# Searching for gravitational waves from compact binary coalescences in Advanced LIGO data

Thesis by
Surabhi Sachdev

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy


CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended October 12, 2018

Surabhi Sachdev
ORCID: 0000-0002-0525-2317

# ACKNOWLEDGEMENTS

Some of the happiest but also the hardest moments of my life happened during graduate school. There are a lot of people who were with me through both the happy and hard times and this thesis would not have happened without them.

First and foremost, I would like to thank my adviser Alan Weinstein for always believing in me, even at times I didn't believe in myself. He gave me the most exciting projects to work on and was always there whenever I needed advice. I would like to thank Tjonnie Li for mentoring me through the first couple years in grad school, for holding my hand through some of the first lines of code I wrote, for supporting and encouraging me, and for always lending his ears when I needed. Stephen Privitera, for being the unexpected friend and also providing academic support. Chad Hanna, Bangalore Sathyaprakash, Tania Regimbau have all provided me with both guidance at work and moral support for which I am extremely grateful. I would also like to thank collaborators Cody Messick, Ryan Magee, Sarah Caudill, Duncan Meacher, Alex Pace, Patrick Godwin, Jolien Creighton, Shasvath Kapadia, Heather Fong, Kipp Cannon, Richard Savage, Sudarshan Karki, Shivaraj Kandhasamy, Jeff Kissel. SURF students Sophia Xiao, Sara Frederick, Ka Tung Lau, Johnathon Lowery, Ka Lok Lo, Li Ka Yue were all amazing students who did exceptional work over summers and made the summers more fun. Office mates Maximiliano Isi, Craig Cahillane, Sarah Gossan, Thomas Callister always kept things fun at work, and conference buddies Karelle Siellez and Ryan Lynch made the trips fun, a big thank you to them for that. Julie, Elizabeth, Nicole, Sophie were always a friendly face and super helpful with any official work.

I would like to thank Ma and Papa for being my first teachers, for always supporting me, for letting me make my own decisions, and for always being there for me. To bhaiyya, bhabhi, and Naina for being my home away from home. Corina Panda for being the absolute best friend, roommate, cooking partner, dancing partner, gossip partner, party partner, and for being with me through it all, I couldn't have done it without her. I would also like to thank Ramya Korlakai Vinayak, Brian Hong, Akshay Sridhar, Maria Sakovsky, Alex Turzillo, Daniel Gonzalez Plaza, Paul Mazur, Arnold Durel for adding fun and silliness to my time at Caltech. I am grateful to the Caltech Tango Club and its members for beautiful dance evenings. And, Michael Buehlmann for showing me love during my last year at Caltech.

I am a member of the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. I carried out most of the work presented here as a member of the LIGO Scientific Collaboration (LSC) and benefited from interaction with many of its members. However, except where otherwise indicated, the descriptions and results presented in this thesis are my own and not necessarily those of the Collaboration.

# ABSTRACT

Advanced LIGO's first observing run marked the birth of gravitational-wave astronomy through the first detection of gravitational waves from coalescing black holes- GW150914. Advanced LIGO's second and Advanced Virgo's first observing run marked the birth of multimessenger astronomy with first joint observations of gravitational and electromagnetic radiation associated with coalescing neutron stars-GW170817. The electromagnetic observations included detection of a burst of gamma rays produced by the merger, and a kilonova powered by the radioactive decay of r-process nuclei synthesized in the neutron star coalescence ejecta. Gravitational waves from compact binary coalescences carry fingerprints of the sources that generated them. Studying them allows us to test Einstein's general relativity in the strongest regimes, where it has never been tested before, and study matter at densities beyond reach of the most powerful laboratories on our planet. Moreover, we can gain insight about the evolution of stars, galaxies and even the Universe as a whole by studying the merger rate of compact objects. Joint electromagnetic and gravitational-wave observations help develop our understanding of the physical processes that occur in such systems, and provide a new method of probing cosmological parameters.

GW170817 was detected by the `GstLAL` pipeline in low-latency making the extensive electromagnetic followup possible. The `GstLAL` pipeline is a matched filtering pipeline that uses compact binary coalescence waveform models to filter the data from gravitational-wave detectors in the time-domain. It can detect gravitational waves from coalescing compact binaries in near real time and provide point estimates for binary parameters. This thesis describes the methods, developments, and the results from the `GstLAL` pipeline over the course of the first two observing runs of Advanced LIGO, focusing on the contributions made by the author. We also present a study about the prospects of observing a cosmological stochastic background which is expected to be buried under the astrophysical background from the population of coalesceing compact binaries with third-generation gravitational-wave detectors.

# PUBLISHED CONTENT AND CONTRIBUTIONS

Chapter 4 is a partial adaptation of [1].

Chapter 5 presents unpublished work that was led by S. Sachdev in collaboration with T. G. F. Li, S. Privitera, S. Xiao, K. Blackburn, P. Schmidt, A. Weinstein, and the GstLAL and CBC working groups of the LIGO-Virgo Collaboration.

Sec. 6.3 is unpublished work that was led by S. Sachdev in collaboration with the GstLAL working group of the LIGO-Virgo Collaboration.

Sec. 6.5 presents some of the results from [2], [3], [4].

Sec. 7.1 - the methods described here are derived from arXiv preprints [5], [6].

Sec. 7.2 presents some of the results from [7], [8].

Chapter 8 - the work described here is from a paper in preparation not yet published.

[1] Messick, C., et al. (2017). "Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data". In: *Phys. Rev. D.* 95, 24700029. DOI: 10.1103/PhysRevD.95.042001.
S.S participated in writing of the manuscript and helped develop some of the methods described in the paper.

[2] Abbott, B. P., et al. (2016). "GW150914: First results from the search for binary black hole coalescence with Advanced LIGO". In: *Phys. Rev. D.* 93, 122003. DOI: 10.1103/PhysRevD.93.122003.
S.S prepared the data for the GstLAL results including calculation of the background models, and made the GstLAL results plot.

[3] Abbott, B. P., et al. (2016). "Binary Black Hole Mergers in the First Advanced LIGO Observing Run". In: *Phys. Rev. X.* 6, 041015. DOI: 10.1103/PhysRevX.6.041015.
S.S prepared the data for the GstLAL results including calculation of the background models, and made the GstLAL results plot.

[4] Abbott, B. P., et al. (2016). "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence". In: *Phys. Rev. Lett.* 116, 241103. DOI: 10.1103/PhysRevLett.116.241103.
S.S co-wrote the section describing the GstLAL analysis. S.S prepared the data for the GstLAL results including calculation of the background models, and made the GstLAL results plot.

[5] Sachdev, S., et al. (2019). "The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO's Second and Advanced Virgo's First Observing Runs". In: *arXiv preprints* 1901.08580.
S.S led the writing of the manuscript and contributed in developing some of the methods presented here.

[6] Mukherjee, D. et al. (2018). "The GstLAL template bank for spinning compact binary mergers in the second observation run of Advanced LIGO and Virgo". In: *arXiv preprints* 1812.05121.
S.S co-wrote the code used to conduct the bank effectualness analysis. S.S also ran several of the analyses and made several figures in the paper.

[7] Abbott, B. P., et al. (2018). "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs". In: *arXiv preprints* 1811.12907.
S.S participated in running some of the GstLAL analyses.

[8] Abbott, B. P., et al. (2017). "GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence". In: *The Astrophysical Journal Letters* 851, 2. DOI: 10.3847/2041-8213/AA9F0C.
S.S prepared the final numbers for the GstLAL pipeline and carried out the GstLAL run.

[9] Karki, S., et al. (2016). "The Advanced LIGO photon calibrators". In: *Review of Scientific Instruments* 87, 114503. DOI: 10.1063/1.4967303.
S.S helped install the photon calibrator system at both ends in Livingston, take measurements, and produce final numbers for the first measurements taken at LIGO Livingston.

# TABLE OF CONTENTS

# LIST OF ILLUSTRATIONS

# LIST OF TABLES

*Chapter 1*

# INTRODUCTION: THE FIRST DETECTION OF GRAVITATIONAL WAVES FROM COMPACT BINARY MERGERS

## 1.1 The age of gravitational-wave science

On September 14, 2015, the observation of gravitational waves [6] by the Advanced LIGO detectors [19] marked the beginning of an era of gravitational-wave astronomy. These waves were emitted during the final fraction of a second of the merger of two black holes to produce one single black hole. This merger of two black holes had been predicted but never observed. This binary black hole merger event was followed by several other such events [9, 10, 16, 17], and a binary neutron star merger event [11] in the course of the first two observing runs of the Advanced LIGO detectors [19] and the first observing run of the Advanced Virgo detector [20].

The idea of gravitational waves was first anticipated by Henri Poincaré [21] in 1905, and they were later predicted by Albert Einstein in 1916 [22] as a direct consequence of his theory of general relativity [23]. The first evidence for the existence of gravitational waves came in 1974 when Russell Alan Hulse and Joseph Hooton Taylor, Jr. discovered a binary system composed of a pulsar in orbit around a neutron star. In 1982, Taylor and Weisberg showed that the orbit of the pulsar was slowly shrinking over time because of the release of energy in the form of gravitational waves [24, 25]. For discovering the pulsar and showing that it would make possible this unique test of general relativity in form of the gravitational-wave measurement, Hulse and Taylor were awarded the 1993 Nobel Prize in Physics. Even though their existence was demonstrated by Hulse and Taylor, Jr. in 1970s and 1980s, the direct measurement of the stretching and squeezing of spacetime caused by the gravitational waves eluded us for about a century after their prediction. In fact this observation came after more than five decades of careful planning and building of the LIGO detectors [19, 26].

The general theory of relativity is a widely accepted theory of gravity that was developed by Albert Einstein between 1907 and 1915, and subsequently published in the year 1916 [23]. According to the theory, gravity is a property of spacetime. Matter and energy curve spacetime, and the paths followed by matter and energy in

absence of any external forces are governed in turn by the curvature of spacetime. Information about a change in the gravitational field (due to a change in the matter-energy distribution) is propagated across long distances in form of gravitational waves which carry energy, momentum, and angular momentum away from their origin at the speed of light.

In general relativity, gravitational waves are transverse waves and they stretch and squeeze spacetime perpendicular to the direction of propagation. The effects of gravitational waves can be described by tidal forces they produce on free test masses. Imagine we have gravitational waves passing through two free test masses. The stretching and squeezing of spacetime due to the gravitational waves will result in the masses moving back and forth relative to each other. If we have localized gravitational waves that are moving across only one of the test masses, they would cause oscillations in the ticking rate of the clock on the test mass that they are passing through. So gravitational waves are ripples in the curvature of spacetime that stretch and squeeze space and also warp time.

Gravitational radiation, however, is extremely weak far from the source. The sources that can produce gravitational waves loud enough to be observed in a controlled experiment on Earth are extremely violent astrophysical events such as the merger of two black holes or a supernova explosion in the nearby Universe. Even for such sources, we need extremely precise experiments to observe them. Advanced LIGO (Laser Interferometer Gravitational-wave Observatory) detectors are arguably the most precise measuring devices ever built. They are highly complex L-shaped Michelson interferometers with 4 km long arms. They have mirrors at the ends of the arms and a beam splitter at the center in order to combine light beams traveling up and down the two arms and create an interference pattern. Any change in arm lengths would change the interference pattern created by the light beams when they combine. The LIGO interferometers are set such that in a null signal case, the light beams from the two arms destructively interfere and cancel each other out. The Advanced LIGO detectors can measure a change in length one-ten-thousandth the diameter of a proton. The LIGO detectors were built by Caltech and MIT with funding from the NSF. LIGO construction was approved in 1990, their construction began in 1994, and the installation and commissioning of LIGO's initial interferometers began in 1999. The LIGO Scientific Collaboration (LSC) which is responsible for organizing and coordinating LIGO's technical and scientific research and data analysis, and includes scientists from all over the world was formed in 1997. The LIGO laboratory

at Caltech, MIT, Hanford and Livingston, remain reponsible for LIGO operations and advanced interferometer R&D. Using the initial LIGO interferometers, gravitational-wave searches were carried out from 2002 to 2010 but found null results. The construction of advanced interferometers was funded by the NSF in 2008, and after 5 years of installation and commissioning from 2010 to 2014, the Advanced LIGO detectors came live in 2015. The Advanced detectors operate in a frequency band of 20 Hz to 2000 Hz, and are sensitive to the binary inspirals from compact stars, neutron stars and black holes. For their decisive contributions to the LIGO detector and the observation of gravitational waves, Rainer Weiss, Barry Barish, and Kip Thorne were awarded the Nobel prize in Physics in 2017.

The coalescence of compact objects plays a unique role in fundamental physics and astrophysics. The gravitational waves carry physical information about their sources and allow us to test Einstein's general relativity in the strong regime [27] and study matter at densities [28] beyond reach of the most powerful laboratories on our planet. Moreover, we can gain insight about the evolution of stars, galaxies, and even the Universe as a whole by studying the merger rate of compact objects [11, 18, 29, 30]. Joint electromagnetic and gravitational-wave observations help develop our understanding of the physical processes that occur in such systems [12, 31], and also provide a gravitational-wave standard siren method to measure the Hubble constant [32, 33].

This thesis tells the story of the first gravitational-waves detections with Advanced LIGO. We discuss the methodology used in analyzing the Advanced LIGO data that led to these detections enabling the start of the field of gravitational wave astronomy. In Chapter 2 we discuss basic theory behind the gravitational-wave phenomena being studied and measured with Advanced LIGO, in Chapter 3 we discuss the compact binary gravitational-wave sources and basic methods for their detection, in Chapter 4 we describe the methods of a low-latency search that can detect these sources, in Chapters 5, 6, and 7 we discuss the developments in these methods and results from the search from early tuning stages to the two observing runs of Advanced LIGO, in Chapter 8 we discuss detecting astrophysical populations of such systems in the case of third-generation detectors with sensitivities more than 10 times better than the final design sensitivity of the Advanced LIGO detectors, and finally, in Chapter 9 we discuss our conclusions and directions for future work.

*Chapter 2*

# GRAVITATIONAL WAVES IN GENERAL RELATIVITY

## 2.1 Introduction

Conceptually, if gravity is sourced by mass and energy, which can move rapidly (close to the speed of light) and thereby change the gravity sensed by a distant observer, the "news" of changing gravity cannot be communicated to the observer instantaneously (as Newtonian gravity implies), but must propagate from the source to the observer at the speed of light or less. Also conceptually, if gravity is described in terms of a dynamically-changing spacetime curvature, responding to rapidly-changing sources, it is natural to expect the "news" of changing gravity to propagate as waves of spacetime curvature. These conceptual foundations are built into general relativity, and gravitational waves are thus a firm prediction of general relativity.

For many years after the publication of Einstein's 1916 paper [23], the reality of gravitational waves were in doubt [34]. The matter was finally settled through the work of Pirani, Feynman, Bondi, and others [34, 35]. It has been shown theoretically that GWs carry energy, momentum, and angular momentum from source to observer, although it took decades for this to be firmly established. Hulse-Taylor observed the loss of energy from GW emission from a compact binary system by measuring the loss of orbital period.

Here we outline the elements of GR that lead to GWs and their fundamental properties. Our discussion below draws from [36, 37].

## 2.2 Einstein's field equations

As discussed in the previous chapter, according to general relativity, spacetime curves in response to the presence of energy and momentum, and the curvature of spacetime determines the motion of free particles. This relationship between the spacetime curvature and energy-momentum is described by Einstein's field equations

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{2.1}$$

Or in units where $G = c = 1$, Eq. 2.1 can be written as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{2.2}$$

All the quantities in these equations are a function of the spacetime basis vectors, $x^\mu$. The indices in the equations run from 0 to 3, for the time and 3 spatial directions. $G_{\mu\nu}$, $R_{\mu\nu}$,$g_{\mu\nu}$, and $g_{\mu\nu}$ are tensors and $R$ is a scalar. These equations relate the geometry of the spacetime (described by $R_{\mu\nu}$, $R$, $g_{\mu\nu}$, left-hand side of the equations) to matter and energy ($T_{\mu\nu}$, right-hand side of the equations). $R_{\mu\nu}$ is the Ricci curvature tensor, $R$ its trace, $g_{\mu\nu}$ is the field metric, and $T_{\mu\nu}$ is the stress-energy tensor. $G_{\mu\nu}$ is known as *Einstein tensor*. These are highly nonlinear equations which can evolve the values of the metric, $g_{\mu\nu}$ from some given initial values. The quantities in the equations are symmetric $4 \times 4$ tensors, and therefore have ten independent components. We will describe these quantities in the next section.

## 2.3 Geometry of spacetime

### Metric tensor

The metric tensor field $\mathbf{g}$ ($g_{\mu\nu}$) is the fundamental geometric construct in general relativity. The metric tensor contains the information about the relative rate at which the clocks run and the distance between points. For two infinitesimally separated points in the spacetime, the proper distance $ds$ between them is given in terms of the metric tensor and the local coordinates as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{2.3}$$

We adopt the Einstein summation convention that repeated indices indicate summation over those indices.

For a flat spacetime, the metric $g_{\mu\nu}$ is given by the Minkowski spacetime metric tensor

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.4}$$

In general, spacetime metric is not flat. But the spacetime metric tensor $\mathbf{g}$ is always given by a covariant, nondegenerate and symmetric tensor. It can be shown that for any point $\mathcal{P}$ on the spacetime manifold (except gravitational singularities like at the centers of black holes), we can always find a coordinate transformation $x^\alpha \to x^{\alpha'}(x)$ such that the metric is flat at the point $\mathcal{P}$ in terms of the new basis coordinates, and is "nearly" flat in the neighborhood of the point $\mathcal{P}$.

$$g_{\alpha'\beta'}(x^\mu(\mathcal{P})) = \boldsymbol{\eta}, \tag{2.5}$$

$$g_{\alpha'\beta'}(x^\mu(\mathcal{P}) + dx^\mu) = \boldsymbol{\eta} + O[(dx^\mu)^2]. \tag{2.6}$$

We will make use of this in the coming sections. The metric tensor is also used to lower and raise indices of tensor fields.

**Covariant derivative**

The *covariant derivative*, $\nabla_\mu$, is given by

$$\nabla_\mu V^\rho \equiv \partial_\mu V^\rho + \Gamma^\rho_{\mu\nu} V^\nu, \tag{2.7}$$

where $\partial_\mu$ is the ordinary spacetime derivative operator, $\partial/\partial x^\mu$, and $\Gamma^\rho_{\mu\nu}$ are the *connection coefficients*, also known as the *Christoffel symbols*. The connection coefficients provide a measure of derivatives of the spacetime basis vectors themselves. The basis unit 4-vectors are given by, $\hat{e}_\mu$. The interpretation of $\Gamma^\rho_{\mu\nu}$ is that it is the $\rho^{th}$ component of $\partial\hat{e}_\mu/\partial x^\nu$ and is given by

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu}). \tag{2.8}$$

### 2.3.1 Riemann tensor, Ricci tensor, Ricci scalar

The *Riemann tensor* (also known as the *curvature tensor*) is given by

$$R^\alpha_{\beta\mu\nu} \equiv \partial_\mu \Gamma^\alpha_{\beta\nu} - \partial_\nu \Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\beta\nu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\beta\mu}. \tag{2.9}$$

It measures the extent to which the metric tensor is not locally isometric to that of Euclidean space, and helps in defining the geometry of spacetime. The spacetime is flat if the Riemann tensor vanishes everywhere. $R^\alpha_{\beta\mu\nu} = 0$.

The *Ricci tensor* is defined by contracting over two indices of the Riemann tensor,

$$R_{\alpha\beta} \equiv R^\mu_{\alpha\mu\beta} = R_{\beta\alpha}, \tag{2.10}$$

and the *Ricci scalar* is defined as

$$R \equiv g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} g^{\alpha\beta} R_{\alpha\mu\beta\nu}. \tag{2.11}$$

The curvature tensor satisfies the following properties:

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}, \tag{2.12}$$

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0, \tag{2.13}$$

and the Bianchi identities

$$\partial_\lambda R_{\alpha\beta\mu\nu} + \partial_\nu R_{\alpha\beta\lambda\mu} + \partial_\mu R_{\alpha\beta\nu\lambda} = 0, \tag{2.14}$$

$$\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0. \tag{2.15}$$

From the second Bianchi identity Eq. 2.14, and the definition of Einstein's tensor, it follows that

$$\nabla_\mu G^{\mu\nu} = 0. \tag{2.16}$$

Consequently, from the Einstein's field equations Eq. 2.1, the stress-energy tensor satisfies

$$\nabla_\mu T^{\mu\nu} = 0. \tag{2.17}$$

This means that energy and momentum are covariantly conserved in general relativity.

The four Bianchi identities reduce the number of independent Einstein's field equations Eq.2.1 from 10 to 6, leaving the metric with four gauge fixing degrees of freedom, which correspond to the freedom to choose a coordinate system.

### 2.4 Linearlized gravity

In general relativity, the Einstein's field equations are covariant under coordinate transformations

$$x^\mu \to x^{\mu'}(x), \tag{2.18}$$

where $x^{\mu'}$ is invertible, differentiable, and has a differentiable inverse. This means that under such transformations, the metric transforms as

$$g_{\mu\nu} \to g_{\mu'\nu'} = \Lambda^\alpha_{\mu'}\Lambda^\beta_{\nu'} g_{\alpha\beta}, \tag{2.19}$$

where $\Lambda^\alpha_{\mu'} = \partial_{\mu'} x^\alpha$. In non-mathematical terms, it means that the physical outcomes of the theory remain unchanged under a change of coordinates.

In the weak-field regime where the spacetime is nearly flat, there exists a reference frame where we can take the spacetime metric to be that of flat Minkowski spacetime, $\eta_{\mu\nu}$, plus a small perturbation $h_{\mu\nu}$

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{2.20}$$

for $\left|h_{\mu\nu}\right| \ll 1$.

(2.21)

However by choosing such a reference frame where Eq.2.20 holds, we break the invariance of general relativity under coordinate transformations. But there is a subset of transformations under which the form of metric given by Eq.2.20 is preserved.

### 2.4.1 Slow-varying gauge transformations

Consider transformations of the form

$$x^{\mu} \rightarrow x^{\mu'}(x) = x^{\mu} + \epsilon^{\mu}(x^{\nu}), \tag{2.22}$$

for $\left|\partial_{\mu}\epsilon_{\nu}\right| \leq \left|h_{\mu\nu}\right|$.

Then the metric transforms as

$$g_{\mu\nu} \rightarrow g_{\mu'\nu'} = \partial_{\mu'}x^{\rho}\partial_{\nu'}x^{\sigma}g_{\rho\sigma}, \tag{2.23}$$

$$\partial_{\mu}x^{\rho'} = \delta^{\rho}_{\mu} + \partial_{\mu}\epsilon^{\rho}, \tag{2.24}$$

$$\partial_{\mu'}x^{\rho} = \delta^{\rho}_{\mu} - \partial_{\mu}\epsilon^{\rho} + O(\partial_{\mu}\epsilon^{\rho})^2, \tag{2.25}$$

$$g_{\mu'\nu'} = \eta_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu} + h_{\mu\nu} + O(\partial\epsilon^2). \tag{2.26}$$

Note that we use $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ for lowering and raising the indices in the linearized theory. We can see from Eq.2.26 that

$$g_{\mu'\nu'} = \eta_{\mu\nu} + h_{\mu'\nu'} \tag{2.27}$$

for

$$h_{\mu'\nu'} = h_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu}. \tag{2.28}$$

Therefore the slowly varying coordinate transformations Eq.2.22 are a symmetry of the linearized Einstein's field equations. In other words, under slowly varying coordinate transformations, the form of linearized Einstein's field equations is preserved, as described below.

### 2.4.2 Lorentz transformations

Another set of transformations that preserve Eq.2.20 are the global (x-independent) Lorentz transformations:

$$x^{\mu} \rightarrow x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu} \tag{2.29}$$

$$\Lambda^{\mu}_{\rho'} \Lambda^{\nu}_{\sigma'} \eta_{\mu\nu} = \eta_{\rho'\sigma'}. \tag{2.30}$$

The metric then transforms as

$$g_{\mu\nu} \rightarrow g_{\mu'\nu'} = \Lambda^{\mu}_{\mu'} \Lambda^{\nu}_{\nu'} g_{\mu\nu} \tag{2.31}$$

$$= \eta_{\mu'\nu'} + \Lambda^{\mu}_{\mu'} \Lambda^{\nu}_{\nu'} h_{\mu\nu}. \tag{2.32}$$

Therefore

$$h_{\mu\nu} \rightarrow h_{\mu'\nu'} = \Lambda^{\mu}_{\mu'} \Lambda^{\nu}_{\nu'} h_{\mu\nu}. \tag{2.33}$$

Under Lorentz transformations, $h_{\mu\nu}$ transforms like a tensor. Linearized theory is invariant under Lorentz transformations.

### 2.4.3 Wave equation

Let us now derive the Einstein's field equations in the linearized theory. To leading order in the metric perturbation, the Christoffel symbol, Riemann tensor, Ricci tensor, and Ricci scalar are given by

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}(\partial_{\beta} h^{\mu}_{\alpha} + \partial_{\alpha} h^{\mu}_{\beta} - \partial^{\mu} h_{\alpha\beta}) \tag{2.34}$$

$$R_{\mu\nu\alpha\beta} = \frac{1}{2}(\partial_{\nu}\partial_{\alpha} h_{\mu\beta} + \partial_{\mu}\partial_{\beta} h_{\nu\alpha} - \partial_{\mu}\partial_{\alpha} h_{\nu\beta} - \partial_{\nu}\partial_{\beta} h_{\mu\alpha}) \tag{2.35}$$

$$R_{\mu\nu} = \frac{1}{2}(\partial_{\nu}\partial_{\alpha} h^{\alpha}_{\mu} + \partial_{\mu}\partial_{\alpha} h^{\alpha}_{\nu} - \partial_{\alpha}\partial^{\alpha} h_{\mu\nu} - \partial_{\mu}\partial_{\nu} h^{\alpha}_{\alpha}) \tag{2.36}$$

$$R = \frac{1}{2}(\partial^{\alpha}\partial^{\beta} h_{\alpha\beta} - \partial_{\beta}\partial^{\beta} h^{\alpha}_{\alpha}). \tag{2.37}$$

We now introduce the trace $h$ given by

$$h = \eta^{\mu\nu} h_{\mu\nu}, \tag{2.38}$$

$$\tag{2.39}$$

and the trace-reversed metric perturbation $\bar{h}_{\mu\nu}$

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h. \tag{2.40}$$

Using these, we can calculate the Einstein's tensor to be

$$G_{\mu\nu} = -\frac{1}{2}(\partial_\alpha\partial^\alpha\bar{h}_{\mu\nu} + \eta_{\mu\nu}\partial^\alpha\partial^\beta\bar{h}_{\alpha\beta} - \partial^\alpha\partial_\nu\bar{h}_{\mu\alpha} - \partial^\alpha\partial_\mu\bar{h}_{\nu\alpha} + O(h_{\mu\nu}^2)). \tag{2.41}$$

**Lorenz gauge**

To simplify the Eq.2.41, we make use of gauge freedom to choose *Lorenz gauge*, also known as the *harmonic gauge*

$$\partial^\nu\bar{h}_{\mu\nu} = 0. \tag{2.42}$$

In the Lorenz gauge, the Einstein's field equation for a weakly perturbed metric Eq.2.41 reduces to

$$-\partial_\alpha\partial^\alpha\bar{h}_{\mu\nu} = -\Box\bar{h}_{\mu\nu} = 16\pi T_{\mu\nu}, \tag{2.43}$$

where $\Box$ is the symbol used for the four-dimensional Laplacian $\partial_\alpha\partial^\alpha$, also known as the *D'Alembertiam* or the wave operator. Eq.2.43 are known as the Einstein's field equations of the weak field, linearized theory.

One can show that for infinitesimal coordinate transformations of the type discussed in Sec.2.4.1, $x \rightarrow x' = x + \epsilon$, where $\epsilon$ is arbitrary and $O(h)$, the trace-reversed metric perturbation transforms as

$$\bar{h}_{\mu\nu} \rightarrow \bar{h}_{\mu'\nu'} = \bar{h}_{\mu\nu} - \partial_\nu\epsilon_\mu - \partial_\mu\epsilon_\nu + \eta_{\mu\nu}\partial_\rho\epsilon^\rho. \tag{2.44}$$

Assuming that the Lorenz gauge holds in the new frame, we get

$$\partial^{\nu'}\bar{h}_{\mu'\nu'} = \partial^\nu\bar{h}_{\mu\nu} - \partial^\nu\partial_\nu\epsilon_\mu = 0 \tag{2.45}$$

$$\Box\epsilon_\mu = \partial^\nu\partial_\nu\epsilon_\mu = \partial^\nu\bar{h}_{\mu\nu}. \tag{2.46}$$

Lorenz gauge is not unique since a homogenous solution of the form

$$\Box\epsilon_\mu = 0 \tag{2.47}$$

can always be added to move to a different Lorenz gauge.

In order to gain insight into the propagation of gravitational waves, we consider the field equations of the linearized theory in vacuum ($T_{\mu\nu} = 0$),

$$\Box \bar{h}_{\mu\nu} = 0. \tag{2.48}$$

Eq.2.48 has complex wave solutions of the form

$$\bar{h}_{\mu\nu} = A_{\mu\nu} e^{ik_\sigma x^\sigma}. \tag{2.49}$$

This gives us

$$\Box \bar{h}_{\mu\nu} = \partial_\alpha \partial^\alpha \bar{h}_{\mu\nu} = -k_\sigma k^\sigma A_{\mu\nu} e^{ik_\sigma x^\sigma}. \tag{2.50}$$

$$\tag{2.51}$$

From Eq.2.51 and Eq.2.48, we get

$$k_\sigma k^\sigma = 0, \tag{2.52}$$

so that the wave vector $k_\mu$ is a null vector, implying that the gravitational waves travel at the speed of light in general relativity. Imposing the harmonic gauge given by Eq.2.42 on the solutions to the vacuum field equations Eq.2.49, we get

$$A_{\mu\nu} k^\nu = 0 \tag{2.53}$$

This shows that the wave vector $k_\mu$ is orthogonal to the wave polarization, $A_{\mu\nu}$. In other words, the perturbation of the metric is transverse to the direction of propagation. Therefore solutions to the vacuum linearized field equations are plane-wave equations with the null wave vectors.

**Transverse traceless gauge**

The metric perturbation $h_{\mu\nu}$ has ten independent components. By introducing the Lorenz gauge, we have imposed four constraints. Lorenz gauge makes the metric perturbation look like a transverse wave. Further specifying gauge using the extra freedom from Eq.2.47 we introduce four extra constraints, and the metric perturbation now has only two independent degrees of freedom remaining. The two degrees of freedom correspond to the two polarizations of the perturbation of the metric caused by the gravitational wave.

We choose this extra freedom within the Lorenz gauge such that the perturbation is traceless and also purely spatial. This is known as the *trasverse traceless* (TT) gauge. In this gauge, the metric perturbation affects only the components of the metric which are orthogonal to the direction of propagation (transverse), and is symmetric about reflection through the direction of propagation (traceless). These extra constraints are given by

$$A_{\mu\nu}u^{\nu} = 0, \tag{2.54}$$

and

$$A^{\mu}_{\mu} = 0, \tag{2.55}$$

where $u^{\nu}$ is a fixed unit four-velocity vector.

Eq.2.55 implies that in the TT gauge, $\bar{h}^{\mu}_{\mu} = 0$, which means that the trace-reversed metric perturbation is equal to the metric perturbation in this gauge. One can then use a Lorentz transformation to make the explicit choice on $u^{\mu}$ to fix the unit four-velocity vector to be $(1, 0, 0, 0)$ to get $\bar{h}_{\mu 0} = h_{\mu 0} = 0$, which means that the metric perturbation only has spatial components. If we choose our coordinates such the wave travels in the +z direction, we can write the metric perturbation as

$$h^{TT}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_{+} & h_{\times} & 0 \\ 0 & h_{\times} & -h_{+} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.56}$$

where $h_{+}$ and $h_{\times}$ are functions of time satisfying the wave equation,

$$\Big(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\Big)h = 0, \tag{2.57}$$

where $h = h_{+}$ or $h = h_{\times}$. The amplitudes $h_{+}$ and $h_{\times}$ correspond to the two polarizations of gravitational waves called *plus* and *cross* polarizaions.

**Polarizations in general relativity**

If the coordinate unit basis vectors in the plane perpendicular to the wave propagation are $\hat{e}_x$ and $\hat{e}_y$, then we can define the basis tensors for the two polarizations as

$$\mathbf{e}_+ = \hat{e}_x \otimes \hat{e}_x - \hat{e}_y \otimes \hat{e}_y \tag{2.58}$$

$$\mathbf{e}_\times = \hat{e}_x \otimes \hat{e}_y + \hat{e}_y \otimes \hat{e}_x. \tag{2.59}$$

In terms of these basis tensors, the metric perturbation tensor can be written as

$$\mathbf{h} = h_+ \mathbf{e}_+ + h_\times \mathbf{e}_\times. \tag{2.60}$$

If the coordinates in the transverse plane are rotated by an angle $\theta$, then we can see that

$$\mathbf{e}'_+ = \cos 2\theta \mathbf{e}_+ + \sin 2\theta \mathbf{e}_\times, \tag{2.61}$$

$$\mathbf{e}'_\times = -\sin 2\theta \mathbf{e}_+ + \cos 2\theta \mathbf{e}_\times, \tag{2.62}$$

and

$$h'_+ = \cos 2\theta h_+ + \sin 2\theta h_\times, \tag{2.63}$$

$$h'_\times = -\sin 2\theta h_+ + \cos 2\theta h_\times. \tag{2.64}$$

Therefore the gravitational wave polarizations are rotated by an angle $2\theta$ when the coordinates are rotated by an angle $\theta$. This shows that gravitational wave frequency is twice the frequency of motion of the source.

For more details, we refer the reader to [38–40].

### 2.4.4 Effect of gravitational waves on a ring of test masses

Consider a ring of test masses lying on the xy plane. The effect of a *plus* and a *cross* polarized wave passing transverse to a plane in the +z direction is to cause a tidal deformation of the circular ring into an elliptical ring with the same area. This is illustrated in Fig. 2.1.

The maximum fractional change in length of the ring is given approximately by the

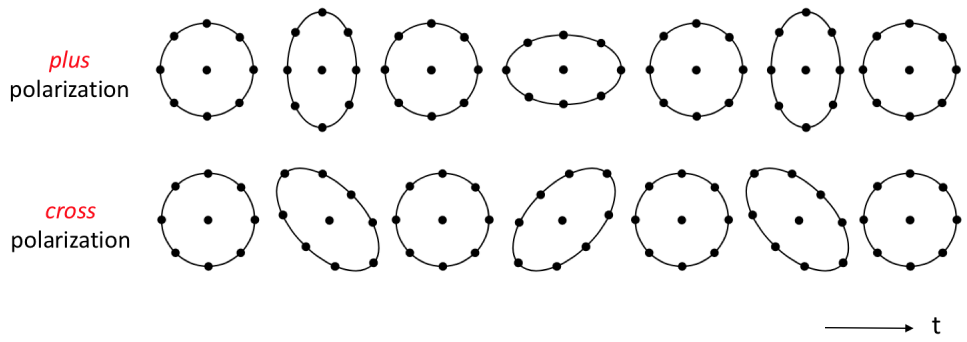



Figure 2.1: Deformation of a ring of test masses caused by the *plus* (+) or *cross* (×) polarized gravitational wave passing perpendicular to the plane of test masses. Fig. taken and modified from [1]

gravitational wave amplitude which is a combination of the *plus* and *cross* amplitudes - $h_+$ and $h_\times$.

$$\frac{\Delta L}{L} \approx |\mathbf{h}| = |h_+ e_+ + h_\times e_\times|. \tag{2.65}$$

## 2.5 Generation of Gravitational Waves

So far we have studied the gravitational waves as perturbations of the flat Minkowski metric. Now we briefly discuss how they are generated. The general, non-vacuum solution of linearized Einstein's field equations (Eq. 2.43) at a point $x^\alpha$ are given by,

$$\bar{h}_{\mu\nu}(x^\alpha) = -16\pi \int G(x^\alpha - y^\alpha) T_{\mu\nu}(y^\alpha) d^4 y, \tag{2.66}$$

where we are integration is performed over all spacetime $y^\alpha$. The Green's function $G(x - y)$ for the d'Alembertian operator $\Box$ is the solution to the wave equation in the presence of a delta function

$$\Box G(x^\alpha - y^\alpha) = \delta(x^\alpha - y^\alpha). \tag{2.67}$$

Analogous to electromagnetism, the Green's function is the retarded Green's function, given by

$$G(x^\alpha - y^\alpha) = -\frac{1}{4\pi \left| x^i - y^i \right|} \delta(t_{ret} - y^0) \theta(x^0 - y^0), \tag{2.68}$$

where the retarded time is given by $t_{ret} = x^0 - \left|x^i - y^i\right|$, $x^i$ and $y^i$ refer to the spatial vectors $(x^1, x^2, x^3)$ and $(y^1, y^2, y^3)$ respectively, and $\theta(x^0 - y^0)$ equals 1 when $x^0 > y^0$ and is 0 otherwise.

Plugging in the Green's function in the non-vacuum solution to the linearized field equations (Eq. 2.66) we get

$$\bar{h}_{\mu\nu}(t = x^0, x^i) = 4 \int d^3y \frac{T_{\mu\nu}(t_{ret}, y^i)}{\left|x^i - y^i\right|}, \tag{2.69}$$

where the integration is performed over the past light cone of the event $(t, x^i)$. The perturbation at the point $(t, x^i)$ is the sum of effects from the energy and momentum of all the sources on the past light cone of the event.

At large distances from the source, $r = \left|x^i\right|$, the denominator in Eq. 2.69 can be expanded as

$$\frac{1}{\left|x^i - y^i\right|} \approx \frac{1}{r} + \ldots \tag{2.70}$$

The solution therefore to first order in $\frac{1}{r}$ is given by

$$\bar{h}_{ij}(t = x^0, x^i) \approx \frac{4}{r} \int d^3y T_{ij}(t - r, y^i). \tag{2.71}$$

The stress-energy tensor must be conserved, which in linearized theory means $\partial^\mu T_{\mu\nu} = 0$. This yields

$$\frac{4}{r} \int d^3y T_{ij} = \frac{2}{r} \frac{\partial^2}{\partial t^2} \int d^3y T_{00} y_i y_j. \tag{2.72}$$

In a standard stress-energy tensor, the tt (or the 00) component denotes the rest-mass energy density of the source $\rho(x)$. We define the second moment of the mass distribution as

$$I_{ij}(t) = \int d^3y \rho(t, y^i) y_i y_j, \tag{2.73}$$

and combining Eqs. 2.71 - 2.73, we get

$$\bar{h}_{ij}(t, x^i) \approx \frac{2}{r} \frac{\partial^2}{\partial t^2} I_{ij}(t - r). \tag{2.74}$$

We now want to project into the TT-gauge using the transverse projection operator

$$P_{ij} = \delta_{ij} - \hat{n}_i \hat{n}_j, \tag{2.75}$$

where $\hat{n}_i = x^i/r$ is the unit vector in the propagation direction (normal to the transverse plane). The TT-gauge solution then becomes (recall that in the TT-gauge, $\bar{h}_{\mu\nu} = h_{\mu\nu}$),

$$h_{ij}^{TT} \approx \frac{2}{r} \frac{\partial^2}{\partial t^2} I_{ij}^{TT}(t - r), \tag{2.76}$$

with

$$I_{ij}^{TT} = P_{ik} I^{kl} P_{lj} - \frac{1}{2} P_{ij} P_{kl} I^{kl}. \tag{2.77}$$

Therefore gravitational waves are generated by accelerated sources, the radiation obeys a $1/r$ fall-off, and is of *quadrupolar* in nature, as the conservation laws do not permit monopole and dipole gravitational radiation [41].

## 2.6 Gravitational-wave sources

In this section we summarize the most important gravitational-wave sources, for which we have dedicated data analysis groups in the LIGO-Virgo Collaboration. For more details, we refer the reader to [36, 37].

### 2.6.1 Compact binary coalescences

So far all the gravitational-wave detections in Advanced LIGO-Virgo have originated from compact binary coalescences (CBCs). Even before the detections, these were believed to be the most promising sources that could be observed with the ground based interferometric gravitational-wave detectors. These binary systems comprise two bound compact objects such as black holes, or neutron stars. As the masses revolve around one another, they lose energy in form of gravitational waves, which leads to the decay of their orbits until they plunge together and merge to form one single object. These sources are the focus of this thesis and are discussed in more detail in the next chapter.

### 2.6.2 Continuous waves

Continuous sources are those sources that continuously emit gravitational waves with approximately constant frequency compared to the observation time. The prime candidates to emit such signals are rapidly rotating neutron stars, known as pulsars, which are non-axisymmetric. The asymmetry is needed to induce a non-zero time-varying quadrupole moment required for the generation of gravitational waves. Although the GW amplitude is generally weaker compared to CBC sources, we can integrate them over longer durations, which means that CW sources may also achieve detectable SNRs. Analysis of known pulsars can be simplified if we know their sky location and rotational frequency by application of targeted searches [42].

### 2.6.3 Burst Sources

Burst sources are associated with astronomical transient phenomena, such as supernovae, instabilities in neutron stars, etc. A supernova occurs during a gravitational collapse of a star, or core of a star, to form a neutron star or black hole. Such sources are typically hard to model because of the complex physics associated with them. Therefore, to search for searches, we employ the use of unmodelled filters, such as sine-Gaussians. Note that high mass CBC sources that are loud and short in duration are also detected by the burst searches. In fact, GW150914 was first detected by such unmodeled burst searches in low-latency [43].

### 2.6.4 Stochastic background

Stochastic gravitational-wave background consists of gravitational-waves that are not localized in either time or frequency. There are two classes of expected sources for stochastic background: primordial gravitational wave background, which consists of radiation from early universe, such as Big Bang; and astrophysical background, which comes from the superposition of the population of astrophysical sources such as CBCs, magnatars, etc. See Sec. 8.1 for a complete list of expected sources for stochastic background.

Since the stochastic background is indistinguishable from instrumental noise, for short observation times, it cannot be detected by a single detector. Instead, we cross-correlate the output of several detectors over long durations. An excess in this cross-correlation can be an indication of stochastic background.

*Chapter 3*

# COMPACT BINARY COALESCENCES

CBCs, by virtue of their extremely strong accelerating quadrupole moment, are the most efficient and the best modeled source of gravitational waves. They are detectable (producing signal in the LIGO detection band) over a broad range of parameters (total masses between $0.4M_{\odot}$ and $500M_{\odot}$), and out to cosmological distances. They provide a laboratory for testing strong-field gravity [44], and for first measurements of fundamental properties of gravitational waves including speed [45, 46], graviton mass [44], polarizations [47]. They are fascinating astrophysical objects, presenting the strongest and most rapidly changing gravitational fields in the universe. Neutron stars are spectacular laboratories for fundamental physics of the strong, weak, electromagnetic and gravitational interactions. Binary neutron stars explore the origin of the heaviest elements [31]. As the endpoints of the evolution of the most massive stars, black holes are tracers of stellar evolution. Therefore binary black holes can provide insight into formation mechanisms [48].

The CBC detection pipelines used by the LVC (there are at least 4 - `GstLAL` [3, 49, 50], `PyCBC` [51–53], `MBTA`, `GstLAL-spiir`, with the `GstLAL` and the `PyCBC` pipelines being run as "flagship" searches) have been under development for many years and are now quite mature. They are even relatively robust against non-Gaussian noise fluctuations. However, there's always the need for continual improvement and feature development. We will describe the `GstLAL` search pipeline in detail in the next few chapters. This chapter describes the basic theory and techniques behind the analysis of Advanced LIGO and Virgo data to look for gravitational-wave signals from compact binary coalescences (CBCs). We begin with a brief description of the CBC signal models that have been developed over the years and discuss their domain of validity. We then describe how these waveforms are used in matched-filtering CBC searches. The matched-filtering algorithm is used to estimate the signal-to-noise ratio (SNR), which is the optimal statistic for the detection of a signal of known form added to stationary Gaussian noise [54–62].

## 3.1 Waveform Models

CBCs are the most extensively modelled sources of gravitational waves. The fact that these sources can be relatively well modeled is exploited by the matched-filtering

searches in Advanced LIGO and Virgo.

