathcal{M}^{\text{det}} [M_{\odot}]$
161202	03:53:44.9	GstLAL	6.00	10.5	1.54
161217	07:16:24.4	GstLAL	10.12	10.7	7.86
170208	10:39:25.8	GstLAL	11.18	10.0	7.39
170219	14:04:09.0	GstLAL	6.26	9.6	1.53
170405	11:04:52.7	GstLAL	4.55	9.3	1.44
170412	15:56:39.0	GstLAL	8.22	9.7	4.36
170423	12:10:45.0	GstLAL	6.47	8.9	1.17
170616	19:47:20.8	PyCBC	1.94	9.1	2.75
170630	16:17:07.8	GstLAL	10.46	9.7	0.90
170705	08:45:16.3	GstLAL	10.97	9.3	3.40
170720	22:44:31.8	GstLAL	10.75	13.0	5.96

Table 3: Marginal triggers from the two matched-filter CBC searches. The search that identified each trigger is given, and the false alarm and network SNR. This network SNR is the quadrature sum of the individual detector SNRs for all detectors involved in the reported trigger. The detector chirp mass reported is that of the most significant template of the search. Reproduced from [13]

Expected Increases

Comparing significant triggers recovered before and after the noise subtraction, there are two ways that a trigger could be elevated into a trigger of interest by the noise subtraction. In order for a trigger to not be identified in the original dataset, the trigger either needs to be recovered in coincidence, but below a ranking statistic of ≈ 8.5 (as to not be considered significant), or not be recovered in coincidence, due to the trigger failing a threshold (such as the requirement that each PyCBC trigger has SNR above 5.5) in a single detector.

In the first case, a trigger with a ranking statistic of 8.5 would receive up to a 15% increase due to the noise subtraction. This would result, at best, in the trigger having a finalized detecting statistic of below 10.0; as a result, the trigger would have a FAR of $\gg 1/100$ years. This was the case for the trigger 170616, which was not identified in analyses of data before the noise subtraction.

For a trigger to be recovered at a much lower FAR, the trigger would need to be in the second category, where the trigger failed a specific threshold. If the threshold was passed after noise subtraction, combined with the trigger being highly significant in the other detector, it may be possible that a trigger would be identified at a FAR $\ll 1/10,000$ years. However, as loud single detector triggers are often manually followed up as part of detector characterization monitoring of the data, it is likely that such a trigger would have already been identified in the data before noise subtraction. This was the case for GW170818, which was initially identified as a significant single-detector trigger, and then identified in coincidence with the use of noise-subtracted data.

We can also use the final results to revisit the increase in sensitivity that was promised from noise subtraction. Based on the expectations discussed in Chapter 5, we should see a 30% increase in the sensitive volume, and hence the recovered number of events. Using the same threshold as before (FAR of 1 per 100 years), there were six gravitational wave events found in the noise-subtracted dataset. Comparing the original and noise subtracted results, one additional event was identified (GW170809). Therefore, the increase in the number of detected gravitational-wave events is 20%. This is well in line with the original prediction, especially with the large variability due to low number statistics that is expected.

Next we look at the recovery of individual events in the catalog, noting how noise subtraction and data quality work impacted their significance. Due to the large ranges in time, template parameters, and false alarm rate of these triggers, these efforts impacted the triggers differently.

GW170104

While GW170104 had already been identified before analysis of the noise-subtracted data set, it was not until the catalog analysis that any noise-subtracted data around this event was implemented. This is in contrast to the other previously released events, which had search results based on data without noise subtraction, but parameter estimation based on noise-subtracted data.

During the month of January, when this event occurred, noise subtraction provided a modest increase in sensitivity at LIGO-Hanford, increasing the range by $\approx 5\%$. Further discussion of this trigger is found alongside parameter estimation results in Section 6.3.

GW170729

The event GW170729 was a newly identified trigger, and was also the largest stellar mass black hole discovered, as well as the furthest identified gravitational-wave event. Due to this large mass, the template duration was extremely short, and hence was found in part of the background that had the highest rate of triggers. This meant that the significance of this event was strongly connected to the SNR distribution and rate of short duration glitches in the detectors.

As previously noted, some of short duration glitches caused higher SNR background triggers in the search after noise subtraction than before. Since these glitches were not related to the broadband noise being subtracted, their overall amplitude was constant, while the broadband noise was reduced. This led to an increased SNR for these glitches. Since the background distribution increased alongside the increase in recovered SNR from the event itself, the significance of GW170729 did not change as much as would be expected from the SNR increase alone.

These competing factors resulted in the network SNR increasing from 9.06 to 9.80, while the FAR only decreased from 1.89 per year to 1.37 per year. Therefore, while

the use of noise subtraction allowed this trigger to be recovered at a higher SNR, it did not increase the significance of the trigger.

GW170809

The event GW170809 was the only new trigger identified by the PyCBC reanalysis that was found with a FAR below 1 per 100 years. This trigger was not found as a significant coincident trigger before noise subtraction due to the requirement that both LIGO-Hanford and LIGO-Livingston recover an SNR of above 5.5. Analyses of the data that did not use this threshold found that the most significant trigger would have been $\text{SNR} \approx 5$ at LIGO-Hanford. Therefore, with the modest increase in SNR due to noise subtraction, the trigger was above threshold in both detectors.

The parallel analysis done with GstLAL used a lower SNR threshold than PyCBC, and was hence able to recover the trigger both before and after noise subtraction.

GW170818

The trigger GW170818 was a gravitational-wave event that was only identified as significant in analyses of noise-subtracted data. This event is also the best localized BBH event, in terms of square degrees of sky area, to date.

This event was identified as a ‘chirp-like’ trigger in investigations of the loudest background triggers at LIGO-Livingston during the analysis of the preliminary dataset, but no counterpart was seen in the LIGO-Hanford data. At the time, no trigger above SNR 4 was identified at LIGO-Hanford, and hence the trigger was not considered of interest. However, in the noise-subtracted data, the trigger experiences the expected 25% increase in signal strength that raised the recovered SNR to 4.2, as can be seen in Figure 56. This change, along with a comparable strength trigger at Virgo, allowed the event to be found at a significant FAR in the noise-subtracted data.

This trigger was still identified as only a single detector trigger by the PyCBC search, as the LIGO-Hanford SNR was below the threshold of 5.5. The GstLAL search, which uses a lower threshold, was able to recover this event as a triple-coincident event in LIGO-Hanford, LIGO-Livingston, and Virgo. A spectrogram of the time around the event for LIGO-Livingston and LIGO-Hanford can be seen in

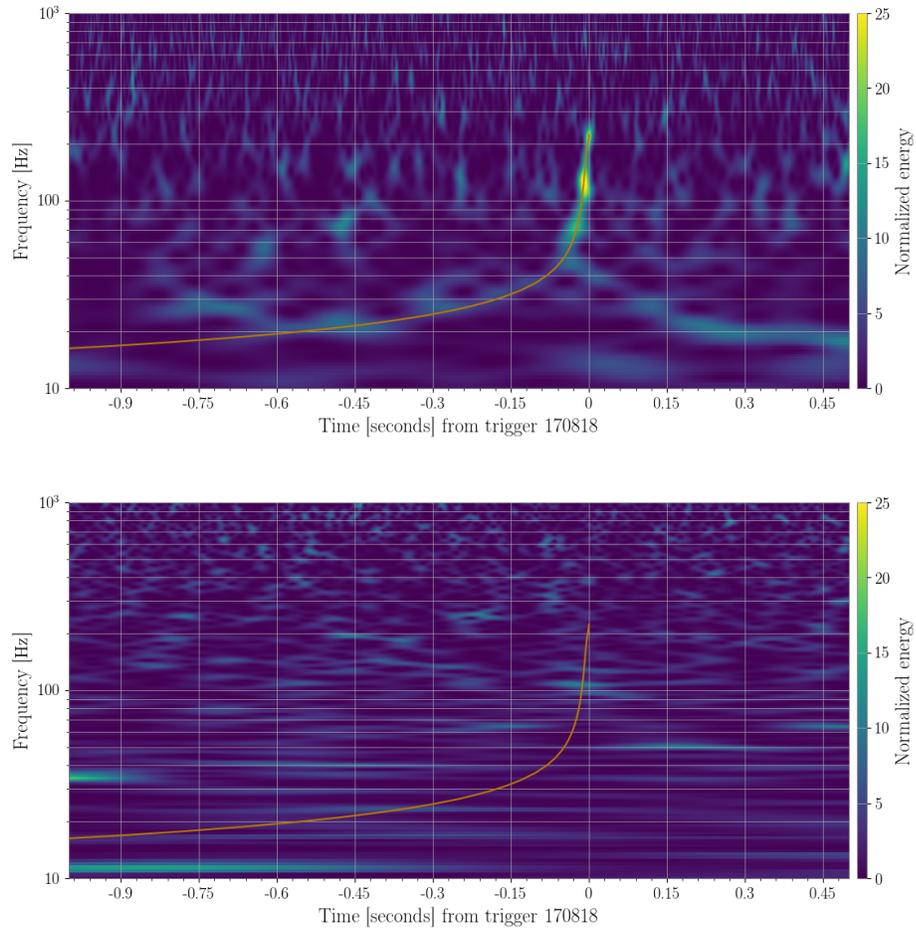


Figure 54: Spectrograms of GW170818 with the inspiral track overlaid. Top: LIGO-Livingston. Bottom: LIGO-Hanford.

Figure 54.

Trigger 170616

The trigger 170616 was the only trigger recovered by PyCBC with a FAR below 1 per month that was not subsequently labeled ‘GW’. This was due to the marginal significance of the trigger, at 1 per 6 months, combined with the low chirp mass of $2.75 M_{\odot}$. This chirp mass and component masses is suggestive of a neutron star - black hole binary system or a low-mass binary black hole system, both of which have yet to be observed via either electromagnetic or gravitational - wave observations. The lack of a previously observed population of signals matching this new trigger (as compared to the population of BBH signals observed) prevents this astrophysical probability from being directly calculated. However, the high FAR combined with the low expected rate of astrophysical events suggests that the chance of this event being astrophysical in origin is low.