The gravitational-wave signal from a CBC depends on a number of parameters. These parameters are classified into two categories: intrinsic parameters, such as the masses and spins of the binary (those that only depend on the source itself), and extrinsic parameters, such as the distance, inclination angle, etc. (those that are related to the position of the source with respect to the observer). There are *at least* fifteen parameters needed to specify the expected gravitational-wave signal from a compact binary coalescence. These are listed in the Table 3.1.

| | |
|---|---|
| $m_1$ & $m_2$ | component masses |
| $\vec{S}_1$ & $\vec{S}_2$ | binary spin vectors, having three components each |
| $\alpha$ & $\delta$ | right ascension and declination, determining the sky position |
| $i$ | inclination of orbital plane of the binary w.r.t. the line of sight |
| $\psi$ | orientation of polarization vector (in radiation frame) w.r.t. detector arms |
| $D$ | luminosity distance |
| $\phi_c$ | coalescence phase |
| $t_c$ | time of coalescense |

Table 3.1: The compact binary parameter space. *At least* fifteen parameters are needed to uniquely specify a compact binary system. The parameters in blue are the ones that are intrinsic to the source, and the parameters in red are the ones that depend on the position of the observer.

In the case where the spins of the binary components are aligned with the orbital angular momentum, all the extrinsic parameters only affect the overall phase and amplitude of the system. However, when the component spins are misaligned with the orbital angular momentum, the inclination and polarization angles become time dependent and lead to non-trivial amplitude and phase modulation of the observed signal.

Note that we restrict our attention to compact binaries on circular orbits, and do not consider eccentricity of the orbit as a parameter. Studies show that the detection efficiency is only negatively impacted for highly eccentric orbits, such systems are expected to be rare. We also ignore any effects related to the matter and internal structure of neutron stars. Matter effects become significant only in the last fraction of second before merger, having almost negligible effect on overall detection SNR. Lastly, we do not include additional parameters that arise when considering theories beyond general relativity. In this section, we briefly describe the techniques that are

used to model the waveforms, and also list the various waveform models that have been developed.

### 3.1.1 Inspiral, merger, and ringdown

Due to the non-linear nature of the theory, the two-body problem in full general relativity becomes extremely non-trivial. When two masses are in an orbit around each other, they undergo a quasi-circular inspiral, losing energy due to emission of gravitational waves, which leads to shrinking of the separation between them, and an increase in the orbital velocity and frequency. The gravitational-wave frequency is given by twice the orbital frequency (as seen in Sec. 2.4.3), so we see an increase in the gravitational-wave frequency with time. Such a signal, in which the frequency monotonically increases (or decreases) with time is called a *chirp*. The gravitational-wave signal from compact binary coalescences is conventionally split into three stages: inspiral, merger, and ringdown (see Fig. 3.1).

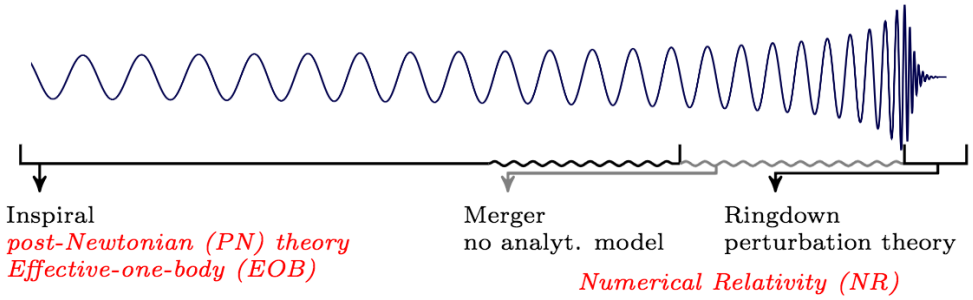



Figure 3.1: The three distinct stages of the binary evolution, and the different approximation schemes used to model these stages. Wavy lines illustrate the regime close to merger where analytical methods have to be bridged by numerical simulations. Figure taken from [2].

The inspiral stage of the waveform is when the two gravitationally bound objects are far apart such that their separation is larger than the innermost stable circular orbit (ISCO). The frequency evolution of the waveform depends on the masses and spins of the two objects in the inspiral. The higher the total mass of the system, the larger the ISCO. The smaller the amount of spin aligned with orbital angular momentum, the larger the ISCO. For a test particle orbiting a Schwarzschild black hole with mass $M$, the ISCO occurs at a radius of $r_{\rm ISCO} = 6GM/c^2$. According to Kepler's law, the speed of the particle oribiting at a radius $R$ is given by $v = \sqrt{\frac{GM}{R}}$, so, the velocity at this ISCO is given by $v_{\rm ISCO} \sim c/\sqrt{6}$. Although Kepler's law is Newtonian, it is approximately valid even in the post-Newtonian regime. For a

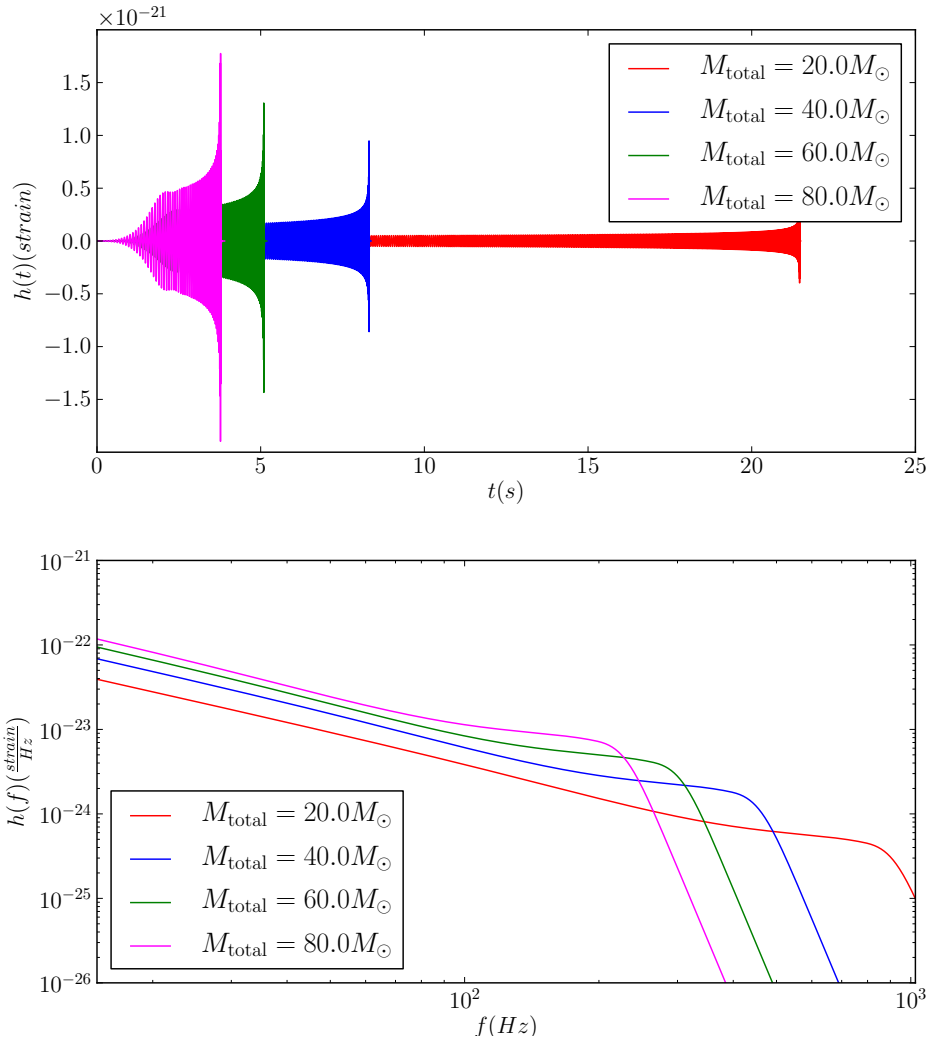



Figure 3.2: Example time-domain (top) and frequency-domain (bottom) waveforms for varying total mass of the system. The mass ratio is kept constant and set to 1. We see that more massive systems are shorter in LIGO band (which was considered as 15 Hz to 1024 Hz for making this figure. We used the model `IMRPhenomD` to generate this waveform.

maximally spinning black hole with spin aligned with the orbital angular momentum, the ISCO is at $r_{\rm ISCO} = GM/c^2$ and $v_{\rm ISCO} \sim c$, and for a maximally spinning black hole with its spin anti-aligned with the orbital angular momentum, the ISCO occurs at $r_{ISCO} = 9GM/c^2$ and $v_{\rm ISCO} \sim c/3$. See Figs. 3.2, 3.3, and 3.4.

This means, for low mass systems, such as neutron star-neutron star binaries, the masses can get extremely close to one another before undergoing merger, and the merger frequencies for such systems is typically $\sim O(kHz)$, and is outside of the frequency bands of ground-based gravitational-wave interferometer detectors. Therefore

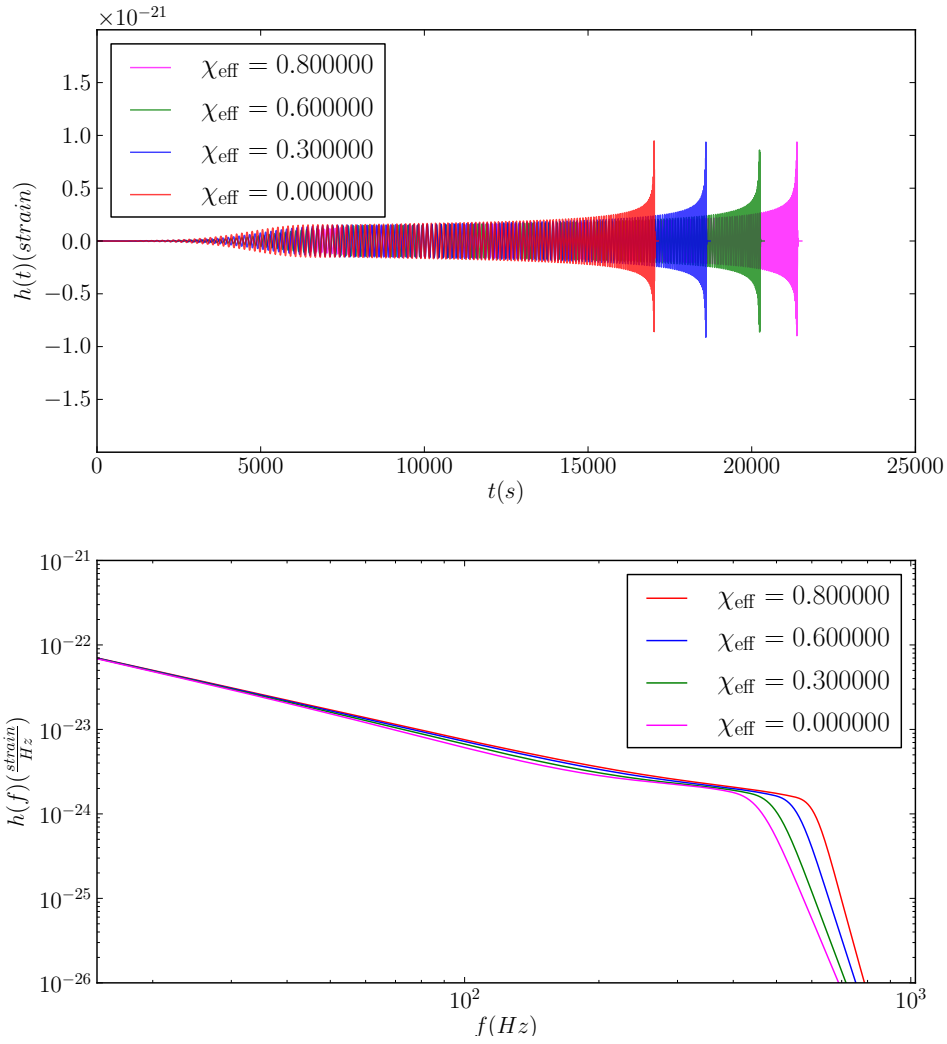



Figure 3.3: Example time-domain (top) and frequency-domain (bottom) waveforms with different $\chi_{\rm eff}$ values but same masses. We see that high spinning systems are longer in band. We used the model `IMRPhenomD` to generate this waveform.

for low-mass systems, we only need to model the inspiral part of the waveform. In the inspiral regime, approximate analytical solutions of Einstein's equations can be found using two main techniques: *post-Newtonian (PN) approximation* [63–67] and the *effective-one-body (EOB)* approach [68, 69]. In the PN approximation, gravitational waves are modeled in terms of an expansion in $\frac{v}{c}$, where $v$ is the characteristic velocity of the binary. This technique breaks down for high mass-ratio systems, and close to merger as $v$ approaches $c$ (see Fig. 3.1). Some inspiral only waveform models used in LIGO data analysis are, in frequency domain `TaylorF2` and `TaylorF2RedSpin` and in time-domain `SpinTaylorT2`, `SpinTaylorT4`, `SpinTaylorT5`, and `TaylorT2`.

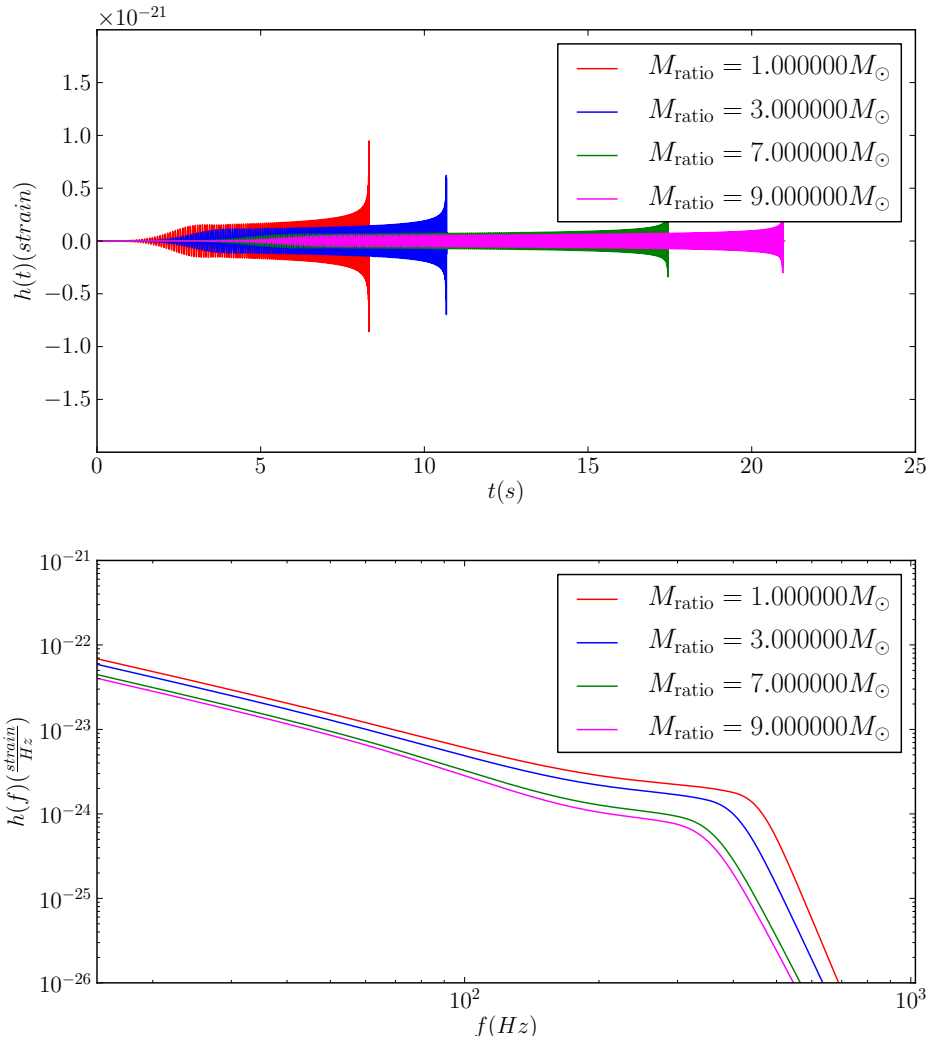



Figure 3.4: Example time-domain (top) and frequency-domain (bottom) waveforms with different mass ratios but a constant total mass. We see that high mass-ratio systems are longer in band, even though they merge at a lower frequency. We used the model `IMRPhenomD` to generate this waveform.

Note the SpinTaylor waveforms are precessing models.

To the leading order, the evolution of waveforms in the inspiral phase is determined by the *chirp mass*, $\mathcal{M}_c$, which is given by

$$\begin{aligned} \mathcal{M}_c &= \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \\ &= \eta^{3/5} M, \end{aligned} \tag{3.1}$$

where $\eta = m_1 m_2/(m_1 + m_2)^2$ is the symmetric mass ratio, and $M = m_1 + m_2$ is the total mass of the binary.

The mass-ratio of the binary system and the spins on the component masses only have an effect at the second order, in terms of, $eta$ and $\chi_{\rm eff}$, where

$$\chi_{\rm eff} = \frac{m_1\chi_1 + m_2\chi_2}{m_1 + m_2}, \tag{3.2}$$

where

$$\chi_i = \vec{S}_i.\hat{L}/m_i^2 \tag{3.3}$$

are the dimensionless spin parameters aligned with the orbital angular momentum of the component masses, $\vec{S}_i$ are the component spin vectors and $\hat{L}$ is the orbital angular momentum vector. The EOB approach is valid for a larger parameter space, including high mass-ratio systems, but is more computationally expensive, as it requires differential-equations to be solved in the time domain for each unique set of intrinsic parameters..

Even before the gravitational-wave detection of binary black holes by Advanced LIGO-Virgo, black holes had been observed in X-ray binaries ranging up to $\approx 20M_\odot$. Furthermore, a number of formation channels had been suggested through which significantly more massive black-hole binaries could form. For such high mass systems, the merger and ringdown part of the waveform would be in the detector band as well. During the later stages of inspiral, close to merger, the approximation methods mentioned above break down, and the Einstein's equations have to be solved in full general relativity. Since no analytical solutions are available for binary spacetimes, the field equations have to be solved numerically on supercomputers.

The gravitational waveform decays exponentially during the ringdown, which can either be approximated by black-hole perturbation theory or one may use the results of numerical simulations.

### 3.1.2 Complete waveform models

For high mass systems ($M = m_1 + m_2 > 12.0$) , the merger of the waveform happens in LIGO band. Therefore to detect such systems, we need waveform models that include merger and ringdown as well, and not just the inspiral. Numerical simulations are computationally expensive to perform, specially for high mass-ratio systems. Only a finite number of such simulations exist, although the number of simulations available are growing in number. Recently a numerical relativity waveform surrogate model for generically precessing spins has been developed [70]. There exist different approaches by which full IMR waveform models are obtained. One such approach is the phenomenological approach [71], where a simple, analytic closed-form expression

for the full IMR model is obtained in terms of binary parameters. This is done by first constructing *hybrid* waveforms in which the merger and ringdown from numerical relativity is attached to an inspiral waveform calculated using the PN theory. This is done by matching the two waveforms in an overlapping time or frequency interval. A number of phenomenological methods are used in LIGO data analysis, and will be referred to in this thesis. These include, `IMRPhenomB` [72], `IMRPhenomD` [73, 74], `IMRPhenomPv2` [75], etc.

Another successful method to create full IMR waveform models to map the two-body problem in General Relativity onto an effective-one-body problem (EOB) to describe the inspiral part of the waveform, This description is then extended to the merger and ringdown using information from numerical theory. Some models that use this strategy that we will refer to later are, in time-domain `SEOBNRv2` [76] and `SEOBNRv4` [77], and the reduced order frequency domain models created from these, `SEOBNRv2_ROM_DoubleSpin` [78] and `SEOBNRv4_ROM`. These waveform models claim to be valid for high mass-ratio systems (upto $m1/m2 \sim 100$) and extremal magnitudes of aligned spins.

### 3.2 Detecting compact binary coalescences

In this section, we describe the signal processing techniques used to detect gravitational waves from coalescing compact binaries buried in detector noise.

The detector output data stream, $s(t)$ can be written as a sum of noise in the detector, $n(t)$, and possibly a gravitational-wave induced strain, $h(t)$.

$$s(t) = \begin{cases} n(t), & \text{noise hypothesis} \\ n(t) + h(t), & \text{signal hypothesis} \end{cases} \tag{3.4}$$

Given the observed data and the predicted form of the signal, our goal is to decide between the signal hypothesis and the noise hypothesis. Optimal methods for extracting signals buried in detector noise are well developed for cases when the statistical properties of detector noise are well characterized. We assume that the noise is both stationary and Gaussian, despite the fact that this is not, in general, true; we deal with non-stationarity and non-Gaussianity through a variety of techniques discussed in the next chapter. Under these assumptions, the one-sided power spectral density (PSD) is sufficient to describe the statistical properties of noise. It is defined as

$$S_n(f) = \lim_{T\to\infty} \frac{2}{T}\left|\int_{-T/2}^{T/2} n(t)e^{-2\pi i f t}dt\right|^2. \tag{3.5}$$

Or equivalently,

$$\frac{1}{2}S_n(f)\delta(f-f') = \langle \tilde{n}*(f')\tilde{n}(f)\rangle\,, \tag{3.6}$$

where $\tilde{n}(f)$ is the Fourier transform of $n(t)$, $\mathcal{F}(n(t))$, defined as,

$$\tilde{n}(f) = \int_{-\infty}^{\infty} n(t)e^{-2\pi i f t}dt. \tag{3.7}$$

The Fourier transform leaves variances in the transformed frequency series $\tilde{n}(f)$. To remove these variances, we can divide by the square-root of the PSD to obtain,

$$\tilde{n}_{white}(f) = \frac{\tilde{n}(f)}{\sqrt{(\frac{1}{2}S_n(f))}}, \tag{3.8}$$

which is $white$ noise and has a unit PSD. We define the inner product on the signal space as [79]

$$\langle a,b\rangle = \int_{-\infty}^{\infty} \frac{\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)}{S_n(f)}df \tag{3.9}$$

In case of gravitational waves from CBCs, we are dealing with signals that are usually of the order of the noise in our detector. Therefore, we have to make use of $matched-filtering$ [54], which is a technique in signal processing used to extract signals buried deep in noise in case the form of signal is known.

### 3.2.1 Matched-filtering

A matched-filter is obtained by cross-correlating a known signal waveform, or $template$, with the data. The output of the matched-filter is the signal-to-noise ratio (SNR), and in our case is just the inner product of the data with the normalized signal template,

$$\rho = \left\langle s,\hat{h}\right\rangle = \int_{-\infty}^{\infty} \frac{\tilde{s}(f)\hat{h}^*(f) + \tilde{s}^*(f)\hat{h}(f)}{S_n(f)}df = 4\mathcal{R}\left(\int_0^{\infty} \frac{\tilde{s}(f)\hat{h}^*(f)}{S_n(f)}df\right),$$

where $\hat{h}^*(f)$ and $\tilde{s}(f)$ are the Fourier transforms of the strain time series, $s(t)$, and the normalized signal time series, $\hat{h}(t)$, respectively. The normalized signal waveform is

$$\hat{h}(t) = \frac{h(t)}{\sqrt{\langle h,h\rangle}}. \tag{3.10}$$

As mentioned in Table 3.1, the signal waveform, $h(t)$, or equivalently, $\tilde{h}(f)$ depends on a number of parameters which are not known *a priori*. To calculate the SNR, therefore, we have to maximize over all possible parameters. In the case where the spins of the binary components are aligned with the orbital angular momentum, the extrinsic waveform parameters are easy to deal with, as they enter only into the overall amplitude and phase of the signal. For these parameters, the maximization step can be performed analytically, or in a computationally efficient efficient way using Fast Fourier Transform (FFT) algorithms[1]. The remaining intrinsic source parameters, however, require the construction of a *template bank* (Sec. 3.2.3), which is a set of discretized points in the intrinsic parameter space. The matched-filter output is then calculated over all of these points to find the maximum overlap.

### 3.2.2 Maximizing the matched-filter output over the extrinsic parameters

In the CBC searches at the moment we only consider waveforms with spins aligned with their orbital angular momentum. This not only reduces the parameter space of the search, but also makes it easy to maximize over the extrinsic parameters. However, efforts are being made to also include precession in our searches [80, 81]. The waveform observed at the detector, $h(t)$, is the sum of the two gravitational-wave polarizations, $h_+$ and $h_\times$, multiplied by the response function of the detector to the respective polarization, $F_+$ and $F_\times$ [82],

$$h(t) = F_+(\alpha, \delta, \psi, t_c)h_+(t) + F_\times(\alpha, \delta, \psi, t_c)h_\times(t). \tag{3.11}$$

The detector response depends on the sky position of the source in the detector frame ($\alpha, \delta, t_c$ : $t_c$ is needed to transform the sky position of the source from the celestial coordinates to the frame of the detector), and the orientation of the polarization vector of the source with respect to the detector arms ($\psi$).

In making the connection between the two waveform polarizations, $h_+(t)$ and $h_\times(t)$, the aligned-spin searches make the assumption that only the dominant $(l, |m|) = (2, 2)$ mode of the waveform contribute to the SNR. In this approximation, the two polarizations are $\pi/2$ phase-shifted from one-another, and can be written as [83],

$$\begin{aligned} h_+(t) &= \frac{1 + \cos^2 i}{2D} A_\zeta(t - t_c)\cos[2(\Phi_\zeta(t - t_c) + \phi_c)], \\ h_\times(t) &= \frac{\cos i}{D} A_\zeta(t - t_c)\sin[2(\Phi_\zeta(t - t_c) + \phi_c)], \end{aligned} \tag{3.12}$$

[1]Only the frequency-domain searches can make use of the inverse FFT to maximize over coalescence time in one shot. In time-domain searches, the maximization of the matched-filter output over coalescence time has to do be done in a brute force way, as we will see in the next chapter.

where $A_\zeta$ and $\Phi_\zeta$ are functions of the intrinsic parameters, $\zeta = (m_1, m_2, \chi_1, \text{and}, \chi_2)$. $\chi_1$ and $\chi_2$ are defined in Eq. 3.3.

Inserting Eq. 3.12 in Eq. 3.11, we get

$$h(t) = \frac{A_\zeta(t - t_c)}{D_{\text{eff}}} \cos[2(\Phi_\zeta(t - t_c) + \phi_0)], \tag{3.13}$$

where

$$D_{\text{eff}} = \frac{D}{\left[F_+^2 \left(\frac{1+\cos^2 i}{2}\right)^2 + F_\times^2 \cos^2 i\right]} \tag{3.14}$$

is the *effective distance*, which is defined as the distance at which an optimally located and oriented binary with the same intrinsic parameters would give the same signal amplitude. It is always greater than the luminosity distance.

And

$$\phi_0 = \phi_c - \frac{1}{2}\tan^{-1}\left(\frac{2F_\times \cos i}{F_+(1 + \cos^2 i)}\right) \tag{3.15}$$

is the *termination phase*, which is the phase of the waveform projected on the detector at coalescence.

Thus, in the aligned-spin case, the waveform h(t) depends on the extrinsic parameters only through combinations $D_{\text{eff}}$ and $\phi_0$, an overall amplitude and phase.

Recall that a time shift $\tau$ in a time-domain signal becomes a phase shift in its Fourier transform,

$$\tilde{h}(f; \tau) = \tilde{h}(f) e^{-2\pi i f \tau}. \tag{3.16}$$

We make the assumption that the stationary phase approximation holds [84], which means that at any moment in time, $\Phi_\zeta(t)$ is dominantly one frequency. This allows us to factorize the dependence of the Fourier domain waveform on $\phi_0$ using the time-translation property of the Fourier transform as

$$\tilde{h}(f) = \frac{1}{D_{\text{eff}}} e^{2i\phi_0} \tilde{h}_0(f; t_c), \tag{3.17}$$

where $\tilde{h}_0(f; t_c)$ is the Fourier transform of $h_0(t - t_c)$, which we have defined as

$$h_0(t - t_c) = A_\zeta \cos[2(\Phi_\zeta(t - t_c)], \tag{3.18}$$

which only depends on the intrinsic parameters, $\zeta = (m_1, m_2, \chi_1, \chi_2)$, and the time of coalescence, $t_c$.

We can now maximize the matched-filter output, $\rho(t_c, \alpha, \delta, \psi, \phi_c, i, \zeta) = \left\langle s, \hat{h} \right\rangle$, over the parameters $\alpha, \delta, \psi, \phi_c, i$ by maximizing over their combinations $D_{\text{eff}}$ and $\phi_0$. Note however, we do not need to maximize over the constant amplitude factor $D_{\text{eff}}$ as we calculate the SNR by filtering with normalized templates.

$$\max_{\alpha,\delta,\psi,\phi_c,i}(\rho(\alpha, \delta, \psi, \phi_c, i, t_c)) = \max_{\phi_0}(\rho(\phi_0, t_c)) = \max_{\phi_0} \left\langle s, \hat{h} \right\rangle, \tag{3.19}$$

$$= \max_{\phi_0} \left\langle s(f), e^{2i\phi_0} \hat{h}_0(f; t_c) \right\rangle, \tag{3.20}$$

$$= \max_{\phi_0} 4\mathcal{R}\left(\int_0^\infty \frac{\tilde{s}^*(f) e^{2i\phi_0} \hat{h}_0(f; t_c)}{S_n(f)} df\right). \tag{3.21}$$

The maximum value of the real part of a complex number on varying its phase is obtained when it's equal to the absolute value. Therefore,

$$\max_{\phi_0}(\rho(\phi_0, t_c)) = \left|\left\langle s(f), e^{2i\phi_0} \hat{h}_0(f; t_c) \right\rangle\right|. \tag{3.22}$$

If we define a complex snr time series, $z$, as

$$z(\phi_0, t_c) = \int_0^\infty \frac{\tilde{s}^*(f) \hat{h}(f; t_c)}{S_n(f)} df, \tag{3.23}$$

then the matched-filter output mazimized over the termination phase can be written as

$$\max_{\phi_0} \rho(\phi_0, t_c) = |z(\phi_0, t_c)|. \tag{3.24}$$

Now we wish to also maximize this output over coalescence time,

$$\rho = \max_{\phi_0, t_c} \rho(\phi_0; t_c) = \max_{t_c} |z(\phi_0, t_c)|. \tag{3.25}$$

Due to the time translation property of the Fourier transform Eq. 3.16, Eq. 3.17 becomes

$$\tilde{h}(f) = \frac{1}{D_{\text{eff}}} e^{2i\phi_0} e^{-2\pi i f t_c} \tilde{h}_0(f). \tag{3.26}$$

Therefore, from Eq. 3.25, we get that

$$\rho = \max_{\phi_0, t_c} z(\phi_0, t_c) = \max_{t_c} \left| \left\langle s(t), \hat{h}(t - t_c) \right\rangle \right| = \max_{t_c} 4 \left| \left( \int_0^\infty e^{2\pi i f t_c} \frac{\tilde{s}(f)\tilde{\hat{h}}^*(f)}{S_n(f)} df \right) \right|. \tag{3.27}$$

Frequency-domain searches can efficiently maximize the SNR over the coalescence time $t_c$ through the use of an inverse Fourier transform. Note that the integral is nothing more than an IFFT, and that fast numerical algorithms can be used to evaluate it [85]. Eq. 3.27 provides a computationally efficient method for maximizing the SNR over time, with a single, fast, inverse Fourier transform. Instead of having to slide our template against the data to calculate the SNR at each sample, we can get the entire SNR time series ($\rho(t)$) for all the time translations of the template against the data by taking the inverse Fourier transform of $\frac{\tilde{s}(f)\tilde{\hat{h}}^*(f)}{S_n(f)}$.

### 3.2.3 Maximizing the matched-filter output over the intrinsic parameters : Template bank

In the previous section, we covered the problem of searching over the extrinsic parameters of the binary source while performing matched-filtering. Here, we discuss searching over the instrinsic parameters of the binary, which are the masses and spins of the binary components. These parameters enter into the phase and amplitude of the gravitational-wave signal in a more complicated way as compared to the extrinsic parameters. Therefore we resort to computing the matched-filter output over the entire instrinsic parameter space and find the maximum overlap. The intrinsic parameter space of the binary is continuous, and searching over it completely requires an infinite amount of computational resources. Thus, we create a *template bank* [86–89], containing a discrete set of waveforms spanning the intrinsic parameter space of the search. These discrete points in the search parameter space are chosen such that the mis-match between any signal and the best matching template from the template bank (arising from the discrete nature of our template bank) is less than a pre-specified tolerance. We will formalize shortly what we mean by "match". We closely follow the dicussion of template banks given in [90].

Given two waveforms $h_1$ and $h_2$, we define the match, $\mathcal{M}$, between them by [79],

$$\mathcal{M} \equiv \max_{\vec{\lambda}_{\rm extr}} \left\langle \hat{h}_1, \hat{h}_2 \right\rangle \ , \tag{3.28}$$

where the inner product is maximized over the extrinsic parameters, $\vec{\lambda}_{\rm extr}$. So the

match measures the maximum overlap between two waveforms given their intrinsic parameters.

We denote the match by $\mathcal{M}(\vec{\lambda}, \Delta\vec{\lambda})$, where $\vec{\lambda} = \vec{\lambda}_{h1}$, $\Delta\vec{\lambda} = \vec{\lambda}_{h1} - \vec{\lambda}_{h2}$ is the difference between the intrinsic parameters of the two waveforms. For the match between two templates with very similar parameters, which are neighboring in the parameters space, we can Taylor-expand it about $\Delta\vec{\lambda} = 0$. There are no linear terms in the expansion, since the match function has its maximum value of unity at $\Delta\vec{\lambda} = 0$ [86].

$$\mathcal{M}(\vec{\lambda}, \Delta\vec{\lambda}) \approx 1 - g_{ij}\Delta\vec{\lambda}^i\Delta\vec{\lambda}^j \tag{3.29}$$

is the Taylor-expansion of the match, where

$$g_{ij} = -\frac{1}{2}\left(\frac{\partial^2\mathcal{M}}{\partial\vec{\lambda}^i\partial\vec{\lambda}^j}\right) = \left\langle \hat{h}_{\lambda^i}, \hat{h}_{\lambda^j} \right\rangle \tag{3.30}$$

can be interpreted as the local metric on a manifold in the coordinates of the binary parameter space. $\hat{h}_{\lambda^i}$and$\hat{h}_{\lambda^j}$ are partial differentials of the normalized waveform $\hat{h}$ with respect to the instrinsic parameters, $\lambda^i$, and $\lambda^j$. The mismatch defined as $1 - \mathcal{M}$, between two neighboring waveforms has the interpretation of proper distance in this manifold [86] and for small $\Delta\vec{\lambda}$ is approximately given by

$$1 - \mathcal{M} \approx g_{ij}\Delta\vec{\lambda}^i\Delta\vec{\lambda}^j. \tag{3.31}$$

The metric on the signal manifold is equal to the Fisher information matrix $\Gamma_{ij}$ [87], scaled down by the square of the SNR, i.e., $g_{ij} = \rho^{-2}\Gamma_{ij}$. Also see Sec. 8.3.

**Fitting factor**

The fitting factor of a template bank towards a particular signal establishes the "effectiveness" of a template bank for recovering the SNR [91]. The fitting factor of a signal waveform with template bank is defined as the maximum match the signal has with the templates in the bank,

$$\mathrm{FF}(\vec{\lambda}_s; B) = \max_{\hat{h}\in B} \mathcal{M}(\vec{\lambda}_s, \vec{\lambda}_s - \vec{\lambda}_h). \tag{3.32}$$

$B$ here denotes the template bank, $s$ the signal, and $h$ the templates in the bank.

We define the effective volume coverage of a template bank with respect to a set of signals as a fraction of the sensitive volume (volume we would have been sensitive to if we had templates exactly matching the signals in our template bank) as the expectation value of the fitting factor,

$$V_{\rm eff} = V_{\rm sens} \left\langle {\rm FF}^3 \right\rangle . \tag{3.33}$$

If we are willing to incur a 10% loss in volume, i.e., $V_{\rm eff} = 0.9 V_{\rm sens}$, this gives $\langle {\rm FF} \rangle = 0.965$. We typically construct template banks with a fitting factor criteria of 0.97 for our searches.

### Template placement

There exist two basic methods for efficiently covering the intrinsic parameter space of the binaries: *lattice placement* [80, 87, 92, 93] and *stochastic placement* [88, 89, 94, 95].

*Lattice placement* methods make use of the metric of the waveforms given by Eq. 3.30 to calculate the match between neighboring waveforms (Eq. 3.31). Eq. 3.31 is only valid for overlaps close to 1, but this method has been used for constructing banks with a minimal match of 0.97. For bank construction, it is required that the metric be constant in parameter space, or at least nearly constant. The metric coefficients are coordinate dependent. For analytic inspiral-only waveform models, one can construct a special set of coordinates in which the metric becomes nearly flat.

In the waveform given by the post-Newtonian expansion upto the second order, for systems with negligible spins, these coordinates are the dimensionless chirp times [87, 92],

$$\theta_0 = \frac{5}{128\eta} \left(\pi M f_{\rm low}\right)^{-5/3} , \tag{3.34}$$

$$\theta_3 = -\frac{\pi}{4\eta} \left(\pi M f_{\rm low}\right)^{-2/3} , \tag{3.35}$$

where $M$ is the total mass of the binary and $\eta = m_1 m_2 / M^2$ is the symmetric mass ratio. $f_{\rm low}$ is the starting frequency of the analysis, with with the match is calculated. However, when merger and ringdown, or higher order post-Newtonian corrections become important for detection, such coordinates are not known. Therefore, we look for different methods to efficiently place templates.

*Stochastic placement* method is a brute force approach to construct template banks [88, 89]. In this method, we randomly select points in the given parameter space, and

place a test template there. We calculate the fitting factor of this test template with the already existing templates. If this fitting factor (which recall is the maximum match over the existing templates) is greater than the fitting factor criteria, we discard the test template, otherwise if the fitting factor is smaller than the minimum fitting factor criteria, we keep it. We keep doing this until a pre-decided convergence condition is reached.

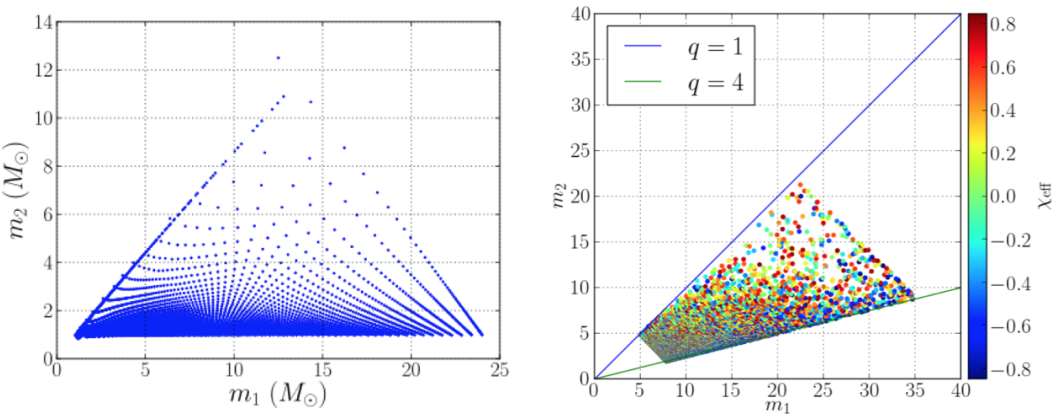



Figure 3.5: A non-spinning bank created using the lattice technique (left). An aligned-spin IMR template bank created using the stochastic method (right).

Fig. 3.5 shows examples of template banks created using the two placement techniques.

In this chapter, we have discussed how to perform matched-filtering in order to detect CBC sources. However, the output of matched-filtering, SNR, is the optimal detection statistic under the condition that noise is Gaussian. Our detector noise often suffers from transient non-Gaussian features we call glitches. So in practice our CBC search pipelines use more complex detection statistics than just SNR to distinguish between noise and signal. At present there are two flagship CBC searches in the LVC, the `PyCBC` (performs filtering in the frequency-domain, applies a signal-consistency test in the frequency domain, uses time slides to establish significances) and the `GstLAL` (time-domain matched-filtering, time-domain signal consistency test, uses a multi-dimensional likelihood-ratio statistic to establish significance of events). The two searches make use of very different methods, and therefore it is useful to have them both to cross-validate results, more on this in Chapter 5. There also exist other CBC searches in the collaboration that are actively being developed, e.g., `GstLAL-spiir`, `MBTA` among others.

*Chapter 4*

# THE GSTLAL-INSPIRAL BASED PIPELINE

## 4.1 Introduction to the GstLAL pipeline

The GstLAL-inspiral based pipeline (henceforth referred to as the GstLAL pipeline) is a matched-filtering analysis pipeline that can detect gravitational waves from compact binary mergers in near real time, and provide point estimates for binary parameters and the significance of events. Matched-filtering involves cross-correlating data against a bank of waveform templates formed using general relativity. The pipeline is built on the GstLAL library, a collection of GStreamer [96] libraries and plug-ins that make use of the LIGO Algorithm Library, LALSuite [97]. It uses the GStreamer library to stream the gravitational strain data in real time, performs matched-filtering in the time-domain [3, 50] as opposed to the more traditional frequency-domain [84] method, and uses a time-domain rather than a frequency-domain signal consistency test [84]. It also employs multi-banding and singular value decomposition on signal templates [3, 4] to reduce the number of filters and samples used in matched-filtering. A multidimensional likelihood ratio statistic is used to rank the gravitational-wave candidates according to the properties of noise and signal [3, 98]. Instead of performing time-slides [84] for background estimation, a technique that is based on tracking the noise distributions [99] and allows for rapid significance estimation is used. In addition, the GstLAL pipeline can be run in an “offline" mode, where it can process data archived on disks. The online pipeline uses background informed from the past data to rank events, whereas the offline pipeline can make use of all available data to inform the background. Furthermore, while analyzing data in the offline mode, the pipeline can benefit from data quality products [100] which are not available online[1].

In this chapter we describe the methods used in the GstLAL pipeline.

### 4.1.1 Why low latency?

GW170817 is a perfect example of why we aim to detect gravitational-wave events in low latency. On August 17, 2017, during the second observing run of the Advanced LIGO detectors and the first observing run of the Advanced Virgo detector, the GstLAL

[1] Although there are plans to have low-latency data quality information for Advanced LIGO’s first and Advanced Virgo’s second observing run (O3).

pipeline identified a BNS merger event, GW170817 [11] in low latency. The Fermi and INTEGRAL gamma-ray telescopes observed the gamma-ray burst GRB170817A 1.7 s after the binary's coalescence time with an inferred sky-location consistent with that measured for GW170817 [45]. The rapid identification of the gravitational-wave event by the `GstLAL` pipeline spurred an extensive follow-up campaign by astronomers all over the world and the optical transient SSS17a/AT 2017gfo was discovered within 11 hours [12], localizing the source of the event to the galaxy NGC 4993. These joint gravitational-wave and electromagnetic observations support the hypothesis that neutron star mergers are progenitors of short gamma-ray bursts and are followed by transient electromagnetic events known as kilonovae [12]. Low-latency gravitational-wave detection serves as early warning triggers for electromagnetic and neutrino follow-up enabling multimessenger astrophysics. Furthermore, even in the absence of a counterpart, low-latency observations can help provide feedback to the detector characterization group about the data quality, and in a timely manner minimize detector changes upon identification of a candidate in order to gather enough data for background estimation.

Upon identification of an event that passes a predetermined significance threshold, it is sent to the Gravitational-Wave Candidate Event Database (GraceDB) [101] which provides an environment for collection of all the events and for recording information about follow-ups.

Now that we have discussed the need for low-latency, we move on to describing the methods of the `GstLAL` pipeline, highlighting the ones that are vital for allowing the pipeline to keep up with the data in real time.

## 4.2 Pipeline methods

In this section we briefly describe the methods used by the `GstLAL` pipeline and provide an overview of its workflow (See Fig. 4.1).

The `GstLAL` pipeline makes use of matched-filtering (Sec. 3.2.1) to extract signals from noisy data by cross-correlating the detector output with a predicted waveform signal. The CBC waveform models are described in Sec. 3.1.2.