The trigger 170616 also benefited from the noise subtraction, only being discovered in coincidence in the subtracted data set. Comparing the SNR recovered before and after subtraction, the SNR increased as expected at LIGO-Hanford from subtraction, as seen in Figure 57.

Further complicating this trigger is the presence of a number of glitches that overlap the inspiral track of the trigger template at LIGO-Hanford. These glitches, along with the inspiral track, can be seen in Figure 55. These transients were found to be due to scattered light artifacts. Subtraction of these glitches with BayesWave [164] was attempted, which was able to subtract some of the excess power without reducing the SNR of the trigger.

While this trigger is unlikely to be astrophysical, it does represent the most significant trigger found to date that lies in the region between typical BBH and NSBH systems. If 170616 is indeed astrophysical in origin, it may be a glimpse of detections of other low-mass systems that may come in the next observing run.

6.3 Impact of Noise Subtraction on Parameter Estimation

As a final check on the performance of the noise subtraction, we can examine how well measured different parameters of a gravitational-wave event are recovered before and

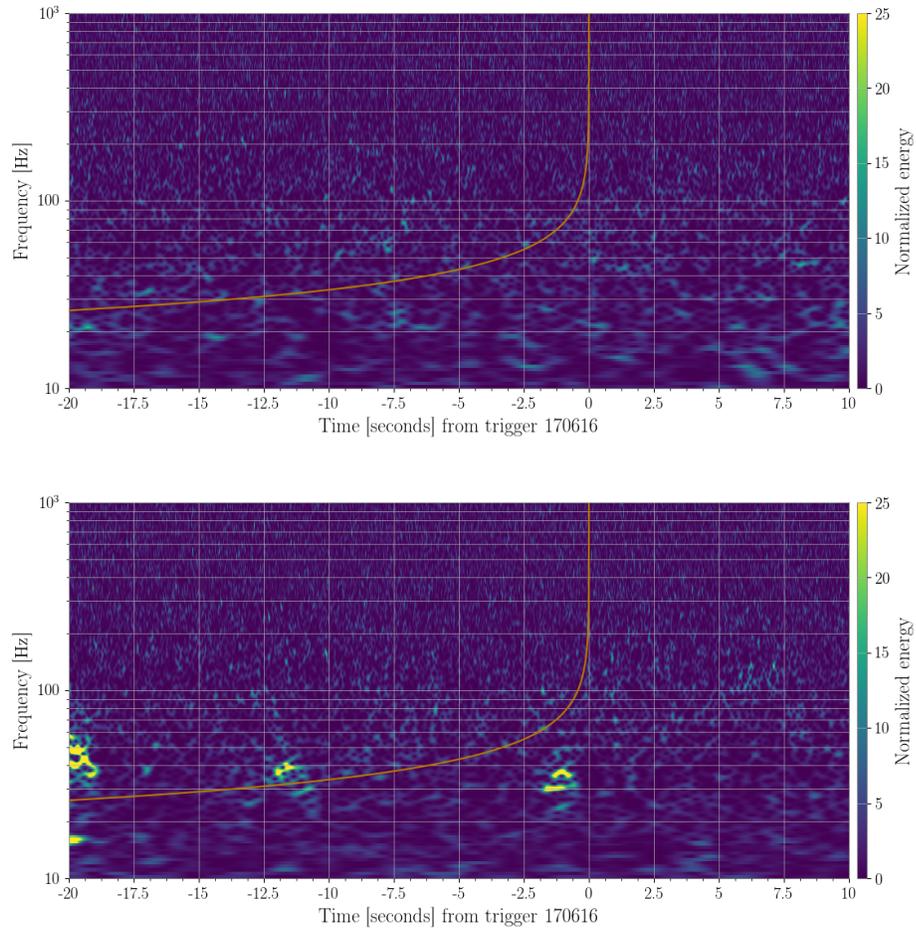


Figure 55: Spectrograms of trigger 170616 with the inspiral track overlaid. Top: LIGO-Livingston. Bottom: LIGO-Hanford.

Trigger	SNR pre-subtraction	SNR post-subtraction	Ratio post/pre
GW170104	8.17	8.67	1.06
GW170608	9.65	12.42	1.28
170616	5.41	5.89	1.09
GW170729	5.45	7.27	1.33
GW170809	4.60	6.15	1.33
GW170814	6.55	9.38	1.43
GW170817	15.54	18.69	1.20
GW170818	3.39	4.23	1.25
GW170823	6.64	6.57	0.99

Table 4: Recovered SNR in LIGO-Hanford for each of the GWTC-1 Triggers before and after subtraction, along with the ratio of the SNR after:before. In all cases, the recovered SNR after subtraction was either higher than or consistent with the recovered SNR before subtraction. On average, the observed increase in SNR was $\approx 20\%$.

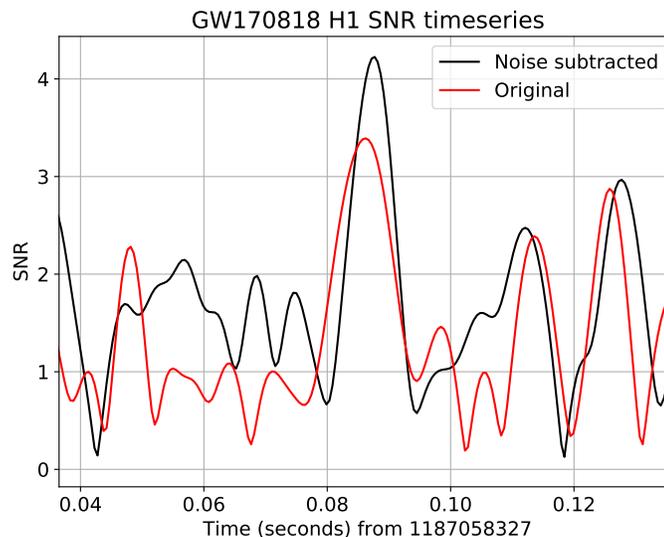


Figure 56: SNR timeseries of GW170818 at LIGO-Hanford based on both before (original) and after noise subtraction. Note the increase in the height of the central peak, indicating that the signal was recovered with an increased SNR.

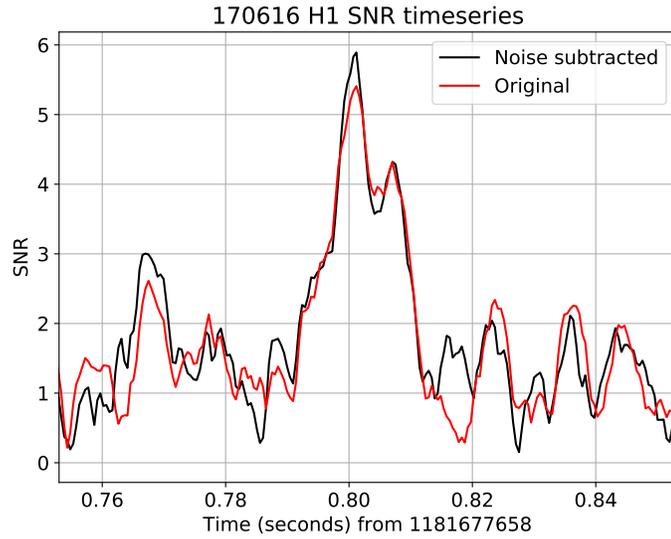


Figure 57: SNR timeseries of trigger 170616 at LIGO-Hanford based on both before (original) and after noise subtraction.

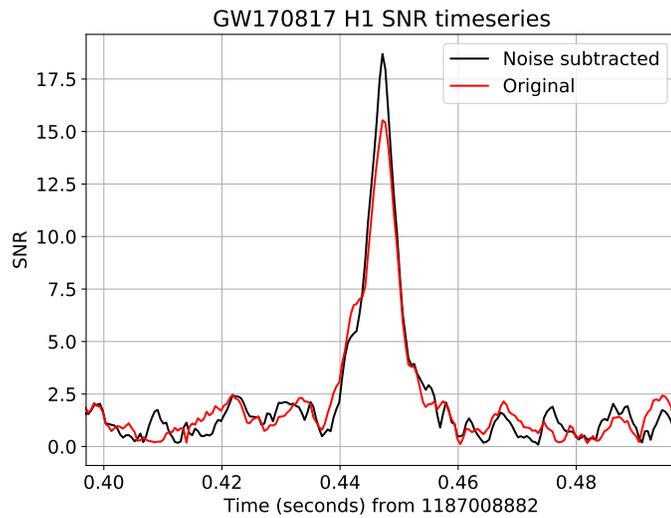


Figure 58: SNR timeseries of GW170817 at LIGO-Hanford based on both before (original) and after noise subtraction.