### 4.2.1 Matched filtering in the time-domain

The output of the matched-filter is the SNR (Sec. 3.2.1), the inner product of the whitened data with the whitened template. To facilitate low latency, it is calculated in the time-domain [3, 50] by the `GstLAL` pipeline:

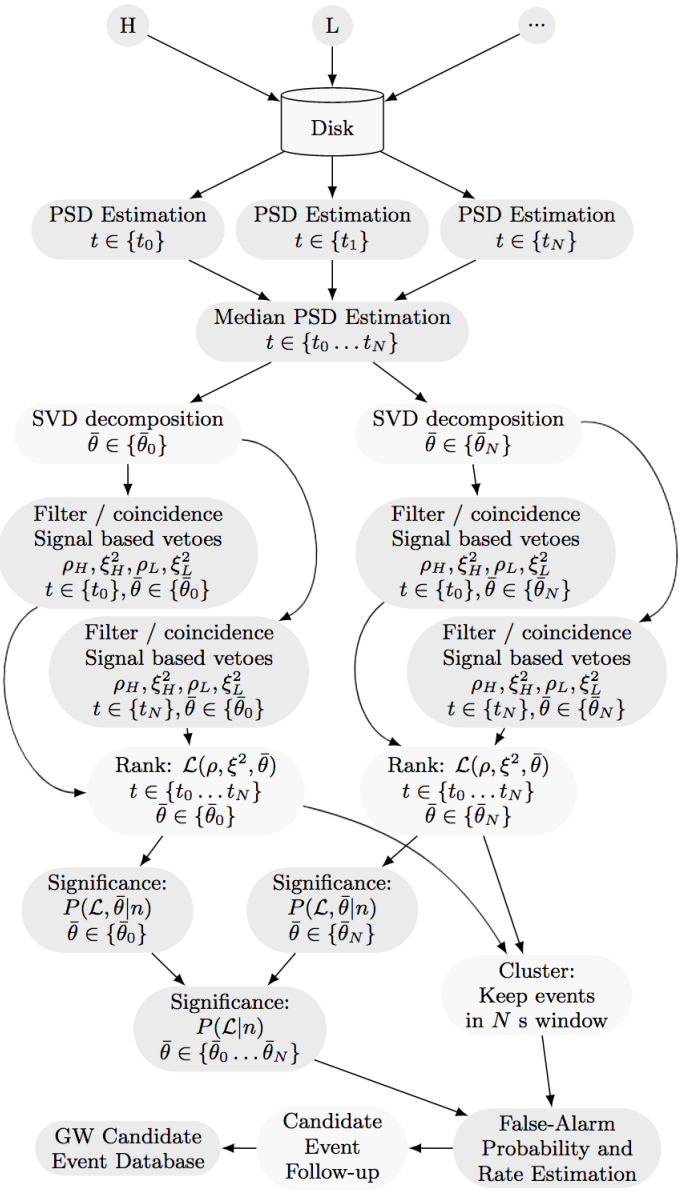



Figure 4.1: Diagram of the workflow of the `GstLAL` pipeline. Figure taken from [3].

$$x_i(t) = \int_{-\infty}^{\infty} \mathrm{d}\tau \hat{h}_i(\tau) \hat{d}(\tau + t), \tag{4.1}$$

where

$$\hat{d}(\tau) = \int_{-\infty}^{\infty} \mathrm{d}f \frac{\widetilde{d}(f)}{\sqrt{S_n(|f|)/2}} e^{2\pi i f \tau} \tag{4.2}$$

is the whitened data and the whitened template $\hat{h}_i(\tau)$ is defined similarly. $S_n(f)$ is the single-sided noise power spectral density (PSD). The stationary noise amplitude spectrum is estimated using a combined median–geometric-mean modification of Welch's method [3]. Piece-wise on overlapping Hann-windowed time-domain blocks are whitened using this estimate of the PSD and subsequently summed together to yield a continuous whitened time series.

The subscript $i$ in Eq. 4.1 runs over each set of template parameters in our template bank. For each set of parameter values in the template bank, there are two real waveforms in the `GstLAL` filter bank, one corresponding to the '+' polarization, and the other corresponding to a 'quadrature phase-shifted +' waveform, which is equal to the '×' polarized waveform barring an overall amplitude factor.

We construct a complex signal-to-noise ratio (SNR) time series, real part of which is the SNR time series from the '+' polarized template ($x_i(t)$), and the complex part is the SNR time series from the 'quadrature-phase shifted +' polarized template ($y_i(t)$). We maximize over the unknown coalescence time and termination phase by maximizing over the absolute value of the complex SNR time series over time,

$$\rho = \max_t |x_i(t) + i y_i(t)|. \tag{4.3}$$

Time-domain filtering is naively more expensive than frequency-domain filtering. Frequency-domain searches can make use of the computationally efficient inverse FFT to calculate the SNR time series and maximize over coalescence time in one shot, which is an advantage that the time-domain searches lack. The `GstLAL` pipeline makes use of certain "tricks" to achieve low-latency computationally efficiently, which we describe in the following section.

### 4.2.2 The LLOID method

The `GstLAL` pipeline makes use of the LLOID (Low Latency Online Inspiral Detection) algorithm to efficiently perform matched filtering in the time-domain and keep up with the data in low latency. This method combines singular value decomposition (SVD) [4, 102] with near-critical sampling to construct a reduced set of orthonormal filters with a reduced number of samples. For performing the SVD, the template bank is first split into partially-overlapping "split-banks" of templates with similar time-frequency evolution. This partitioning is performed based on the intrinsic template parameters, as depicted in Fig. 4.2. In O1, this division was done

on the basis on the chirp mass, $\mathcal{M}_c$, and the effective aligned-spin, $\chi_{\rm eff}$ parameter. $\mathcal{M}_c$ and $\chi_{\rm eff}$ are the leading order terms that describe the phase evolution of the inspiral part of the waveform according to the Post Newtonian expansion.

The chirp mass $\mathcal{M}$ is defined as

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}. \tag{4.4}$$

The effective spin parameter is given by

$$\chi_{\rm eff} \equiv \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2}. \tag{4.5}$$

$2N_T$ real templates are places in each split-bank, where $N_T$ is typically $\mathcal{O}(100)$. The factor of 2 is a result of having two orthogonal real-valued templates (corresponding to the two polarizations) in place of 1 complex-valued template (Sec. 4.2.1). The adjacent bins are made to have some overlap in templates to mitigate the boundary effects from the SVD. Overlapping templates are "clipped" after reconstruction so that the output has no redundant waveforms. The waveforms are then whitened using the PSD, and each split-bank is decomposed via the LLOID method as described below. In addition to being used for the LLOID decomposition, split-banks are also used as bins for estimating the background and the signal model for calculating the likelihood-ratio detection statistic (Sec. 4.2.5). These are referred to as $\bar{\theta}$ bins.

Each split-bank is divided into various time slices after prepending the templates with zeros such that every template has the same number of sample points; this allows for an efficient sampling of the different regions of the waveforms with the appropriate Nyquist frequency instead of over-sampling the low-frequency regions of the waveform with the Nyquist frequency required for the high-frequency regions. The SVD is then performed on each time slice of each split-bank. We retain only the most important basis waveforms returned by the SVD algorithm, as measured by the match between the original templates and the reconstructed waveforms [3, 50]. See Fig. 4.3.

### 4.2.3 Signal-based vetoes

Detector data often contain glitches, which can produce high peaks in the SNR time series. Because of that, SNR is not sufficient to distinguish noise from transient signals in presence of non-Gaussian data. Therefore, in addition to recording the peaks in the SNR time series, the pipeline performs a signal consistency check

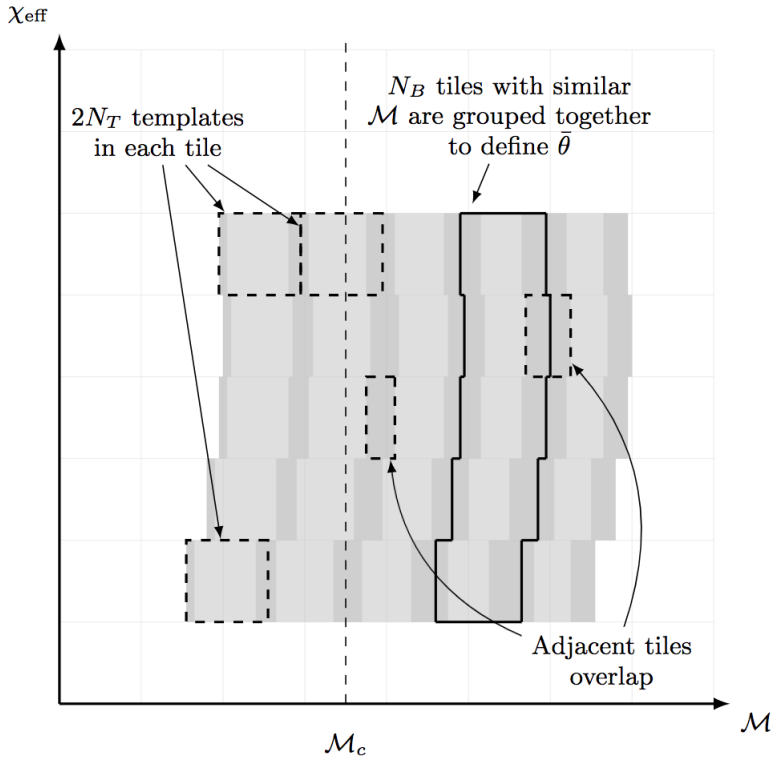



Figure 4.2: A visual representation of how the physical parameter space is tiled into sub regions upon which the SVD and time slicing is performed. Figure taken from [3]

whenever it records an SNR above a certain threshold. This is done by determining how similar the SNR time series of the data around the peak value is to the SNR time series expected from a real signal. The SNR time series is predicted by calculating the auto-correlation between the complex template waveform and itself, and scaling it by the peak complex SNR. This predicted SNR time series is equal to the SNR time series under the assumption that the signal matches the template waveform exactly in absence of any detector noise. This signal consistency test value, $\xi^2$, is computed by integrating the amplitude squared of the difference between the complex SNR time series and the predicted SNR time series over a $\delta t$ time window around the peak, and normalizing it appropriately [3],

$$\xi^2 = \frac{\int_{-\delta t}^{\delta t} dt |z(t) - z(0)R(t)|^2}{\int_{-\delta t}^{\delta t} dt (2 - 2|R(t)|^2)}. \tag{4.6}$$

Here $z(t)$ is the complex SNR time series, $z(0)$ its peak, and $R(t)$ is the auto-correlation series. Fig. 4.4 shows an example of the $\xi^2$ test being performed for the GW170608

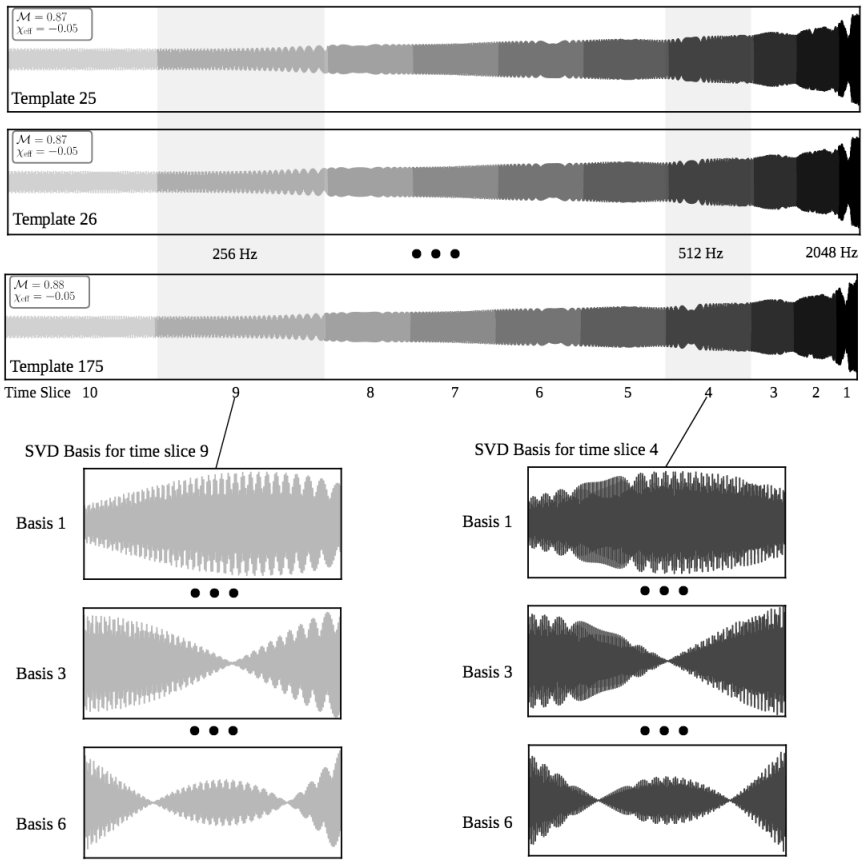



Figure 4.3: An example of LLOID decomposition taken from [3]. In this example, the number of templates in the split-bank is $2N_T - 390$. These are whitened binary inspiral waveforms with $\mathcal{M}_c$ between 0.87 and 0.88. The waveforms are prepended with zeros to ensure they all contain the same number of sample points. They are then decomposed into 30 time slices containing sample rates ranging between 128 Hz and 2048 Hz. A basis filter set was created from each time slice is constructed after performing the SVD [4]. We find that only 6-10 basis filters per slice was needed to reconstruct the 390 input waveforms to an accuracy of better than 99.9%.

event. Fig. 5.13 shows how the $\xi^2$ test is helpful in distinguishing non-Gaussian noise from signals.

### 4.2.4 Triggers

The raw SNR time series, typically sampled at 2 kHz, is discretized into “triggers" before being stored to disk. This is done by maximizing the SNR over time in one-second windows and recording the peak if it crosses a predetermined threshold. This threshold was set to SNR = 4 for the Advanced LIGO detectors in the Engineering

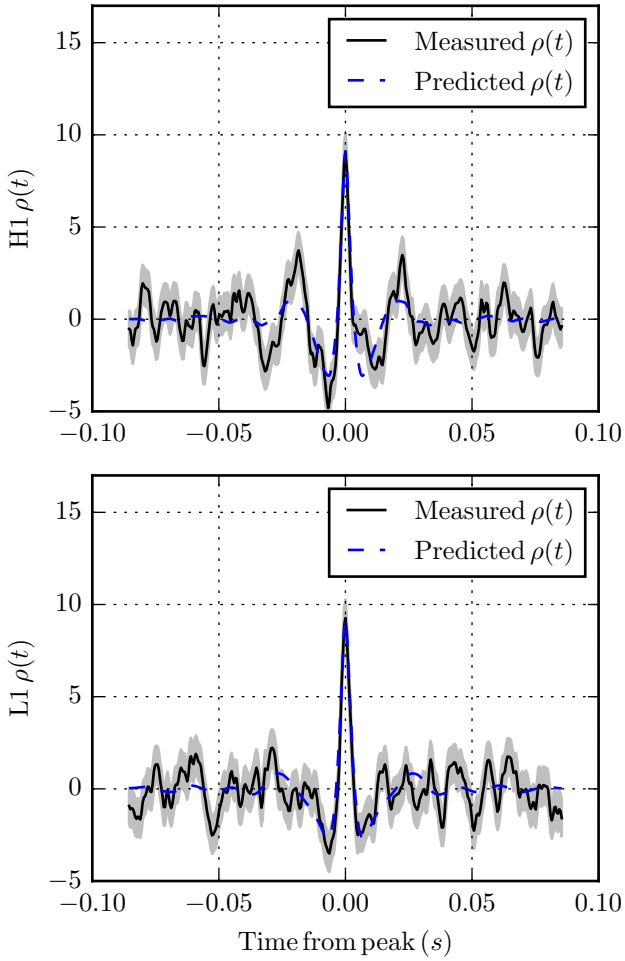



Figure 4.4: The two panels show the SNR time series near GW170608, a binary black hole event that was detected by Advanced LIGO during the second observing run (O2) on June 8$^{\text{th}}$, 2016 along with the predicted SNR time series computed from the autocorrelation of the template. Subtracting these two time series and integrating their squared magnitude provides a signal consistency test, $\xi^2$.

Runs 5 through 10 (ER5 - ER10), and the Observing Runs 1 and 2 (O1 & O2). Whenever the pipeline records a peak in the SNR time-series that is greater than the preset threshold, it records the SNR and $\xi^2$, the template that returned those values upon matched filtering (masses, spins of the binary, time of coalescence, phase of coalescence, SNR, and $\xi^2$). Together these quantities form a 'trigger'. These triggers are divided into the $\bar{\theta}$ bins corresponding to the bin that the template belongs to, as described in the previous section. Given a trigger in one detector, we look for triggers from the same template in other detectors within an appropriate time window, which takes into account the maximum light travel time between detectors and statistical

fluctuations in the measured event time due to detector noise. The latter is given by the sum of the reciprocal bandwidths of the detectors. These are called coincident triggers and are the candidates for the search[2]. After coincidences are formed, it is necessary to rank the candidates (coincident events) and assign a significance to each. The set of triggers that did not participate in a coincidence are used to characterize noise for their respective $\bar{\theta}$ bin.

### 4.2.5 Significance estimation

After performing matched filtering, it is important that we assign significances to the candidates of the search. The `GstLAL` pipeline uses the likelihood-ratio [98, 103] as the detection statistic to rank coincident events by their SNR, $\xi^2$, the instantaneous sensitivity of each detector, expressed as their horizon distance, $\{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}$), and the detectors involved in the coincidence $\{\mathrm{H1}, \mathrm{L1}, \dots\}$,

$$\begin{aligned}
&\mathcal{L}\left(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}, \{\mathrm{H1}, \mathrm{L1}, \dots\}, \rho_{\mathrm{H1}}, \xi^2_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \xi^2_{\mathrm{L1}}, \dots, \bar{\theta}\right) \\
&\quad = \mathcal{L}\left(\cdots \mid \bar{\theta}\right) \mathcal{L}\left(\bar{\theta}\right), \\
&\quad = \frac{P\left(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}, \{\mathrm{H1}, \mathrm{L1}, \dots\}, \rho_{\mathrm{H1}}, \xi^2_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \xi^2_{\mathrm{L1}}, \cdots \mid \bar{\theta}, \mathrm{s}\right)}{P\left(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}, \{\mathrm{H1}, \mathrm{L1}, \dots\}, \rho_{\mathrm{H1}}, \xi^2_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \xi^2_{\mathrm{L1}}, \cdots \mid \bar{\theta}, \mathrm{n}\right)} \mathcal{L}\left(\bar{\theta}\right). \quad (4.7)
\end{aligned}$$

The numerator and denominator in Eq. 4.7 are factored into products of several terms as described in [98], assuming that the noise distributions for each interferometer are independent of each other. The numerator is given as

$$\begin{aligned}
&= P(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}) P(\{\mathrm{H1}, \mathrm{L1}, \dots\} \mid \{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}, \mathrm{s}) \\
&\quad \times P(\rho_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \cdots \mid \{D_{\mathrm{H1}}, D_{\mathrm{L1}}, \dots\}, \{\mathrm{H1}, \mathrm{L1}, \dots\}, \mathrm{s}) \prod_{\mathrm{inst} \in \mathrm{H1,L1}} P(\xi^2_{\mathrm{inst}} \mid \rho_{\mathrm{inst}}, \mathrm{s}).
\end{aligned} \quad (4.8)$$

The signal model PDFs are constructed by performing a Monte Carlo integration by generating sources that are uniformly distributed in volume, inclination angle and sky location [98, 103]. The distribution of $\xi^2$ is obtained by assuming a maximum of 10% loss in SNR due to waveform mismatch [103, 104]. A detailed description of how each of these terms is calculated can be found in [98]. The denominator of

[2]Note that starting in O2, the `GstLAL` pipeline also treated the non-coincident triggers as candidates in the online search. Refer to Ch. 7.

Eq. 4.7 is factored as

$$P\left(\{D_{\rm H1}, D_{\rm L1}, \dots\}, \{\rm H1, L1, \dots\}, \rho_{\rm H1}, \xi^2_{\rm H1}, \rho_{\rm L1}, \xi^2_{\rm L1}, \cdots \mid \bar{\theta}, \rm n\right) \propto \prod_{\rm inst \in \{H1, L1, \dots\}} P\left(\rho_{\rm inst}, \xi^2_{\rm inst} \mid \bar{\theta}, \rm n\right). \tag{4.9}$$

The (SNR, $\xi^2$) distributions for the noise hypothesis, $P\left(\rho_{\rm inst}, \xi^2_{\rm inst} \mid \bar{\theta}, \rm n\right)$, are obtained from histograms of non-coincident triggers, which are then normalized and smoothed by a Gaussian smoothing kernel. The time dependence of these PDFs is not tracked[3]. Examples of these PDFs used to assign the likelihood-ratio statistic can be found in Figs. 6.22, 6.23, and 6.24.

### Trigger Clustering

Signals can produce triggers associated with several templates at the same time. While assigning significances we wish to ensure that we only consider the most likely candidate associated with a given signal. To do so, in the offline pipeline, we cluster the candidate events by keeping the set of triggers with the highest likelihood-ratio across the entire template bank within a ±4 s window. The online pipeline does not cluster events across different $\bar{\theta}$ bins in order to save time. The online pipeline, however, does cluster triggers by keeping the trigger-set with the highest likelihood-ratio in each $\bar{\theta}$ bin within a ±1 s window.

### False-alarm probability and false-alarm rate

Now that we have obtained a clustered candidate-event list with likelihood-ratios assigned, we would like to translate these numbers into physical significances: a false-alarm probability and a false-alarm rate (FAR). False-alarm probability is the probability that noise would produce an event with a ranking statistic greater or equal to the ranking statistic of the event under consideration. Conventionally the false-alarm probability (FAP or, the p-value) has been measured by performing time-slides, where a set of time-shifts that are much larger than the gravitational-wave travel time between the sites are performed in the data, and the result is used to construct a background. But the bookkeeping of time slides is not desirable for a low-latency search [105, 106]. In order to compute the FAP, we first need to construct the likelihood-ratio PDF for the noise model, $P(\mathcal{L} \mid n)$. The likelihood distribution

[3]The dependence of these PDFs on time was added in O2. Refer to Ch. 7.

for the noise hypothesis given a $\bar{\theta}$ bin is

$$P(\mathcal{L} \mid \bar{\theta}, \mathrm{n}) = \int_{\Sigma(\mathcal{L})} P(\cdots \mid \bar{\theta}, \mathrm{n}) \mathrm{d}^{n-1}\Sigma, \tag{4.10}$$

obtained by integrating the noise PDF for all $n$ parameters in the ranking statistic over $n-1$ dimensional surfaces $\Sigma$, of constant $\mathcal{L}$. However, we do not know the $\Sigma$ surfaces, therefore we construct an approximation of $P(\mathcal{L} \mid \bar{\theta}, \mathrm{n})$ by sampling points in the n-dimensional parameter space and computing $\mathcal{L}$ and $P(\cdots \mid \bar{\theta}, \mathrm{n})$ at each point, constructing a histogram of $\mathcal{L}$ with each sample weighted by $P(\cdots \mid \bar{\theta}, \mathrm{n})$ [98, 103].

Next, we marginalize over the template bins to get

$$P(\mathcal{L} \mid \mathrm{n}) = \int P(\mathcal{L} \mid \bar{\theta}, \mathrm{n}) \mathrm{d}\bar{\theta} \tag{4.11}$$

The FAP is then given by

$$P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n}) = \int_{\mathcal{L}^*}^{\infty} P(\mathcal{L} \mid \mathrm{n}) \mathrm{d}\mathcal{L}. \tag{4.12}$$

In case of M observed events, this can be written as the complement of the binomial distribution,

$$\begin{aligned} P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n}_1, \ldots, \mathrm{n}_\mathrm{M}) &= 1 - \binom{M}{0} P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n})^0 [1 - P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n})]^M, \\ &= 1 - (1 - P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n}))^M. \end{aligned} \tag{4.13}$$

In order to calculate the FAR, we assume that the noise events follow Poisson statistics [103]. For a Poisson process with a mean $\lambda$, the probability of observing $N$ or more events is

$$P(N \mid \lambda) = 1 - e^{-\lambda} \sum_{j=0}^{N-1} \frac{\lambda^j}{j!}. \tag{4.14}$$

Setting $N = 1$ in Eq. 4.14 and combining with Eq. 4.13, we obtain the mean number of noise events with likelihood ratio $\mathcal{L} \geq \mathcal{L}^*$,

$$\lambda(\mathcal{L} \geq \mathcal{L}^*) = -M \ln(1 - P(\mathcal{L} \geq \mathcal{L}^* \mid \mathrm{n})). \tag{4.15}$$

The quantity FAR is then given by

$$\mathrm{FAR} = \frac{\lambda}{T}, \tag{4.16}$$

where $T$ is the time of the experiment, and the inverse of false alarm rate which tells us the waiting time before we see such an event from noise is given by

$$\text{IFAR} = \frac{1}{\text{FAR}} \tag{4.17}$$

### 4.2.6 Estimating the sensitivity of the pipeline

Simulated gravitational wave signals known as “injections" are added in a duplicated strain data stream in order to ascertain how our pipeline would fare against signals. Injections are considered “found” if a coincident event is found above a threshold significance in a certain time window around the injection. This threshold significance requirement is set to $IFAR \leq 30$ d. The sensitive volume space of the pipeline $V_{\text{sens}}$ is computed as

$$V_{\text{sens}} = 4\pi \int_0^\infty \text{dr}\epsilon(r) r^2. \tag{4.18}$$

Here $\epsilon(r)$ is the efficiency parameter given by the fraction of found injections. Range, which is the average furthest distance a signal can be detected from, is defined by the radius corresponding to the sensitive volume,

$$R = \left(\frac{3V_{\text{sens}}}{4\pi}\right)^{1/3}. \tag{4.19}$$

In Chapters 5, 6, and 7, we describe searches carried out using these methods. We also explain some figures of merit in Chapter 5 which are useful in understanding how well the pipeline is performing. For a detailed description of these methods we refer the reader to [3, 4, 50, 98].

*Chapter 5*

# BBH GSTLAL

Before the start of the first observing run of Advanced LIGO, the LIGO-CBC group, as a legacy from the Initial LIGO's science runs was divided into 3 main search groups, targeting the three most anticipated source types: neutron star-neutron star binaries (BNS), neutron star-black hole binaries (NSBH), and black hole-black hole binaries (BBH). There were three search pipelines at the time: GstLAL (running both in low-latency and in an offline configuration), pycbc offline, and MBTA online. This chapter documents the developments in the BBH GstLAL pipeline that were made before the start of O1.

## 5.1 Motivation to develop a low-latency BBH search using GstLAL

There were already plans to run the PyCBC search (Advanced LIGO version of the ahope pipeline [83, 84] used to detect CBC signals in Initial LIGO). At that time, however PyCBC could only run in high latency[1]. We wanted to develop a low-latency BBH search using GstLAL. The main motivation for low-latency searches is to send out prompt alerts to the broader astronomy community in order to provide an opportunity of multi-messenger observations [50]. Even if we don't expect any electromagnetic counterparts to BBH mergers, it would be nice to confirm that (or even better, contradict by detecting light from a BBH merger!). Furthermore a low-latency matched-filtering search would prove useful to provide early feedback to the Detector Characterization group about short-duration instrumental glitches which will potentially trigger the short duration BBH templates. In case of GW150914, its detection by the low-latency burst searches helped alert the detector commissioners to freeze the commissioning activities. This allowed the searches to collect sufficient data for background estimation before the configuration of the detectors was changed. Since then, low-latency matched filtering searches have been successfully deployed and have been analyzing data in real time using template banks covering a broad parameter space that includes BNS, NSBH, and BBH regions. There have been no highly significant multimessenger observations of BBH signals. GW170817, the first BNS merger observed by Advanced LIGO, was observed in electromagnetic over several bands.

[1] Now the PyCBC pipeline can also detect CBC signals in real time.

The GstLAL pipeline and the PyCBC pipeline use different methods of filtering and background and significance estimation. Therefore, running both could provide a means of cross-checking the analyses results for clear loud events. In fact, studies have shown that marginal events from the two pipelines can be very different, so running both would mean that we could detect more marginal events. A study was conducted with a common template bank and common sets of injections (simulated signals) to compare the performance of the multiple pipelines. We briefly discuss those results at the end of this chapter.

#### Parameter space considerations

Now that we have decided to develop a BBH search, it is necessary to also define the search parameter space. The ringdown frequency of the final black hole is inversely proportional to the total mass, and is approximately given by

$$f_{\text{ringdown}} = (1.5251 - 1.1568(1-j)^{0.1292})\frac{c^3}{(2.0\pi GM)}. \tag{5.1}$$

Here $M$ is the mass and $j$ is the reduced Kerr spin parameter of the final black hole. For neutron star-neutron star binary, the ringdown frequency is typically higher than 1 kHz and therefore outside of the LIGO band. For a non-spinning black hole, if the mass is greater than $\sim 12M_{\odot}$, the ringdown frequency enters the LIGO band. Therefore, for such systems a significant fraction of the SNR comes from the merger and ringdown.

Such a search should use IMR templates (Sec. 3.1.2 ), ideally including at least aligned spin. Studies have shown increase in sensitive volume of the search on including the effect of spin in the tempplates [95].

It should cover a parameter range extending from where the low-latency low-mass, inspiral-only search(es) end (at the time envisioned as masses of 3-3 $M_{\odot}$) up to as high as is sensible ($100M_{\odot}$ or more), and extend down to frequencies as low as 10-20 Hz.

### 5.2 Mock data challenge

To tune and test our pipelines for O1, and compare the pipelines among each other, a mock data set was created by recoloring two calendar months of Initial LIGO's sixth science run (S6) data to the anticipated sensitivity of the Early Advanced LIGO detectors [5]. See Fig. 5.1.

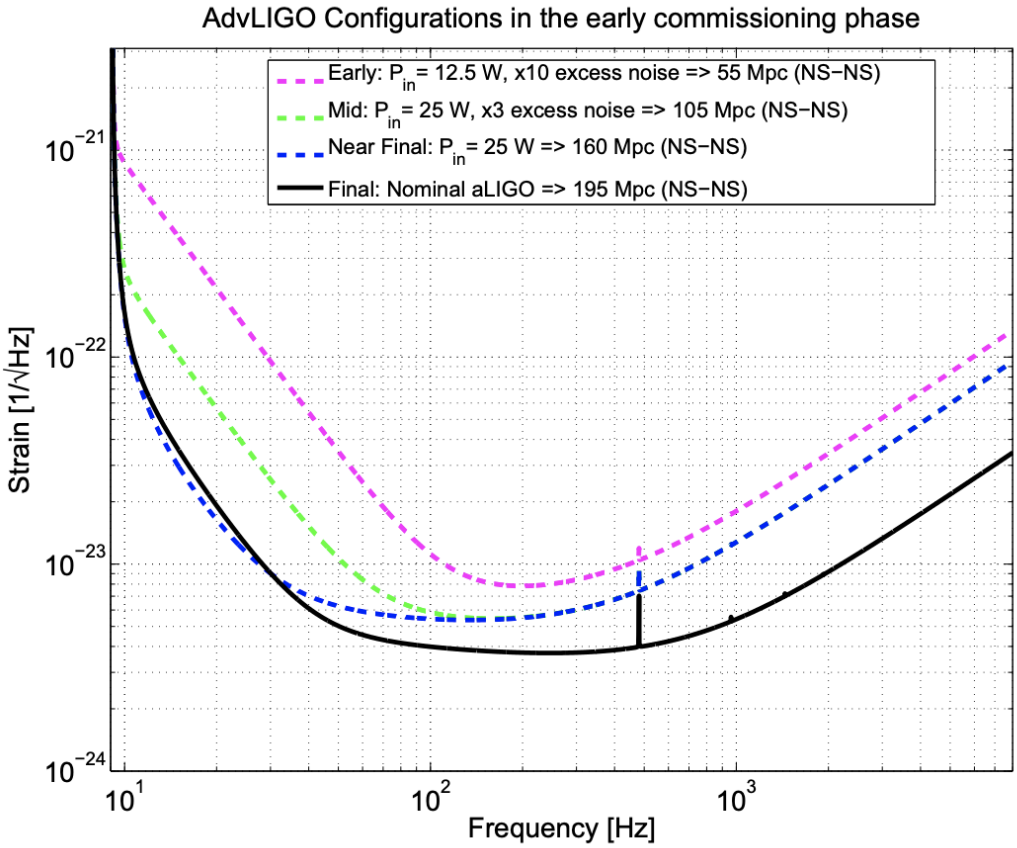



Figure 5.1: Plausible sensitivity curves in the early phase of the aLIGO commissioning. The pink curve was used to recolor the S6 data used in this study. Figure taken from [5].

We started testing the `GstLAL` BBH pipeline by creating a template bank (Fig. 5.2 of full inspiral-merger-ringown waveforms that went from $3M_{\odot}$-$3M_{\odot}$ (component masses) to $110M_{\odot}$ (total mass). The spins were restricted between [-0.9, 0.9]. We considered a starting frequency of 30Hz. The bank was created in three stages using the program `lalapps_cbc_sbank` that lives in the LALSuite [97]: the "low-mass bank", which went up to $9.8M_{\odot}$ in total mass was created using the lattice placement method by using the `TaylorF2RedSpin` metric. Note that, even though we used the `TaylorF2RedSpin` metric to create the template bank, while filtering we used the `IMRPhenomB` waveforms. Since low-mass templates are longer in the relevant frequency band (30 Hz onwards), it is more computationally expensive to perform matches in the low-mass region, and to make use of the stochastic placement technique. Since the merger and ringdown of the low-mass waveforms fall outside the frequency band, it is sufficient to use inspiral-only models in the region only for

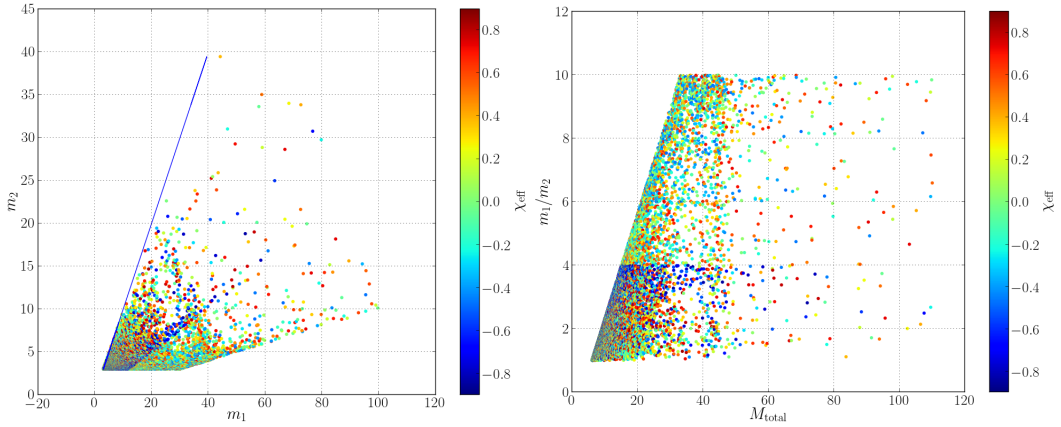



Figure 5.2: Visual illustration of the template bank used for developing the BBH GstLAL search and to analyze the ER6 data.

placement purposes.

The "intermediate-mass bank", which went from $9.8\mathrm{M}_\odot$ to $44\mathrm{M}_\odot$ in total mass was constructed using the stochastic placement method in multiple parallel processes using a Condor DAG [107]. DAG, which stands for Directed Acyclic Graph, is a job control system. The intermediate waveforms are long enough in the band that the generation of the bank requires chopping up the parameter space in different $\mathcal{M}_c$ slices and placing templates in parallel in these slices. We need to be careful while dividing the parameter space into several bank placement jobs while using the stochastic placement technique because it leads to overcoverage in the boundaries.

The density of the template banks decrease with the increase in the masses of the binaries. The waveform of a system with smaller masses is longer in the frequency band Advanced LIGO and Virgo detectors are the most sensitive in. Therefor even a small change in masses can lead to a big mismatch in this region since there are a lot of cycles in band for which the match is being performed. Very few templates are required in the high mass region, and therefore they can be placed easily without any parallelization of jobs. Finally, the "high-mass" bank, which went from $44\mathrm{M}_\odot$ to $110\mathrm{M}_\odot$ in total mass, was constructed using the stochastic method in one single job.

### 5.2.1 Challenges - new waveforms

Just before the start of the first Observing run (O1) of the Advanced LIGO detectors, a new waveform model, `SEOBNRv2` [76], became available. It allowed for the waveforms to be generated with component spins, $\chi_1, \chi_2 \in [-1, 0.99]$, and mass ratios, $\frac{m_1}{m_2} \in [1, 100]$. Even though there are not many waveforms from numerical

relativity to tune to at high mass ratios, the EOB formalism [68, 69] is designed to work well for high mass-ratio systems. The only drawback was that the waveforms are slow to generate and thus difficult to use for filtering purposes. Soon after a frequency domain reduced order model based on the `SEOBNRv2` waveform model, `SEOBNRv2_ROM_DoubleSpin` [78] was developed. This model was shown to be thousands of times faster for typical system masses with a faithfulness mismatch of better than 0.1% than the original model [78]. This speedup in waveform generation allowed us to use make use of this waveform model for filtering, and therefore allowed us to increase the parameter space of our search, previously restricted to $\chi_1, \chi_2 \in [-0.9, 0.9]$ and mass ratios, $\frac{m_1}{m_2} \in [1, 10]$, to $\chi_1, \chi_2 \in [-1, 0.99]$, and $\frac{m_1}{m_2} \in [1, 100]$.

Subsequently, a new template bank was created to cover the extended BBH parameter space. It is shown in Fig. 5.3. We started testing the `GstLAL` BBH pipeline with this new template bank.

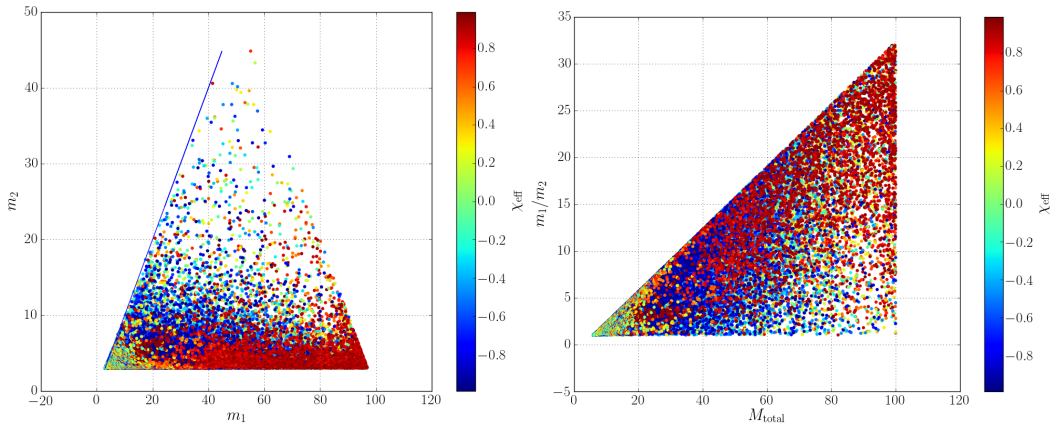



Figure 5.3: Visual illustration of the template bank created with increased BBH parameter space, no mass ratio cut, component spins upto 0.99.

It was found that the pipeline didn't find any of the software injections while filtering with the new bank of waveforms. Software injections are simulated gravitational waveforms that are used to assess the pipeline response to real gravitational-wave signals. Furthermore, it was found that the problem remained if we replaced the waveform model in the template bank from `SEOBNRv2_ROM_DoubleSpin` to that of the previously tested `IMRPhenomB`. So it was concluded that the problem was not with the way the pipeline handled the new waveform, but with the way the pipeline handled the extended parameter space. Recall that, in order to be computationally efficient, the `GstLAL` pipeline uses SVD'ed, multi-banded filters (Sec. 4.2.2). On

visualizing the reconstructed waveforms from the SVD filter banks, the culprit for the missing of injections was found. The reconstructed filter waveforms were clearly suffering from under sampling, Fig. 5.4.

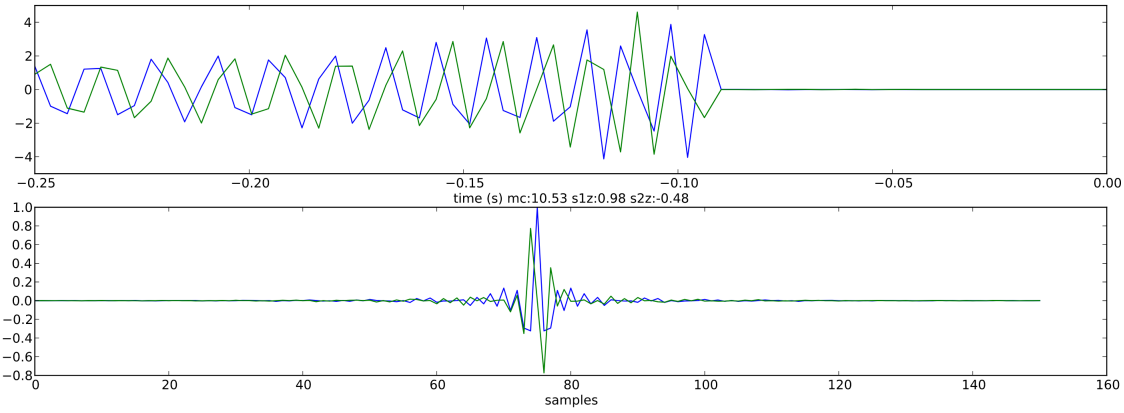



Figure 5.4: Reconstructed SEOBNRv2ROM_DoubleSpin waveform - undersampling issues. The blue and green waveforms are the reconstructed $h_+$ and $h_\times$ waveforms (top). The bottom figure shows the autocorrelation time series of the reconstructed plus and cross waveforms.

### 5.2.2 Issue with the time slicing

In the `GstLAL` pipeline, each split-bank (Sec. 4.2.2) is divided into various time slices before performing the SVD, such that all the templates in a split-bank have the same number of samples in each time slice. This is done by calculating the number of samples in each time slice starting from the highest sampling frequency. The number of samples in a particular time slice depend on the waveform that is the longest in the frequency band corresponding to $f_{n-1}/(2 \times \text{padding})$ and $(f_n \times \text{padding})/2$, where $f_n/2$ is the sampling frequency of the current time slice, and $f_{n-1}/2$ is the next highest sampling frequency coressponding to the next time slice (Fig. 5.5). "Padding" is used to ensure that that the waveform doesn't cross the Nyquist rate in the next time slice.

The undersampling issues we saw earlier were happening because sampling rates were being skipped over. As we can see in the snippet above, an exception is made when `imrchirptime` raises an error. In the exception, however, the code simply skips over a sampling frequency and continues. The idea behind that was to try a smaller starting frequency so that `imrchirptime` returns a non-negative inspiral time, but this had the unintended side effect of also reducing the current sample rate.

```
./gstlal/gstlal-inspiral/python/templates.py

300 try:
301   longest_chirp =
max(spawaveform.imrchirptime(row.mass1,row.mass2,this_flow,spawaveform.c
omputechi(row.mass1, row.mass2, row.spin1z, row.spin2z)) for row in
sngl_inspiral_rows)
302 except ValueError as e:
303   print "Continuing decomposition by moving to next slice
despite:\n\t", e
304   continue
```

Figure 5.5: Number of samples in a time slice is given by the duration of the waveform that spends the longest amount of time in the frequency band $f_{n-1}/(2 \times \text{padding})$ and $(f_n \times \text{padding})/2$.

```
./lalsuite/pylal/pylal/spawaveform.py

85 if inspiral_time < 0:
86   raise ValueError("Inspiral time is negative: m1 = %e, m2 = %e, flow
= %e, chi = %e" % (m1, m2, fLower, chi)) # demand positive inspiral
times
```

Figure 5.6: `imrchirptime` raises an error whenever the inspiral time returned is negative.

`imrchirptime` computes the inspiral time of a waveform in a given frequency band according to 3.5 order in the Post Newtonian expansion. But the PN expansion breaks down for high mass-ratios. Therefore `imrchirptime` returns nonsensical results, i.e., negative times in a frequency band. Fig. 5.7 shows the points in the new template bank that returned negative times for a frequency band and therefore had some sample rates skipped over.

To fix the undersampling issue, we replaced the call to the imrchirptime function, with a function that overestimates the IMR time. The inspiral time is calculated using the 2 PN formula, but dropping all the negative terms. The merger time is overestimated by assuming the merger and the plunge takes one orbit at the maximum ISCO radius. The maximum ISCO is for orbits about a Kerr black hole with maximum dimensionless aligned spin and equals $9GM/c^2$ (Sec. 3.1.1). This corresponds to an orbital speed of $\approx c/3$. The ringdown time is overestimated by overestimating the final spin of the black hole, and calculating time for 11 e folds.

Using the new function solved the undersampling issue that we were encountering in

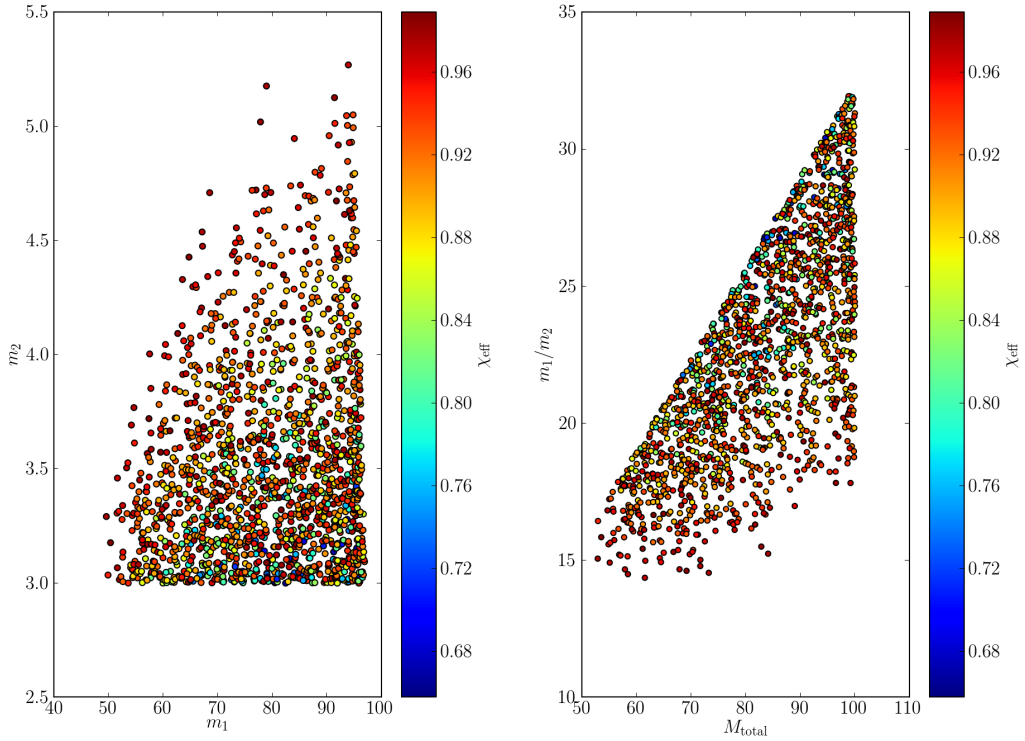



Figure 5.7: Points in the new template bank (Fig. 5.3) that `imrchirptime` returns negative times for.

high mass-ratio part of the parameter space (see Fig. 5.8). But we ended up with an inspiral-only waveform. The merger and ringdown was being cut-off.

After investigations, it was found that the merger and the ringdown of the waveform were being wrapped to the front of the buffer by the new waveform approximant, `SEOBNRv2_ROM_DoubleSpin`. We needed to fix that, since we need to sample the merger and the ringdown of the waveform with the highest sampling frequency. We also wanted to align the IMR waveforms before performing the SVD so as to get maximum compression. We use the epoch information returned by the waveform generator to find the merger time, and align them at merger.

### 5.2.3 Epoch returned by the `LALSimulation` waveform generator

Epoch information tells us the time by which we need to shift a given waveform to get the merger at the beginning (or equivalently, at the end) of the buffer. We use `SimInspiralFD`, which is a waveform generator in `LALSimulation` that conditions

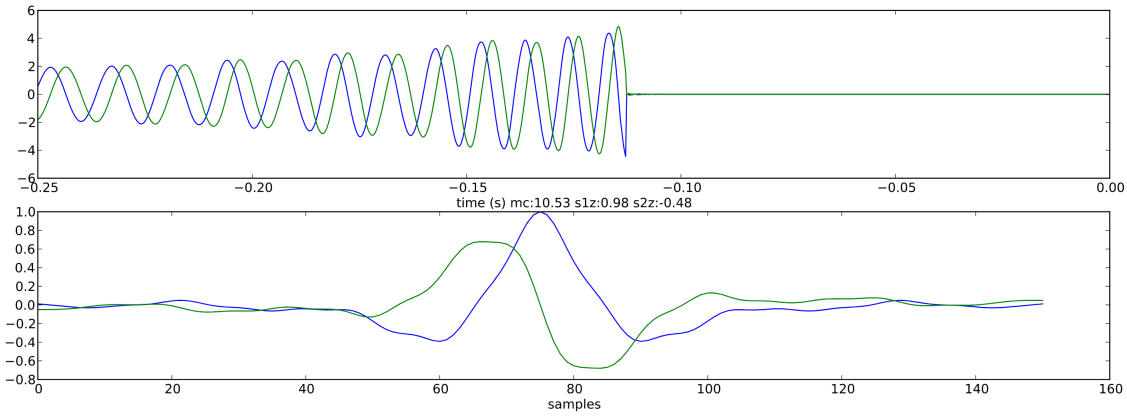



Figure 5.8: Reconstructed SEOBNRv2ROM_DoubleSpin waveform - sampling issue fixed, but the merger and ringdown is being cut off.The blue and green waveforms are the reconstructed $h_+$ and $h_\times$ waveforms (top). The bottom figure shows the autocorrelation time series of the reconstructed plus and cross waveforms.

waveforms properly (both TD and FD). We find the position of peak amplitude using the epoch returned by `SimInspiralFD`. We then align the peaks according the maximum ringdown time (for a given split bank) from the end of the buffer, before performing the SVD.

This fixed all the issues we were seeing with the use of new waveforms, and going upto extreme mass ratios in the parameter space. Fig. 5.9 shows the final fixed reconstructed `SEOBNRv2_ROM_DoubleSpin` waveform that we looked at before.

### 5.2.4 Results

Now that we have fixed the issues with waveform reconstruction from SVD basis filters, we ran our BBH pipeline on the MDC data. We chose to run on a month's worth of MDC data, which was S6 data recolored to early Advanced LIGO sensitivity.

We will use these results to go over some figures of merit, which will come up in the next chapters.

#### Injection accuracy

Injections are simulated signals that are injected in the data that we use to measure the performance of the pipeline. We examine the pipeline's ability to recover the injected parameters. Recall that the purpose of a search pipeline is to detect signals, and not perform detailed parameter estimation (which is done coherently across the

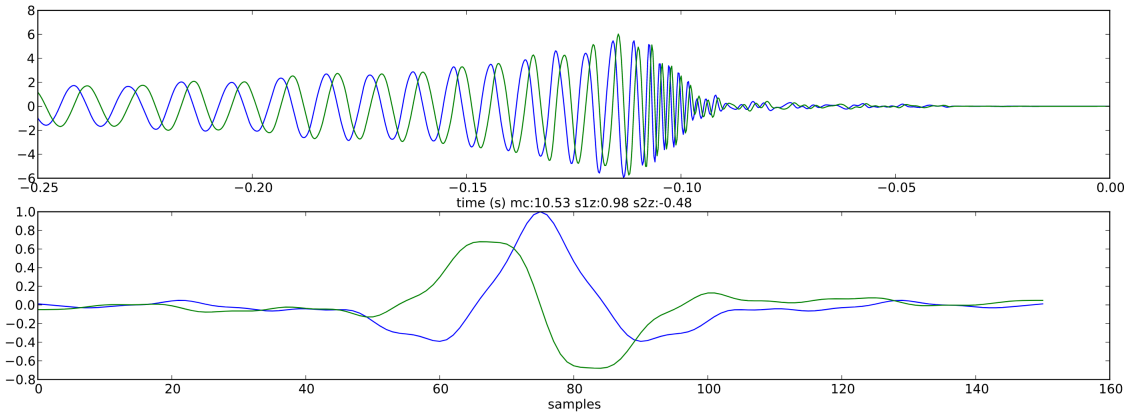



Figure 5.9: Reconstructed SEOBNRv2ROM_DoubleSpin waveform - sampling issue fixed, merger and ringdown not being cut off. The blue and green waveforms are the reconstructed $h_+$ and $h_\times$ waveforms (top). The bottom figure shows the autocorrelation time series of the reconstructed plus and cross waveforms.

detector network using a full multidimensional parameter space of the considered models with a fine (stochastic) sampling). Still, these results are useful to trace potential bugs, and they also provide us with an idea about the recovery we should expect for high mass systems.

**Recovery of chirp mass and effective spin**

Chirp mass parameters for the low mass systems are recovered better than the chirp mass parameters for the high mass systems (Fig. 5.10) most of the SNR comes from the inspiral, the phase evolution for which is governed by the chirp mass. For high mass systems, most of the SNR comes from the merger and ringdown; the merger frequency is proportional to $M_{\rm tot}$. High mass systems also contain merger and ringdown in the band, and those phases are not as well parametrized by the chirp mass. Degeneracy with chirp mass causes some outliers in recovering spins. See Figs. 5.11 and 5.12. For the injections that have their spins aligned with the orbital angular momentum, the recovery of the effective spin parameter is good. For precessing injections (With spins arbitrartily aligned), the recovery of the effective spin parameter becomes worse because we do not model the precession effects in the templates. Chirp mass recovery can be seen in Fig. 5.10 and the effective spin parameter recovery can be seen in Fig. 5.11.

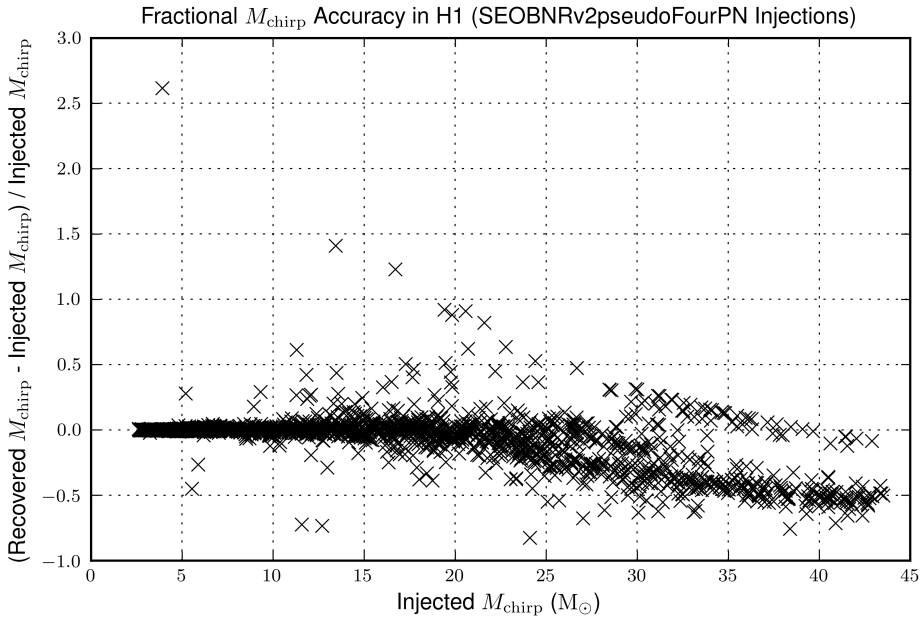



Figure 5.10: $\mathcal{M}_c$ recovery. We can see that the chirp mass recovery of low-mass systems is better than that of high mass as expected.

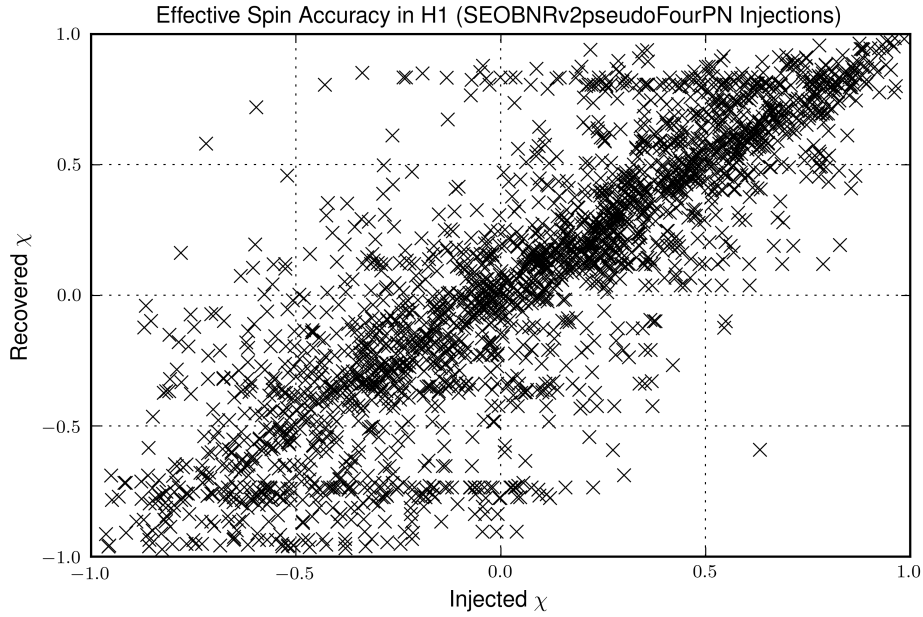



Figure 5.11: $\chi_{\rm eff}$ recovery. Majority of injections show good $\chi_{\rm eff}$ recovery.

### $\xi^2$ vs. SNR distribution for background and injections

SNR is insufficient to distinguish signal from background in the presence of non-Gaussian noise. A good $\xi^2$ should be able to help separate background from injections

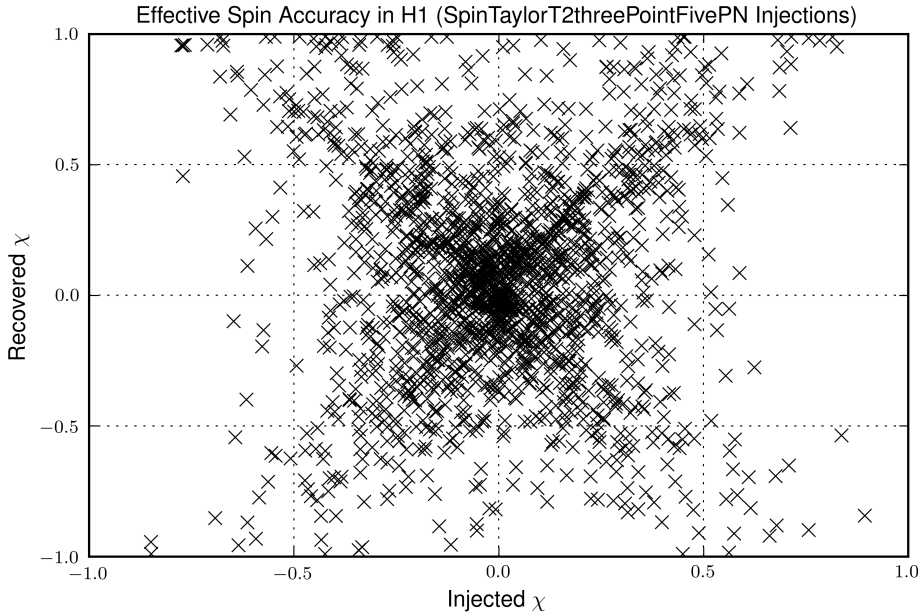



Figure 5.12: $\chi_{\rm eff}$ recovery. Note that these are precessing injections.

in presence of non-Gaussian noise. We just show the figure for one of the detectors for reference. See Fig. 5.13. Notice that if we had just used SNR as a detection statistic, the background points would have completely overlapped with the injections, ruining our ability to find signals whatsoever.

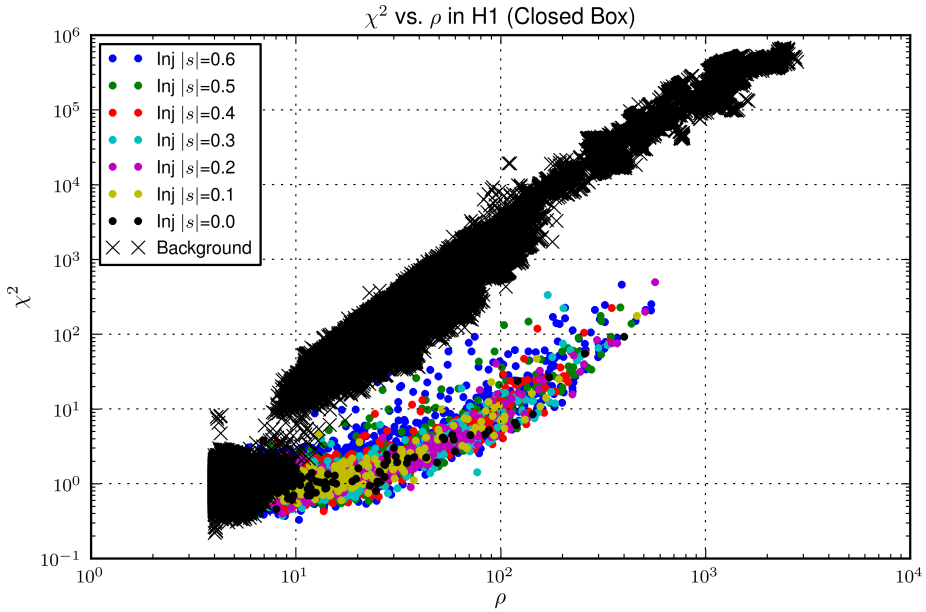



Figure 5.13: $\xi^2$ vs. SNR distribution for background and injections in Hanford. The black crosses are background coincident events (these are coincident events found in time-shifted data between the two detectors), and the colored points are injections, with the colors showing the spin range of the injections. A good $\xi^2$ should be able to help separate background from injections in presence of non-Gaussian noise. We see two distinct background populations, one that centered around a $\xi^2$ value of 1 and low SNRs is the Gaussian population, and the other population with high SNRs are due to glitches. Notice that if we had just used SNR as a detection statistic, the background points would have completely overlapped with the injections, ruining our ability to find signals whatsoever.

### Search Sensitivity

We want to present the results of our searches in terms of the rate of binary mergers per unit volume of space to which we are sensitive (in $Mpc^3$), assuming that this "rate density" is constant in the local universe (although it will change at cosmological distances, when the universe was much younger; see chapter 8). Therefore, we need to know the sensitive volume to which our pipeline can detect signals. Note that this depends on the masses of the system being considered, since heavier systems are intrinsically louder. In the pipeline, it is calculated by injecting signals uniform in volume and uniform in sky location (see Sec. 4.2.6). It is defined as $V_{\text{sens}} = \frac{4}{3}\pi R^3$ where $R$ is the range defined as the average distance to which our detcetors (and the pipeline) can detect signals. It is a function of the FAR. At a higher FAR, we can detect signals further away but the contamination fraction of noise incresaes. Refer Fig. 5.14.

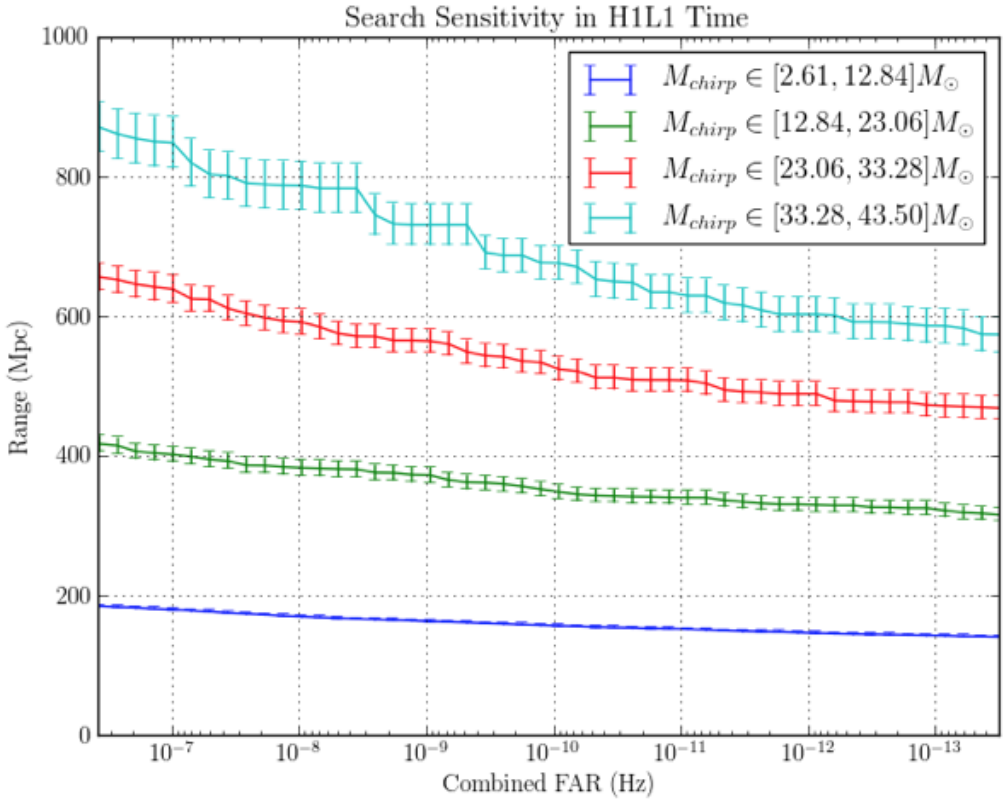



Figure 5.14: Mean sensitive distance as calculated by the pipeline for recovered S6 data recolored to Early Advanced LIGO sensitivity.

**Main results ("Money plots")**

Finally, we want to see results of the run. In order to ensure that the background model is working, we perform exactly one time-slide analysis in `GstLAL`. It also serves for the purpose of performing a blind analysis as discussed in Sec. 6.2.3 below. In Fig. 5.15, we plot IFAR (Eq. 4.17), which has a monotonic mapping with the likelihood ratio statistic on the x-axis, and the cumulative counts (or cumulative expected counts from noise in case of background model) on the y-axis. Note that the expected background model (dashed line) is precisely of the form $dN_{\rm cum}/d\text{IFAR} = 1/(\text{IFAR} * T_{\rm obs})$, with Poisson errors on $N_{\rm cum}$ shown as grey bands. We want the time-shifted result to match with the background model within Poisson errors, because we don't expect to see any signals in the time-shifted analysis. When we look at the un-shifted analysis, if there is a positive excess in the un-shifted result compared to the background model, it could indicate a signal. Notice that this allows us to be able to detect a population of signals at low IFAR values, which would not be significant if taken individually, but collectively could indicate a sub-threshold population.

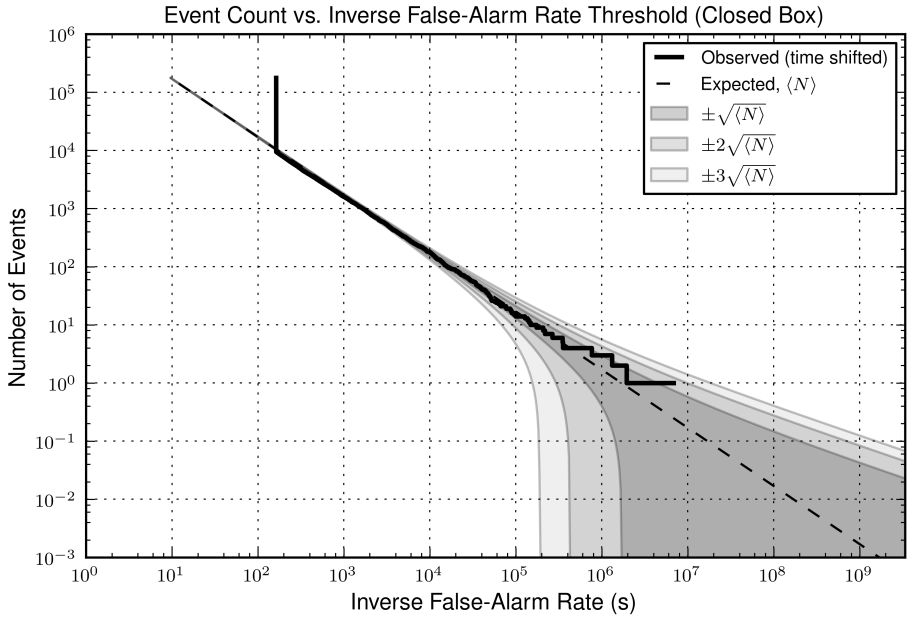



Figure 5.15: Cumulative candidate counts vs. IFAR. The time-shifted result lies on top of the background model as expected.

In the next chapter, we will see how the `GstLAL` BBH pipeline was successfully used in O1.

## 5.3 Appendix: Spin vs Non-spinning BNS Bank MDC

As described in Sec. 5.2, a mock data set was created in advance of O1 to test our pipelines, and evaluate their efficiency as a function of source distance, masses, and component spins. The BNS group wanted to test two banks, one that included aligned-spin waveforms (we will call this the aligned-spin bank) and one that only had non-spinning binary waveforms (we will call this the non-spin bank) against common injection sets: one set that contained only aligned-spin waveforms (we will call this the aligned-spin injection set), and another set that included precessing waveforms with the spins of the binary components isotropically distributed (we will call this the isotropic injection set). While the aligned-spin bank trivially covers a larger parameter space as compared to the non-spin bank, we would like to quantify these numbers to ascertain if the extra volume covered is worth the additional computational costs.

### The banks

1. *The non-spin bank* included inspiral only waveform templates describing GW emission from systems with component masses in the range $1M_\odot$ to $3M_\odot$. It contained $\sim$ 8,000 templates generated using the TaylorF2 waveform model [63–67]. It was constructed using a geometric placement technique (Sec. 3.2.3).

2. *The aligned-spin bank* also included inspiral only waveform templates describing GW emission from systems with component masses in the range $1M_\odot$ to $3M_\odot$. In addition, it allowed the components to have a dimensionless spin in the range $-0.4 \leq \chi_1, \chi_2 \leq 0.4$. It contained $\sim$ 80000 templates generated using the TaylorF2 waveform model [63–67]. It was also constructed using a geometric placement method that allows for spin [93].

### The injections

1. *The aligned-spin injection set* included waveform templates describing GW emission from systems uniformly distributed in component masses in the range $1M_\odot$ to $3M_\odot$, with their spins restricted to be aligned with their orbital angular momentum and uniformly distributed in magnitude from -0.4 to 0.4. The injection set had $\sim$ 37000 injections.

2. *The isotropic-spin injection set* also included systems uniformly distributed in component masses in the range $1M_{\odot}$ to $3M_{\odot}$. The spin magnitudes were also required to be less than 0.4, however their direction was not restricted to be aligned with the orbital angular momentum; they were isotropically distributed. This injection set also contained *sim*37000 injections.

We perform a bank effectualness study with the two banks and the two injection sets that were created by calculating the fitting factors for all the signals in the injection sets against the two banks, and compute the effective sensitive volume covered by the two banks against the two injection sets (Eq. 3.33).

### 5.3.1 Non-spin Bank - aligned-spin injection set

We compute the fitting factor for all the aligned-spinning injections with the non-spin bank. The results are shown in Fig. 5.16 and Fig. 5.17. We see that 66% of the injections have a fitting factor of less than 0.97, but there is good chirp mass parameter agreement between the injections and templates.

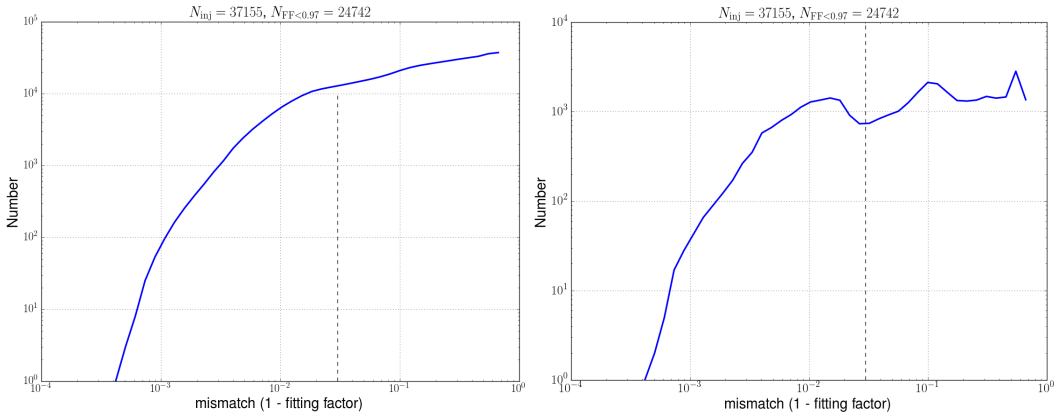



Figure 5.16: Histogram of mismatch calculated for the aligned-spin injections with the non-spin bank (left). The cumulative histogram is shown on the right. 24,742 out of the total 37,155 injections are found to have a fitting factor of less than 0.97.

The effective volume covered by the non-spin bank for the aligned-spin injection set using Eq. 3.33 is found to be $V_{\rm eff} = 0.66 \times V_{\rm sens}$.

### 5.3.2 Non-spin Bank - isotropic spin injection set

We compute the fitting factors for all the injections in the isotropic-spin set with the non-spin bank. The results are shown in Figs. 5.18 and 5.19. We find that $\sim 49\%$ of the injections are found to have a fitting factor less than 0.97. So, the isotropic-spin

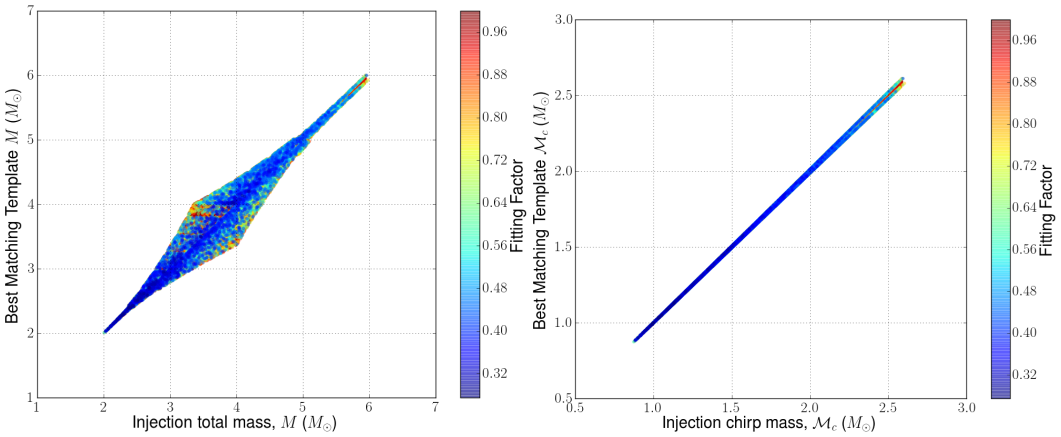



Figure 5.17: Best matching template total mass vs. injected total mass (left). Best matching template chirp mass vs. injected chirp mass (right). We see an excellent recovery of the chirp mass as expected, as it is the leading order term that describes the inspiral part of the waveform. We will see later that such an agreement is not seen in the high-mass parameter space of binaries.

injections are recovered better than aligned spin injections, with the non-spin bank. This is because the effect of spin on a waveform is described by the effective spin parameter to a leading order. The spins in the plane of orbit cause sub-dominant amplitude and phase modulations. Since the isotropic injections have their spin directions uniformly distributed, their effective spin component is on average less than that for the aligned injection set (see Fig. 5.20).

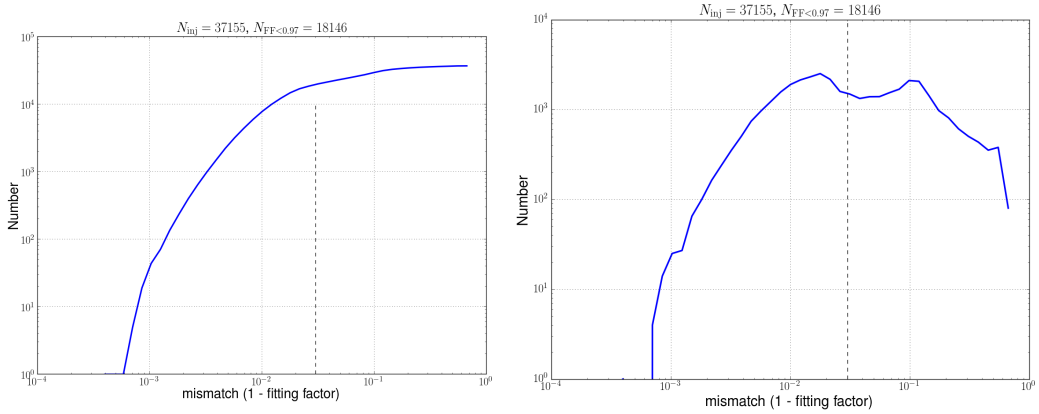



Figure 5.18: Differential (left) and cumulative (right) histogram of mismatch between the isotropic injections and the non-spin bank. 18,146 out of the 37,155 injections are found to have a fitting factor of less than 0.97.

The effective volume covered by the non-spin bank for the isotropic injection set using Eq. 3.33 is found to be $V_{\rm eff} = 0.83 \times V_{\rm sens}$.

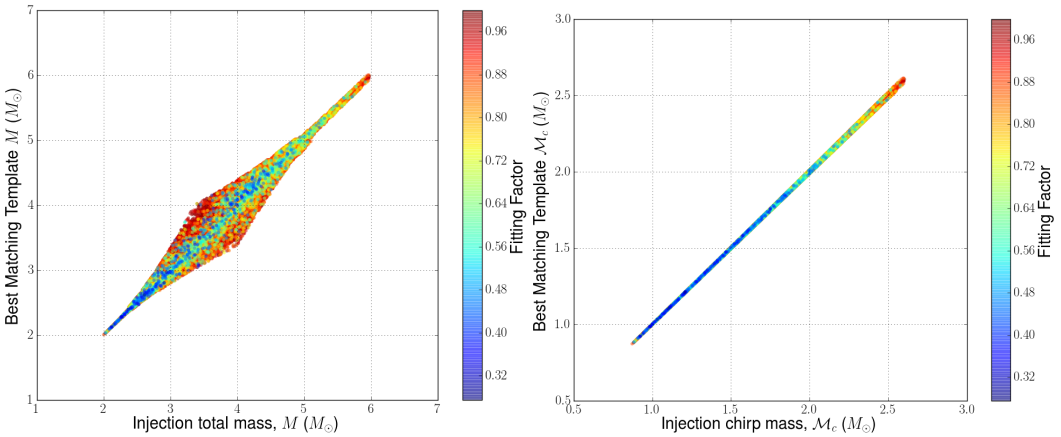



Figure 5.19: Best matching template total mass vs. injected total mass (left). Best matching template chirp mass vs. injected chirp mass (right). As before, we see an excellent recovery of the chirp mass.

### 5.3.3 Aligned-spin Bank - aligned-spin injection set

We see that almost all the aligned-spin injections have a fitting factor greater than 0.97 with the aligned-spin bank. The results are shown in Figs. 5.21 and 5.22. This is expected if the bank was constructed correctly.

The effective volume covered by the aligned-spin bank for the aligned injection set is found to be $V_{\text{eff}} = 0.97 \times V_{\text{sens}}$.

### 5.3.4 Aligned-spin Bank - isotropic injection set

We see that the aligned-spin bank covers the isotropic spin space relatively well, and more than 90% of the injections have a fitting factor greater than 0.97. Refer to Figs. 5.23 and 5.24.

The effective volume covered by the aligned-spin bank for the isotropic injection set is found to be $V_{\text{eff}} = 0.95 \times V_{\text{sens}}$.

In conclusion we find that the aligned-spin bank covers the aligned-spin and isotropic spin space considerably better than the non-spin bank. Given the considerably extra computational cost (and much higher trials factor) associated with the use of a full isotropic-spin template bank, which would have an order of magnitude more templates, we choose to make use of the aligned-spin bank for O1.

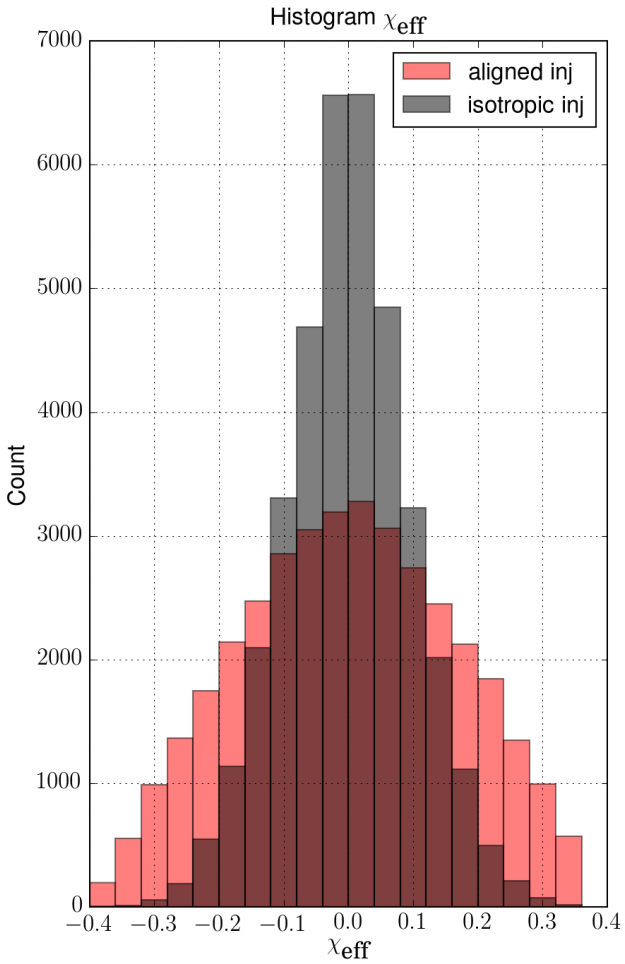



Figure 5.20: Effective spin distributions for aligned and isotropic injections. We see that the aligned-spin injections have a higher effective spin parameter on average.

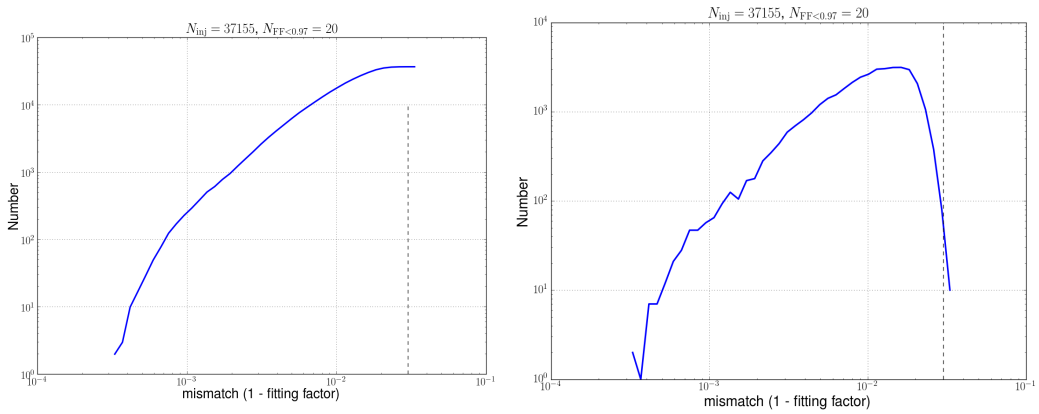



Figure 5.21: Differential (left) and cumulative (right) histogram of mismatch between the aligned-spin injections and the aligned-spin bank. Only 20 out the 37,155 aligned-spin injections have a fitting factor of less than 0.97 with the aligned-spin bank.

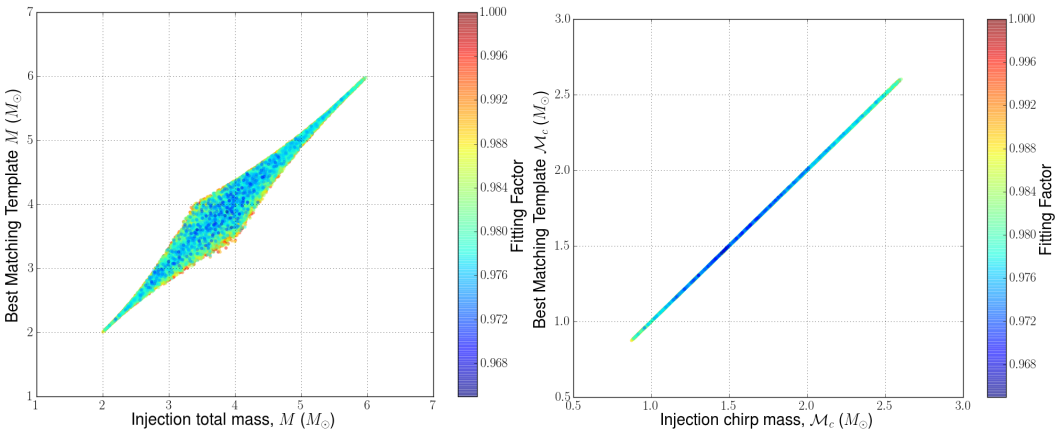



Figure 5.22: Best matching template total mass vs. injected total mass (left). Best matching template chirp mass vs. injected chirp mass (right).

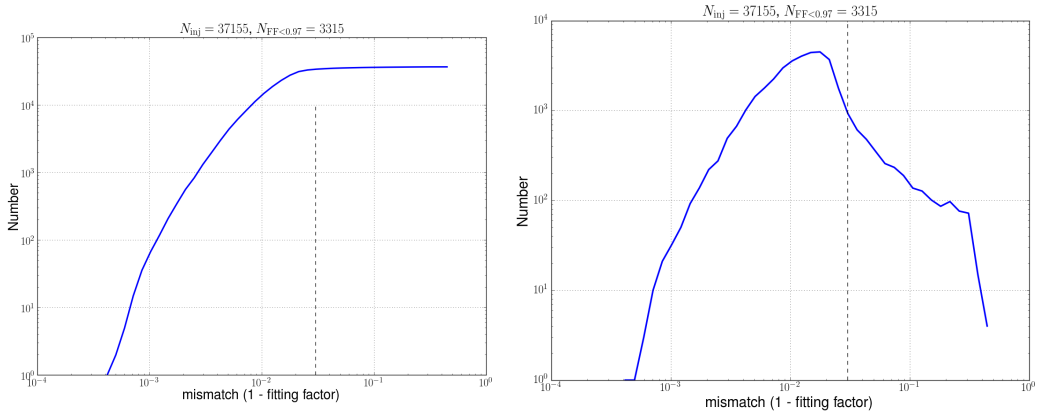



Figure 5.23: Differential (left) and cumulative (right) histogram of mismatch between the isotropic-spin injections and the aligned-spin bank.

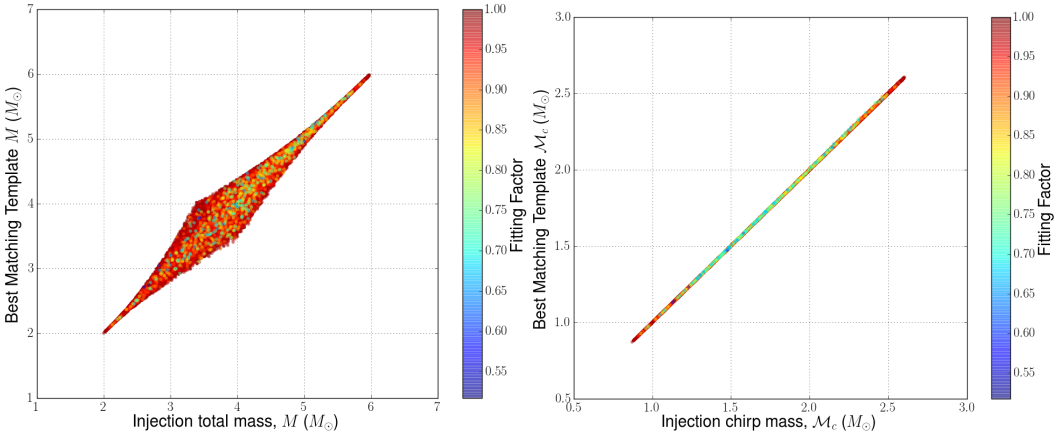



Figure 5.24: Best matching template total mass vs. injected total mass (left). Best matching template chirp mass vs. injected chirp mass (right).