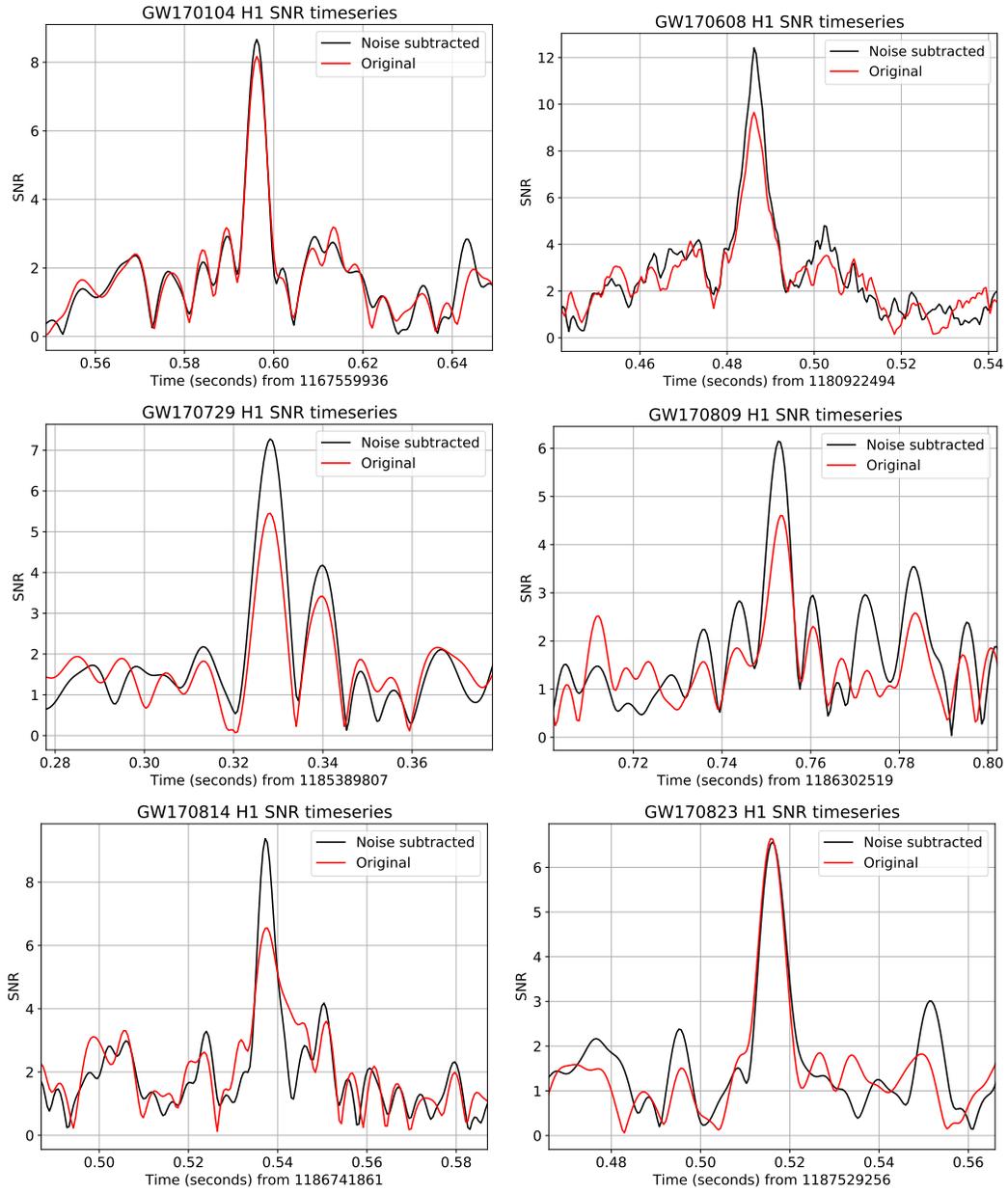


Figure 59: SNR Timeseries of O2 BBH signals at LIGO-Hanford. In all cases, the signals were recovered at either a consistent or increased SNR after subtraction.

Parameter	Discovery Paper	GWTC-1
Primary mass m_1/M_\odot	$31.2^{+8.4}_{-6.0}$	$31.0^{+7.2}_{-5.6}$
Secondary mass m_2/M_\odot	$19.4^{+5.3}_{-5.9}$	$20.1^{+4.9}_{-4.5}$
Chirp mass \mathcal{M}/M_\odot	$21.1^{+2.4}_{-2.7}$	$21.5^{+2.1}_{-1.7}$
Effective inspiral spin χ_{eff}	$-0.12^{+0.21}_{-0.30}$	$-0.04^{+0.17}_{-0.20}$
Source redshift z	$0.18^{+0.08}_{-0.07}$	$0.19^{+0.07}_{-0.08}$

Table 5: Recovered properties of GW170104 before and after subtraction. Note that the 90% credible regions are smaller by $\approx 20\%$ after subtraction, indicating increased precision in the measurement. Values reproduced from [108, 13].

after subtraction. Parameter estimation is performed with the LALInference software library [141]. As parameters are more easily resolved as the signal becomes louder, the noise subtraction process should increase the precision of parameter estimation. In this section, we will be quoting results based on the 90% credible region, which indicates the interval that has a 90% chance of containing the true value.

As noise subtraction was available for short time periods in O2 before the introduction of the noise subtraction pipeline discussed here, the majority of gravitational wave events previously announced were analyzed with noise subtracted data. However, as mentioned before, GW170104 was identified before noise subtraction was available. Therefore, this event provides an opportunity to directly compare results for parameter estimation before and after noise subtraction².

Using the results from [108] and [13], we compare the derived source property distributions and median values. In Figure 60, the credible region for the component masses of the system are plotted alongside each other. Between the two examples, the recovered values are consistent with each other, with the total area spanned by the credible region smaller after noise subtraction. Furthermore, if we examine Table 5, we can see that the credible regions are smaller by $\approx 20\%$ with the noise subtracted results.

²It is important to note that the calibration version between these two datasets also is different, which may also create unquantified differences. However, these differences are expected to be lower than those due to noise subtraction based on [168]

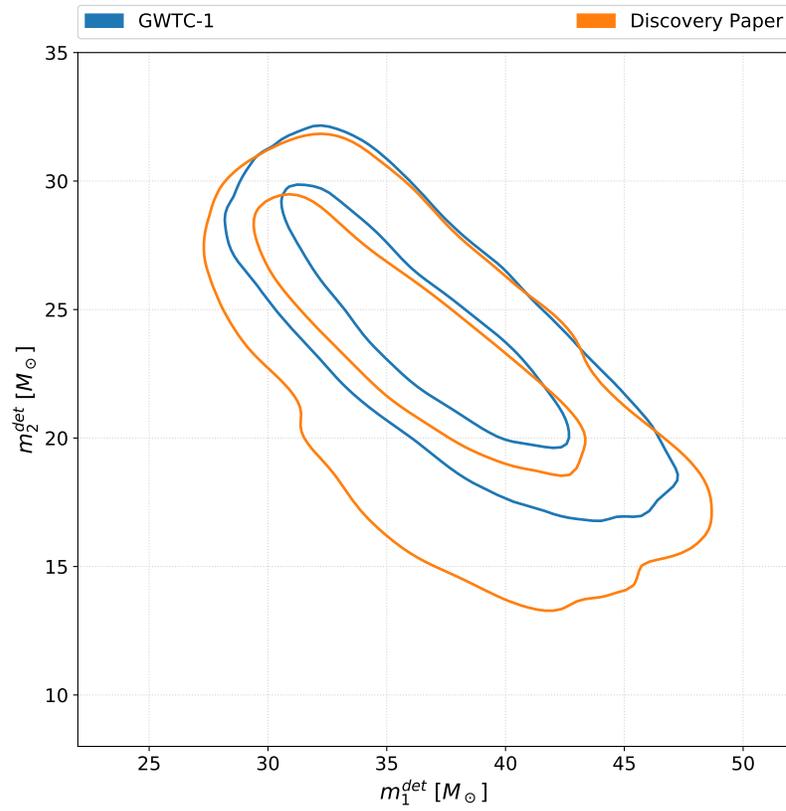


Figure 60: Comparison of measured component masses for GW170104 before (Discovery Paper) and after (GWTC-1) subtraction. Contours are drawn for the 50% and 90% credible regions. Note that the recovered parameters are consistent in both cases, with a smaller credible region after subtraction.

6.4 Impact of Noise Transients on Search Result

Now we turn to a summary of the detector characterization information for each marginal trigger, indicated in Table 6. Following a subset of procedures used for previous gravitational-wave detections [63], we evaluated the possibility that artifacts from instrumental or environmental noise could have caused each of the marginal triggers. Using auxiliary sensors at each detector, as well as the gravitational-wave strain data, we evaluated the state of the detectors at the time of each marginal trigger, identified and investigated any artifacts in the data due to noise, and tested whether any identified artifacts might explain the excess SNR observed in the analysis. Of the marginal triggers presented in this catalog, 8 have excess power from known sources of noise that overlaps the inspiral track of matched-filter template of the trigger. For 4 of these cases, the observed instrumental artifact overlaps the signal region, and accounts for the observed trigger. These classifications, along with the specific artifact class related to each trigger, can be found in Table 6.

To determine whether artifacts identified as noise ‘accounts for’ marginal triggers there are two metric used: 1) whether the type of noise had been previously shown to produce an excess of triggers consistent with the properties of the trigger present and 2) whether the noise artifact is able to account for the presence of the trigger as reported by that search, including SNR and time-frequency evolution, without the presence of an astrophysical signal.

A significant trigger due to noise is not inconsistent with the false alarm rates reported by the searches. It is expected that a substantial fraction of marginal events at the false-alarm rate values reported are caused by noise, given the background of the searches and the rate of signals. However, the complete dataset of available auxiliary channels is not utilized by the search. Including information from these sources can help interpret the results of the searches and identify triggers that are related to instrumental noise.

It is also important to note that the absence of evidence that a trigger is due to noise (either due to instrumental artifacts or Gaussian noise) is not evidence of the absence of such an impact. Therefore triggers that are not related to known sources of noise are not guaranteed to be astrophysical.

Date	Data Quality	Artifact Class
161202	Artifacts account for	Scratchy
161217	Artifacts account for	Scattering
170208	Artifacts present, not related	Scattering
170219	No artifacts	-
170405	Artifacts present, not related	Short-Duration
170412	Artifacts can account for	Scratchy
170423	No artifacts	-
170616	Artifacts present, not related	Scattering
170630	Artifacts present, not related	Short-Duration
170705	No artifacts	-
170720	Artifacts account for	Scattering

Table 6: Marginal triggers from the two matched-filter CBC searches along with a data quality statement for each trigger. In the case of an instrumental artifact present, the specific artifact class in question is listed. Adapted from [13].

The marginal triggers in this section are discussed based on the type of instrumental artifact that the triggers occurred in time coincidence with and the overall level of impact. Further discussion on the causes and mitigation strategies for these artifacts can be found in Chapter 2.

No noise artifacts present: 170219, 170423, 170705

Investigations into this set of marginal triggers have identified no instrumental artifacts in time coincidence with the triggers.