*Chapter 6*

# OBSERVING RUN 1

*"Ladies and gentlemen. We have detected gravitational waves. We did it!"*

*- David Reitze*

We could not have asked for a better start to our first observing run. The first observing run (O1) of the Advanced LIGO detectors started on September 12, 2015, and continued until January 19, 2016 (see Fig 6.1). On September 14, 2015 at

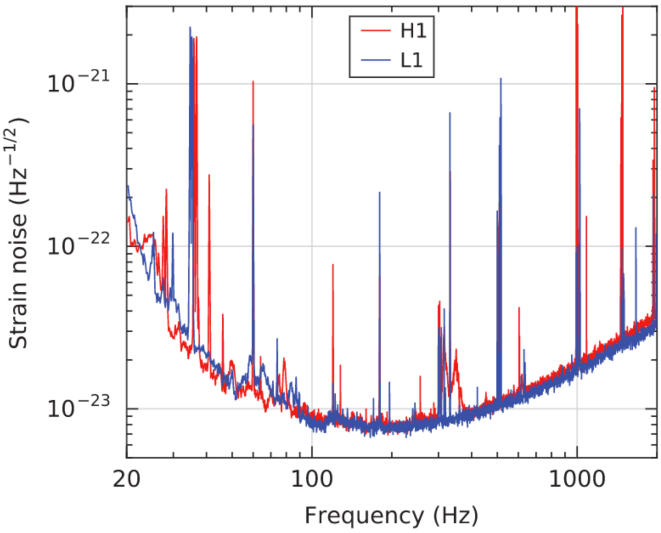



Figure 6.1: Reference strain sensitivity of the two Advanced LIGO detectors during O1, which lasted from September 12, 2015 to January 19, 2016. Not yet operating at their design sensitivity, these detectors were 3-4 times more sensitive than the initial detectors in the frequency band 100 to 300 Hz. Figure taken from [6].

09:50:45 UTC, four days before the start of O1, during the Engineering Run 8 (ER8), the two LIGO detectors simultaneously observed a transient gravitational-wave signal [6], refer to Fig. 6.2. But by then, the detectors were stable and were being left undisturbed to observe the sky. The signal, GW150914, which sweeps upwards in frequency from 35 to 250 Hz with a peak gravitational-wave strain of $1.0 \times 10^{-21}$ matches that of the waveform predicted by general relativity for the inspiral and merger of a pair of black holes and the ringdown of the resulting merged

black hole. This detection was a result of the hard work of thousands of people spread across multiple experiments and decades. The initial detection of the event was made by low-latency searches meant for generic gravitational-wave transients, coherent Waveburst (`cWB`) and omicron-LALInference-Bursts (`oLIB`) [43]. Besides GW150914, the matched filtering pipelines identified a possible (LVT151012), and another confirmed (GW151226) gravitational-wave signal from binary black hole mergers [9, 18, 108] during O1. The `GstLAL` pipeline analyzed the O1 data both in an online and an offline mode[1]. GW151226 was detected by the `GstLAL` pipeline within 70s [9] of its arrival at the detectors, and became the first event to be identified by a matched-filtering search in low latency. While GW150914 was loud enough, and localized in time enough to be clearly visible in a time-frequency representation [109] of the strain data [6] (Fig. 6.2) and to be detected by a generic gravitational-wave transient search, GW151226 required the accumulated SNR from matched filtering for the signal to become apparent in both detectors [9] (Fig. 6.3). In this chapter, we report the results from the `GstLAL` analysis of the O1 data. Sec. 6.1 briefly describes the Advanced LIGO instruments, Sec. 6.2 describes the details of the O1 `GstLAL` search, Sec. 6.3 discusses some of the challenges we encountered during O1, and finally in Sec. 6.5 we describe the results from the `GstLAL` search.

[1] In the beginning of O1, the `GstLAL` pipeline was only looking for low-mass binary mergers (BNSs and NSBHs) online. It wasn't until $22^{nd}$ December, 2015, that the pipeline started uploading events from the full uberbank search (Sec. 6.2, Fig. 6.5) to the GraceDB (Gravitational Wave Candidate Event Database [101], just 4 days before the detection of GW151226.

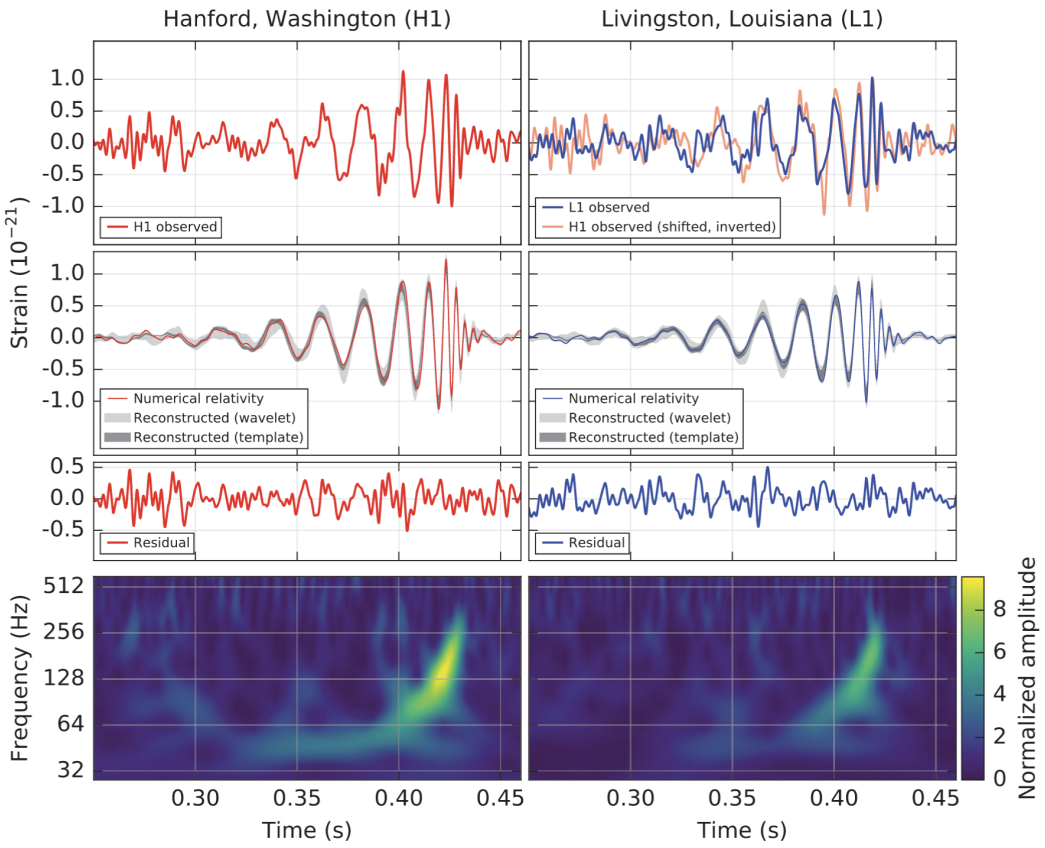



Figure 6.2: The top row shows the observed strain for GW150914 after band-passing and notching out instrumental spectral lines in each detector. Note that in the top right plot, the Hanford data has been time shifted and inverted (to account for the detectors' relative positions and orientations). The second row shows the gravitational-wave strain signal projected onto the detectors according to a numerical relativity waveform (red) [7] for a system consistent with GW150914 parameters. The dark and the light gray shaded areas show 90% credible regions for binary black hole template waveform reconstruction and a reconstruction that uses a linear combination of sine-Gaussian wavelets respectively. The third row shows the residuals after subtracting the filtered numerical relativity waveform from the filtered detector time series. The bottom row shows a time-frequency representation of the strain data in which the signal is clearly visible as a chirp with its frequency increasing over time. Figure taken from [6].

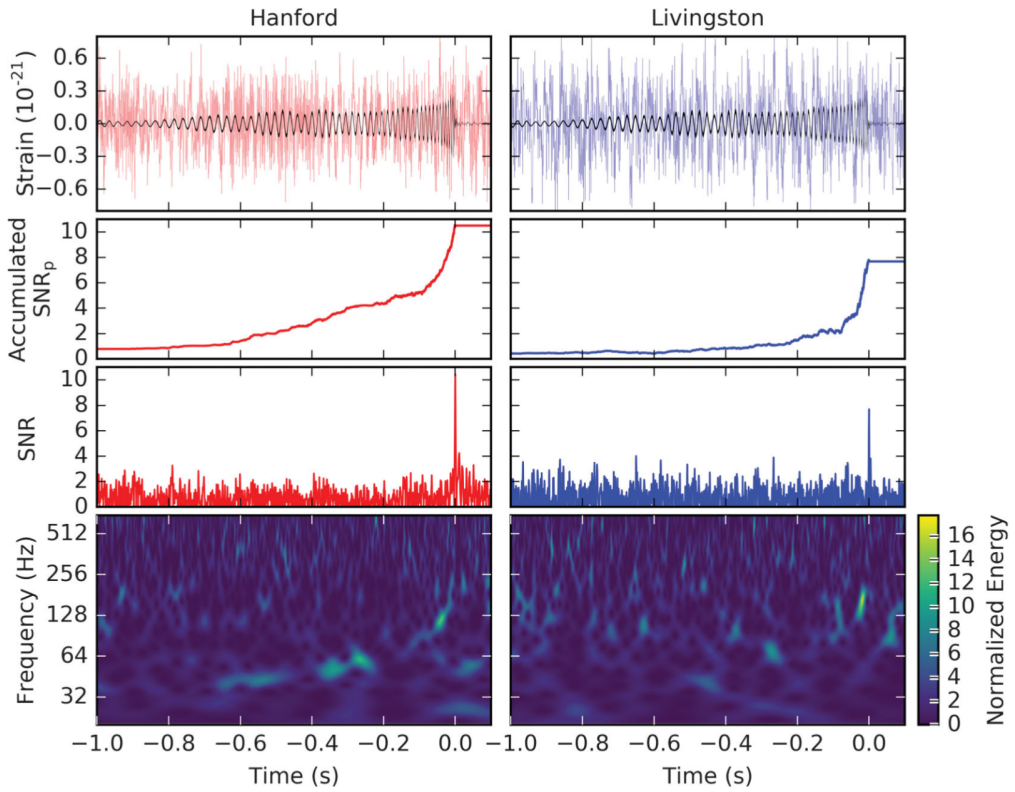



Figure 6.3: The top row shows the best matching template that identified GW151226 buried in the detector noise. The modulations we see in the waveform are due to the conditioning of the data and the template with a 30-600 Hz bandpass filter, which was done to suppress the fluctuations outside this range and to remove strong instrumental spectral lines [8]. The second row shows how the SNR is accumulated over time by matched filtering and the third row shows the SNR time series with a clear peak that happens when the coalescence time of the template matches the signal. The fourth row shows the time-frequency representation of the signal. We can see that in contrast to GW150914 (Fig. 6.2), this signal is not easily visible, and needs matched-filtering for its detection. Figure taken from [9].

## 6.1 The Advanced LIGO detectors

The Advanced LIGO detectors are modified Michelson interferometers. Each arm contains a resonant optical cavity, formed by its two test mass mirrors, that multiplies the effect of a gravitational wave on the light phase by a factor of $\sim$ 300 [110]. A partially transmissive power-recycling mirror at the input provides additional resonant buildup of the laser light in the interferometer as a whole and in O1 20 W of laser input is increased to 700 W incident on the beam splitter, which is further increased to 100 kW circulating in each arm cavity [19]. A partially transmissive signal-recycling mirror is used to extract the gravitational-wave signal by broadening the bandwidth of arm cavities. A schematic diagram of the detector is shown in Fig. 6.4.

A passing gravitational wave alters the arm lengths of the interferometers, such that the measured difference is $\Delta L(t) = \partial L_x(t) - \partial L_y(t) = h(t)L$, where h is the gravitational-wave strain amplitude projected on to the detector (Eq. 3.11). Calibration of the interferometers is performed by inducing test mass motion using photon pressure from a modulated calibration laser [111, 112]. Calibrated strain data is produced in the time domain using both a low-latency, online procedure and a high-latency, offline procedure [113]. The searches use the low-latency calibration to get preliminary results, and the analyses are repeated once we have a better version of calibrated data. The methods to estimate the uncertainty in calibration and the numbers for both O1

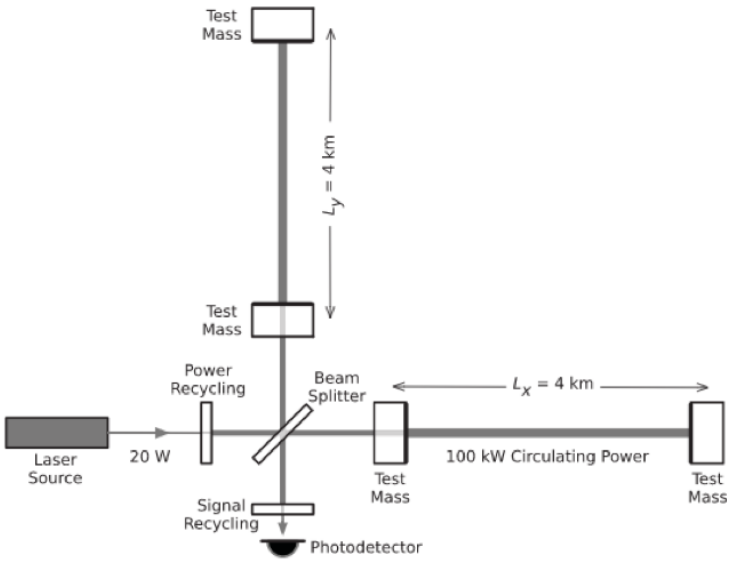



Figure 6.4: Simplified diagram of an Advanced LIGO detector. Figure taken from [6].

and O2 can be found in [114].

## 6.2 Search description

For the first observing run, the two main matched-filtering searches, GstLAL and PyCBC, both used a common template bank to analyze the data. They looked for systems with component masses in the range $1M_\odot$ to $99M_\odot$. The minimum component masses of the templates used in the search were decided based on the lowest expected neutron star mass at the time, which was $1M_\odot$ [115]. Since then, we have also searched for ultracompact binaries with component masses between $0.2M_\odot$ - $1.0M_\odot$ in the O1 data, Sec. 6.5.2. There is no known number for the maximum black hole mass, however the search we report results from here was limited to binaries with total mass less than $M = m_1 + m_2 \leq 100M_\odot$. The LIGO detectors are sensitive to higher mass binaries, however; the results of searches for binaries that lie outside this search space (conducted separately) are briefly described in Sec. 6.5.2. The spins on the component objects were restricted to be aligned to the orbital angular momentum (Sec. 3.2.2). The masses between $1M_\odot$ and $2M_\odot$ were assumed to be neutron stars, and those that were greater than $2M_\odot$ were assumed to be black holes. The magnitude of the dimensionless aligned spin (Eq. 3.3) for the neutron stars was assumed to be between $-0.05 \leq \chi \leq 0.05$. The shortest observed pulsar period in a double neutron star system is 22 ms, corresponding to a spin of 0.02. The fastest-spinning millisecond pulsar is known to have a spin of 0.4. At the current detector sensitivity, limiting the spin magnitudes of the templates to $\chi_{1,2} \leq 0.05$ does not reduce the search sensitivity for sources containing neutron stars with spins up to 0.4. Therefore for the template bank, we limit the spins on the neutron star to be less than 0.05. For black holes this value was allowed to be $0 \leq \chi \leq 0.9895$, since the observations of X-ray binaries indicate that astrophysical black holes may have near extremal spins [116].

### 6.2.1 O1 template bank

As described in Sec. 3.2.3, since the parameters of signals are not known in advance, each detector's output is filtered against a discrete bank of templates that span the search target space. Fig. 6.5 shows the template bank that was used by the matched-filtering pipelines including the GstLAL pipeline in O1. The placement of templates depends on the shape of the power spectrum of the detector noise, Sec. 3.2.3 and Fig. 6.1. We see in Fig. 6.1 that the noise spectrum rises sharply at lower frequencies; therefore, we used a low-frequency cutoff of 30 Hz for this search. The average noise power spectral density of the LIGO detectors was measured over the period September 12, 2015 to September 26, 2015, and the harmonic mean of

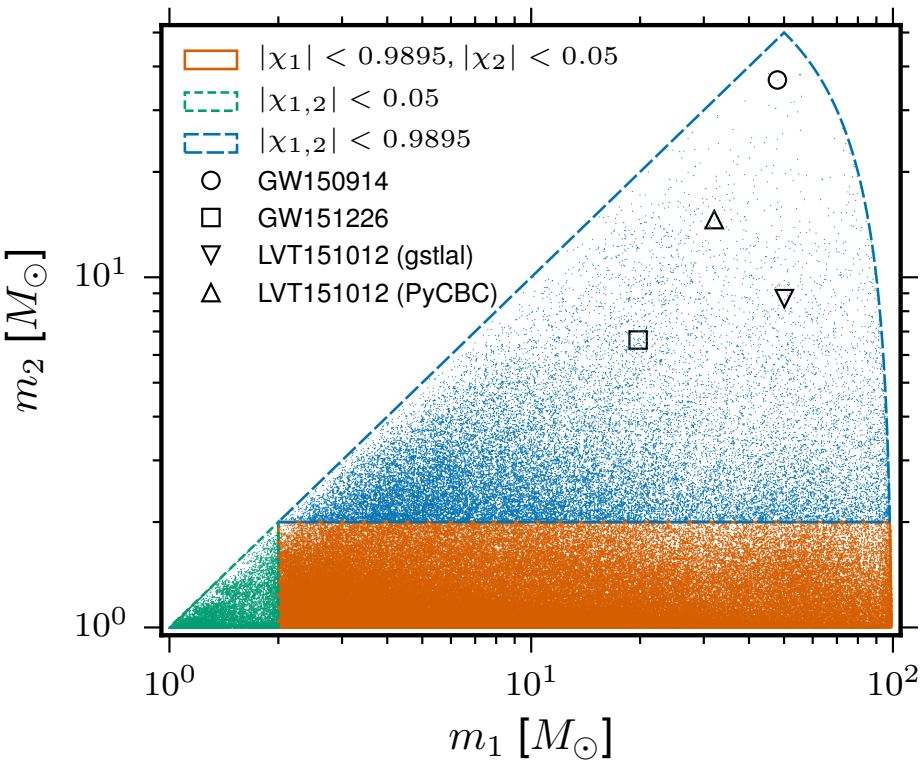



Figure 6.5: The four-dimensional search parameter space covered by the O1 template bank shown projected into the component-mass plane, using the convention $m_1 > m_2$. Each point in the figure is a template in the bank. The different colors indicate the different source types, and the spin limits considered for that part of the parameter space. The symbols identify that the templates that found the 3 events in O1. LVT151012 was identified by different templates in `PyCBC` and `GstLAL`. The bank used in O1 shown here is fondly known as the uberbank.

these noise spectra was used to place a template bank that was used in the search. Recall that we require exact template match between the two detectors for forming candidates, therefore it is imperative that we use a common template bank for the two detectors.

In Sec. 3.2.3 we talked about measuring the performance of the template bank using the fitting factor [91]. Fig. 6.6 shows the distribution of fitting factors obtained for the template bank against a set of injections spanning the same parameter space as the template bank. The mismatch between signals and templates is less than 3% for more than 99% of the $10^5$ simulated signals.

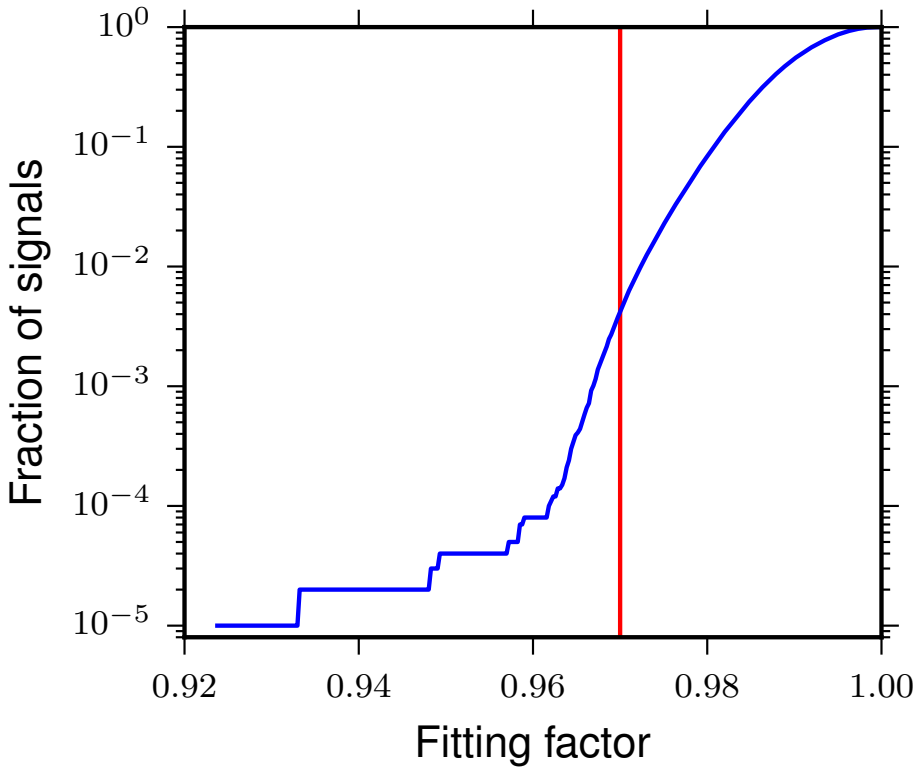



Figure 6.6: Cumulative distribution of the fitting factors obtained with the O1 template bank for a population of simulated signals spanning the same parameter space as the templates. Less than 1% of the signals have a mismatch of more than 3% against the template bank.

### 6.2.2 Data conditioning

Matched filtering is optimal under the assumption that the data are both stationary and Gaussian. Unfortunately, in reality that is not true. Non-stationarity over long time scales can be handled by tracking the PSD. But short noise transients, commonly referred to as “glitches", can cause high-SNR matched filter outputs mimicking signal detections. Therefore the pipelines cannot rely solely on SNR to detect signals.

#### Data quality vetoes

We have a dedicated team working towards the characterization of the noise trasients that are specifically harmful to the CBC and burst searches. Several environmental and instrumental monitors are maintained on sites which output information to auxiliary channels that can ascertain times of clear coupling between local transient noise sources. Data quality products are created based on these couplings, which can then be applied by the pipelines on the data as “vetoes”. There are two different types

of data quality products: *data quality flags* and *data quality triggers*. *Data quality flags* are created whenever a reproducible criterion with known noise couplings are met [117–120] and are typically have durations on the order of seconds to hours. The data quality flags are further divided in two veto categories: CAT1 and CAT2. CAT1 vetoes indicate times when the data should not be analyzed at all due to a critical issue with a key detector component. CAT1 vetoes are identically applied to all transient searches to the input data. CAT2 vetoes indicate times when a noise source with a known physical coupling to the strain channel is active. These are typically applied after the initial processing of data for a specific search. Most search pipelines require a certain minimum stride of unbroken data, and applying CAT2 vetoes after the initial processing has been done allows more data to be used by the pipelines. This minimum segment length of data was 512 s for the `GstLAL` pipeline, 2048 s for the `PyCBC` pipeline, and 620 s for the burst searches. *Data quality triggers* are short duration vetoes associated by algorithms that identify significant statistical correlations between a transient in the strain channel with one in an auxiliary channel [121–124]. Data quality triggers are applied as CAT3 vetoes after the initial processing of the data by the pipelines, similar to CAT2 vetoes. These are typically times when the coupling between the strain and auxiliary channels are statistically generated but the coupling mechanism is not understood. For more details about the data characterization process, we refer the reader to [100].

On top of applying the data quality vetoes, the `GstLAL` pipeline combats the glitches present in the data in two ways. One is self-vetoing the data whenever the whitened data exceeds a certain number of pre-decided standard deviations; this is referred to as auto-gating; see below. The other is using signal consistency checks which have been described in Sec. 4.2.3.

#### Auto-gating

Glitches can sometimes produce matched-filter outputs that have considerably higher amplitudes than any expected output from a compact binary signal. These can therefore safely be removed from the data through a process called gating. Whitened data should have unit variance. If a momentary excess, greater than a pre-decided number of standard deviations, $\sigma$, is observed in the whitened data, then the gating process zeros this excess with a 0.25 s padding on each side. See Fig. 6.7.

For O1, this pre-decided number of standard deviations was $50\sigma$. The lower this threshold is set to, the more glitches this method can potentially gate. However, we

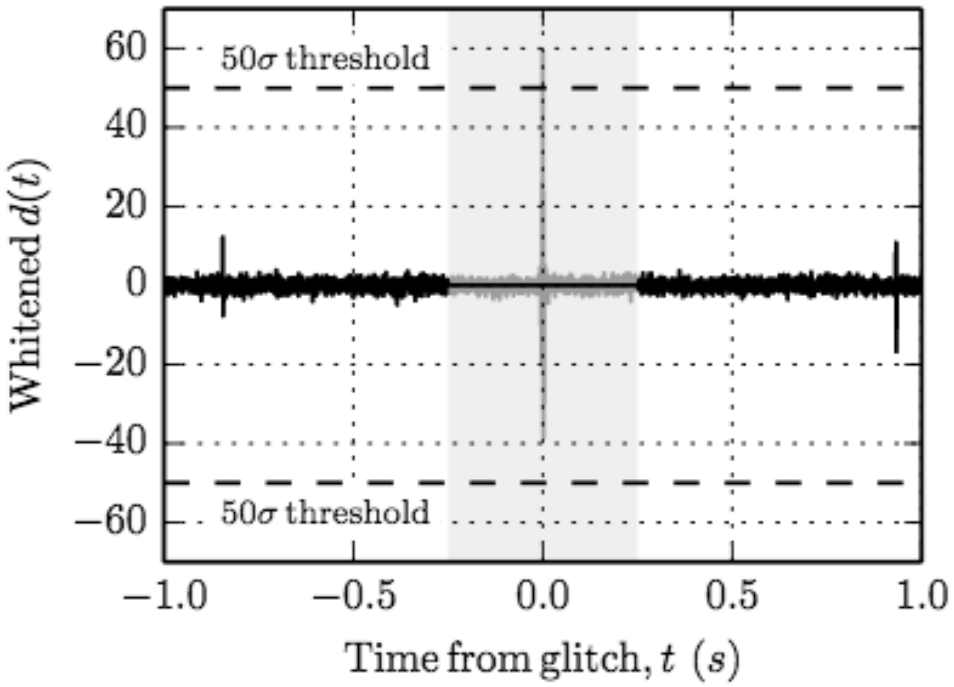



Figure 6.7: Example of gating. The figure shows 2 s of LIGO S6 data, showing 3 glitches, but only one of them was loud enough to surpass the 50$\sigma$ threshold and therefore gated. Figure taken from [3].

don't want it to be so low where we start gating real signals. The number 50$\sigma$ was decided based on the results of several injection studies conducted before O1 using the ER7 and ER8 data and different gate thresholds. It was the lowest threshold among our trials at which we were not gating out injections. See also Sec. 7.1.3.

**O1 data**

The GstLAL pipeline used coincident observations between the two LIGO detectors in O1 (See also Sec. 7.1.5), from September 12, 2015 to January 19, 2016. During the 129.7 days of operation, the detectors recorded 51.5 days of coincident data.

Unstable instrumental operation and hardware failures in form of CAT1 vetoes affected 2.9 days of these coincident observations. These data were discarded and the remaining 48.6 days of data were used as input to the pipelines [18].

The `GstLAL` pipeline takes this data and only analyzes continuous stretches that are longer than 512 s. This condition reduces the available coincident data to 48.3 days.

Note that the PyCBC pipeline applied a different minimum length criteria in O1[2] of 2048 s, and therefore analyzed 46.1 days of data.

### 6.2.3 Blind analysis

There are a number of tuning parameters in the search pipelines, and to eliminate any bias, conscious or subconscious, both in producing results and upper limits, we blind ourselves to the results of the search until the tuning configurations have been frozen. But to decide on the tuning choices it is important that we look at the performance of the pipeline in terms of its search sensitivity, ability to separate signal and background, ability to predict the background model, *etc.* We study and tune the performance of the pipeline to real signals, using software injections of waveforms. We study the background (accidental coincidence of detector noise fluctuations) using real data (as opposed to simulated Gaussian detector noise). This is done by looking at the results of a time-shifted analysis, in which the data streams of the instruments have been shifted by a constant time with respect to one another. The candidates of the time-shifted analysis, which are the coincidences found in that analysis are not real candidates, since they would not have been coincident in the real "zerolag" unshifted analysis[3]. We call the results from the time-shifted analysis "closed-box" results. We want to see a good agreement between the closed-box results and the background predicted by the pipeline.

Before the start of O1, it was decided that the offline pipelines will wait until five days of coincident data had been accumulated by the two detectors before analyzing those data, and repeating this process for every 5-coincident-day "chunk". This time period was decided based on two competing requirements; we wanted it to be long enough to get a good background estimate (if we use too little time then the noise PDFs don't converge), but short enough that the character of the data or the sensitivity of the instrument doesn't change too much[4]. These five-day chunks of data were called "analysis chunks". We had nine such chunks in O1. For each chunk, the pipelines had to prepare the closed box results along with a "checklist" whose goal was to ensure that everything was working as expected, that all the data quality products (Sec. 6.2.2) had been applied correctly, that the pipeline recovered the injections that

[2]The PyCBC analysis also used a minimum segment length of 512 s in O2.

[3]This is only true when the pipeline is only looking at coincident candidates. A "single-detector search" where one can rank non-coincident events is being actively developed for both the GstLAL (Sec. 7.1.5) and PyCBC [125] searches. It is yet unclear as to how to conduct a blind analysis in such a case.

[4]The GstLAL pipeline now adapts itself to change in instrument sensitivity; see Sec. 7.1.8.

were within the sensitivity range of the instruments, and that the background model built by the pipeline is able to explain the closed box results. If an injection with a high expected SNR was missed, there needed to be a satisfactory explanation for it before we could open the box. Expected SNR is calculated assuming an exact matching template is present in the template bank. An example of a satisfactory explanation for an injection being missed is the presence of a glitch (see Fig. 6.8) on the top or near the signal in one or both detectors (which would throw off the $\xi^2$ statistic), high precession of the injection making it fall outside the parameter space being covered by our bank, etc. These results were presented to the entire data analysis group and the boxes were to be opened once we had the go ahead from everyone. The consequence of the unblinding is that there should be no further tuning of the results. However, we do allow for a possibility for some things to change after the unblinding of the analysis, for example if we find an error in the analysis due to an accidental misconfiguration which is not the result of a deliberate tuning choice or an unavoidable technical error due to a bug. We also allow for the possibility to rerun on updated calibrated data. During the six science runs of initial LIGO, scientists waited patiently and eagerly for the boxes opened, beseeching for a signal. In O1, even before the first closed box analyses had been prepared for the first O1 chunk, we knew that it contained a signal since GW150914 was detected by the online coherent wave burst pipeline [43] and was loud enough to be seen in an omegascan (time-frequency spectrogram). So it was not possible to conduct an entirely blind analysis of these data, but we did the best we could to get an unbiased estimate of the significance of GW150914 by freezing the tuning choices of the pipelines when it was detected online.

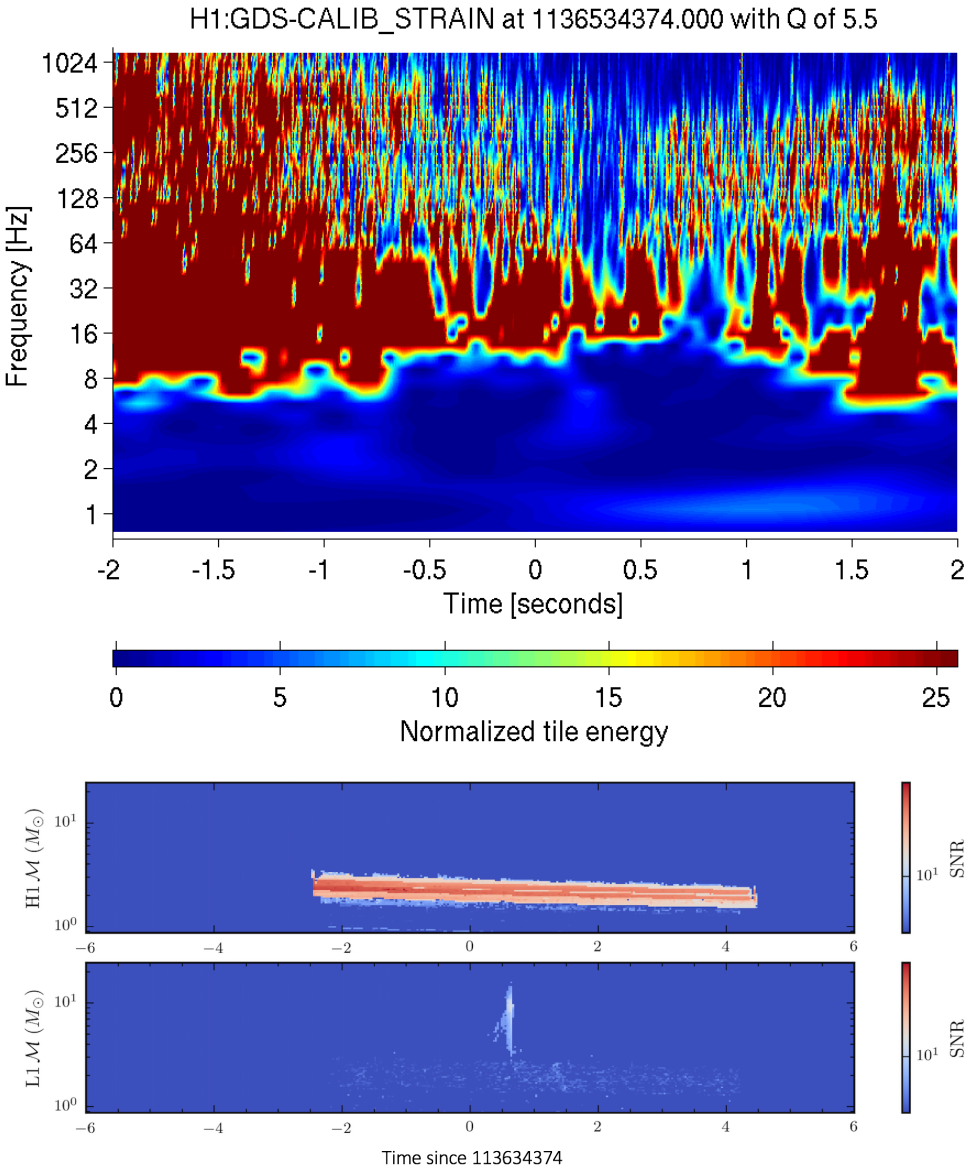



Figure 6.8: A glitch in the Hanford detector which caused the pipeline to miss an injection made at the GPS time 113634374 (top). The heat map for SNRs recovered by the pipeline in both detectors (bottom) as a function of time and chirp mass of the templates. We see that there is no clear peak in the SNR time series for any of the templates in the Hanford detector such as we expect for a signal, for example in the Livingston detector.

## 6.3 Challenges in Observing run 1

### 6.3.1 Case of the missing templates - GW150914

Both the `GstLAL` and the `PyCBC` pipelines prepared the closed boxes for O1 chunk 1 that contained GW150914 following the guidelines that were set before O1 and

described in the previous section.

On opening of the boxes, GW150914 was found as the loudest event in both the analyses as expected. However, it was noted that it was found with significantly different SNRs in PyCBC and GstLAL, Table 6.1.

| Pipeline | H1 SNR | L1 SNR | $m_1$ | $m_2$ | $\mathcal{M}_c$ |
|---|---|---|---|---|---|
| PyCBC | 19.71 | 13.28 | 47.93 | 36.6 | 36.40 |
| GstLAL | 16.02 | 9.79 | 95.34 | 4.42 | 14.97 |

Table 6.1: Discrepancy in original recovered SNRs for GW150914 between the two pipelines

The reason the two pipelines recovered such a different SNR for the same event is because they recovered it with a different template. The GstLAL pipeline recovered it with a template (let's call it Tmplt2) with a chirp mass of $\mathcal{M}_c = 14.97$, whereas the chirp mass of the template that the PyCBC (let's call it Tmplt1) pipeline recovered it with $\mathcal{M}_c = 36.4$. Both the pipelines use a different ranking statistic and a different method to estimate significance, so it is not impossible or even unlikely for the two to recover something with different SNRs. But still the there was the question as to why the GstLAL pipeline ranked a template that picked up a fraction of SNRs in both the detectors highest.

We found that the reason behind this discrepancy was that the Tmplt1 was missing from the GstLAL filter banks. Recall that the GstLAL pipeline uses the LLOID algorithm for matched filtering (Sec. 4.2.2). Instead of performing the matched filtering with physical templates, it performs matched filtering with a reduced set of basis templates and then reconstructs the SNRs for the original physical templates. To do this it divides the template bank into several smaller split banks. This is done by first diving the template bank into $\chi_{\text{eff}}$ bins, and then diving each $\chi_{\text{eff}}$ bin into several $\mathcal{M}_c$ bins, such that the neighboring bins have some overlap in the templates they contain. The reconstruction of templates at the boundaries of split banks from the basis filters is suboptimal, and therefore we borrow some templates from neighboring bins which are discarded after performing the SVD (see Fig. 6.9). The problem was that there was no overlap between the extreme $\chi_{\text{eff}}$ bins but the waveforms were still being clipped after the SVD. This bug resulted in our filter bank containing 1170 templates less than the original template bank which included Tmplt1. These templates are shown in Fig. 6.10.

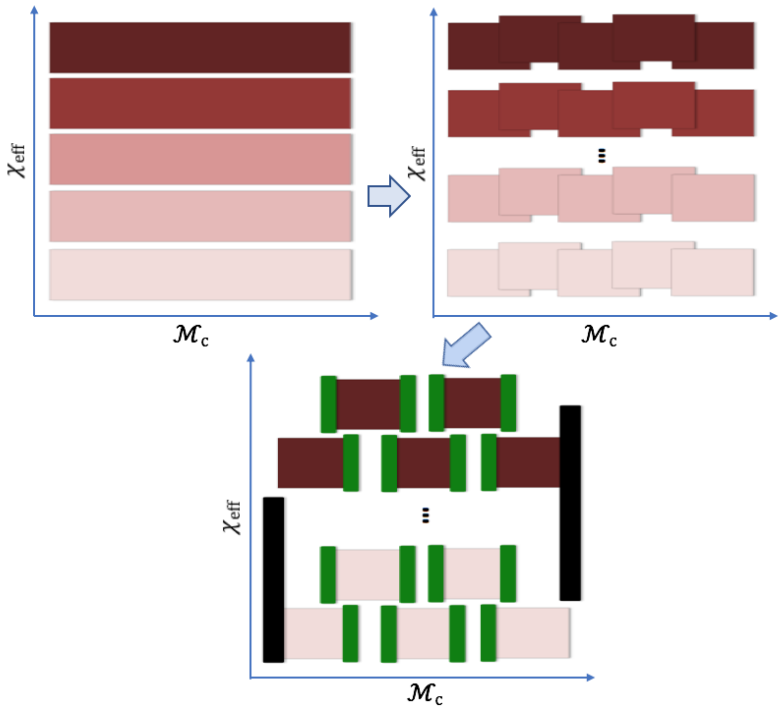



Figure 6.9: Cartoon illustration of the bank binning process in the `GstLAL` pipeline. The black bars show the waveforms that are not repeated because there was erroneously no overlap being clipped off. The green bars depict the waveforms that were present twice in a bin because of an overlap being properly clipped off. Also refer back to Fig. 4.2.

This bug was fixed by padding the split banks at the edge of $\chi_{\rm eff}$ bins with duplicate templates from themselves so that when we perform clipping we get rid of those redundant templates. We reran the analysis with this fix, and the new recovered SNRs matched with the `PyCBC` analysis, Table 6.2.

| Pipeline | H1 SNR | L1 SNR | $m_1$ | $m_2$ | $\mathcal{M}_c$ |
|---|---|---|---|---|---|
| `PyCBC` | 19.71 | 13.28 | 47.93 | 36.6 | 36.40 |
| `GstLAL` | 20.08 | 13.35 | 47.93 | 36.6 | 36.40 |

Table 6.2: The recovered SNRs for GW150914 agree between the two pipelines after fixing of the "missing templates" bug. They also find GW150914 with the same template.

The fixed analysis also showed improvement in the accuracy in $\mathcal{M}_c$ parameter (Fig. 6.11), signal and noise separation (Fig. 6.12), and the search sensitivity of the pipeline (Fig. 6.13).

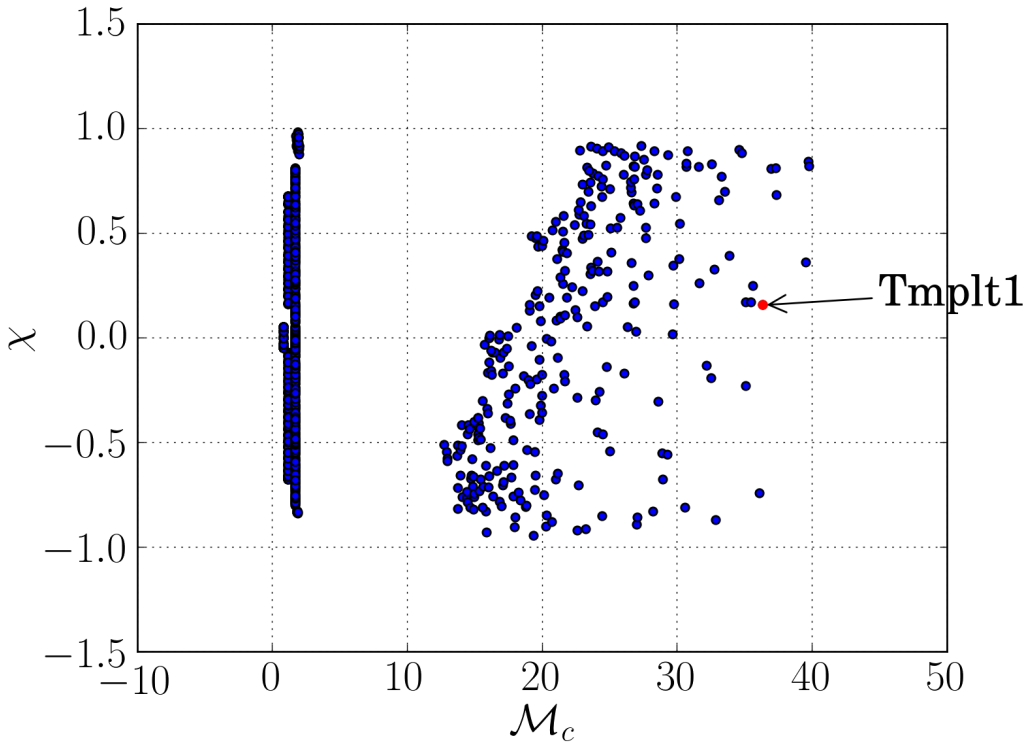



Figure 6.10: Templates missing from `GstLAL` filter bank, in $\mathcal{M} - \chi_{\text{eff}}$ space. The templates shown in this figure correspond to the black bars in the cartoon illustration in Fig. 6.9.

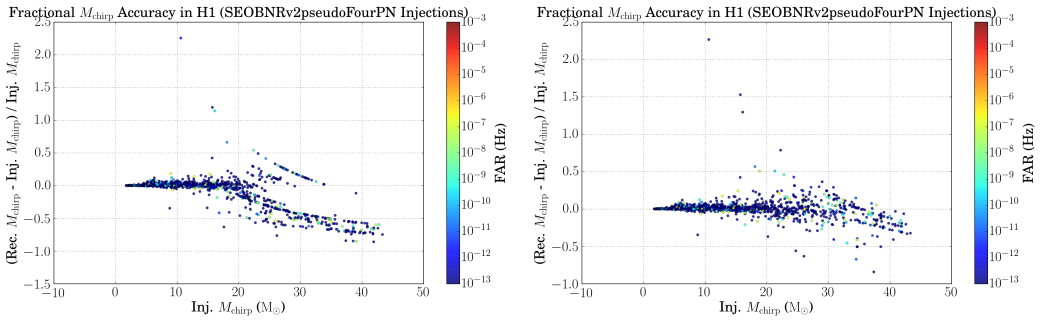



Figure 6.11: Comparison of $\mathcal{M}_c$ recovered by the pipeline for the same set of injections before (left) and after (right) fixing the bug of missing templates. We can see that the recovery has improved significantly specially in the high chirp mass region. These figures are identical for the two detectors, since we only rank coincident events; therefore, we only show the figure for Hanford.

Overall, the bug seemed to have affected the high mass region of the parameter space in particular. The density of templates in the high mass region is low, and each template covers a much larger space in masses as compared to the low-mass

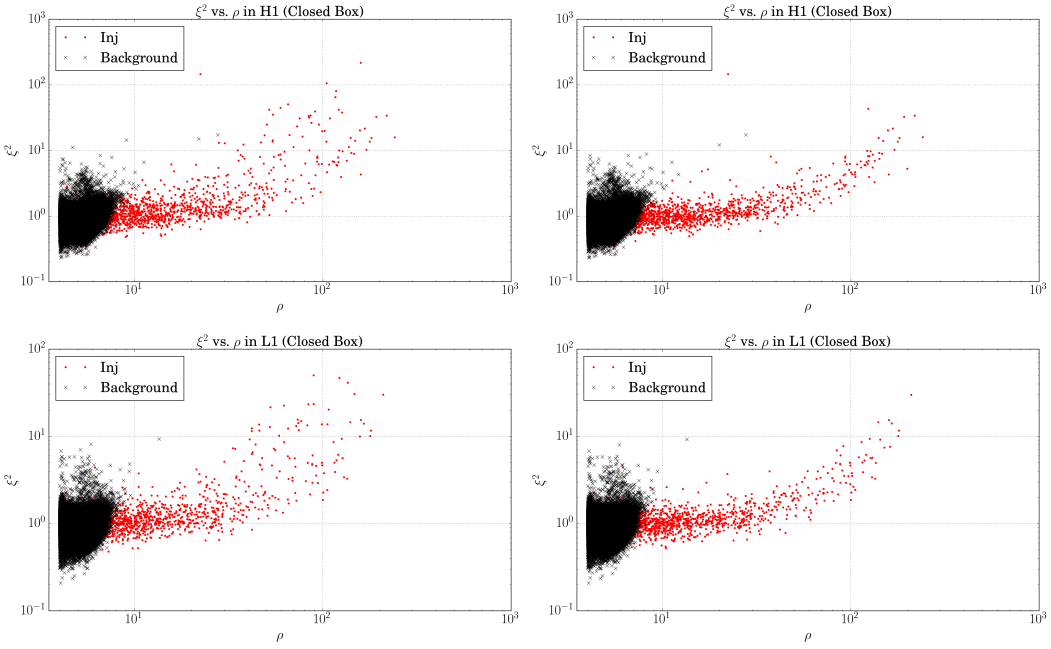



Figure 6.12: Comparison of signal and background separation in SNR-$\xi^2$ plane before (left) and after (right) fixing the bug of missing templates. We see that the $\xi^2$ values for injections have shifted closer to 1, and the separation between the signal and background has improved.

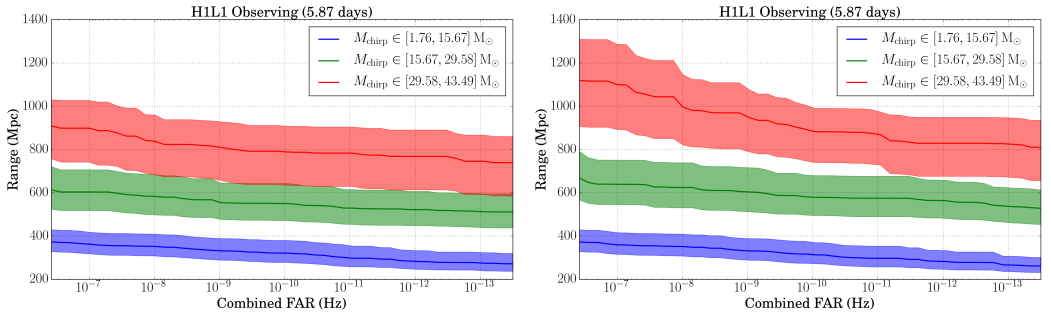



Figure 6.13: Comparison of sensitive distance as measured by the same set of injections made in the pipeline before (left) and after (right) fixing the bug of missing templates. The range has improved significantly, particularly for high mass injections.

region. The fixing of the bug didn't require any changes to the core DAG or analysis (Sec. 5.2); we just changed the input template filter banks used for matched-filtering.

### 6.3.2 Case of the missing injections

After fixing the bug that affected O1 chunk 1, the `GstLAL` offline pipeline continued analyzing the data for the following chunks as they became available. Recall that for opening the box, we are required to prepare a checklist in which among other things

we followed up on any injections with high expected SNRs that were being missed.

While following up on missed injections from chunk 2, we stumbled upon three injections (Table 6.3) that could not be explained by the usual acceptable reasons. They were not being gated by the pipeline, they were not being injected in vetoed data, they were not precessing, and there were no glitches in either detector at the time of these injections. These injections belonged in the parameter space covered

| Number | H1 expected SNR | L1 expected SNR | H1 recovered SNR | L1 recovered SNR |
|---|---|---|---|---|
| INJ1 | 39.82 | 22.41 | 5.50 | 6.71 |
| INJ2 | 16.70 | 8.38 | 4.73 | 6.69 |
| INJ3 | 16.58 | 14.45 | 6.86 | 6.83 |

Table 6.3: Missed injections from O1 chunk2 with high expected SNRs. These injections were generated using the `SEOBNRv2` model. We see that the pipeline is only recovering a small fraction of the injected SNRs.

by the bank. The template bank has been shown to be effectual (Fig. 6.6) in this parameter space, which means that any mismatch between the injection and the template should only account for a maximum of 10% loss in SNR. To explicitly quantify this number, we calculate the fitting factors of the three injections with the template bank using a local PSD. We report the results in Fig. 6.14.

```
1. m1: 51.46   m2: 43.17   spin1z: -0.10   spin2z: -0.59
        best matching template:   m1: 63.42   m2: 19.12   spin1z: -0.60
spin2z: -0.44
        best match:  0.975447
```

```
2. m1: 9.98   m2: 85.33   spin1z: -0.38   spin2z: -0.70
        best matching template:   m1: 85.28   m2: 11.54   spin1z: -0.56
spin2z: -0.42
        best match:  0.952091
```

```
3. m1: 54.97   m2: 16.22   spin1z: -0.69   spin2z: 0.81
        best matching template:   m1: 47.67   m2: 23.27   spin1z: -0.53
spin2z: 0.15
        best match:  0.979924
```

Figure 6.14: Best-matching template for each injection along with the fitting factor. We see that the template bank has good fitting factors with all three injections.

After digging in the output databases of the pipeline, we found that the highest likelihood-ratio triggers returned by the pipeline for the three injections corresponded to templates much lower in masses as compared to the best-matching templates.

Now the question is why were these triggers with low SNRs being assigned a higher detection statistic as compared to the triggers from the best-matching templates. To answer this question, we looked at unclustered outputs of the pipeline. Our investigations are reported in Figs. 6.15, 6.16, and 6.17.

**INJ1**

| | Injection | Final trigger from the pipeline | Trigger for the best match template |
|---|---|---|---|
| **mass 1** | 51.45628 | 9.0589972 | 63.422466 |
| **mass 2** | 43.17348 | 1.6343842 | 19.123806 |
| **spin 1z** | -0.100697 | 0.97785664 | -0.59504938 |
| **spin 2z** | -0.594858 | 0.046404183 | -0.43602765 |
| **H1 snr** | 39.8248017296 (expected) | 5.5032654 | 39.230526 |
| **L1 snr** | 22.4052987615 (expected) | 6.7108722 | 23.138092 |
| **H1 chisq** | - | 3.7102835 | 1.3657118 |
| **L1 chisq** | - | 1.5735655 | 0.91684389 |
| **H1 end_time** | 1128069141 | 1128069141 | 1128069141 |
| **H1 end_time_ns** | 485711097 | 720375435 | 490640890 |
| **L1 end_time** | 1128069141 | 1128069141 | 1128069141 |
| **L1 end_time_ns** | 480910548 | 710926042 | 485829333 |
| **svd-bank-number** | - | 163 | 235 |
| **H1 event_id** | - | 1136378 (in raw trigger file) | 1433663 |
| **L1 event_id** | - | 1136177 (in raw trigger file) | 1434266 |
| **coinc_event_id** | - | 6768 | ?? |
| **Likelihood** | - | 4.69604647823 | ?? |

Figure 6.15: Recovered output from the pipeline for INJ1 with the best-matching template vs. that saved by the pipeline. We see that both the SNRs and $\xi^2$s for the maximum overlap template are better as compared to those for the template returned by the pipeline. There has been no likelihood-ratio value assigned to the triggers for the best-match template.

**INJ2**

| | Injection | Final trigger from the pipeline | Trigger for the best match template |
|---|---|---|---|
| **mass 1** | 9.982027 | 17.381136 | 85.283157 |
| **mass 2** | 85.33270 | 1.6890575 | 11.535029 |
| **spin 1z** | -0.383017 | -0.93537635 | -0.56100571 |
| **spin 2z** | -0.703889 | 0.0078926645 | -0.42154437 |
| **H1 snr** | 16.7028041486 (expected) | 4.7322683 | 15.397376 |
| **L1 snr** | 8.37821059356 (expected) | 6.6859555 | 8.4837809 |
| **H1 chisq** | - | 1.0523698 | 1.1242108 |
| **L1 chisq** | - | 0.69529456 | 0.97120833 |
| **H1 end_time** | 1128145227 | 1128145228 | 1128145227 |
| **H1 end_time_ns** | 475859661 | 393173472 | 482111557 |
| **L1 end_time** | 1128145227 | 1128145228 | 1128145227 |
| **L1 end_time_ns** | 481995138 | 401719547 | 487746512 |
| **svd-bank-number** | - | 197 | 235 |
| **H1 event_id** | - | 2831712 (in the raw trigger file) | 1849738 |
| **L1 event_id** | - | 2832081 (in the raw trigger file) | 1849537 |
| **coinc_event_id** | - | 24421 | ?? |
| **Likelihood** | - | 6.09406482104 | ?? |

Figure 6.16: Recovered output from the pipeline for INJ2 with the best-matching template vs. that saved by the pipeline. We see that both the SNRs and $\xi^2$s for the maximum overlap template are better as compared to those for the template returned by the pipeline. There has been no likelihood-ratio value assigned to the triggers for the best-match template.

**INJ3**

| | Injection | Final trigger from the pipeline | Trigger for the best match template |
|---|---|---|---|
| **mass 1** | 54.969820 | 29.577032 | 47.673481 |
| **mass 2** | 16.222320 | 9.4891233 | 23.272543 |
| **spin 1z** | -0.694287 | 0.9382658 | -0.52876478 |
| **spin 2z** | 0.809538 | 0.51627928 | 0.14958972 |
| **H1 snr** | 16.5812009372 (expected) | 6.8581052 | 13.834707 |
| **L1 snr** | 14.4537016221 (expected) | 6.8335752 | 14.95764 |
| **H1 chisq** | - | 2.1856968 | 1.6855433 |
| **L1 chisq** | - | 2.6643603 | 1.0076261 |
| **H1 end_time** | 1128227635 | 1128227635 | 1128227635 |
| **H1 end_time_ns** | 506769512 | 536559108 | 511395377 |
| **L1 end_time** | 1128227635 | 1128227635 | 1128227635 |
| **L1 end_time_ns** | 508891630 | 549427551 | 513316199 |
| **svd-bank-number** | - | 247 | 239 |
| **H1 event_id** | - | 2781900 (in the raw trigger file) | 3341880 |
| **L1 event_id** | - | 2782302 (in the raw trigger file) | 3342081 |
| **coinc_event_id** | - | 53096 | ?? |
| **Likelihood** | - | 6.46002909196 | ?? |

Figure 6.17: Recovered output from the pipeline for INJ3 with the best-matching template vs. that saved by the pipeline. We see that both the SNRs and $\xi^2$s for the maximum overlap template are better as compared to those for the template returned by the pipeline. There has been no likelihood-ratio value assigned to the triggers for the best-match template.

It turned out that the pipeline wasn't assigning the coincident label to these "good" triggers, so they weren't treated as candidates or assigned a likelihood ratio.

The `GstLAL` pipeline finds coincident events in streaming manner, using a program called `streamthinca`. It throws away single detector triggers that are older than a time up to which it considers that the coincident event list should be complete. While tuning the SVD for high mass waveforms, we had aligned them at their mergers (Sec. 5.2.3). This offset was not being taken into account properly, and therefore for some of the high mass templates, where the shift in the waveforms was a lot, the pipeline failed to perform the coincidence test accurately. Once the bookkeeping for coincident candidates was fixed, and the analysis rerun, the pipeline recovered these missing injections. Once again, we compare the accuracy in $\mathcal{M}_c$ parameter (Fig. 6.18), signal and noise separation (Fig. 6.19) and the search sensitivity of the pipeline (Fig. 6.20) before and after the fixing of coincident event candidate list.

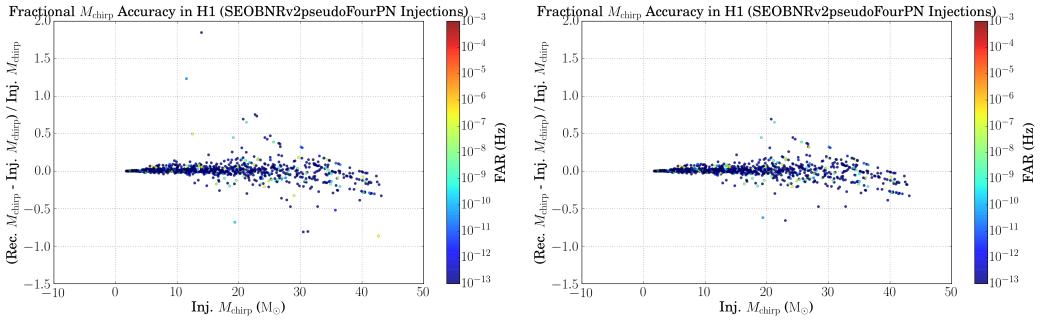



Figure 6.18: Comparison of $\mathcal{M}_c$ recovered by the pipeline for the same set of injections before (left) and after (right) fixing the bug of missing coincident events. We see fewer outliers in the $\mathcal{M}_c$ recovery after fixing the list of candidates.

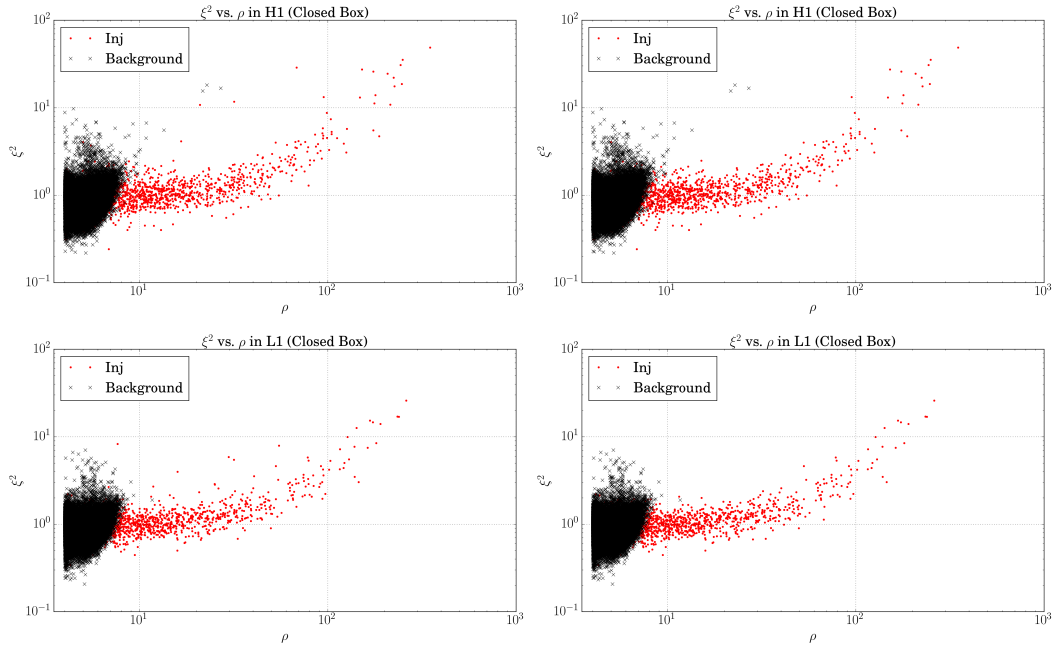



Figure 6.19: Comparison of background and signal separation before (left) and after (right) fixing the bug of missing coincident events. We see less scatter (fewer outliers) in $\xi^2$ distribution after the bug fix.

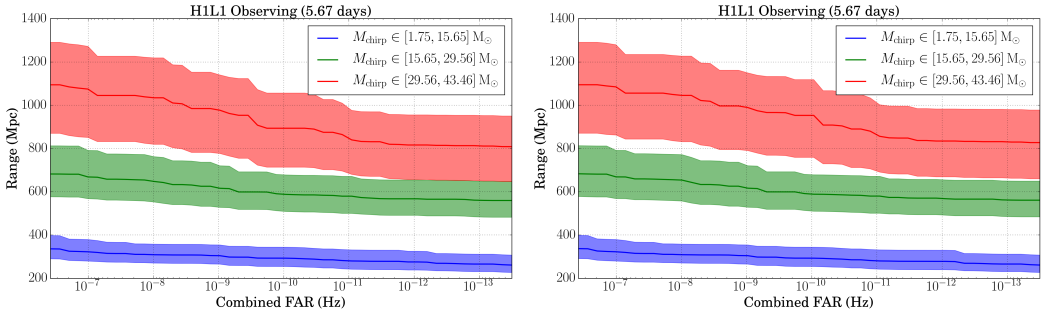
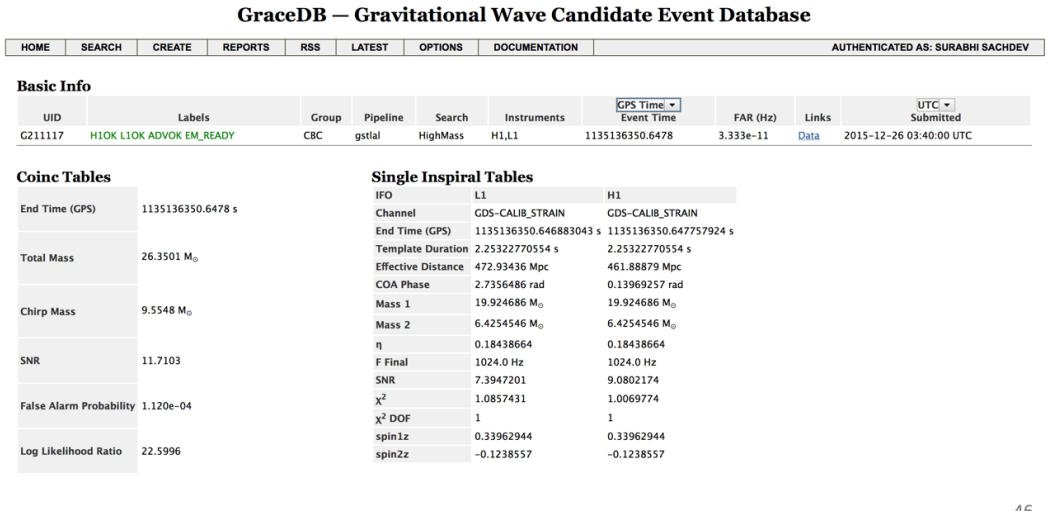



Figure 6.20: Comparison of sensitive distance as measured by the same set of injections made in the pipeline before (left) and after (right) fixing the bug of missing candidates from the list. The range has improved slightly, particularly for high mass injections.

## 6.4 First online detection

On December 26, 2015, the gravitational-wave candidate GW151226, fondly known as the “Boxing day event" was identified within 70 s by the low-latency `GstLAL` pipeline (Fig. 6.21), and was a huge success for the pipeline and GW science in general. This detection came only a few days after the pipeline started uploading events from the full uberbank (Fig. 6.5). Up until then, the pipeline was only running on an EM-bright bank which contained BNS and NSBH templates which are expected to have an EM counterpart signal and which therefore needed to be identified with low latency in order to alert EM telescope operators. The pipeline was also running a blind analysis on the full uberbank for testing purposes.



**GraceDB — Gravitational Wave Candidate Event Database**

HOME | SEARCH | CREATE | REPORTS | RSS | LATEST | OPTIONS | DOCUMENTATION | AUTHENTICATED AS: SURABHI SACHDEV

**Basic Info**

| UID | Labels | Group | Pipeline | Search | Instruments | GPS Time Event Time | FAR (Hz) | Links | UTC Submitted |
|---|---|---|---|---|---|---|---|---|---|
| G211117 | H1OK L1OK ADVOK EM_READY | CBC | gstlal | HighMass | H1,L1 | 1135136350.6478 | 3.333e-11 | Data | 2015-12-26 03:40:00 UTC |

**Coinc Tables**

| | |
|---|---|
| End Time (GPS) | 1135136350.6478 s |
| Total Mass | 26.3501 $M_\odot$ |
| Chirp Mass | 9.5548 $M_\odot$ |
| SNR | 11.7103 |
| False Alarm Probability | 1.120e-04 |
| Log Likelihood Ratio | 22.5996 |

**Single Inspiral Tables**

| IFO | L1 | H1 |
|---|---|---|
| Channel | GDS-CALIB_STRAIN | GDS-CALIB_STRAIN |
| End Time (GPS) | 1135136350.646883043 s | 1135136350.647757924 s |
| Template Duration | 2.25322770554 s | 2.25322770554 s |
| Effective Distance | 472.93436 Mpc | 461.88879 Mpc |
| COA Phase | 2.7356486 rad | 0.13969257 rad |
| Mass 1 | 19.924686 $M_\odot$ | 19.924686 $M_\odot$ |
| Mass 2 | 6.4254546 $M_\odot$ | 6.4254546 $M_\odot$ |
| $\eta$ | 0.18438664 | 0.18438664 |
| F Final | 1024.0 Hz | 1024.0 Hz |
| SNR | 7.3947201 | 9.0802174 |
| $\chi^2$ | 1.0857431 | 1.0069774 |
| $\chi^2$ DOF | 1 | 1 |
| spin1z | 0.33962944 | 0.33962944 |
| spin2z | -0.1238557 | -0.1238557 |

Figure 6.21: A screenshot of the uploaded GW151226 event in the GraceDB database. The green labels are a sign that a notice/circular should be sent off to our partner EM, neutrino, and other observatories alerting them of this event.

Besides being our second confirmed event, the boxing day event was the first event where matched filtering proved to be crucial to the detection since the signal has a smaller strain amplitude with energy spread over a longer time interval than GW150914. GW151226 was also the first signal to provide a strong evidence for at least one of its components having a nonzero spin [9].

## 6.5 Results

The fact that we had a very successful O1 is indisputable. Even though we saw some problems that negatively affected the high mass parameter space with the `GstLAL` pipeline early on, they were identified and fixed quickly.

In this section we present the final results from the offline GstLAL pipeline on the entirety of O1 data with the final calibration. These results have been presented in [9, 18]. Since then we have reanalyzed the data from O1 using the template bank that was constructed for O2, which covers a bigger parameter space, and an improved version of the GstLAL pipeline. These will be presented in the next chapter.

### 6.5.1 Assigning significances

For the final results, we divided the data from O1 into 3 analysis blocks for running the GstLAL pipeline. The 3 analysis blocks each corresponded to 3 original chunks made for box-openings.

In O1, the GstLAL pipeline was not tracking the changes in detector sensitivity (as quantified with horizon distances) over time. In practice it has been found that a week or more of data is required before the non-coincident SNR-$\xi^2$ histograms used to model the background PDFs converge. Therefore it is not possible to monitor the changes in the instruments' horizon distances on time scales less than $\sim O(1 \text{ week})$. The way the pipeline incorporated the time-dependence of, for example, the SNR-$\xi^2$ histograms was by analyzing data in blocks, so that the events are ranked based on the data from just that block.

By dividing O1 data into 3 analysis blocks, we ensured that the assignment of the likelihood-ratios to the candidates was being done based on the data from their respective blocks. Finally in order to transform the likelihood-ratio numbers to FAP or FAR, the construction of $P(\mathcal{L} \mid \text{n})$ was done by sampling the background PDFs marginalized over all 3 blocks of data.

#### Background estimation: Candidates in or out?

You will recall that in the GstLAL pipeline, we use the statistics of non-coincident triggers to construct the background PDFs which are used to assign the likelihood-ratio to the candidates. See Sec. 4.2.5 for details. By default, we leave the coincident detection candidates of our search out of this calculation. If the candidates are just noise events, they should follow the same distribution as the non-coincident noise events, and the PDFs should not change significantly by their inclusion or exclusion, since they are only a tiny fraction ($< 1\%$ in O1) of the non-coincident events. But if they are gravitational-wave signals, their inclusion in the histograms that inform the background PDFs would have the effect of contaminating the background PDF with something signal-like which would mean that we would underestimate the

significance of a real signal. There is one more possibility, however extremely unlikely. If it so happened that there were noise events in the candidate list which were different from all the other noise events (an extremely tiny probability) and they accidentally ended up in the pile of coincident events ($< 1\%$), then not including these in the background PDFs would mean we could end up overestimating the significance of the candidates.

To err on the side of producing conservative significance estimates we decided to include the coincident detection candidates in our noise model, but only partially; after all these were the very first detections. The way this was done was, we still assigned the likelihood-ratios to the candidates based on background PDFs constructed out of the non-coincident triggers. This means, in Eq. 4.7, the denominator is informed by only non-coincident triggers. But while sampling the background PDFs to construct $P(\mathcal{L} \mid \mathrm{n})$, we used the PDFs which included the coincident candidate events in them, which means that in Eq. 4.10, the integration is performed over the PDFs that were informed using all the triggers of the search.

Technically, this meant that we could keep the results of `gstlal_inspiral_calc_likelihood` from the default runs, but we had to recompute the likelihood distribution for each $\bar{\theta}$ bin, which is the output of `gstlal_inspiral_calc_rank_pdfs`. This was done using a DAG. DAG (Sec. 5.2) refers to a job scheduler system which defines parent-child dependencies between different jobs.

Figs. 6.22, 6.23, and 6.24 show the various signal and noise PDFs marginalized over all the template bins.

Fig. 6.25 shows the cumulative distribution of likelihood-ratio for O1, which tells us how to map $\mathcal{L} : FAP(\mathcal{L})$.

We convert the FAP to single-sided Gaussian standard deviations according to $-\sqrt{2}\mathrm{erf}^{-1}(1 - 2(1 - FAP))$, where $\mathrm{erf}^{-1}$ is the inverse function.

Now we have all the ingredients to report the final results (almost!). Recall that the effect of putting in candidates that are real gravitational-wave events has the effect of producing under-estimated significances. We get very different results for the significance of the boxing day event if we use the candidates for GW150914 to inform the background PDFs vs. if we take them out. We proceed as follows:

1. We assign FAP to all the events based on the background PDFs that contain all the events.

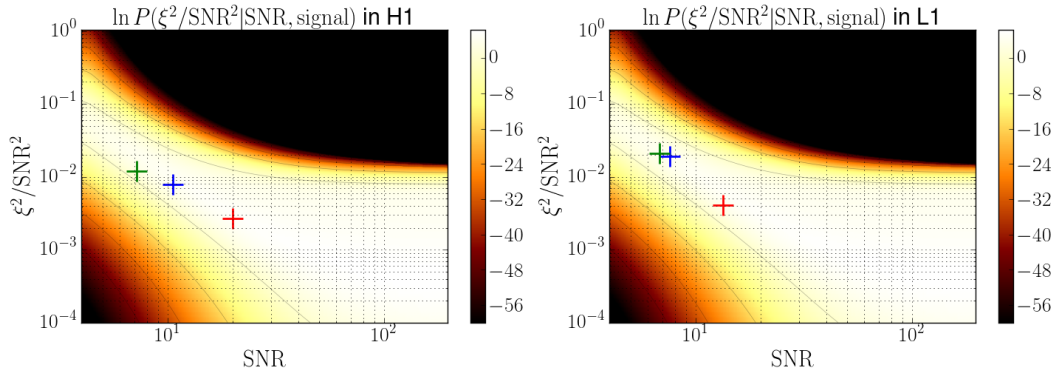



Figure 6.22: $\ln P(\xi^2/\mathrm{SNR}^2 \mid \mathrm{SNR}, \mathrm{s})$ for all of O1 in Hanford (left) and Livingston (right). The red, blue, and green plus signs show the events GW150914, GW151226, and LVT151012 for reference. Note that the PDFs have been marginalized over all the $\bar{\theta}$ bins so they are not the same as the ones used to rank the events, but they give us a pretty good idea about the (SNR-$\xi^2$) position of the events in signal space.

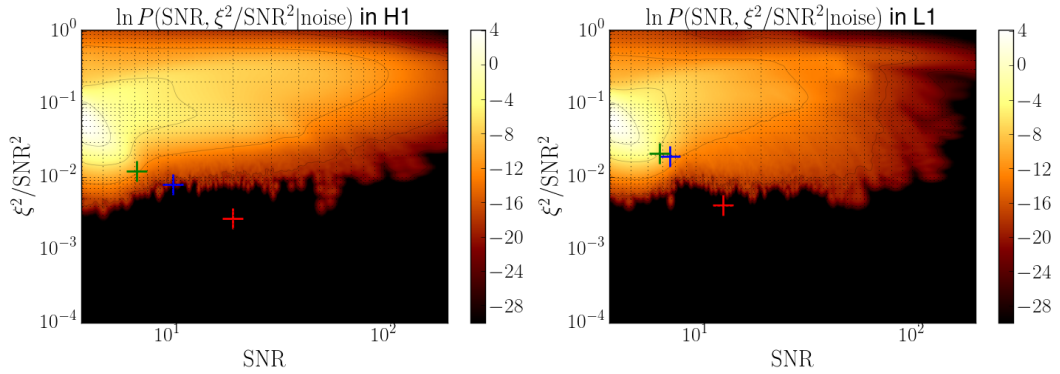



Figure 6.23: $\ln P(\mathrm{SNR}, \xi^2/\mathrm{SNR}^2, \mathrm{n})$ for all of O1 in Hanford (left) and Livingston (right). The red, blue, and green plus signs show the events GW150914, GW151226, and LVT151012 for reference. Note that the PDFs have been marginalized over all the $\bar{\theta}$ bins so they are not the same as the ones used to rank the events, but they give us a pretty good idea about the (SNR-$\xi^2$) position of the events noise space.

2. We note that the significance of GW150914 is $> 6.7\sigma$, even when including the candidates in background PDFs.

3. We decide that GW150914 must be a gravitational-wave signal, and thus, we remove the candidates corresponding to the time of event and form a second background model which we then use to assign FAPs and FARs to all the other events. See Fig. 6.26.

4. With this new background PDF which does not contain GW150914 candidates, but still contains GW151226 candidates, we find the significance of GW151226

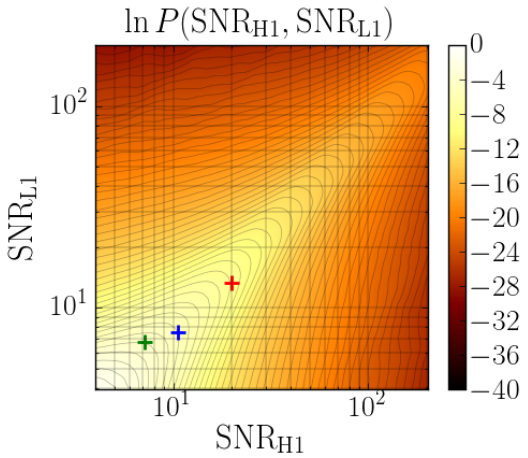



Figure 6.24: Joint SNR PDF for the two detectors under the signal hypothesis. Note that for O1, the pipeline assumed both detectors to have the same horizon distance in order to compute this PDF.

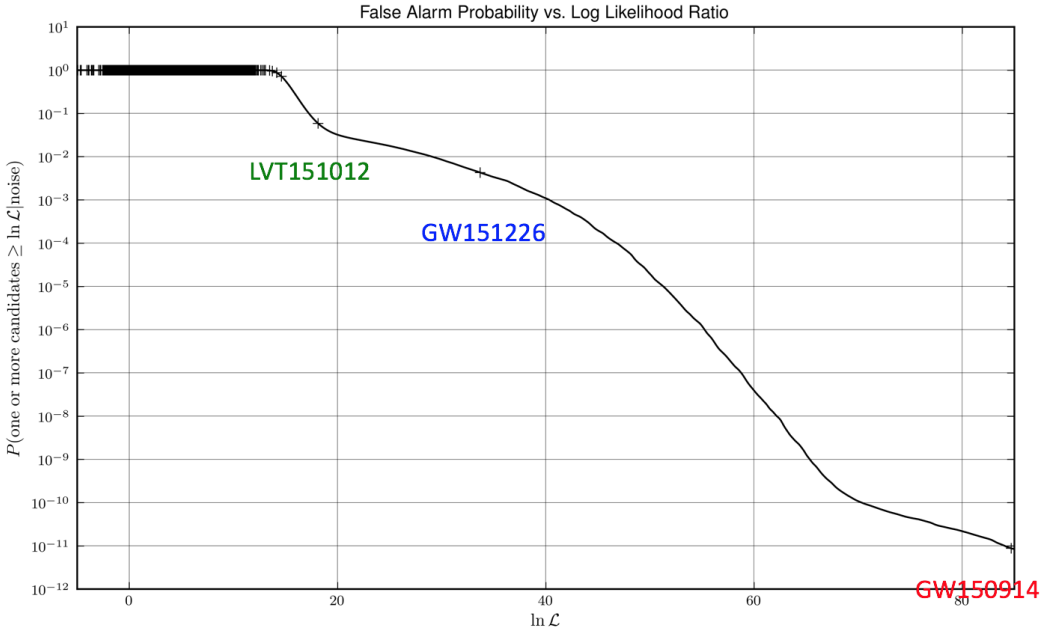



Figure 6.25: The figure shows the mapping between $\ln\mathcal{L}$ and FAP($\mathcal{L}$). This mapping was constructed using background PDFs with all the candidates included.

to be $4.5\sigma$. See Fig. 6.27.

5. Finally, we construct a third background model using PDFs that don't have candidates from either GW150914 or GW151226 and use it to assign the significance of LVT151012. However, we note that at the $\ln\mathcal{L} = 18.1$, the background models do not change significantly by addition or removal of the

loud candidates from GW150914 and GW151226. The addition and removal of these loud events affects the likelihood-ratio distribution for high values of likelihood-ratio the most.

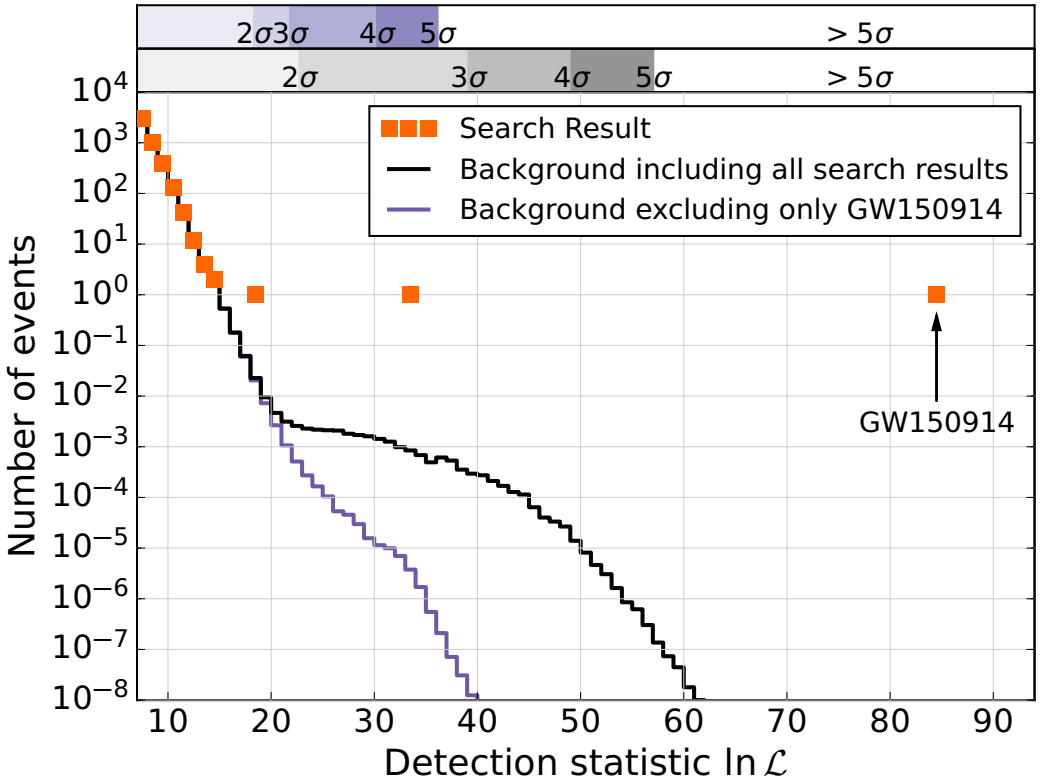



Figure 6.26: Number of events per log likelihood-ratio bin (width = 1) for two background models, one that includes all the candidates (black), and one without the candidates from GW150914 (purple). Orange boxes show the zerolag search results. The bars on the top show the significance levels based on the two backgrounds. We see that the significance of GW150914 is $> 5\sigma$ using both the models. The significance of GW151226, however, increases when we use the background without the candidates from GW150914 to $4.5\sigma$.

We also show these results with cumulative counts on the y-axis, so we do not have to pick an arbitrary bin width in terms of $\ln\mathcal{L}$. And the significance bars on the top correspond directly to the values on the y-axis based on the background model we are using. See Fig. 6.28.

We summarize the results from the `GstLAL` pipeline in Table 6.4.

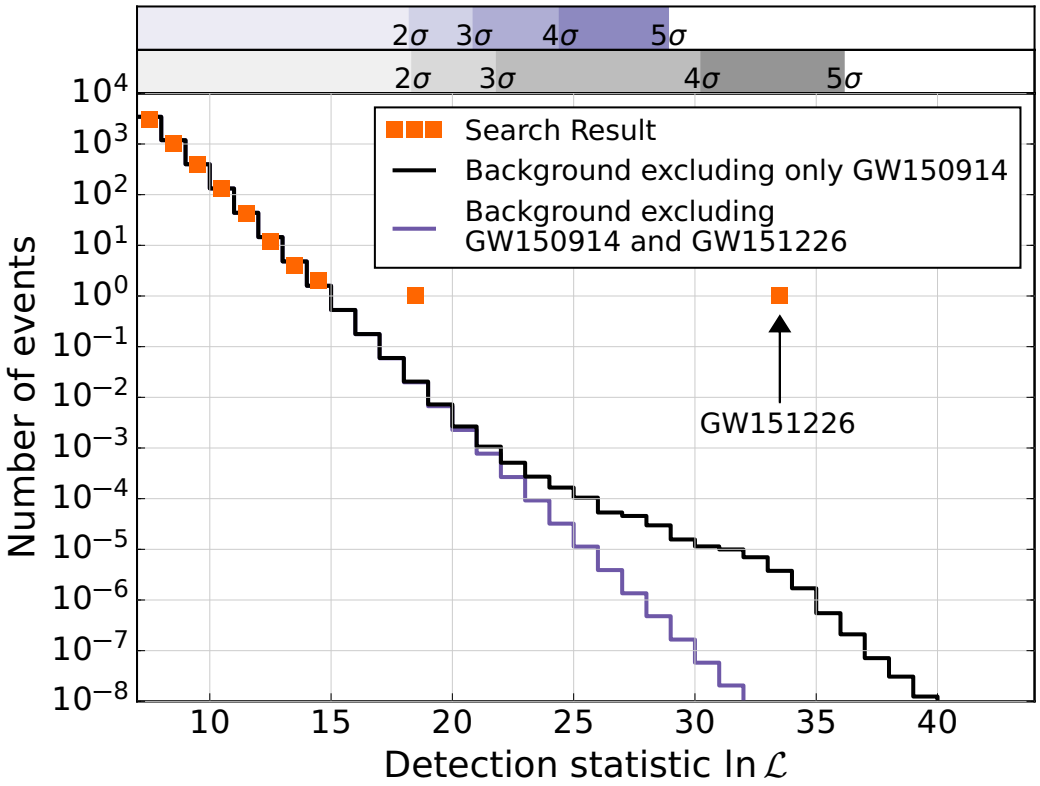



Figure 6.27: Number of events per log likelihood-ratio bin (width = 1) for two background models, one that includes all but GW150914 candidates (black), and one without the candidates from either GW150914 or GW151226 (purple). Orange boxes show the zerolag search results. The bars on the top show the significance levels based on the two backgrounds. We see that the significance of GW151226 is $> 5\sigma$ using both the purple model. The significance of LVT151012, however, remains at the $2\sigma$ level using either background model.

| Event | FAP | $\ln\mathcal{L}$ | H1 SNR | H1 $\xi^2$ | L1 SNR | L1 $\xi^2$ |
|---|---|---|---|---|---|---|
| GW150914 | $8.8 \times 10^{-12}$ | 84.7 | 19.96 | 1.07 | 13.39 | 0.73 |
| GW151226 | $3.5 \times 10^{-6}$ | 33.7 | 10.51 | 0.86 | 7.59 | 1.11 |
| LVT151012 | 0.026 | 18.1 | 7.16 | 0.61 | 6.79 | 0.97 |

Table 6.4: Summary of O1 results from the `GstLAL` pipeline.

### 6.5.2 Other `GstLAL` searches on O1 data

Apart from the main search that covered the space of total masses between $2M_\odot$ and $100M_\odot$, searches were also conducted in higher (intermediate-mass black hole, or IMBH) and lower mass (sub-solar mass) space for O1. The IMBH search also detected GW150914, however with a lower significance. The sub-solar mass search reported null results.

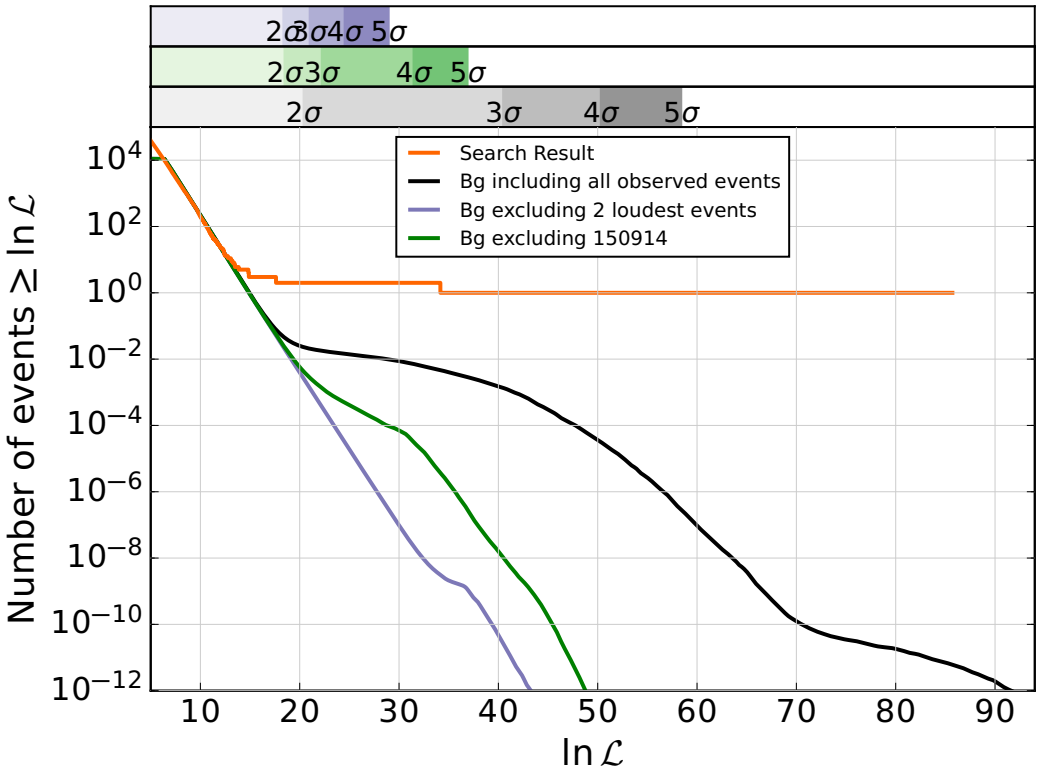



Figure 6.28: Cumulative counts of events with a detection statistic greater than or equal to the one on x-axis. We show the three background models described above, with the corresponding $\sigma$ bars on the top. Significance of GW150914 is greater than $5\sigma$ using all the background models. The significance of LVT151012 is $\sim 2\sigma$ using all the models. The significance of GW151226 changes a lot based on the model considered. We report $4.5\sigma$ based on the model that includes the candidates from GW151226, but excludes those from GW150914.

The IMBH search covered the parameter space of $50M_\odot \leq M = m_1 + m_2 \leq 600M_\odot$. In addition a maximum mass-ratio cut of 10 ($1 < \frac{m_1}{m_2} < 10$) was imposed. The spins on the black holes were restricted to be aligned to the orbital angular momentum but were allowed to have up to near maximal values of 0.99 ($-0.99 < s_1, s_2 < 0.99$). The IMBH search space overlapped with the main search space, and detected the three events found by the main search but with a lower significance. Since parameter-estimation studies have placed these events outside of the IMBH mass range [126, 127], these triggers were removed from the IMBH analysis. The IMBH search reported no other significant events [128].