Light scattering accounts for: 161217, 170720

All marginal triggers in this class and the next are in time coincidence with artifacts from scattered light in one of the detectors. Scattered light leads to excess power at low frequencies that appear in time-frequency spectrograms as arch-like shapes. In some cases the frequencies affected are above the minimum frequency used in the analysis. When this happens, scattered light transients can create significant triggers

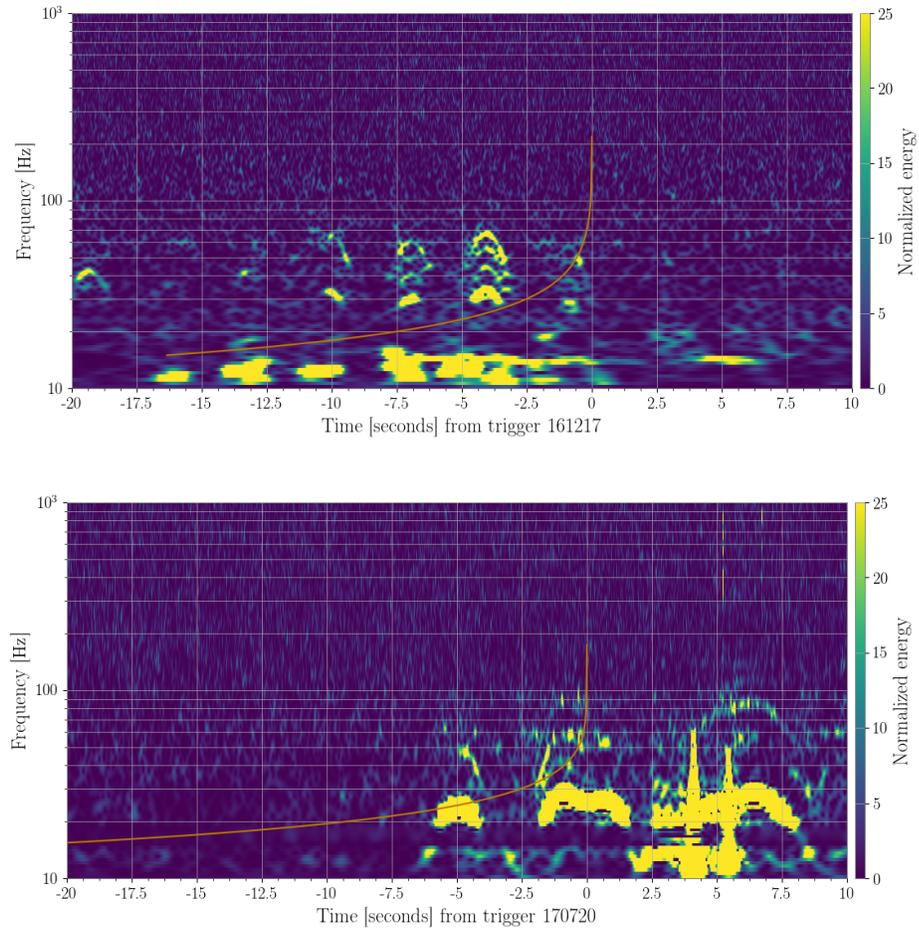


Figure 61: Spectrograms of striggers impacted by scattered light. Top: trigger 161217 at LIGO-Livingston. Bottom: trigger 170720 at LIGO-Livingston.

in matched-filter searches [83, 99, 100].

The two marginal triggers 161217 and 170720 occurred during periods of scattered light affecting frequencies up to 80 Hz with high-amplitude arches. In both cases, significant overlap with the trigger template and the excess power from scattering was observed. Investigations into the status of the observatories at the times in question identified high amplitude ground motion correlated with the scattering.

The marginal trigger 161217 occurred during a period of high-amplitude ground motion at LIGO-Livingston caused by storm activity. During this storm activity, the LIGO-Livingston detector was not able to operate continuously for longer than 10 minutes. The presence of intense scattering artifacts contributed to the unstable state of the interferometer and accounts for the SNR of the marginal trigger. As each lock stretch has a different distribution of triggers related to artifacts, it is difficult to accurately measure the background distribution in such a time. Furthermore, the factors that led to the detector not being capable of sustaining lock are also likely to render the available data less stable than a typical period of science-mode-quality data. Because of the short observing duration, this time period was not analyzed by the PyCBC search.

Within 20 seconds of trigger 170720, excess ground motion from earthquakes forced the LIGO-Livingston detector to drop out of its nominal mode of operation. Before the detector dropped out of the observing state, the data was heavily polluted with scattering artifacts that accounts for the SNR of the trigger. A data quality flag indicating the presence of an earthquake was active within 10 seconds of this trigger. However, this specific flag was constructed with a resolution of 60 seconds. If instead, the ground motion was evaluated at a cadence of 1 second, the time of trigger would have met the standard for the data quality flag. Artifacts related to scattered light were also observed at LIGO-Hanford at this time.

The PyCBC search does not consider times near the edges of observing periods, as time stretches close to lock losses are likely to be corrupted by the factors that lead to the detector losing lock. Therefore, this time period was also not analyzed by the PyCBC search.

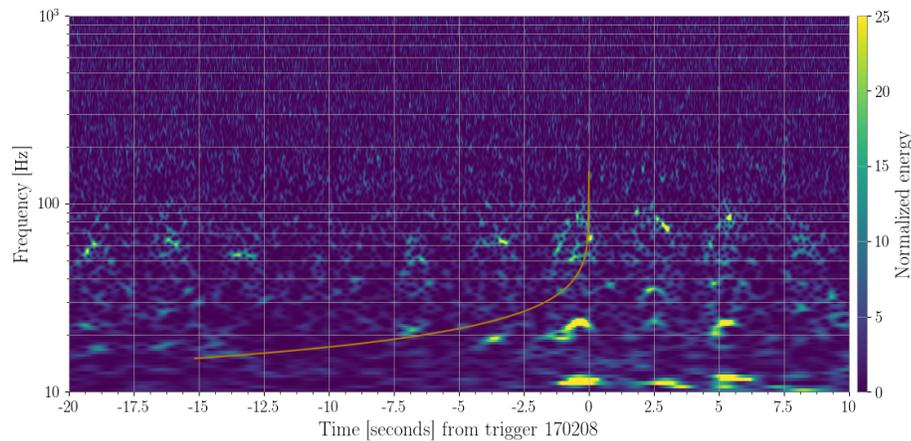


Figure 62: Spectrogram of trigger 170208 at LIGO-Hanford.

Light scattering present: 170208, 170616

Investigations into triggers 170208 and 170616 have found that light scattering does not introduce significant power above 30 Hz that overlaps the inspiral track of the triggers. In the case of these triggers, a slight overlap with excess power from scattering was observed. Multiple efforts, including BayesWave [164] glitch subtraction and gating [77], were used to mitigate the scattered light artifacts. After subtraction of the noise artifacts, the data was reanalyzed to evaluate whether the excess power subtracted could have accounted for the trigger. In both cases, the marginal trigger remained with similar significance, suggesting that the observed scattering artifacts could not have accounted for the SNR of the marginal trigger.

Scratchy glitches account for: 161202, 170412

This class of marginal triggers occurred during periods of transient noise artifacts corresponding to scratchy glitches. Previous work [83] has shown that periods of scratchy glitching can cause significant triggers in the searches, impacting the ability of searches to accurately measure the noise spectrum of the data and contributing excess noise to matched-filter searches. Triggers 161202 and 170412 demonstrate significant overlap with excess power from the nonstationarity noise. BayesWave [164] glitch subtraction was unable to completely mitigate the relevant scratchy glitch due to its long duration.

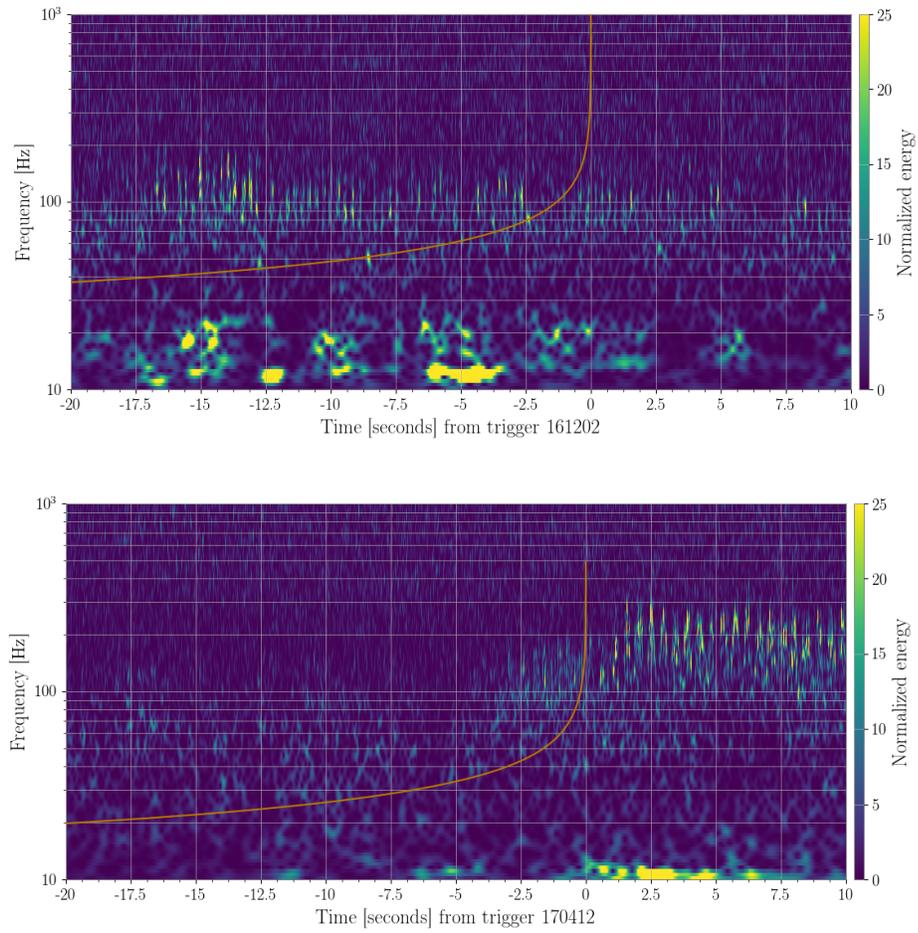


Figure 63: Spectrograms of triggers impacted by scratchy glitches with the inspiral track of the relevant trigger template plotted in orange. Top: Trigger 161202 at LIGO-Hanford. Bottom: Trigger 170412 at LIGO-Hanford.