The sub-solar mass search covered the parameter space of ultracompact binary systems with component masses between $0.2M_\odot$ and $1.0M_\odot$. They used non-

spinning templates, and a starting frequency of 40 Hz to restrict the size of the template bank. Their null search result provided a constraint on the rate of such objects to be less than $1.9 \times 10^4 \mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ (at 90% confidence) [129, 130].

The next chapter describes the developments made in the `GstLAL` pipeline before and during O2, and reports the results from the combined O1 and O2 data.

*Chapter 7*

# OBSERVING RUN2

After our successful first observing run O1 (September 12, 2015 - January 12, 2016), the Advanced LIGO detectors were upgraded to increase their sensitivity for the second observing run, O2, (November 30, 2016 - August 26, 2017). The Advanced Virgo detector joined the second observing run on August 1, 2017. Fig. 7.1 shows the strain sensitivity of the three Advanced detectors during O2. We discuss the updates that were applied to the `GstLAL` pipeline in this before and during O2, and then present the results from the O2 search.

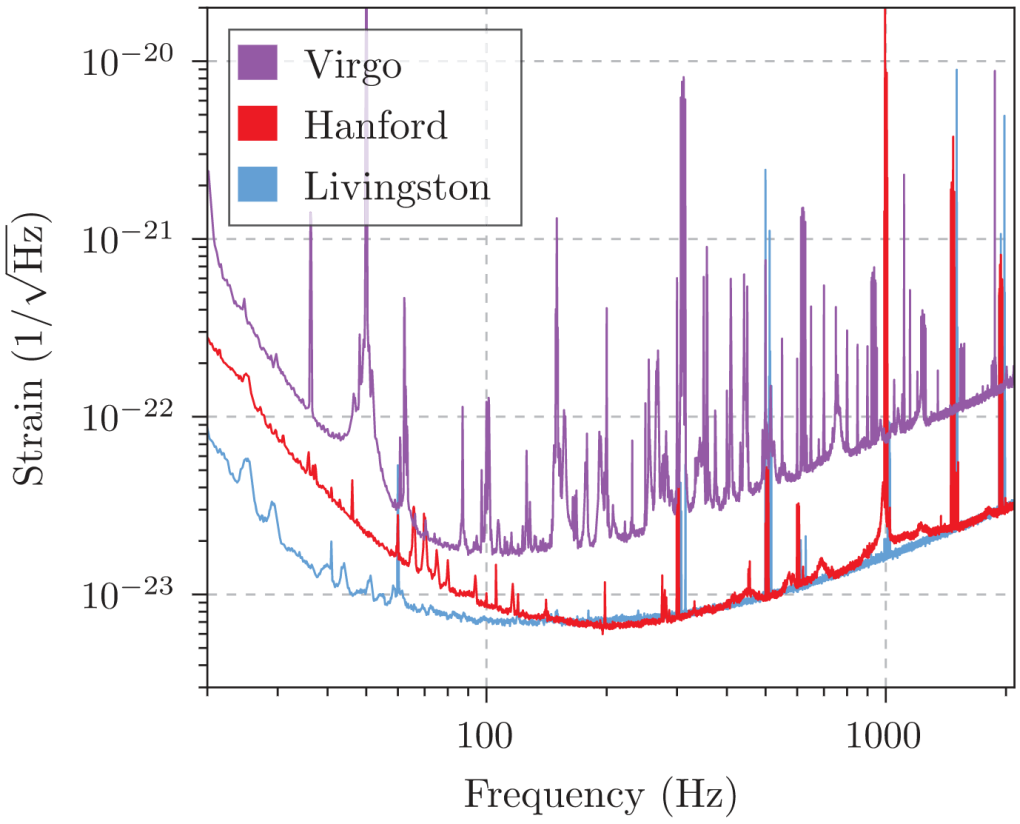



Figure 7.1: Amplitude spectral density of strain sensitivity of the Advanced LIGO-Virgo network. Figure taken from [10].

## 7.1 Developments in the GstLAL pipeline for O2

In Sec. 4.2, we discussed the methods used in the GstLAL pipeline for the O1 search. In this section, we describe the updates that were made in the pipeline before or during the O2 search. These updates include deployment of a zero-latency whitening filter to reduce the over-all latency of the pipeline by up to 32 seconds, incorporation of the Virgo data stream in the analysis, introduction of a single-detector search to analyze data from the periods when only one of the detectors is running, addition of new parameters to the likelihood ratio ranking statistic, increase in the parameter space of the search, and introduction of a template mass-dependent glitch-excision thresholding method.

### 7.1.1 Template bank and template bank splitting

As described in Sec. 3.2.3, template banks are a discrete set of waveforms that ensure that the SNR loss due to signal and template waveform mismatch will not be greater than a pre-specified threshold.

For Advanced LIGO's second observing run, the parameter space of the template bank was increased from a total mass of $2M_{\odot}$-$100M_{\odot}$ to $2M_{\odot}$-$400M_{\odot}$. Neutron stars were assumed to have masses less than $2M_{\odot}$ with dimensionless spin parameters in the range (-0.05, 0.05). Black holes were assumed to have masses greater than $2M_{\odot}$ with dimensionless spins in the range (-0.999, 0.999). Individual masses of the systems lie in the range 1-399 $M_{\odot}$ with mass ratios in the range 1-97.989. The templates were placed in two stages, below a total mass of 4 $M_{\odot}$ the inspiral-only templates of Post Newtonian approximation, `TaylorF2`, [131] were layed down using a geometric technique [93], while above a total mass of $4M_{\odot}$ the full inspiral-merger-ringdown templates of effective-one-body approximation hybridized with numerical relativity and black hole perturbation theory, `SEOBNRv4_ROM`, [77] were placed using stochastic methods [89]. Refer [132] for a detailed description of the construction of the bank and effectualness studies. The original templates that were placed are shown in green (BNS), red (NSBH), and blue (BBH) in Fig. 7.2. The colors denote the different source classes and the range of spins covered by the template bank for each. The purple and the black points show the templates that were needed to be added in order to get a good background estimation in the high mass parameter space. The details are discussed in the following paragraphs.

For analyzing the O2 data in an offline mode, we waited until the Hanford and Livingston detectors had recorded coincident data amounting to 5 days before

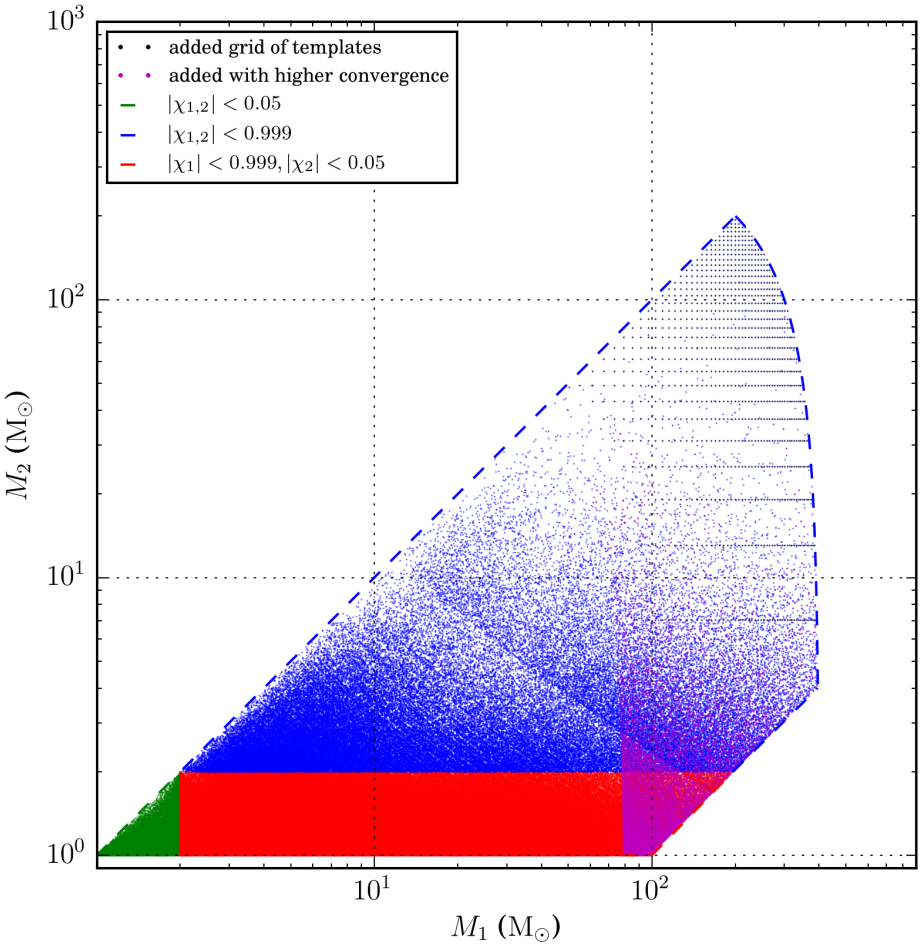



Figure 7.2: A visual representation of the bank used in the second observing run in the component mass space. Each point here represents a template in the bank. All regions shown in the figure are discrete points, the continuous regions being the highly dense regions. The bank used for O2 shown here is known as the hyperbank.

conducting an analysis. This is similar to what was done in O1, see Sec. 6.2.3. We call these the "offline chunks". In O2, we had a total of 22 such chunks of data. It wasn't until Chunk-5 that we started seeing problems in the closed box money plots (Sec. 5.2.4). Fig. 7.3 shows the money plot for the GstLAL analysis of Chunk-5. The closed box result should closely follow the background model when pipeline is only treating coincident events as gravitational-wave candidates. However, we see an excess in the closed box result compared to the background predicted by the pipeline. This is bad because this means that the results from this analysis cannot be trusted because the background model is insufficient to explain the null result. We

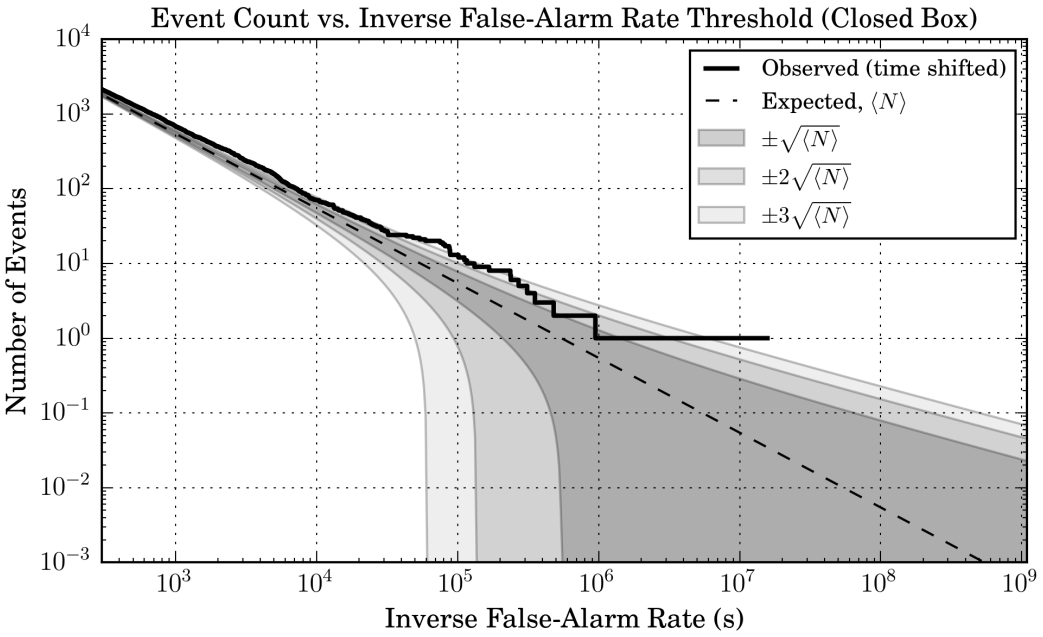



Figure 7.3: Cumulative candidate count vs. IFAR plot for the GstLAL analysis of O2, Chunk-5. We expect the closed box result to follow the noise model, which it does not here.

saw a similar trend in the closed box results of the next few chunks. Opening of the GstLAL boxes was put on hold until we could identify and fix this problem.

Thus began a hunt for the cause of discrepancy between the background model and the closed box result. After several investigations, it was found that the template bank, together with the grouping scheme of templates that was used in the first observing run of Advanced LIGO detectors [3] for collecting noise statistics was inadequate to construct a sufficiently accurate background model for the high mass region that was added to the parameter space for this run. As a result, a different grouping scheme was introduced in the run, and the template bank was also modified as described below.

Recall that, for the purpose of background estimation in the pipeline, we first divide the entire template bank into several sub banks, the " $\bar{\theta}$ bins", containing "similar" waveforms, based on their intrinsic parameters. We want to group templates that respond similarly to noise together. Noise statistics are collected for all templates in a bin as a whole. These are typically formed by combining a certain number of "split banks", which were formed for performing the SVD (Sec. 4.2.2). In O2, we were grouping 2 split banks containing 500 templates each, in each bin. We

want the binning method to (a) group together similar templates to use the LLOID method (Sec. 4.2.2) for computationally-efficient time-domain searches, and (b) group together templates with similar noise statistics for accurate FAR estimates. Prior to the run, the pipeline used two composite parameters, which are a combination of the four intrinsic parameters, in order to group the waveforms into split banks: the chirp mass and the effective spin. The chirp mass and the effective spin parameter are the leading order terms that describe the phase evolution of the inspiral part of the waveform according to the Post Newtonian expansion (Sec. 3.1.1).

In Fig. 7.2 we see that the density of the templates decrease with the increase in the masses of the binaries. This is a typical feature seen in all template banks. The waveform of a system with smaller masses is longer in the frequency band that Advanced LIGO and Virgo detectors are most sensitive in (see Fig. 7.1). In O2, the `GstLAL` pipeline filtered the data from 15 Hz to 1024 Hz. Therefore even a small change in masses can lead to a big mismatch between two waveforms in this region since there are more cycles in band for which the match has to be performed. This means that we need more densely-spaced waveforms to cover the lower mass region of the parameter space.

The much lower density of the templates in the high mass region of the bank, which was added for the run caused templates vastly different in $\mathcal{M}_c$, but similar in $\chi_{\rm eff}$ to be grouped together in the low template density region, since the number of templates in each bin was required to be the same ($\sim 1000$). This in turn led to sub-optimal estimation of noise properties for these groups, which led to inaccurate significance estimation of noise events. These are the highest $\mathcal{M}$ groups in each $\chi_{\rm eff}$ bin, in Fig. 7.4 where we show the grouping of templates used for background estimation at the beginning of O2. In Fig. 7.5, we show the distribution of non-coincident triggers in SNR-$\xi^2$ space for two of these bins that span a large $\mathcal{M}_c$ range. The pipeline uses these distributions to construct the background model used to assign the likelihood-ratio statistic and ultimately the IFAR of candidates. For comparison we also show some well-behaved background bins. The features marked in white in the PDF happen when there are templates present in a bin whose properties are not well represented by the rest of the templates in the bin.

We tried several new binning methods to solve the issue of background estimation in the high-mass region. Here we document the the grouping scheme that was used in the end. We changed the grouping scheme only for these high mass templates, with total mass $> 80M_\odot$. Instead of using $\mathcal{M}$ and $\chi_{\rm eff}$, we use template duration in

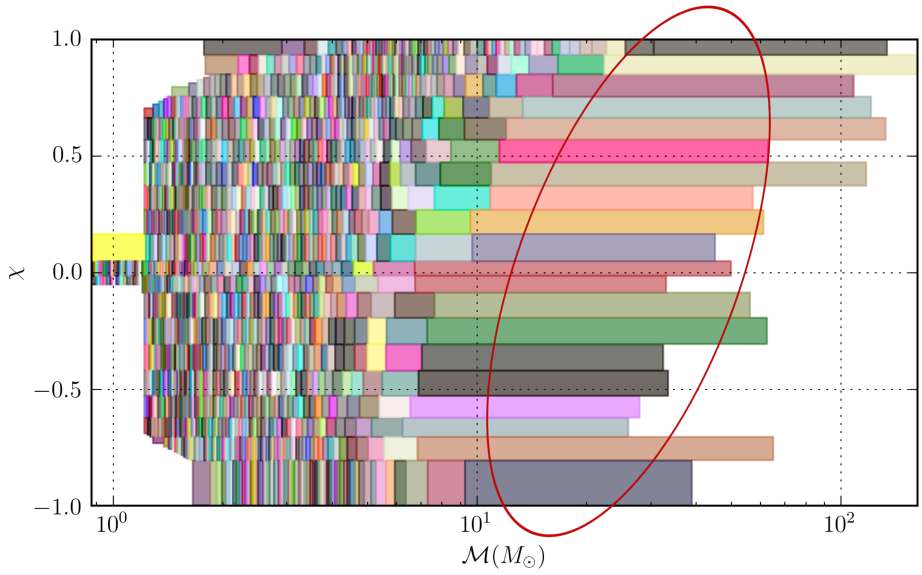



Figure 7.4: A visual representation of the old grouping of the templates in $\mathcal{M}$ - $\chi_{\text{eff}}$ space. The colors in this plot have no intrinsic meaning; they are chosen randomly in order to distinguish the different groupings into bins.

the LIGO sensitivity band (starting frequency 15 Hz) for grouping these high mass templates. $\mathcal{M}$ and $\chi_{\text{eff}}$ are the leading order parameters that describe the inspiral part of the waveforms, but not the merger and ringdown which have duration of only a small fraction of a second, for all templates. However, for the short-duration high mass systems, most of the power in the LIGO band comes from merger and ringdown phases of the coalescence; therefore $\mathcal{M}$ and $\chi_{\text{eff}}$ are no longer the best parameters to group the templates. The number of templates in each of these high-mass bins was also reduced to account for the sparse density of templates in this region. The highest mass bins now contain 200, 400, or 800 templates instead of 1000 templates. This helps in grouping only those templates that have the same noise properties together. 15,665 extra templates were also added to the bank above a total mass of 80 to increase the density of the high mass, short duration templates, thus helping to collect noise statistics in this region. This was done by increasing the convergence threshold from 0.97 to 0.99, and by adding a grid of templates, to increase the template density in that region. The templates that were added by increasing the convergence threshold of the stochastic placement process are shown in magenta, and those that were added as a grid are shown in black in Fig. 7.2. This brought the total number of templates of the final O2 bank to 677,000.

Finally, in Fig. 7.6, we show the money plot from the fixed GstLAL analysis of O2,

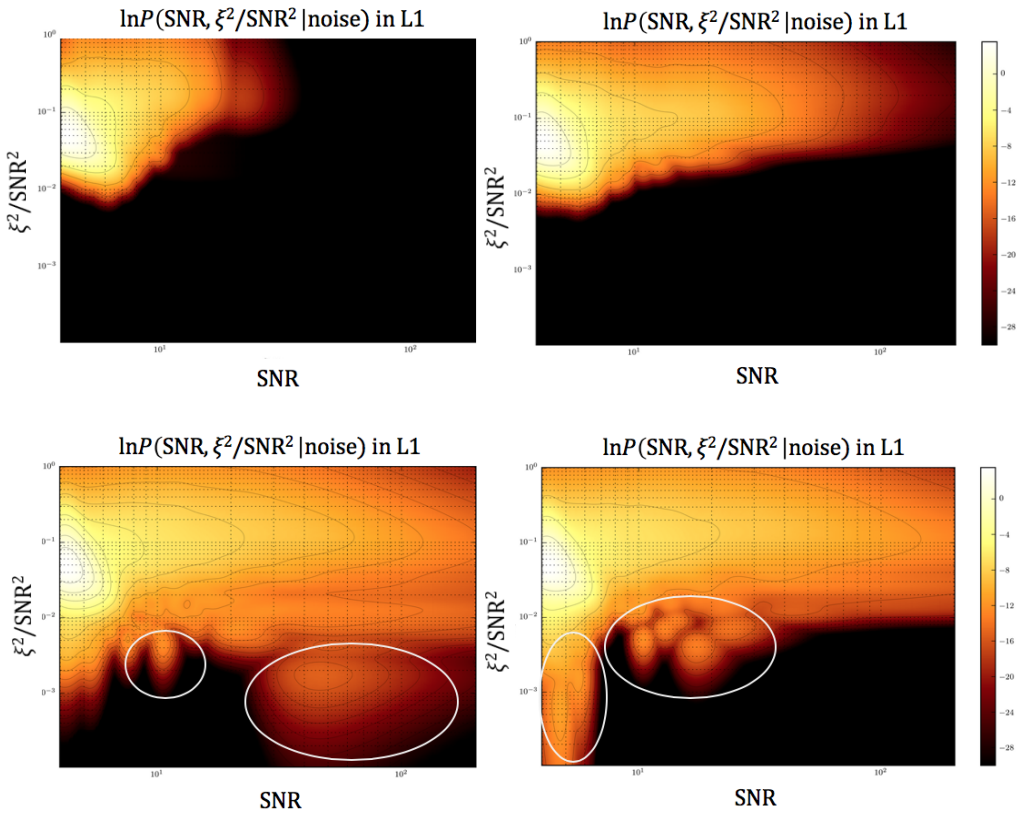



Figure 7.5: Example of well behaved SNR-$\xi^2$ bins (top). The PDF on top-left corresponds to a low-mass bin containing long duration templates which do not suffer from the non-Gaussian glitches in the detector as much as the high-mass, short duration templates. The PDF on the top-left shows an example of a relatively well behaved background high-mass bin. The PDFs on the bottom correspond to some of the "bad" background bins. We draw the reader's attention to some of the island features inside the white circles. These occur when the pipeline doesn't have enough statistics about the behavior of some of the templates in the bin that are different from the majority.

Chunk-5 that uses the new binning method and the modified bank.

We could now continue analyzing the O2 data, and look at the open box results.

### 7.1.2 Zero latency whitener

The original whitening filter employed in GstLAL contributes to one of the bottleneck processes in the pipeline's latency, adding up to 32 seconds to the total latency time [3]. The zero latency whitener was introduced in order to reduce the latency of the pipeline [133].

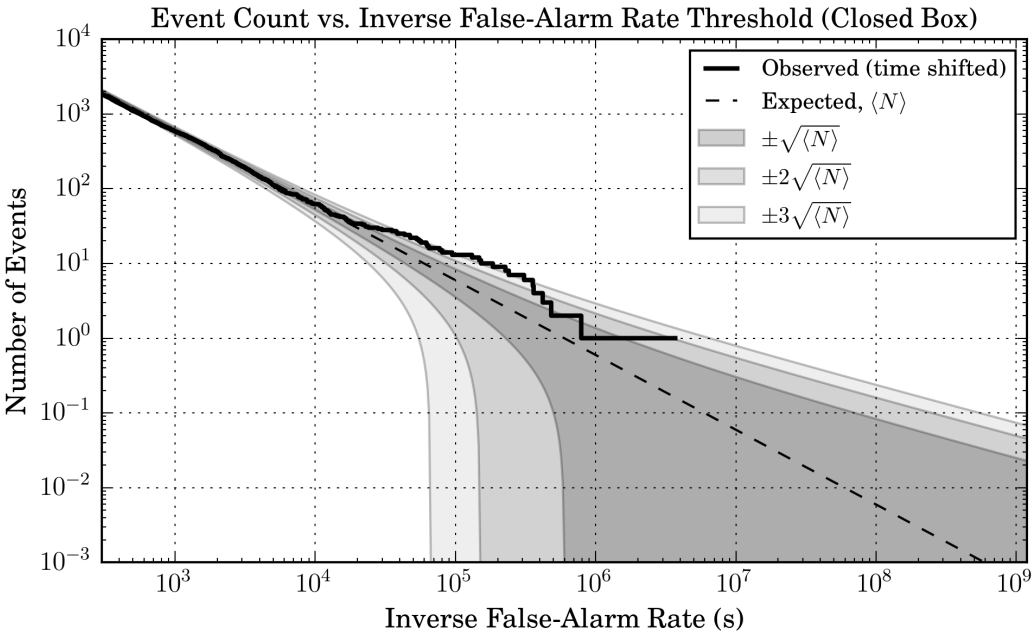



Figure 7.6: Cumulative candidate count vs. IFAR plot from the fixed GstLAL analysis of O2, Chunk-5 that uses the new binning method, and the modified bank.

As described in Sec. 3.2.1, a matched filter is the inner product between a template waveform and the strain data. However, since the strain data are strongly colored by the frequency-dependent noise of the detector, we 'whiten' the inner product by weighting both the data and the template waveform by a factor of $1/\sqrt{S_n(f)}$ each, where $S_n(f)$ is the single-sided power spectral density of the detector noise. The method of a frequency-domain whitening filter described in [3] has discrete Fourier transforms and window functions applied to 32-second blocks of input data, with the PSD being updated every 16 seconds. Since a 32-second block is processed every 16 seconds, this filter has a latency between 16 to 32 seconds, which is almost half of the pipeline's total latency.

To reduce the latency, we need to whiten in the time-domain. To this end, a Finite Impulse Response (FIR) filter-based algorithm to the frequency-domain whitening is introduced. The square root of inverse PSD of a given strain data is used to construct the FIR of a linear phase filter. This filter still requires 16 seconds of data from the future for its evaluation [133]. It is not possible to further reduce the latency of this filter without changing the whitening transformation. Therefore an approximation to the original filter is introduced [133], which derives a minimum-phase approximation of the desired filter [134]. The matched filter output depends on $1/S_n(f)$ therefore is

insensitive to the error in phase response of the whitening filter, i.e., error in the phase response of $1/\sqrt{S_n(f)}$. The minimum-phase whitening filter accurately approximates the amplitude response of the FIR-based filter, introducing most errors only in the phase response. This filter does not use any information from future samples for its evaluation, and is therefore called a "zero-latency whitening filter". A new windowing process has also been implemented for the new whitening method. The PSD transition is now allowed to occur continuously, and the resulting filter is a linear combination of the newest and next newest filters during their transition. This is recursively applied to the zero-latency algorithm when a new whitening filter becomes available. It is shown in [133] that the causal FIR filter approximation successfully whitens the data, producing zero mean, unit variance, white Gaussian noise. Matched-filter outputs (SNR and $\xi^2$, Sec. 3.2.1) produced using the zero-latency whitener are compared to those produced using the frequency-domain whitener for both noise and simulated signal triggers, and we see a good agreement for both. The details of the filter, and the consistency checks with the old filter are described in [133]. Fig. 7.7 shows the comparison of the time taken by the `GstLAL` online pipeline to upload an event in O1 vs. in O1. We can see that the latency of the pipeline has reduced by $\sim 40$ s as compared to O1, in which the zero-latency whitener played a significant role.

### 7.1.3 Data conditioning

The output of the matched filter is the SNR (Sec. 3.2.1), which is the optimal detection statistic under the assumption that noise is stationary and Gaussian. For offline analyses, where the pipeline processes archival gravitational-wave data, we use data quality vetoes (Sec. 6.2.2) to flag poor data. However, such information is not available for online analyses. Gating on the whitened strain data, whitened $h(t)$, is one of the techniques adopted by the pipeline to eliminate short transient instrumental noise fluctuations. These fluctuations, also known as glitches, can cause unreasonably high values of SNRs in the data, mimicking gravitational-wave signals, and causing false alarm triggers.

In presence of glitches, the whitened $h(t)$ may have values higher than the expected values from the coalescences of binary systems that the pipeline is aiming to detect. By construction, whitened $h(t)$ should have a unit variance. Whenever the magnitude of whitened $h(t)$ is momentarily greater than a threshold value, set as some multiple of the standard deviation $\sigma$ of $h(t)$, we gate that piece of data by setting the samples around the peak with a window of 0.25 s on each side to zero [3]. Refer Fig. 6.7.

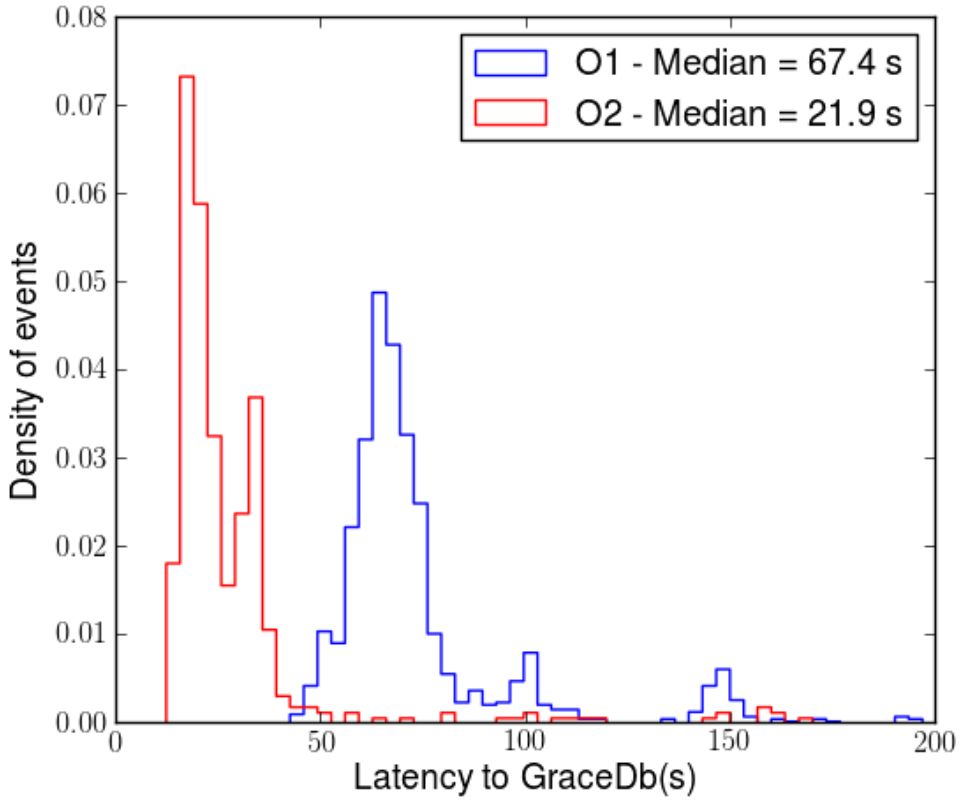



Figure 7.7: Comparison of the latency of detections by the `GstLAL` online pipeline in O1 (blue) vs. in O2 (red). We can see that the latency of the pipeline has been reduced by $\sim 40$ s. The application of the zero-latency whitener played a significant contribution in decreasing the latency of the pipeline.

The amplitude of a signal at a fixed effective distance increases with the chirp mass $\mathcal{M}$ of a binary system. Therefore we set the threshold based on the highest masses we are sensitive to, such that it is higher than the whitened strain amplitude we expect from such systems. We want the threshold to be such that it removes the maximum number of glitches from our data without gating out real signals. During the first observing run, the threshold value was set to $50\sigma$. At lower threshold values, it was seen that the pipeline started to gate some of the high mass and/or nearby (small $d_L$) simulated BBH signals that were injected in the data.

As described in Sec. 7.1.1, in the second observing run, the parameter space of our search was increased from a maximum total mass of $100M_\odot$ to $400M_\odot$. In order to avoid gating the highest mass signals, we would have to increase our value of the gate threshold. This would cause an increase in the number of glitches that pass through

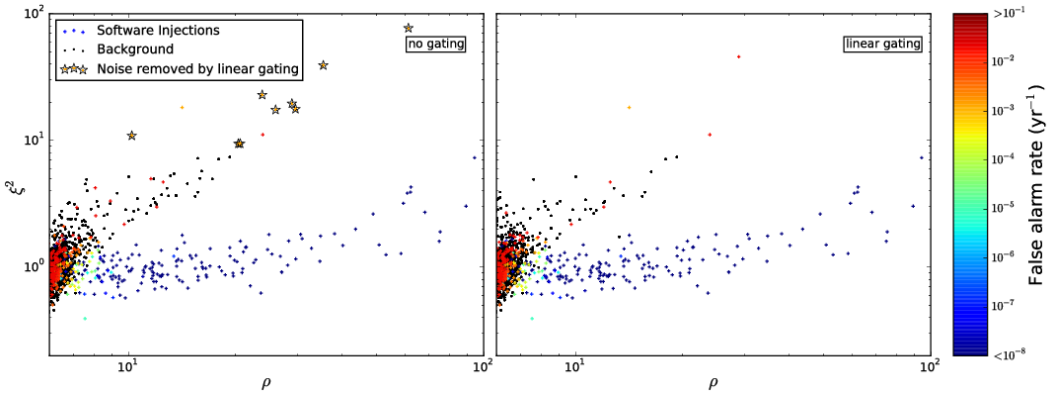



Figure 7.8: Comparison results between a run with constant gating scheme (left) and linear gating scheme (right). The yellow stars highlight the additional background glitches present in the constant gating run that have been removed in the run with the linear gating scheme applied. No software injections were gated.

without being gated, and therefore an increase in the number of false alarm triggers. Thus a linear gating scheme, in which the gate threshold value is a linear function of chirp mass instead of a constant for all masses, was introduced. The pipeline computes an appropriate threshold value according to the linear scale provided by the user and the highest chirp mass template in a sub-bank. Fig. 7.8 shows that linear gating scheme helps in removing more glitches as compared to the constant gating scheme, while still recovering all the simulated signals injected in the data.

### 7.1.4 Likelihood-ratio statistic

Once all the candidate events are identified by the pipeline, each of them is assigned a likelihood-ratio statistic in order to determine its significance. The likelihood ratio as defined in [3, 98] was used in the analysis of data from the first observing run of Advanced LIGO, during which only the Hanford and Livingston detectors were operating. It is the ratio of probability of certain observables given the signal hypothesis versus the noise hypothesis. Included observables were terms to account for the set of detectors involved in the coincidence {H1, L1}, the detector horizon distances (for a $1.4M_\odot$ - $1.4M_\odot$ system) at the time of coincidence $\{D_{\rm H1}, D_{\rm L1}\}$, the SNRs for each trigger $\{\rho_{\rm H1}, \rho_{\rm L1}\}$, and the $\xi^2$-signal-based-veto values for the each trigger $\{\xi^2_{\rm H1}, \xi^2_{\rm L1}\}$. For Advanced LIGO's second observing run, the likelihood-ratio statistic was modified to allow the pipeline to rank single detector events when only one detector is operational and to include additional parameters when both advanced LIGO facilities are operational. These modifications were introduced before the start of O2. The ranking of single detector events is an active area of research

(although it has already proven to be successful in the case of GW170817!), and has not been used while analyzing the data offline as yet. It was used by the online pipeline sparingly in O2. The inclusion of time and phase delay terms between the two detectors as introduced before O2 (let's call it LR1) was used by the online pipeline, and the offline pipeline for preparing the closed box results and in the results reported in [16, 17]. However, after the data taking for O2 had finished, changes were introduced in the likelihood-ratio statistic to include time and phase delay terms while generalizing it to an arbitrary number of detectors. We will refer to this generalized likelihood-ratio statistic as LR2.

### 7.1.5 Single detector events

In Advanced LIGO's first observing run, the pipeline could only identify gravitational-wave events when both the detectors were operating. Gravitational-wave candidates were formed by demanding coincidence between the two detectors. This meant that we were blind to signals occurring during single-detector time (defined as time when only one detector is operational) even though there was science quality data available from that detector. For O2, the online analysis also looked at the non-coincident candidates, and these were also assigned a log likelihood ratio statistic.

Non-coincident triggers found during single-detector time (defined as time when only one detector is operational) were now excluded from informing the background model, since these could potentially be loud signals that were found as non-coincident triggers in absence of data from multiple detectors. Since we use an SNR threshold of 4 for triggers, there are too many non-coincident triggers to write to disk. Therefore the non-coincident triggers are first assigned a preliminary log likelihood-ratio, and only those that have $\ln\mathcal{L}_{prelim} > 2$ are considered as gravitational-wave candidates. The likelihood-ratio defined in [3, 98] and Sec. 4.2.5 (Eq. 4.7) is still valid for single detector events,

$$\begin{aligned}\mathcal{L}\left(\{D_{\mathrm{IFOnet}}\},\{\mathrm{IFO}\},\rho_{\mathrm{IFO}},\xi^2_{\mathrm{IFO}},\bar{\theta}\right) &= \mathcal{L}\left(\{D_{\mathrm{IFOnet}}\},\{\mathrm{IFO}\},\rho_{\mathrm{IFO}},\xi^2_{\mathrm{IFO}} \mid \bar{\theta}\right)\mathcal{L}\left(\bar{\theta}\right)\\ &= \frac{P\left(\{D_{\mathrm{IFOnet}}\},\{\mathrm{IFO}\},\rho_{\mathrm{IFO}},\xi^2_{\mathrm{IFO}} \mid \bar{\theta},\mathrm{s}\right)}{P\left(\{D_{\mathrm{IFOnet}}\},\{\mathrm{IFO}\},\rho_{\mathrm{IFO}},\xi^2_{\mathrm{IFO}} \mid \bar{\theta},\mathrm{n}\right)}\mathcal{L}\left(\bar{\theta}\right).\end{aligned} \tag{7.1}$$

Here, $\{D_{\mathrm{IFOnet}}\}$ is the set of horizon distances for all instruments in the network at the time the event is observed, {IFO} is the detector that produced the non-coincident trigger, $\rho_{\mathrm{IFO}}$, and $\xi^2_{\mathrm{IFO}}$ are the SNR and $\xi^2$ values of the trigger. The numerator and

denominator of the fraction in Eq. 7.1 are factored in the same way as described in [3, 98]. In particular the factorization leads to a form,

$$\mathcal{L}\left(\{D_{\mathrm{IFOnet}}\},\{\mathrm{IFO}\},\rho_{\mathrm{IFO}},\xi^2_{\mathrm{IFO}},\bar{\theta}\right) = \cdots \times \frac{P\left(\{\mathrm{IFO}\} \mid \{D_{\mathrm{IFOnet}}\},\bar{\theta},\mathrm{s}\right)}{P\left(\{\mathrm{IFO}\} \mid \bar{\theta},\mathrm{n}\right)} \times \cdots . \tag{7.2}$$

The probability that a signal yields a trigger above-threshold in only one of the detectors depends on the horizon distances of all the detectors operating at the time, and the duty cycles of all the detectors in the network. The probability that noise yields an above-threshold trigger in one of the detectors is computed from the trigger rates, coincidence window size, and the duty cycles of all the detectors in the network. If more than one detector is operating, and the signal is only seen above-threshold in one of the detectors then that becomes a constraint on the SNR distribution. The single-variable SNR distribution is computed using a Monte Carlo generation of samples described in [98]. In the case where only one detector is operating, the SNR distribution reduces to $P(\rho) \propto \rho^{-4}$. Even though we were still testing the ranking of single-detector events, including this feature in the online `GstLAL` analysis proved to be the biggest success of O2. On August 17, 2018, at 12:41:04 UTC the Advanced LIGO and Advanced Virgo gravitational-wave detectors made their first observation of a binary neutron star inspiral. The signal, GW170817, was detected by the `GstLAL` online search as a single-detector event in Hanford, because of the presence of an instrumental noise transient in the Livingston detector 1.1 s before the coalescence time of GW170817. Refer Fig. 7.9. Because of the glitch, the online pipeline did not register a coincident candidate between Hanford and Livingston at the time of the event, and had it not been for the single detector search, we would not have been able to detect the BNS event in low-latency and possibly missed out on the followups that led to the detection of an optical transient counterpart.

### 7.1.6 Inclusion of phase and time delay terms in the likelihood statistic

As discussed above, the inclusion of time and phase delay terms in the likelihood statistic was done in two different ways. The first (LR1) assumes only two detectors, Hanford and Livingston, and was used by the online and the offline pipeline during O2, and in the early O2 detections. The other method (LR2) was introduced after O2, in order to include Virgo in the detection statistic as well, and was used by the offline `GstLAL` pipeline in the re-analysis of combined O1 and O2 cleaned data using the final version of calibration. These results will be reported in [135].

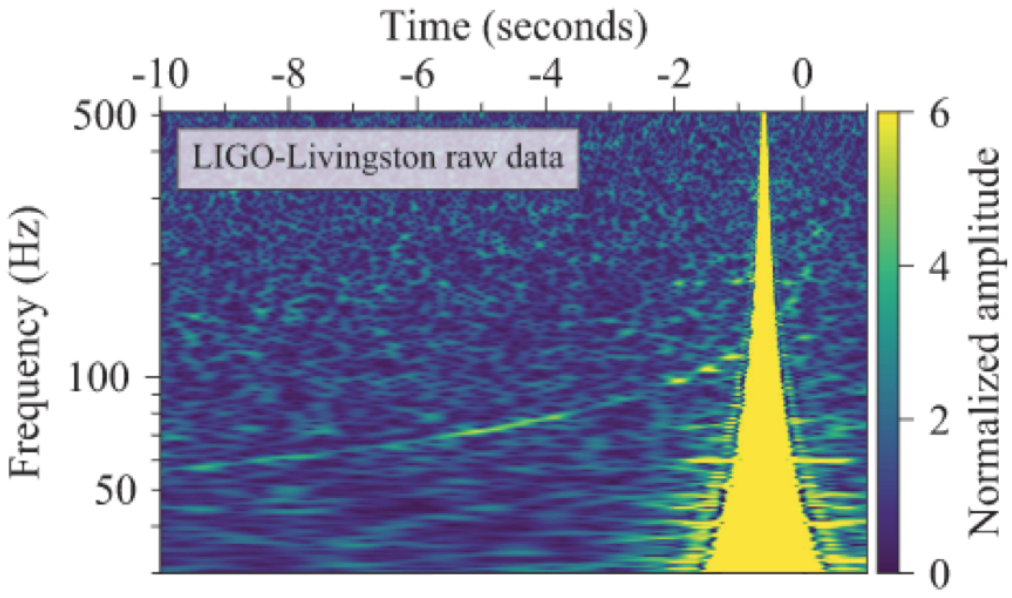



Figure 7.9: Time-frequency representation of the raw Livingston data at the time of GW170817. The 0 in time marks the coalescence time of GW170817. There is a loud glitch at -1.1 s in the figure because of which the online pipelines analyzing the data at the time did not recover the event as an H1L1 double. The time-frequency track of GW170817 is clearly visible despite the presence of the glitch. Figure taken from [11].

**LR1**

Two additional parameters were added to the likelihood-ratio statistic for the case where the Hanford (H1) and Livingston (L1) advanced LIGO detectors are both operational: $\Delta t$, the difference in end times between the H1 and L1 triggers, and $\Delta\phi$, the difference in coalescence phase between the H1 and L1 triggers. We require $\Delta\phi \in [-\pi, \pi]$ and compute the modulus $\Delta\phi \,(\mathrm{mod}\, 2\pi)$ to enforce a cyclic distribution.