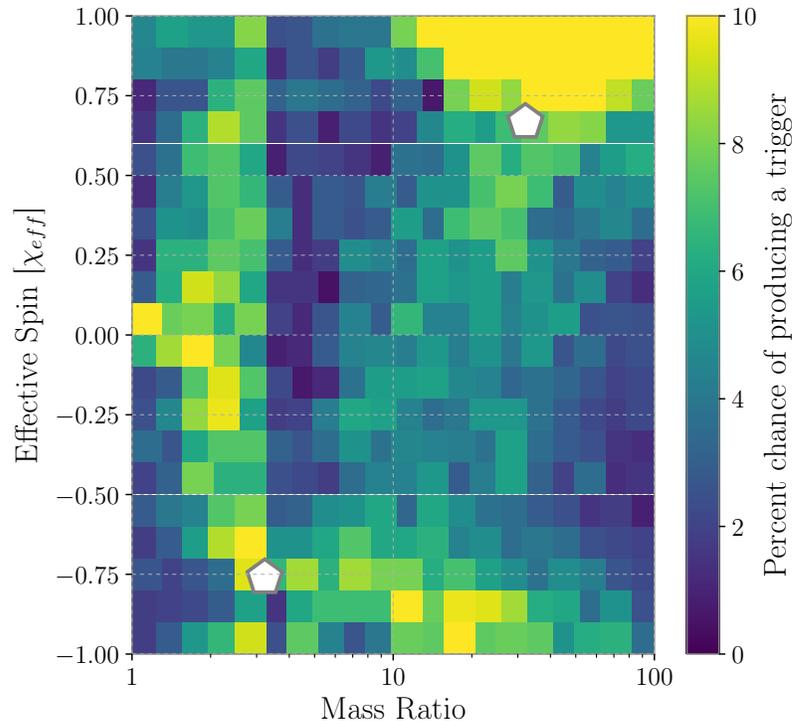


Figure 64: Probability of PyCBC triggers in O2 during scratchy glitches, as a function of mass ration and effective spin. The two catalog triggers related to scratchy glitches are marked with white pentagons. Coloring corresponds to the fraction of scratchy glitches in O2 that produced a trigger in that parameter bin above a detection statistic of 7.0.

These triggers can also be compared against the most likely paramters of triggers overlapping scratchy glitches.³ Plotting the PyCBC triggers identified alongside scratchy glitches in O2, we also add marks to indicate the parameters of these two triggers, as seen in Figure 64. The two triggers have vastly different parameters, yet still fit into regions of the parameter space where scratchy glitches are known to produce triggers. One trigger is found with anti-aligned spin and a modest mass ratio, while the other is found with aligned spin and an extreme mass ratio. This further motivates the statement that these triggers are accounted for by their temporal coincidences with a scratchy glitch.

³Additional details on this study can be found in Chapter 4.

Short-duration, high-amplitude artifacts present: 170405, 170630

The marginal triggers in this class occur in time coincidence with short-duration, high-amplitude noise transients that are removed in the data-conditioning step of the search pipelines [77] with gates, as discussed in Chapter 3. The times surrounding these transients do not demonstrate an elevated trigger rate after the transient has been removed with a gate. Trigger 170405 is in coincidence with this class of transient at LIGO-Hanford, and trigger 170630 is in coincidence with this class of transient at LIGO-Livingston. As triggers 170405 and 170630 were identified as significant after removal of the short-duration transients, the presence of noise artifacts cannot account for the SNR of these marginal triggers.

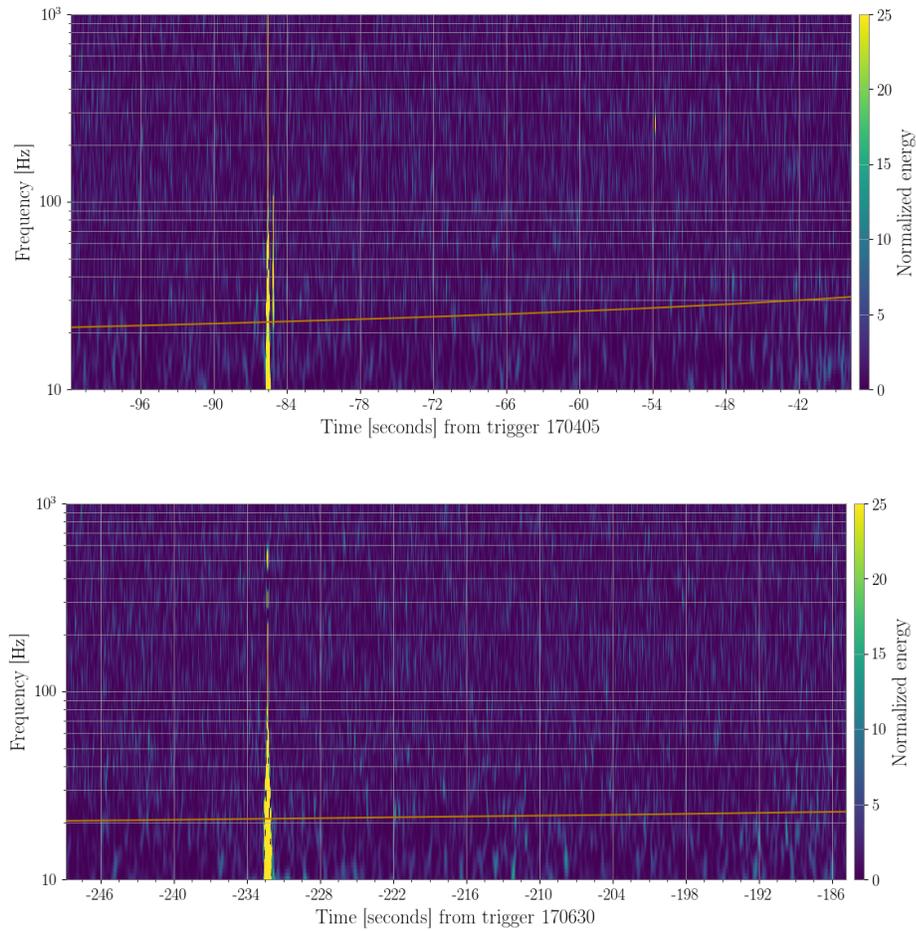


Figure 65: Spectrograms of triggers impacted by short-duration, high amplitude glitches. In each case, the inspiral track of the template is plotted in orange. Top: trigger 170405 at LIGO-Hanford. Bottom: trigger 170630 at LIGO-Livingston.

Chapter 7

Future Prospects

The second observing run with Advanced LIGO and Advanced Virgo marks a number of important milestones in the field of gravitational-wave astronomy. This is the first instance of a three-detector network made up of second generation interferometers. It also included, to date, the most massive gravitational-wave event, the least massive event, and the most distant event, as well the first event observed with both gravitational and electromagnetic waves. All of these steps have resulted in turning gravitational-wave observations from a promised future into an everyday occurrence. At the time of the publication of this dissertation, aLIGO and aVirgo have already begun their third observing run, which is expected to result in an event rate of over 1 per week. Publicly released alerts, indicating the presence of a significant gravitational-wave candidate, are facilitating worldwide follow up campaigns of these triggers, with the hope of repeating and extending the observations of GW170817.

Looking forward, it is now relevant to ask what additional discoveries are on the horizon with gravitational-wave observations. Many kinds of compact binary coalescences have yet to be detected, including events with high spins, high mass ratios, and objects from mass regimes not expected to occur as a product of typical stellar evolution, such as black holes below $5M_{\odot}$ or above $50M_{\odot}$. However, these regions are perhaps the most problematic from a data quality perspective. This means that continued exploration of the methods discussed in this dissertation will be an important aspect of ensuring that analyses of gravitational-wave data are both capable of identifying novel events and confirming that these events are astrophysical in origin.

These efforts will be complicated, going forward, by the extremely high event rate that is expected in current and future observing runs. Instead of focusing on individual events, detector characterization work must now be able to handle a large number of concurrent events, stressing the need for both automated follow ups and quantitative metrics for interpreting the data. Many of the techniques discussed in this dissertation will be an important part of these future validation procedures.

Even as there is a drive to work efficiently to handle regular data quality issues, there will always be novel issues that need human input and expertise as a part analyses of gravitational-wave events. The first and second observing runs of aLIGO have shown that even with a low event rate, there is a high likelihood of an event occurring near a significant, unexpected data quality issue that requires intervention. As the event rate increases, this chance only grows. In order for the field of gravitational-wave astrophysics to take full advantage of the highly sensitive interferometers that are now online, there will always be a need to think carefully about how the physical realities of these instruments differ from their idealized models, and how these differences impact our understanding of the gravitational-wave events they observe.

Bibliography

- [1] Albert Einstein. On the General Theory of Relativity. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1:778–786, 1915. [Addendum: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*1915,799(1915)].
- [2] K. Thorne C. Misner and J. Wheeler. *Gravitation*. WH Freeman & Co, 1973.
- [3] Frank Watson Dyson, Arthur Stanley Eddington, and Charles Davidson. IX. A determination of the deflection of light by the sun’s gravitational field, from observations made at the total eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 220(571-581):291–333, 1920.
- [4] Kazunori Akiyama, Antxon Alberdi, Walter Alef, Keiichi Asada, Rebecca Azulay, Anne-Kathrin Baczko, David Ball, Mislav Baloković, John Barrett, Dan Bintley, et al. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *The Astrophysical Journal Letters*, 875(1):L1, 2019.
- [5] C. W. F. Everitt et al. Gravity Probe B: Final Results of a Space Experiment to Test General Relativity. *Phys. Rev. Lett.*, 106:221101, May 2011.
- [6] R Abuter, A Amorim, N Anugu, M Bauböck, M Benisty, JP Berger, N Blind, H Bonnet, W Brandner, A Buron, et al. Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole. *Astronomy & Astrophysics*, 615:L15, 2018.
- [7] Anne M Archibald, Nina V Gusinskaia, Jason WT Hessels, Adam T Deller, David L Kaplan, Duncan R Lorimer, Ryan S Lynch, Scott M Ransom, and

- Ingrid H Stairs. Universality of free fall from the orbital motion of a pulsar in a stellar triple system. *Nature*, 559(7712):73, 2018.
- [8] A. Einstein. Approximative Integration of the Field Equations of Gravitation. *Sitzungsber. K. Preuss. Akad. Wiss.*, 1:688, 1916.
- [9] R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astroph. J. Lett.*, 195:L51–L53, January 1975.
- [10] J. H. Taylor and J. M. Weisberg. A new test of general relativity : gravitational radiation and the binary pulsar PSR 1913 +16. *The Astrophysical Journal*, 253:908, 1982.
- [11] J Aasi et al. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, 2015.
- [12] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, February 2016.
- [13] B P Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. 2018. arXiv:1811.12907.
- [14] F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Class. Quantum Grav.*, 32(2):024001, 2015.
- [15] B P Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119:161101, 2017.
- [16] Benjamin P Abbott, Richard Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. Multi-messenger observations of a binary neutron star merger. *Astrophys. J. Lett*, 848(2):L12, 2017.
- [17] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 1. Signal-to-noise for inspiral, merger, and ring-down. *Phys. Rev. D*, 57:4535–4565, 1998.

- [18] Giles Hammond, Stefan Hild, and Matthew Pitkin. Advanced technologies for future laser-interferometric gravitational wave detectors. *Journal of Modern Optics*, 61, 02 2014.
- [19] Peter R. Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors*. World Scientific, Hackensack, NJ, 2nd edition, 2017.
- [20] C Cutler and E Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658, 1994.
- [21] K. Schwarzschild. On the gravitational field of a mass point according to Einstein’s theory. *Sitzungsber. K. Preuss. Akad. Wiss.*, 1:189–196, 1916.
- [22] Frank Elavsky. Masses in the Stellar Graveyard: GWTC-1. <https://dcc.ligo.org/LIGO-G1802255>, 2018.
- [23] A. Heger, C. L. Fryer, S. E. Woosley, N. Langer, and D. H. Hartmann. How Massive Single Stars End Their Life. *Astrophysical J.*, 591:288–300, July 2003.
- [24] Nihan Pol, Maura McLaughlin, and Duncan R. Lorimer. Future prospects for ground-based gravitational-wave detectors: The galactic double neutron star merger rate revisited. *The Astrophysical Journal*, 870(2):71, Jan 2019.
- [25] S. Chandrasekhar and S. Detweiler. The quasi-normal modes of the Schwarzschild black hole. *Proc. Roy. Soc. Lond.*, A344:441–452, 1975.
- [26] Feryal Ozel, Dimitrios Psaltis, Ramesh Narayan, and Antonio Santos Villarreal. On the Mass Distribution and Birth Masses of Neutron Stars. *Astrophys.J.*, 757:55, 2012.
- [27] BP Abbott, R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. GW170817: Measurements of neutron star radii and equation of state. *Physical Review Letters*, 121(16):161101, 2018.
- [28] Soumi De, Daniel Finstad, James M Lattimer, Duncan A Brown, Edo Berger, and Christopher M Biwer. Tidal Deformabilities and Radii of Neutron Stars

- from the Observation of GW170817. *Physical Review Letters*, 121(9):091102, 2018.
- [29] Jens Hjorth, Andrew J Levan, Nial R Tanvir, Joe D Lyman, Radosław Wojtak, Sophie L Schröder, Ilya Mandel, Christa Gall, and Sofie H Bruun. The distance to ngc 4993: the host galaxy of the gravitational-wave event gw170817. *The Astrophysical Journal Letters*, 848(2):L31, 2017.
- [30] B. P. Abbott et al. Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817a. *The Astrophysical Journal*, 848(2):L13, Oct 2017.
- [31] M. R. Burleigh, M. A. Barstow, Howard E. Bond, J. B. Holberg, I. Hubeny, and D. Koester. Hubble Space Telescope spectroscopy of the Balmer lines in Sirius B. *Monthly Notices of the Royal Astronomical Society*, 362(4):1134–1142, 10 2005.
- [32] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. *Astron. and Astrophys.*, 375:890–898, September 2001.
- [33] Viktoriya Morozova, David Radice, Adam Burrows, and David Vartanyan. The gravitational wave signal from core-collapse supernovae. *The Astrophysical Journal*, 861(1):10, 2018.
- [34] David Radice, Viktoriya Morozova, Adam Burrows, David Vartanyan, and Hiroki Nagakura. Characterizing the gravitational wave signal from core-collapse supernovae. *The Astrophysical Journal Letters*, 876(1):L9, 2019.
- [35] Vincent Roma, Jade Powell, Ik Siong Heng, and Raymond Frey. Astrophysics with core-collapse supernova gravitational wave signals in the next generation of gravitational wave detectors. *Phys. Rev. D*, 99:063018, Mar 2019.
- [36] BP Abbott, R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, C Adams, RX Adhikari, VB Adya, C Affeldt, et al. Search for transient gravitational wave signals associated with magnetar bursts during Advanced LIGO’s second observing run. 2019.

- [37] Bence Kocsis, Merse Előd Gaspar, and Szabolcs Marka. Detection rate estimates of gravity waves emitted during parabolic encounters of stellar black holes in globular clusters. *The Astrophysical Journal*, 648(1):411–429, sep 2006.
- [38] Weidong Li et al. Nearby supernova rates from the Lick Observatory Supernova Search - III. The rate-size relation, and the rates as a function of galaxy Hubble type and colour. *Monthly Notices of the Royal Astronomical Society*, 412(3):1473–1507, 04 2011.
- [39] K Riles. Recent Searches for Continuous Gravitational waves. *Phys. Rev. A*, 32:1730035, 2017.
- [40] T Regimbau. The astrophysical gravitational wave stochastic background. *Research in Astronomy and Astrophysics*, 11(4):369, 2011.
- [41] J. Weber. Gravitational-wave-detector events. *Phys. Rev. Lett.*, 20:1307–1308, Jun 1968.
- [42] R Weiss. Electromagnetically coupled broadband gravitational wave antenna. *Quarterly Progress Report, Research Laboratory of Electronics of MIT*, 105:54, April 1972.
- [43] Albert A Michelson and Edward W Morley. On the relative motion of the earth and of the luminiferous ether. *Sidereal Messenger*, vol. 6, pp. 306-310, 6:306–310, 1887.
- [44] LIGO-Caltech. Basic michelson interferometer, 2019.
- [45] Hermann A Haus. *Waves and fields in optoelectronics*. Prentice-Hall,, 1984.
- [46] D. V. Martynov et al. Sensitivity of the advanced ligo detectors at the beginning of gravitational wave astronomy. *Phys. Rev. D*, 93:112004, Jun 2016.
- [47] Donald R. Herriott and Harry J. Schulte. Folded optical delay lines. *Appl. Opt.*, 4(8):883–889, Aug 1965.
- [48] Kiwamu Izumi. *Multi-Color Interferometry for Lock Acquisition of Laser Interferometric Gravitational-wave Detectors*. PhD thesis, University of Tokyo, Tokyo, Japan, 2012.

- [49] Brian J Meers. Recycling in laser-interferometric gravitational-wave detectors. *Physical Review D*, 38(8):2317, 1988.
- [50] Chris L Mueller, Muzammil A Arain, Giacomo Ciani, Ryan T DeRosa, Anamaria Effler, David Feldbaum, Valery V Frolov, Paul Fulda, Joseph Gleason, Matthew Heintze, et al. The advanced LIGO input optics. *Review of Scientific Instruments*, 87(1):014502, 2016.
- [51] Tobin T Fricke, Nicolás D Smith-Lefebvre, Richard Abbott, Rana Adhikari, Katherine L Dooley, Matthew Evans, Peter Fritschel, Valery V Frolov, Keita Kawabe, Jeffrey S Kissel, et al. DC readout experiment in Enhanced LIGO. *Classical and Quantum Gravity*, 29(6):065005, 2012.
- [52] Carlton M Caves. Quantum-mechanical radiation-pressure fluctuations in an interferometer. *Physical Review Letters*, 45(2):75, 1980.
- [53] Carlton M Caves. Quantum-mechanical noise in an interferometer. *Physical Review D*, 23(8):1693, 1981.
- [54] A. Gillespie and F. Raab. Thermally excited vibrations of the mirrors of laser interferometer gravitational-wave detectors. *Phys. Rev. D*, 52:577–585, Jul 1995.
- [55] Yu. Levin. Internal thermal noise in the LIGO test masses: A direct approach. *Phys. Rev. D*, 57:659–663, Jan 1998.
- [56] Gregory M Harry, Andri M Gretarsson, Peter R Saulson, Scott E Kittelberger, Steven D Penn, William J Startin, Sheila Rowan, Martin M Fejer, D R M Crooks, Gianpietro Cagnoli, Jim Hough, and Norio Nakagawa. Thermal noise in interferometric gravitational wave detectors due to dielectric optical coatings. *Classical and Quantum Gravity*, 19(5):897–917, Feb 2002.
- [57] Gregory M. Harry, Helena Armandula, Eric Black, D. R. M. Crooks, Gianpietro Cagnoli, Jim Hough, Peter Murray, Stuart Reid, Sheila Rowan, Peter Sneddon, Martin M. Fejer, Roger Route, and Steven D. Penn. Thermal noise from optical coatings in gravitational wave detectors. *Appl. Opt.*, 45(7):1569–1574, Mar 2006.
- [58] Peter R Saulson. Thermal noise in mechanical experiments. *Physical Review D*, 42(8):2437, 1990.

- [59] Gabriela I González and Peter R Saulson. Brownian motion of a mass suspended by an anelastic wire. *The Journal of the Acoustical Society of America*, 96(1):207–212, 1994.
- [60] Gabriela González. Suspensions thermal noise in the LIGO gravitational wave detector. *Classical and Quantum Gravity*, 17(21):4409, 2000.
- [61] F Matchard, B Lantz, R Mittleman, K Mason, J Kissel, B Abbott, S Biscans, J McIver, R Abbott, S Abbott, et al. Seismic isolation of Advanced LIGO: Review of strategy, instrumentation and performance. *Classical and Quantum Gravity*, 32(18):185003, 2015.
- [62] D M Macleod, S Fairhurst, B Hughey, A P Lundgren, L Pekowsky, J Rollins, and J R Smith. Reducing the effect of seismic noise in LIGO searches by targeted veto generation. *Classical and Quantum Gravity*, 29(5):055006, feb 2012.
- [63] B. P. Abbott et al. Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. *Class. Quant. Grav.*, 33(13):134001, 2016.
- [64] M. W. Coughlin, J. Harms, J. Driggers, D. J. McManus, N. Mukund, M. P. Ross, B. J. J. Slagmolen, and K. Venkateswara. Implications of Dedicated Seismometer Measurements on Newtonian-Noise Cancellation for Advanced LIGO. *Phys. Rev. Lett.*, 121:221104, Nov 2018.
- [65] J C Driggers, M Evans, K Pepper, and R X Adhikari. Active noise cancellation in a suspended interferometer. *Review of Scientific Instruments*, 83(2):024501, 2012.
- [66] M Coughlin, N Mukund, J Harms, J Driggers, R Adhikari, and S Mitra. Towards a first design of a Newtonian-noise cancellation system for Advanced LIGO. *Classical and Quantum Gravity*, 33(24):244001, 2016.
- [67] Sheila Dwyer et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=32994>.

- [68] V Necula, S Klimenko, and G Mitselmakher. Transient analysis with fast Wilson-Daubechies time-frequency transform. In *Journal of Physics: Conference Series*, volume 363, page 012032. IOP Publishing, 2012.
- [69] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis. Multiresolution techniques for the detection of gravitational-wave bursts. *Class. Quantum Grav.*, 21:S1809–S1818, 2004.
- [70] BP Abbott et al. Observing gravitational-wave transient GW150914 with minimal assumptions. *Phys. Rev. D*, 93:122004, Jun 2016.
- [71] S. Klimenko et al. Coherent method for detection of gravitational wave bursts. *Class. Quantum Grav.*, 25:114029, 2008.
- [72] R Lynch, S Vitale, R Essick, E Katsavounidis, and F Robinet. An information-theoretic approach to the gravitational-wave burst detection problem. 2015.
- [73] Visual comparison of convolution, cross correlation and autocorrelation of two signals. https://upload.wikimedia.org/wikipedia/commons/2/21/Comparison_convolution_correlation.svg, 2016.
- [74] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev. D*, 85:122006, 2012.
- [75] B. P. Abbott et al. Tests of general relativity with GW150914. *Phys. Rev. Lett.*, 116(22):221101, 2016.
- [76] B P Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119:141101, 2017.
- [77] Samantha A Usman, Alexander H Nitz, Ian W Harry, Christopher M Biwer, Duncan A Brown, Miriam Cabero, Collin D Capano, Tito Dal Canton, Thomas Dent, Stephen Fairhurst, et al. The PyCBC search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33(21):215004, 2016.

- [78] C Messick et al. Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data. *Phys. Rev. D*, 95:042001, Feb 2017.
- [79] T Adams, D Buskulic, V Germain, GM Guidi, F Marion, Matteo Montani, B Mours, Francesco Piergiovanni, and G Wang. Low-latency analysis pipeline for compact binary coalescences in the advanced gravitational wave detector era. *Classical and Quantum Gravity*, 33(17):175012, 2016.
- [80] Qi Chu. *Low-latency detection and localization of gravitational waves from compact binary coalescences*. PhD thesis, The University of Western Australia, Perth, Australia, 2017.
- [81] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. A New Search Pipeline for Compact Binary Mergers: Results for Binary Black Holes in the First Observing Run of Advanced LIGO. 2019. arXiv:1902.10341.
- [82] T Dal Canton and I W Harry. Designing a template bank to observe compact binary coalescences in Advanced LIGO’s second observing run. 2017. arXiv:1705.01845.
- [83] B P Abbott et al. Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO’s first observing run. *Classical and Quantum Gravity*, 35(6):065010, 2018.
- [84] Marissa Walker, Alfonso F Agnew, Jeffrey Bidler, Andrew Lundgren, Alexandra Macedo, Duncan Macleod, TJ Massinger, Oliver Patane, and Joshua R Smith. Identifying correlations between LIGO’s astronomical range and auxiliary sensors using lasso regression. *Classical and Quantum Gravity*, 35(22):225002, 2018.
- [85] Bruce Allen. LIGO calibration accuracy. Technical Report LIGO-T960189-00-Z, LIGO Project, 1996.
- [86] Alexander H Nitz. Distinguishing short duration noise transients in LIGO data to improve the PyCBC search for gravitational waves from high mass binary black hole mergers. *Classical and Quantum Gravity*, 35(3):035016, 2018.

- [87] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and improving the sensitivity of the PyCBC search. *The Astrophysical Journal*, 849(2):118, nov 2017.
- [88] Alexander H Nitz, Collin Capano, Alex B Nielsen, Steven Reyes, Rebecca White, Duncan A Brown, and Badri Krishnan. 1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data. *The Astrophysical Journal*, 872(2):195, 2019.
- [89] A H Nitz, T Dal Canton, D Davis, and S Reyes. Rapid detection of gravitational waves from compact binary mergers with pycbc live. *Phys. Rev. D*, 98:024050, Jul 2018.
- [90] Collin Capano, Thomas Dent, Chad Hanna, Martin Hendry, Chris Messenger, Y-M Hu, and John Veitch. Systematic errors in estimation of gravitational-wave candidate significance. *Physical Review D*, 96(8):082002, 2017.
- [91] J. McIver et al. Mitigating noise artifacts in gravitational-wave detector data in the era of open public alerts. 2019. In Preparation.
- [92] Miriam Cabero, Andrew Lundgren, Alexander Harvey Nitz, Thomas Dent, David Barker, Evan Goetz, Jeffrey Kissel, Laura K Nuttall, Paul Schale, Robert Schofield, and Derek Davis. Blip glitches in Advanced LIGO data. *Classical and Quantum Gravity*, 2019.
- [93] T. J. Massinger et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=42851>.
- [94] L K Nuttall et al. Improving the Data Quality of Advanced LIGO Based on Early Engineering Run Results. *Class. Quantum Grav.*, 32(24):245005, 2015.
- [95] R. Abbott et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=26227>.
- [96] T. Accadia et al. Noise from scattered light in Virgo’s second science run data. *Class. Quant. Grav.*, 27(19):194011, 2010.

- [97] David J Ottaway, Peter Fritschel, and Samuel J. Waldman. Impact of upconverted scattered light on advanced interferometric gravitational wave detectors. *Opt. Express*, 20(8):8329–8336, Apr 2012.
- [98] Guillermo Valdes, Brian O’Reilly, and Mario Diaz. A Hilbert–Huang transform method for scattering identification in LIGO. *Class. Quant. Grav.*, 34(23):235009, 2017.
- [99] L. K. Nuttall. Characterizing transient noise in the LIGO detectors. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 376(2120), 2018.
- [100] J. McIver et al. Data quality studies of enhanced interferometric gravitational wave detectors. *Class. Quant. Grav.*, 29(12):124010, 2012.
- [101] O. Patane et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=43177>.
- [102] M Zevin, S Coughlin, S Bahaadini, E Besler, N Rohani, S Allen, M Cabero, K Crowston, A K Katsaggelos, S L Larson, T K Lee, C Lintott, T B Littenberg, A Lundgren, C Østerlund, J R Smith, L Trouille, and V Kalogera. Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science. *Classical and Quantum Gravity*, 34(6):064003, 2017.
- [103] F. Robinet. Omicron: An Algorithm to Detect and Characterize Transient Noise in Gravitational-Wave Detectors. <https://tds.ego-gw.it/ql/?c=10651>, 2015.
- [104] J R Smith, T Abbott, E Hirose, N Leroy, D MacLeod, J McIver, P Saulson, and P Shawhan. A hierarchical method for vetoing noise transients in gravitational-wave detectors. *Classical and Quantum Gravity*, 28(23):235005, 2011.
- [105] Rahul Biswas, Lindy Blackburn, Junwei Cao, Reed Essick, Kari Alison Hodge, Erotokritos Katsavounidis, Kyungmin Kim, Young-Min Kim, Eric-Olivier Le Bigot, Chang-Hwan Lee, John J. Oh, Sang Hoon Oh, Edwin J. Son, Ye Tao, Ruslan Vaulin, and Xiaoge Wang. Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. *Phys. Rev. D*, 88:062003, Sep 2013.

- [106] Duncan Macleod, Alex L. Urban, Maximiliano Isi, Thomas Massinger, paulaltin, Matt Pitkin, and Alex Nitz. gwpy/gwsumm: 0.1.4, April 2019.
- [107] J. Areeda et al. LigoDV-web: Providing easy, secure and universal access to a large distributed scientific data store for the LIGO Scientific Collaboration. 2016.
- [108] B P Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118:221101, 2017.
- [109] B P Abbott et al. GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence. *The Astrophysical Journal Letters*, 851(2):L35, 2017.
- [110] T. J. Massinger et al. Detector characterization in the second observing run of Advanced LIGO. 2019. In Preparation.
- [111] Michele Vallisneri, Jonah Kanner, Roy Williams, Alan Weinstein, and Branson Stephens. The LIGO open science center. *Journal of Physics: Conference Series*, 610:012021, May 2015.
- [112] Gravitational-Wave Open Science Center. <https://www.gw-openscience.org>.
- [113] Maggie Tse et al. PEM Channel Info. <https://http://pem.ligo.org>.
- [114] T Hardwick et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=32651>.
- [115] Seiji Kawamura and Michael E Zucker. Mirror-orientation noise in a Fabry–Perot interferometer gravitational wave detector. *Applied optics*, 33(18):3912–3918, 1994.
- [116] B. Abbott et al. Search for intermediate mass black hole binaries in the first and second observing runs of the advanced ligo and virgo network. 2019. arXiv:1906.08000.
- [117] Chris Pankow, Katerina Chatziioannou, Eve A Chase, Tyson B Littenberg, Matthew Evans, Jessica McIver, Neil J Cornish, Carl-Johan Haster, Jonah

- Kanner, Vivien Raymond, et al. Mitigation of the instrumental noise transient in gravitational-wave data surrounding GW170817. *Physical Review D*, 98(8):084016, 2018.
- [118] Beverly K. Berger. Identification and mitigation of Advanced LIGO noise sources. *Journal of Physics: Conference Series*, 957(1):012004, 2018.
- [119] B P Abbott et al. Low-Latency Gravitational Wave Alerts for Multi-Messenger Astronomy During the Second Advanced LIGO and Virgo Observing Run. 2019. arXiv:1901.03310.
- [120] A H Nitz et al. PyCBC Software. <https://github.com/ligo-cbc/pycbc>, 2019.
- [121] S Coughlin et al. Updated Gravity Spy Data Set, 2018.
- [122] L White et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=43249>.
- [123] O Patane et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=43177>.
- [124] G Vajente et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=35073>.
- [125] J Smith et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=32580>.
- [126] S Coughlin et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=30804>.
- [127] A Lundgren et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=27138>.
- [128] A Lundgren et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=44194>.
- [129] J Smith et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=42402>.

- [130] T Massinger et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=42851>.
- [131] P Meyers et al. aLIGO LLO Logbook. <https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=35195>.
- [132] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [133] B P Abbott et al. Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO. *Phys. Rev. D*, 96:022001, Jul 2017.
- [134] James Healy and Carlos O. Lousto. Hangup effect in unequal mass binary black hole mergers and further studies of their gravitational radiation and remnant properties. *Phys. Rev. D*, 97:084002, Apr 2018.
- [135] J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, C. Affeldt, and et al. Search for gravitational waves from intermediate mass binary black holes. *Phys. Rev. D*, 85(10):102004, May 2012.
- [136] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, K Ackley, C Adams, T Adams, P Addresso, RX Adhikari, et al. Upper limits on the rates of binary neutron star and neutron star–black hole mergers from advanced LIGO’s first observing run. *The Astrophysical Journal Letters*, 832(2):L21, 2016.
- [137] Kipp Cannon et al. Toward early-warning detection of gravitational waves from compact binary coalescence. *The Astrophysical Journal*, 748(2):136, 2012.
- [138] R DeRosa, J C Driggers, D Atkinson, H Miao, V Frolov, M Landry, J A Gaiame, and R X Adhikari. Global feed-forward vibration isolation in a km scale interferometer. *Classical and Quantum Gravity*, 29(21):215008, 2012.

- [139] V Tiwari, M Drago, V Frolov, S Klimenko, G Mitselmakher, V Necula, G Prodi, V Re, F Salemi, G Vedovato, and I Yakushin. Regression of environmental noise in LIGO data. *Classical and Quantum Gravity*, 32(16):165014, 2015.
- [140] G D Meadors, K Kawabe, and K Riles. Increasing LIGO sensitivity by feed-forward subtraction of auxiliary length control noise. *Classical and Quantum Gravity*, 31(10):105014, 2014.
- [141] J Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Phys. Rev. D*, 91:042003, Feb 2015.
- [142] J C Driggers et al. Offline noise subtraction for Advanced LIGO. *Technical document LIGO-P1700260*, 2017.
- [143] JC Driggers, S Vitale, AP Lundgren, M Evans, K Kawabe, SE Dwyer, K Izumi, RMS Schofield, A Effler, D Sigg, et al. Improving astrophysical parameter estimation via offline noise subtraction for Advanced LIGO. *Physical Review D*, 99(4):042001, 2019.
- [144] Derek Davis, Thomas Massinger, Andrew Lundgren, Jennifer C Driggers, Alex L Urban, and Laura Nuttall. Improving the sensitivity of Advanced LIGO using noise subtraction. *Classical and Quantum Gravity*, 36(5):055011, 2019.
- [145] B Allen et al. Automatic cross-talk removal from multi-channel data. 1999. arXiv:gr-qc/9909083.
- [146] B. Allen et al. Observational limit on gravitational waves from binary neutron stars in the galaxy. *Phys. Rev. Lett.*, 83:1498, 1999.
- [147] E Deelman et al. Pegasus: a workflow management system for science automation. *Future Generation Computer Systems*, 46:17–35, 2015. Funding Acknowledgements: NSF ACI SDCI 0722019, NSF ACI SI2-SSI 1148515 and NSF OCI-1053575.
- [148] P Kwee et al. Stabilized high-power laser system for the gravitational wave detector advanced LIGO. *Opt. Express*, 20(10):10617–10634, May 2012.

- [149] R Schofield. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=30290>.
- [150] G Vajente. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=302734>.
- [151] S Dwyer. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=34154>.
- [152] A Brooks. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=34853>.
- [153] A F Brooks et al. Overview of Advanced LIGO adaptive optics. *Appl. Opt.*, 55(29):8256–8265, Oct 2016.
- [154] S Karki et al. The Advanced LIGO photon calibrators. *Review of Scientific Instruments*, 87:114503, 2016.
- [155] C Biwer et al. Validating gravitational-wave detections: The Advanced LIGO hardware injection system. *Phys. Rev. D*, 95:062002, Mar 2017.
- [156] A D Viets et al. Reconstructing the calibrated strain signal in the Advanced LIGO detectors. *Classical and Quantum Gravity*, 35(9):095015, 2018.
- [157] C Cahillane et al. Calibration uncertainty for Advanced LIGO’s first and second observing runs. *Phys. Rev. D*, 96:102001, Nov 2017.
- [158] P B Covas et al. Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of Advanced LIGO. *Phys. Rev. D*, 97:082002, 2018.
- [159] J Kissel et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=37846>.
- [160] Tyson B Littenberg, Ben Farr, Scott Coughlin, Vicky Kalogera, and Daniel E Holz. Neutron Stars versus Black Holes: Probing the Mass Gap with LIGO/Virgo. *The Astrophysical Journal Letters*, 807(2):L24, 2015.

- [161] Will M Farr, Niharika Sravan, Andrew Cantrell, Laura Kreidberg, Charles D Bailyn, Ilya Mandel, and Vicky Kalogera. The mass distribution of stellar-mass black holes. *The Astrophysical Journal*, 741(2):103, 2011.
- [162] Feryal Özel, Dimitrios Psaltis, Ramesh Narayan, and Jeffrey E. McClintock. The black hole mass distribution in the galaxy. *The Astrophysical Journal*, 725(2):1918, 2010.
- [163] T Massinger et al. aLIGO LHO Logbook. <https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=47199>.
- [164] N J Cornish and T B Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *Class. Quant. Grav.*, 32(13):135012, 2015.
- [165] Bernard F. Schutz. Networks of gravitational wave detectors and three figures of merit. *Class. Quantum Grav.*, 28:125023, 2011.
- [166] Shasvath J Kapadia, Sarah Caudill, Jolien DE Creighton, Will M Farr, Gregory Mendell, Alan Weinstein, Kipp Cannon, Heather Fong, Patrick Godwin, Rico KL Lo, et al. A self-consistent method to estimate the rate of compact binary coalescences with a poisson mixture model. *arXiv preprint arXiv:1903.06881*, 2019.
- [167] S. Klimenko, G. Vedovato, M. Drago, G. Mazzolo, G. Mitselmakher, C. Pankow, G. Prodi, F. Re, V. Salemi, and I. Yakushin. Localization of gravitational wave sources with networks of advanced detectors. *Phys.Rev.*, D83:102001, 2011.
- [168] S Vitale et al. Calibration requirements for CBC parameter estimation in O2. *Technical document LIGO-T1600409*, 2016. <http://dcc.ligo.org/LIGO-T1600409/public>.

Derek Davis

Department of Physics
Syracuse University
Syracuse, NY 13244 USA

Dissertation:

- Improving the Sensitivity of Advanced LIGO Through Detector Characterization

Previous Degrees:

- B.A., Physics, University of Chicago, 2015
- B.S., Mathematics, University of Chicago, 2015

Awards and Fellowships:

- LIGO Laboratory Award for Excellence in Detector Characterization and Calibration
- Henry Levinstein Fellowship

Teaching Experience:

- Instructor - PHY 211 (Summer 2016)
- Teaching Assistant - AST 101 (Fall 2016)
- Teaching Assistant - PHY 211 (Fall 2015/Spring 2016)

Selected Presentations:

- *The Impact of Detector Noise on Gravitational-Wave Astrophysics* - Astrophysics with Gravitational Wave Populations Workshop in Aspen, Colorado, February 2019
- *Detector Characterization in the Era of Gravitational-Wave Astronomy* - LIGO Laboratory Award Seminar at California Institute of Technology, February 2019

- *The Role of Detector Characterization in Gravitational-Wave Astronomy* - University of Wisconsin- Milwaukee, February 2019
- *Discovering New Gravitational Waves in Noisy LIGO Data* - Marquette University, January 2019
- *The Impact of Detector Noise on Multi-messenger Gravitational-Wave Astrophysics* - JSI Gravitational Wave Physics and Astronomy Workshop in College Park, MD, December 2018
- *Gravitational Wave Detection: Challenges in the Noise* - Portsmouth University, Portsmouth, UK, September 2018
- *Detector Characterization for the O1-O2 CBC Catalog* - LIGO-Virgo Conference in Maastricht, Netherlands, September 2018
- *O2 LIGO Noise Subtraction* - LIGO-Virgo Conference in Sonoma, California, March 2018
- *The Impact of Detector Characterization on Gravitational Wave Searches* - California Institute of Technology, October 2017
- *The Most Troublesome LIGO Noise Sources for the Searches* - LIGO-Virgo Conference in Geneva, Switzerland, September 2017

Selected Publications:

- Derek Davis, Thomas Massinger, Andrew Lundgren, Jennifer C Driggers, Alex L Urban, and Laura Nuttall. Improving the sensitivity of Advanced LIGO using noise subtraction. *Classical and Quantum Gravity*, 36(5):055011, 2019
- A H Nitz, T Dal Canton, D Davis, and S Reyes. Rapid detection of gravitational waves from compact binary mergers with pycbc live. *Phys. Rev. D*, 98:024050, Jul 2018
- Miriam Cabero, Andrew Lundgren, Alexander Harvey Nitz, Thomas Dent, David Barker, Evan Goetz, Jeffrey Kissel, Laura K Nuttall, Paul Schale, Robert Schofield, and Derek Davis. Blip glitches in Advanced LIGO data. *Classical and Quantum Gravity*, 2019

- B P Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. 2018. arXiv:1811.12907
- B P Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119:161101, 2017
- B P Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119:141101, 2017
- B P Abbott et al. GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence. *The Astrophysical Journal Letters*, 851(2):L35, 2017
- B P Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118:221101, 2017