This new likelihood-ratio statistic is defined as follows:

$$\mathcal{L} = \frac{P\left(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}\}, \{\mathrm{H1}, \mathrm{L1}\}, \rho_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \xi^2_{\mathrm{H1}}, \xi^2_{\mathrm{L1}}, \Delta\phi, \Delta t \middle| \,\mathrm{signal}\right)}{P\left(\{D_{\mathrm{H1}}, D_{\mathrm{L1}}\}, \{\mathrm{H1}, \mathrm{L1}\}, \rho_{\mathrm{H1}}, \rho_{\mathrm{L1}}, \xi^2_{\mathrm{H1}}, \xi^2_{\mathrm{L1}}, \Delta\phi, \Delta t \middle| \,\mathrm{noise}\right)}. \tag{7.3}$$

This statistic only supports ranking coincidences found with the H1L1 network. Future work for the third observing run will add support for the H1L1V1 network with a goal towards a generalized N-detector network statistic. Additionally, we make several assumptions when factoring the dependencies of the probability density functions in Eq. 7.3. We assume that the noise distributions of $\Delta t$ and $\Delta\phi$ are independent of each other. As in [3], we assume that the $\xi^2$ statistic is dominated by

instrumental noise, thus the $\xi^2$ term reduces to its previous form. We expect that the signal (but not the noise) distributions for $\Delta t$ depend on trigger SNRs as well as on detector sensitivities. The SNR ratio from the two detectors for a signal depends on the position source with respect to the detectors, and the inherent sensitivities of the detectors. $\Delta t$ for a signal depends only on the position of the source and the location of the detectors. Thus we model the $\Delta t$ distributions as a function of a ratio of SNRs normalized by horizon distances, to factor out the inherent sensitivities of the detectors from the SNR ratios,

$$\rho_{\text{ratio}} = \frac{\rho_1/D_1}{\rho_2/D_2}. \tag{7.4}$$

On the other hand, we do not consider dependence on the detector sensitivities when modeling the $\Delta\phi$ distributions. We only consider dependence of $\Delta\phi$ on $\Delta t$ and network SNR, defined as

$$\rho_{\text{network}} = \sqrt{\rho_1^2 + \rho_2^2}. \tag{7.5}$$

With these assumptions, we can factor Eq. 7.3 as

$$\begin{aligned} \mathcal{L} &\propto \frac{P\left(\Delta t|\,\{\rho\}, \{\text{Horizon}\}, \text{signal}\right) P\left(\Delta\phi|\,\Delta t, \{\rho\}, \{\text{Horizon}\}, \text{signal}\right)}{P\left(\Delta t|\,\text{noise}\right) P\left(\Delta\phi|\,\text{noise}\right)} \\ &\approx \frac{P\left(\Delta t|\,\{\rho_{\text{ratio}}\}, \text{signal}\right) P\left(\Delta\phi|\,\Delta t, \{\rho_{\text{network}}\}, \text{signal}\right)}{P\left(\Delta t|\,\text{noise}\right) P\left(\Delta\phi|\,\text{noise}\right)}. \end{aligned} \tag{7.6}$$

In order to construct distributions for $P\left(\Delta t|\{\rho_{\text{ratio}}\}, \text{signal}\right)$ and $P\left(\Delta\phi|\Delta t, \{\rho_{\text{network}}\}, \text{signal}\right)$, we performed an injection (simulated signal injected in the data to test the performance of the pipeline) campaign in the data from the first observing run covering the parameter space of binary neutron star and binary black hole systems. We discarded any injection recovered with $|\Delta t| > 0.01$, assuming that this regime is dominated by accidental coincidences. Smoothed, normalized $\Delta t$ histograms were modeled as a function of $\rho_{\text{ratio}}$ using Chebyshev polynomials. Fig. 7.10 shows the logarithm of the PDF for $P\left(\Delta t|\{\rho_{\text{ratio}}\}, \text{signal}\right)$ as defined over a range of $\Delta t$ and $\rho_{\text{ratio}}$. Normalized histograms of $\Delta\phi$ were modeled using von Mises distributions which is a continuous probability distribution on a circle since $\Delta\phi$ is cyclic, plus a uniform noise background as a function of $\Delta t$ and $\rho_{\text{network}}$. In Fig. 7.11, we see the logarithm of the PDF for $P\left(\Delta\phi|\Delta t, \{\rho_{\text{network}} = 14\}\right)$ as defined over a range of $\Delta t$ and $\Delta\phi$.

In order to construct the distributions for $P\left(\Delta t|\text{noise}\right)$ and $P\left(\Delta\phi|\text{noise}\right)$, we simply assume $\Delta t$ and $\Delta\phi$ are uniformly distributed for noise triggers that form false coincidences.

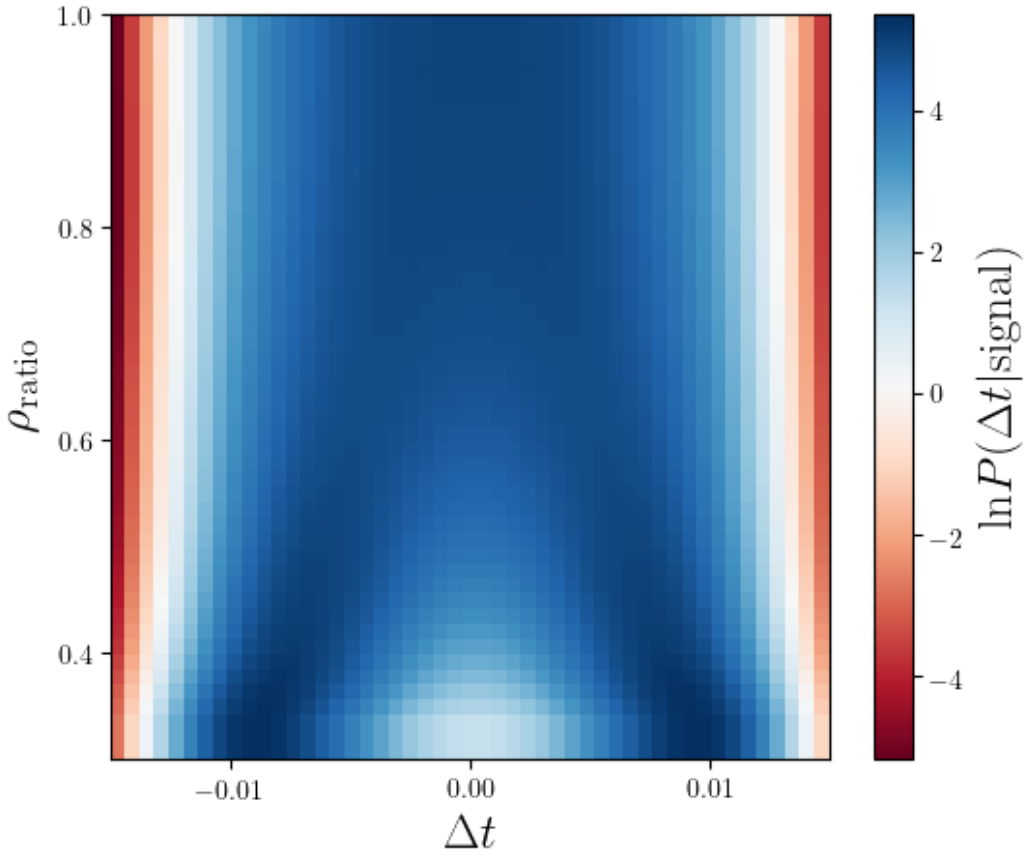



Figure 7.10: Distribution of $P(\Delta t|\{\rho_{\mathrm{ratio}}\}, \mathrm{signal})$

### 7.1.7 Introducing Virgo: LR2

The Advanced Virgo [20] detector joined the second observing run of the Advanced LIGO detectors on August 1st, 2017. The Advanced Virgo detector operated at a lower sensitivity relative to the Advanced LIGO detectors (see Fig. 7.1). During O2, the online search filtered over the Virgo data stream, but did not use the outputs of the matched-filtering in the likelihood-ratio statistic. A lower SNR threshold of 3 for triggers (Sec. 4.2.4) was used for Virgo (the threshold is 4 for the LIGO detectors) due to its lower sensitivity. In the first observing run, gravitational-wave candidates were formed by demanding coincidence (both in time and template) between the LIGO Hanford and LIGO Livingston triggers [3], Sec. 4.2. However in O2, due to the lower sensitivity of Virgo in the second observing run, if Virgo participated in a coincidence with either one of the two detectors, the set of triggers was still considered as a single-detector, non-coincident event. In other words H1V1 and L1V1 doubles were treated as H1 and L1 singles respectively. And if Virgo participated in a triple event (H1L1V1 triple), it was still treated as an $HL$ double.

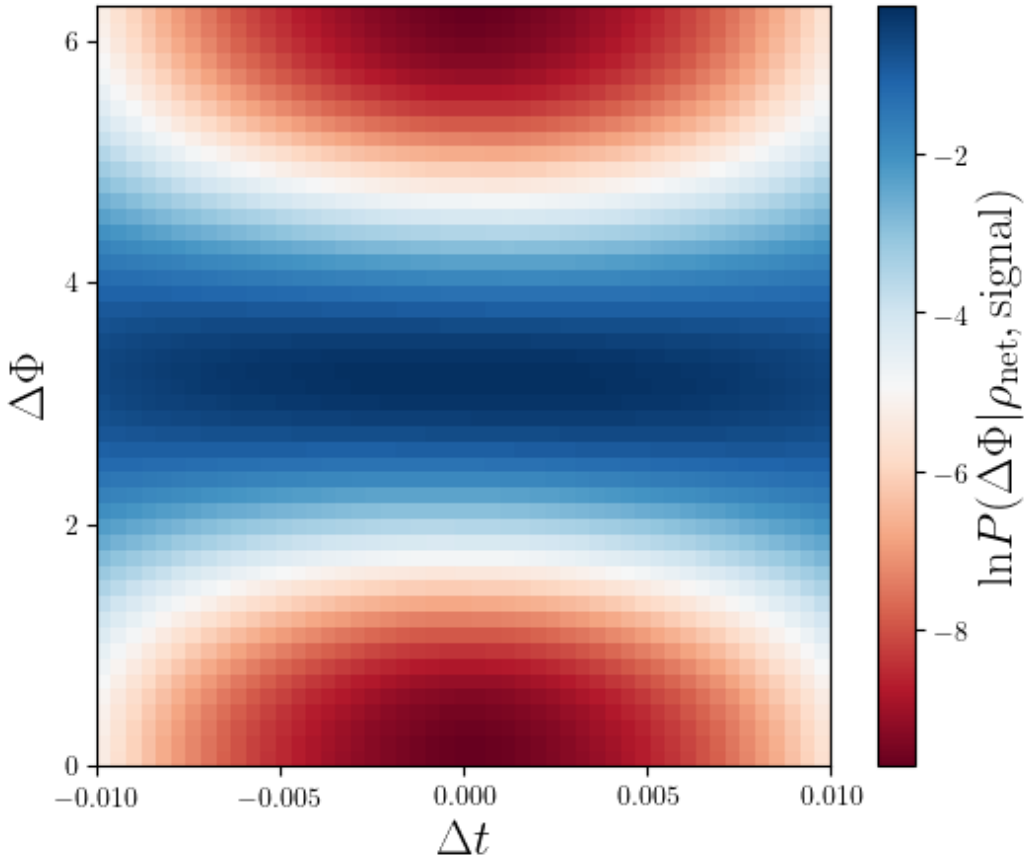



Figure 7.11: Distribution of $P\left(\Delta\phi|\Delta t, \{\rho_{\rm network}\}\right)$, where we have set $\rho_{\rm H1} = 10$ and $\rho_{\rm L1} = 10$.

Nonetheless a network of three detectors improves the sky localization of the source. GW170814 was the first gravitational-wave event that had a significant SNR in Virgo. Including Virgo in the analysis helps in reducing the area of the 90% credible region from 1160 deg$^2$ when using only the two LIGO detectors to 60 deg$^2$ when using all three detectors for the binary black hole event GW170814 [10]. Using the fact that Virgo did not pick up GW170817 with appreciable SNR also helped reduce the 90% credible region in the skymap from 190 deg$^2$ to 31 deg$^2$ [11, 12].

Before the reruns of O1 and O2 cleaned and recalibrated data, modifications were made to the likelihood-ratio statistic to also include the outputs of matched-filtering with Virgo as terms for the likelihood-ratio statistic. The denominator in the likelihood-ratio statistic, Eq. 7.3, is still computed like as in the LR1 method. We

also begin factoring the numerator in a similar manner,

$$P\left(\vec{D}_H, \vec{O}, \vec{\rho}, \vec{\xi}^2, \vec{\Delta\phi}, \vec{\Delta t}|\text{signal}\right) = P\left(\vec{D}_H|\text{signal}\right) P\left(\vec{O}, \vec{\rho}, \vec{\xi}^2, \vec{\Delta\phi}, \vec{\Delta t}|\vec{D}_H, \text{signal}\right)$$
$$= P\left(\vec{D}_H\right) P\left(\vec{O}|\vec{D}_H, \text{signal}\right) P\left(\vec{\rho}, \vec{\xi}^2, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right). \tag{7.7}$$

But we choose a different factorization for the $P\left(\vec{\rho}, \vec{\xi}^2, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right)$ term in the numerator:

$$P\left(\vec{\rho}, \vec{\xi}^2, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right) = P\left(\vec{\rho}, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right) P\left(\vec{\xi}^2|\vec{\rho}, \vec{\Delta\phi}, \vec{\Delta t}, \vec{O}, \vec{D}_H, \text{signal}\right)$$
$$= P\left(\vec{\rho}, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right) P\left(\vec{\xi}^2|\vec{\rho}, \vec{O}, \vec{D}_H, \text{signal}\right), \tag{7.8}$$

where we have assumed that the distribution of the $\xi^2$ values for signals is independent of $\vec{\Delta t}$ and $\vec{\Delta\phi}$. The term $P\left(\vec{\rho}, \vec{\Delta\phi}, \vec{\Delta t}|\vec{O}, \vec{D}_H, \text{signal}\right)$ is calculated using a Monte Carlo integration of signals, and the term $P\left(\vec{\xi}^2|\vec{\rho}, \vec{O}, \vec{D}_H, \text{signal}\right)$ is calculated as before in LR1 or the O1 model, which was used in O1.

To compare the different likelihood models - O1, LR1, and LR2, we generate 10,000 noise and signal samples, and we assign the likelihood-ratio values using the different methods. The results are shown in Figs. 7.12, 7.13, and 7.14.

### 7.1.8 Computing joint SNR PDF for different horizon distance ratios

The horizon distance $D_h$ is the effective distance at which a binary system is observed with a nominal SNR of 8 [84]. This means that horizon distance is a measure of a detector's sensitivity to a particular system. The horizon distances included in the ranking statistic of the pipeline are computed for a $1.4M_\odot - 1.4M_\odot$ binary neutron star system, and their fluctuations reflect fluctuations in the noise spectrum.

One of the factors in the numerator of the likelihood-ratio ranking statistic is the joint SNR PDF given a set of horizon distances of all the detectors at the time of the event, and the set of detectors that observed the event with an SNR above the threshold. [98]. For candidates that arise from genuine signals, the joint SNR PDF depends only on the ratios of the horizon distances. In the first observing run, the pipeline was limited to two detectors and assumed a fixed joint SNR PDF, corresponding to equal

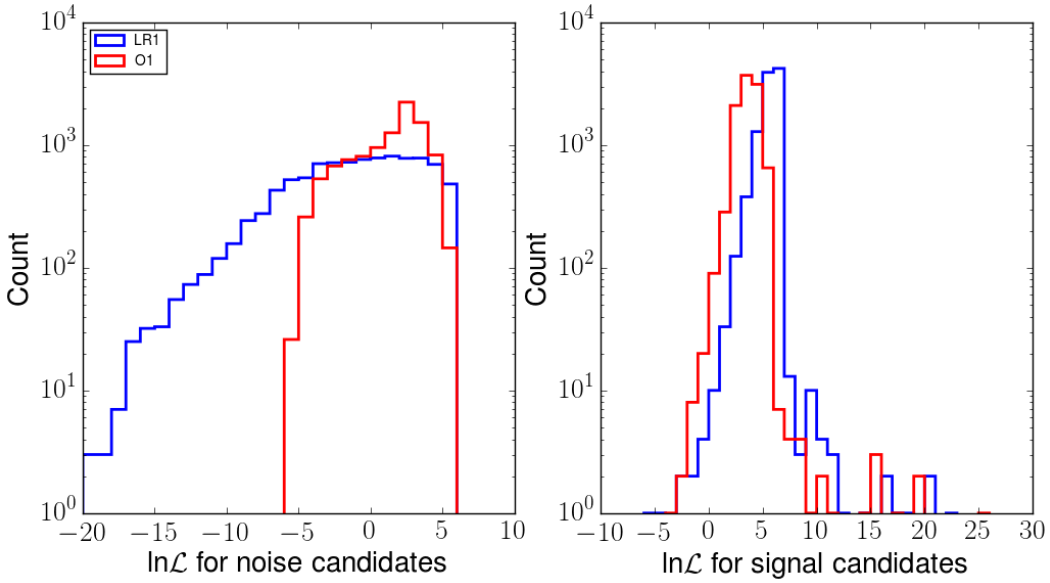



Figure 7.12: Histograms of likelihood-ratio assigned by O1 model (red) and LR1 model (blue) for noise (left) and signal (right) samples. We see that the LR model assigns lower likelihood-ratio values for noise samples, and higher likelihood-ratio values for signal samples as compared to the O1 model.

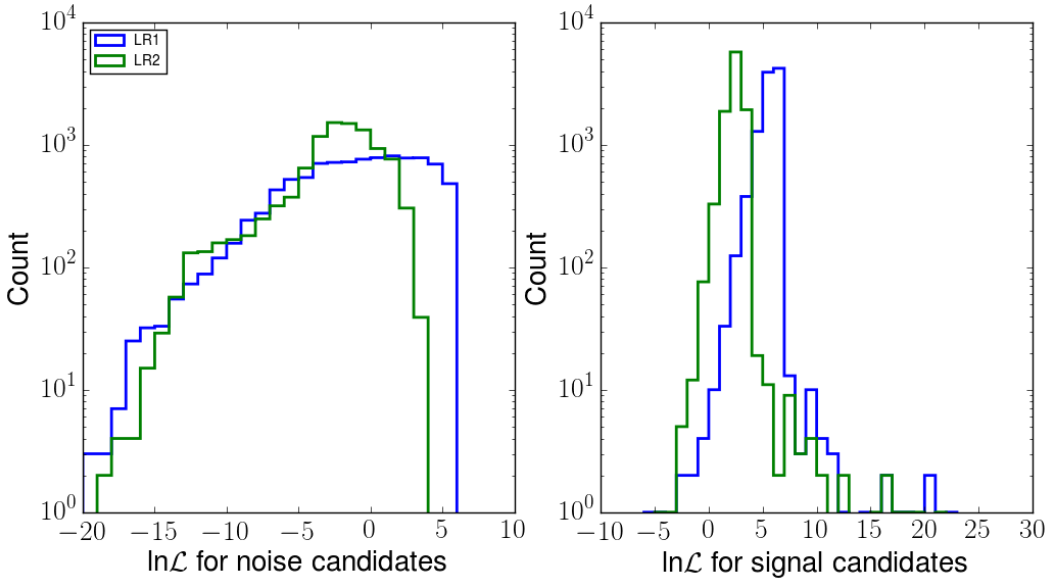



Figure 7.13: Histograms of likelihood-ratio assigned by LR1 model (blue) and LR2 model (green) for noise (left) and signal (right) samples. It is hard to tell from these histograms, but Fig. 7.14 shows that the LR2 model is better at separating signal from noise.

horizon distances. In the second observing run, this assumption was relaxed by pre-computing joint SNR PDFs for a collection of discrete horizon distance ratios.

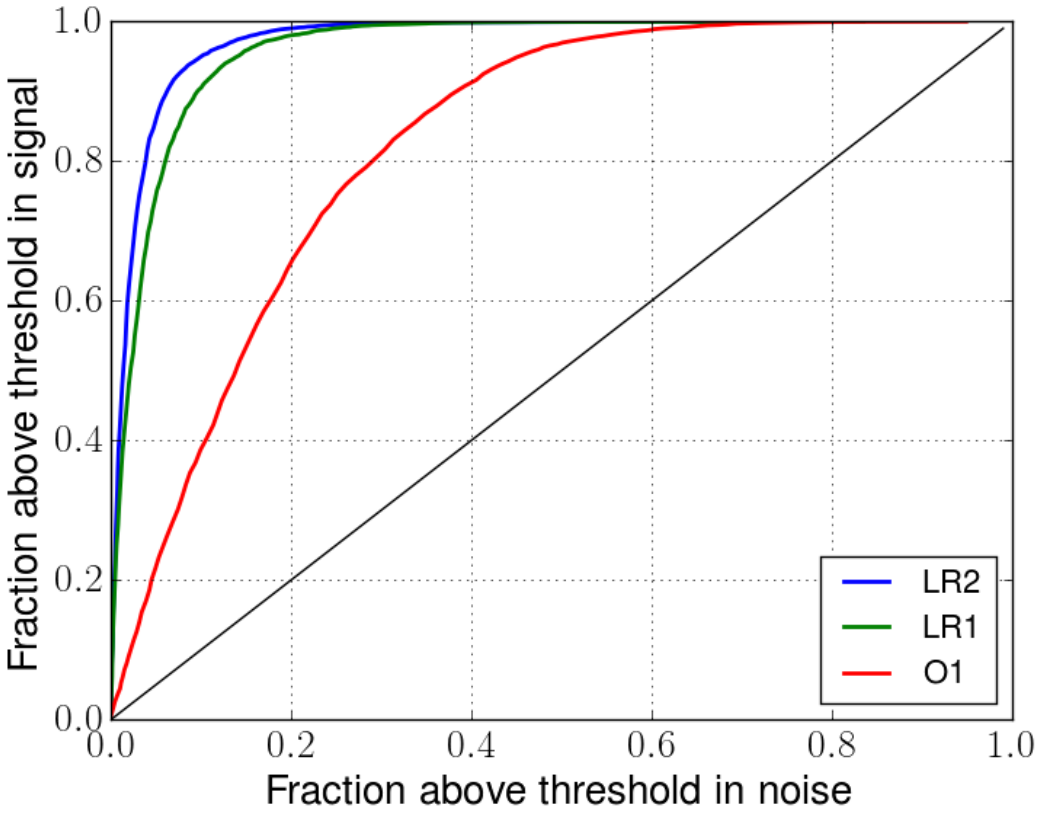



Figure 7.14: Receiver Operating Characteristic (ROC) curve for the three different likelihood-ratio models using the same set of samples for noise and signal. We see that LR2 performs the best out of the models considered here.

In this section we discussed all the changes to the `GstLAL` pipeline since O1. We discuss the final results from the `GstLAL` search using these methods over O2 (and O1) in the next section.

## 7.2 Results

Perhaps the greatest success of the `GstLAL` pipeline was the detection of GW170817 in low-latency, found in coincidence with a GRB detected by Fermi and INTEGRAL. The timely alert sent out by the pipeline made possible the extensive followup carried out across the electromagnetic spectrum which led to the detection of the associated optical transient (SSS17a/AT 2017gfo) in the galaxy NGC4993 [12]. This is exactly what the pipeline was built to do, to send out prompt alerts to the broader astronomy community making it possible to observe the onset of electromagnetic emission from CBC [50]. The joint gravitational-wave and electromagnetic observations of GW170817 support the hypothesis that it was produced by the merger of two neutron

stars in the galaxy NGC 4993, followed by a GRB with a time-delay of ~1.7s and a kilonova/macronova powered by the radioactive decay of r-process nuclei synthesized in the ejecta [12].

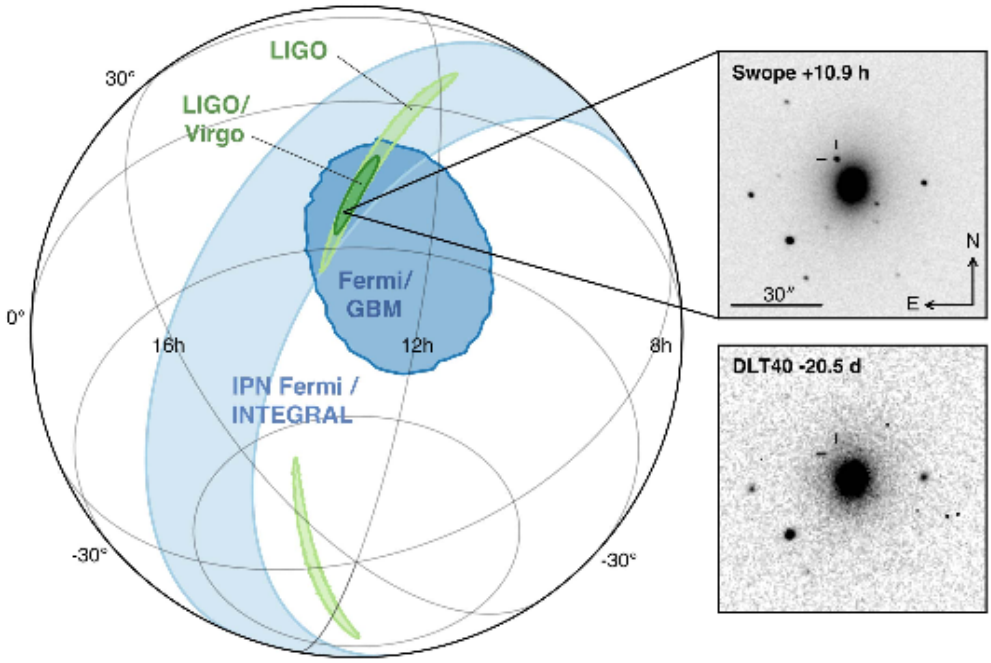



Figure 7.15: Localization of GW170817 from gravitational-wave, gamma-ray, and optical observations. The light green region shows the initial localization using only the data from the Hanford and Livingston detectors, the dark green region shows the updated localization using also the data from the Virgo detector. The blue and the dark blue regions respectively show the IPN triangulation for the GRB using the time-delay between Fermi and INTEGRAL observations, and the skymap from Fermi-GBM. The inset shows the location of the apparent host galaxy NGC 4993 in the Swope optical discovery image at 10.9 hr after the merger (top right) and the DLT40 pre-discovery image from 20.5 days prior to merger (bottom right). Figure taken from [12].

Using the gravitational-wave and electromagnetic data for GW170817, numerous studies have been conducted, including but not limited to gravitational-wave standard siren measurement of the Hubble constant, implications for the stochastic background, measurement of neutron star radii and equation of state, estimation of the contribution of dynamical ejecta in the kilonova associated with the event, and many others.

### 7.2.1 Offline analysis of final cleaned and calibrated data

In addition to reanalyzing the final cleaned and calibrated O2 data, we also use the improved methods and the bank covering a bigger parameter space than in

O1 described in this chapter to reanalyze data from O1. This was done so we could properly combine the results. The analyses were conducted separately in the original 5 day chunks (Sec. 6.2.3). The results from different analysis chunks were combined using an analogous procedure to that used in O1, see Sec. 6.5.1, i.e., the candidates were assigned likelihood-ratios based on their respective chunks, but the likelihood-ratio distribution for noise was obtained by sampling background distributions marginalized over the entire dataset. However, we don't adapt the procedure of including candidates in the background, and their hierarchical removal as was done in O1, rather we just compute a single background model constructed from non-candidates. The results are shown in Fig. 7.16, and we list the loudest candidates in Table 7.1. The full results will be published soon in [135].

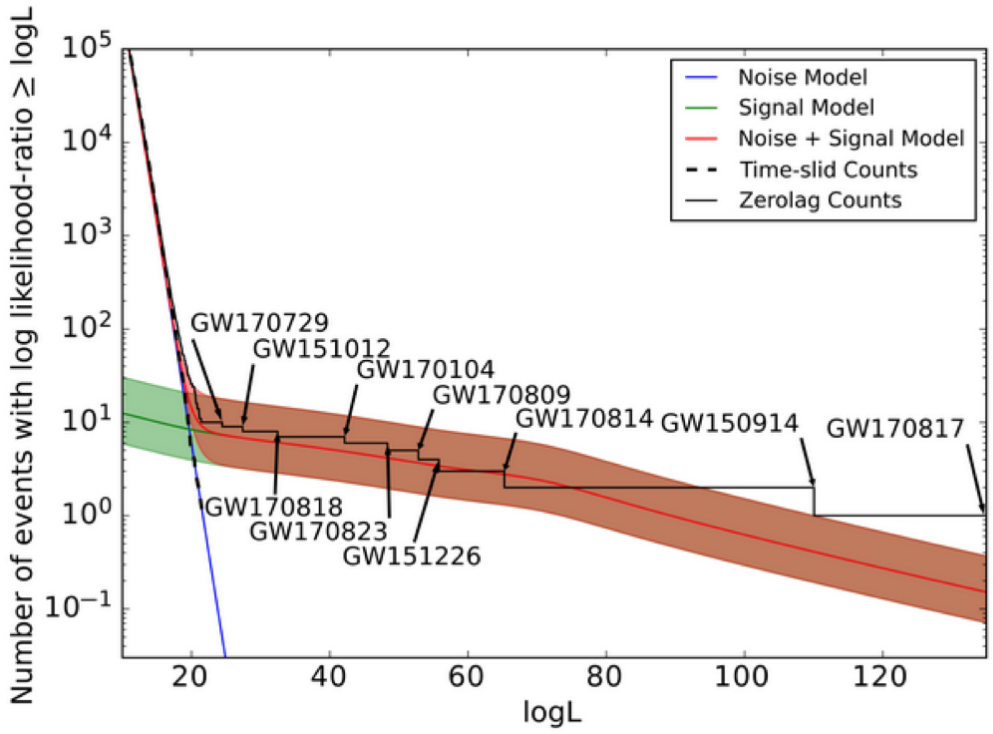



Figure 7.16: Cumulative count of candidate events as a function of $\ln\mathcal{L}$. We show the noise model (blue), signal model (green), signal + noise model (red), and the error bars on the signal + noise model. The solid black line shows the results of the search. The dashed black line is the result of time-slid closed box analysis which is consistent with the noise model. The top ten loudest detections are marked with their names. Note however the absence of GW170608 from this plot. Due to the unusual analysis that needed to be carried out for GW170608 which included analyzing Hanford data that was not marked science quality and using a different starting frequency of 30 Hz for Hanford, we do not combine its results with the other analyses.

| Event | FAR | ln$\mathcal{L}$ | H1 SNR | H1$\xi^2$ | L1 SNR | L1$\xi^2$ | Notes |
|---|---|---|---|---|---|---|---|
| GW170817 | $4.48 \times 10^{-51}$ | 134.68 | 19.11 | 0.70 | 26.91 | 1.63 | BNS, latency = 6 min |
| GW150914 | $6.30 \times 10{-}40$ | 110.05 | 20.31 | 0.91 | 13.54 | 0.77 | loudest BBH, first detection from O1 |
| GW170814 | $7 \times 10^{-20}$ | 65.25 | 9.14 | 1.23 | 13.04 | 1.05 | First BBH detected in coincidence with Advanced LIGO and Virgo. latency = 32.5 s |
| GW170608 | $5.27 \times 10^{-17}$ | 56.87 | 11.88 | 1.53 | 8.99 | 1.11 | Result from a special analysis, which included non-science H1 segments, and a starting frequency of 30 Hz for H1. |
| GW151226 | $1.10 \times 10^{-15}$ | 55.79 | 10.51 | 0.85 | 7.82 | 1.11 | Boxing day event. First online detection by the pipeline. latency = 70 s |
| GW170809 | $2.6 \times 10^{-14}$ | 52.80 | 6.59 | 1.12 | 10.53 | 0.57 | latency = 32.5 s |
| GW170823 | $2.62 \times 10^{-12}$ | 48.35 | 6.40 | 1.05 | 9.51 | 0.70 | latency = 32.5 s |
| GW170104 | $1.42 \times 10^{-9}$ | 42.13 | 8.94 | 0.93 | 9.42 | 0.97 | |
| GW170818 | $2.42 \times 10^{-5}$ | 32.45 | 4.13 | 1.07 | 9.66 | 0.96 | detected by the online pipeline in coincidence with L1 and V1.latency = 37 s |
| GW151012 | $4.60 \times 10^{-3}$ | 27.38 | 7.33 | 0.60 | 6.76 | 1.00 | |
| GW170729 | 0.11 | 24.42 | 7.90 | 1.08 | 7.34 | 0.89 | |

Table 7.1: Loudest events from the combined O1+O2 run. Most of these events were also detected by the `GstLAL` online search. Exceptions include events for which one of the detectors was marked in non-science mode (GW170104 [16] and GW170608 [17]) and events from O1 when the online search was only looking at the low-mass em-bright search space (GW150914 [6] and GW151012 [18]).

In the loudest candidate table, Table 7.1, we see 10 binary black holes and 1 binary neutron star detected in the combined O1+O2 analysis. The loudest event so far is the BNS, GW170817, and the loudest BBH so far is the first detection, GW150914. Most of the events were also identified online (see the latencies reported in Table 7.1). In addition, the pipeline detected several marginal events which are being collected in [135].

In addition to calculating the FAPs and the FARs (Sec. 4.2.5), we also assign a $P_{\text{category}}$, which is the probability that a candidate event belongs to a certain category. The categories we considered in O2 are - background (or, terrestrial) and astrophysical (which is subdivided into BNS, BBH, and NSBH). This is done by fitting for number of events in each category. The background model is derived from the background distributions from the pipeline, and the signal model is derived from the distribution of ranking statistics over astrophysical source categories. Of all the events listed in the Table 7.1, GW170729, GW151012, and GW151226 have a non-zero $P_{\text{NSBH}}$ eqaual to 0.01, 0.02, and 0.18 respectively. The O! search identified 151012 as a marginally significant event, but the improved O2 pipeline analysis identified it as more significant ($\text{P}_{\text{astrophysical}} = 1$), so LVT modifier became GW.

*Chapter 8*

# ASTROPHYSICAL POPULATIONS OF COMPACT BINARY SYSTEMS

## 8.1 Introduction

With the continued detections of gravitational waves from binary black hole coalescences [6, 9, 10, 16, 17] and the first direct detection of gravitational waves from a binary neutron star system [11], the LIGO Scientific and Virgo Collaborations have kept up to their promise of taking us into the era of gravitational-wave astronomy. In addition to these loud and nearby sources that are seen as isolated transient events, there is a population of weak, unresolved sources at higher redshifts [136–140]. The superposition of these sources is expected to be the main contributor to the stochastic background which may be detectable in the next few years as the Advanced LIGO [19] and Virgo detectors [20] reach their design sensitivity and accumulate more data [141, 142]. Assuming the most probable rate for compact binary mergers ($103^{+110}_{-63}\mathrm{Gpc}^{-3}\mathrm{yr}{-}1$ [30] for BBH and $1540^{+3200}_{-1220}\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ [11] for BNS), it has been shown that the total background may be detectable with a signal-to-noise-ratio of 3 after 40 months of total observation time, based on the expected timeline for Advanced LIGO and Virgo to reach their design sensitivity [142]. The astrophysical background potentially contains a wealth of information about the history and evolution of a population of point sources, but it is a confusion noise background that obscures the observation of the primordial gravitational-wave background (PGWB) produced in the very early stages of the Universe. Proposed theoretical cosmological models include the amplification of vacuum fluctuations during inflation [143–145], pre Big-Bang models [146–148], cosmic (super)strings [149–152], or phase transitions [153–155]. For more comprehensive texts on cosmological backgrounds from gravitational waves, we refer the reader to [156, 157].

Detection of the primordial gravitational-wave background would create a unique window on the very first instants of the Universe, up to the limits of the Planck era, and on the physical laws that apply at the highest energy scales. Needless to say that such a detection would have a profound impact on our understanding of the evolution of the Universe.

In addition to the astrophysical background from unresolved compact binary mergers,

a contribution is expected to result from the superposition of several other unresolved sources [158], such as cosmic strings [150], core collapse supernovae to neutron stars or black holes [159–162], rotating neutron stars [163, 164], including magnetars [165–168], phase transitions [169], or initial instabilities in young neutron stars [170–172].

With the increased sensitivity of the third generation gravitational-wave detectors, such as the Einstein Telescope (ET) [173] and the Cosmic Explorer (CE) [174], it will be possible to detect almost all of the binary black hole mergers, even the ones at high redshifts. In this work, we explore the possibility of probing the cosmological gravitational-wave background with the third generation detectors, after removing the astrophysical background from compact binary mergers from the data. This work is an extension to [15], where the authors have shown the level at which we can expect amplitude of background from unresolved, sub threshold CBC signals for various detector network cases. We extend the previous study to also provide of an estimate of errors we introduced while subtracting the signals above threshold for the most optimistic network of detectors considered in [15].

In Sec. 8.2, we describe the basic method that we use to calculate the gravitational-wave spectrum from the error introduced by imperfect subtraction of CBC signals. In Sec. 8.3, we describe the framework used to estimate the deviations of the estimated parameters of the CBC sources from their true values. We discuss the results from the simulation of a population of binaries in Sec. 8.4, and finally we present our conclusions in Sec. 8.5.

## 8.2 Method

The energy-density spectrum in gravitational waves is described by the dimensionless quantity [175],

$$\Omega_{\rm GW}(f) = \frac{f}{\rho_c}\frac{d\rho_{\rm GW}}{df}, \tag{8.1}$$

where $d\rho_{\rm GW}$ is the energy density in the frequency interval $f$ to $f + df$, $\rho_c = 3H_0^2c^2/8\pi G$ is the closure energy density, and $H_0$ is the Hubble constant whose current best value is $67.8 \pm 0.9$ km/c/Mpc [176].

The gravitational-wave energy spectrum density can be written as a sum of contribution from the astrophysical and cosmological energy densities,

$$\Omega_{\rm GW} = \Omega_{\rm astro} + \Omega_{\rm cosmo}. \tag{8.2}$$

Taking the contribution of the compact binary coalescences out of the astrophysical background, and writing it explicitly, we have

$$\Omega_{\rm GW} = \Omega_{\rm astro,\,r} + \Omega_{\rm cosmo} + \Omega_{\rm cbc}. \tag{8.3}$$

Here $\Omega_{\rm astro,\,r}$ is the remaining astrophysical background after taking out the contribution from the CBC sources.

When estimating the parameters of a binary source, by using Monte Carlo methods, or nested sampling, we invariably end up with parameters that deviate from the true values because of the noise in the detector. Therefore when we subtract the recovered CBC signals from the data, we introduce an additional background due to the error in subtraction, $\Omega_{\rm error}$.

$$\begin{aligned}\Omega_{\rm GW} = \Omega_{\rm cbc,\,rec} + \Omega_{\rm error} + \\ \Omega_{\rm cbc,\,unres} + \Omega_{\rm cosmo} + \Omega_{\rm astro,\,r},\end{aligned} \tag{8.4}$$

where $\Omega_{\rm cbc,\,rec}$ is the background from the recovered CBC sources that we can subtract from our data, $\Omega_{\rm error}$ is the contribution to the background because of the error introduced from such a subtraction, $\Omega_{\rm cbc,\,unres}$ is the background from the unresolved CBC sources which are not detected as resolvable "foreground" events. Let us assume that we have an experiment where we have detected a list of CBC sources and subtracted them from the data. Now we are left with the gravitational-wave backgrounds $\Omega_{\rm error}$ and $\Omega_{\rm cbc,\,unres}$, on top of the cosmological and astrophysical (from sources other than the CBCs) backgrounds. We want to answer the question of whether the cosmological or astrophysical backgrounds from sources other than CBCs can stand above the residue background after removal of the CBC sources. That is,

$$\Omega_{\rm error} + \Omega_{\rm cbc,\,unres} \;?\; \Omega_{\rm cosmo} \;?\; \Omega_{\rm astro,\,r}. \tag{8.5}$$

In order for us to be able to detect the gravitational-wave background from cosmological sources or that from different astrophysical sources, we would need $\Omega_{\rm residue} = \Omega_{\rm error} + \Omega_{\rm cbc,\,unres}$ to lie below these.

The gravitational-wave energy density from a population of compact binary sources is given by [15]

$$\Omega_{\rm cbc} = \frac{1}{\rho_c c} f F(f), \tag{8.6}$$

where $F(f)$ is the total flux, computed as the sum of individual contributions

$$F(f) = T^{-1} \frac{\pi c^3}{2G} f^2 \sum_{k=1}^{N} (\tilde{h}^2_{+,k}(f) + \tilde{h}^2_{\times,k}(f)), \tag{8.7}$$

where $N$ is the number of sources in the Monte Carlo sample, and $T^{-1}$ assures that flux has the correct dimension, $T$ being the length in time of the data sample. $\tilde{h}_{+,k}(f)$ and $\tilde{h}_{\times,k}(f)$ are the Fourier domain waveforms for the two polarizations, and the index $k$ runs over all the sources. We calculate $\Omega_{\rm error}$ as

$$\Omega_{\rm error} = \frac{1}{\rho_c c} f F_{\rm error}(f), \tag{8.8}$$

where

$$F_{\rm error}(f) = T^{-1} \frac{\pi c^3}{2G} f^2 \times \tag{8.9}$$

$$\times \sum_{k=1}^{N} \left[ (\tilde{h}^{\rm true}_{+,k}(f) - \tilde{h}^{\rm rec}_{+,k}(f))^2 + (\tilde{h}^{\rm true}_{\times,k}(f) - \tilde{h}^{\rm rec}_{\times,k}(f))^2 \right]. \tag{8.10}$$

To get an estimate of $\Omega_{\rm error}$, we need to estimate the quantities, $\tilde{h}^{\rm rec}_{+,k}(f)$ and $\tilde{h}^{\rm rec}_{\times,k}(f)$.

### 8.3 Estimating the deviation from the true value of parameters

Ideally we want the full Bayesian posteriors to estimate the deviation from the true value of parameters. However, at present it is computationally unfeasible to compute the full posterior probability distribution functions of all 15 binary parameters for the hundreds of thousands of sources that we simulate upto a redshit of 10 in the following section. The Fisher matrix provides a computationally cheap method to estimate the errors in the case when the posteriors are Gaussian, which is unfortunately not true in general. Nevertheless, for high SNRs, the Fisher matrix can give us a good enough idea of the errors while estimating the value of parameters.

We follow the framework described in [177] to calculate the errors in estimating the parameters of the compact binary system using the Fisher matrix method. According

to the Post-Newtonian expansion formalism [131], the gravitational-wave waveform from a compact binary coalescence in frequency domain is given by

$$\tilde{h}(f) = \mathcal{A} f^{-7/6} e^{i\Psi(f)}, \tag{8.11}$$

$$\tag{8.12}$$

where $A$ is the amplitude of the waveform, and $\Psi(f)$ is the phase given by

$$\Psi(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128\eta\nu^5} \sum_{k=0}^{N} \alpha_k \nu^k. \tag{8.13}$$

Here $t_c$ is the time of coalescence, $\phi_c$ is the coalescence phase, $\nu = (\pi M f)^{1/3}$, M is the total mass ($M = m_1 + m_2$), $\eta$ is the symmetric mass ratio ($\eta = m_1 m_2 / M^2$) of the system, and the $\alpha_k$ terms are known as the Post-Newtonian coefficients.
In the following discussion, we restrict ourselves to $\alpha_0$ terms. We will justify this choice later in this chapter.

For the fisher matrix study, we choose a set of independent parameters $\theta$ for describing the gravitational waveform,

$$\theta = (f_0 t_c, \phi_c, ln\mathcal{M}), \tag{8.14}$$

where $f_0$ is a reference frequency needed to keep the parameters for the Fisher matrix dimensionless. $\mathcal{M}$ is the dimensionless chirp mass, and is defined as $\mathcal{M} = \eta^{3/5} M / M_\odot$.

Writing the phase of the waveform in terms of these parameters, we have

$$\Psi(f) = 2\pi \frac{f}{f_0}(f_0 t_c) - \phi_c - \frac{\pi}{4} + \frac{3}{128}(\pi \mathcal{M} f)^{-5/3}, \tag{8.15}$$

$$\tag{8.16}$$

$$\Psi(f) = 2\pi \frac{f}{f_0}\theta_1 - \theta_2 - \frac{\pi}{4} + \frac{3}{128}\left(\frac{\pi e^{\theta_3} f G M_\odot}{c^3}\right)^{-5/3}. \tag{8.17}$$

In going from Eq. (8.16) to Eq. (8.17), we have truncated the expansion at $\alpha_0$ term (i.e., at $k = 0$), plugged in the value $\alpha_0 = 1$, and we have introduced the Newtonian constant , G, the speed of light, c, and solar mass $M_\odot$, explicitly to keep all quantities in the Eq. (8.16) unitless, and define masses in solar mass units.

The Fisher matrix elements are given by

$$\Gamma_{ij} = 2\int_{f_L}^{f_H} \frac{\tilde{h}^*_{\theta_i}(f)\tilde{h}_{\theta_j}(f) + \tilde{h}_{\theta_i}(f)\tilde{h}^*_{\theta_j}(f)}{S_n(f)}\mathrm{d}f, \tag{8.18}$$

where

$$\tilde{h}_{\theta_i}(f) = \frac{\partial}{\partial\theta_i}\tilde{h}(f), \tag{8.19}$$

are the partial derivatives of the waveform with respect to $\theta_i$, the parameters of the waveforms, and $S_n(f)$ is the single-sided power spectral density of the detector. The partial derivatives of the waveform can be calculated analytically:

$$\tilde{h}_{\theta_1}(f) = \frac{2\pi f A}{f_0} f^{-7/6} e^{i(\Psi(f)+\pi/2)},$$

$$\tilde{h}_{\theta_2}(f) = A f^{-7/6} e^{i(\Psi(f)-\pi/2)},$$

and

$$\tilde{h}_{\theta_3}(f) = A f^{-7/6} e^{i(\Psi(f)-\pi/2)} \frac{5}{128}\left(\frac{\pi e^{\theta_3} f G}{c^3}\right)^{-5/3}. \tag{8.20}$$

The fisher matrix is then calculated by performing the integration in Eq. (8.18) numerically. For a network of detectors, the fisher matrix is the sum of fisher matrices for individual detectors,

$$\Gamma^{\mathrm{net}}_{ij} = \sum_{\mathrm{det}} \Gamma^{\mathrm{det}}_{ij}. \tag{8.21}$$

The variance-covariance matrix, or simply the covariance matrix, defined as the inverse of the Fisher information matrix, is given by

$$\Sigma_{ij} = (\Gamma^{-1})_{ij}. \tag{8.22}$$

Once we have the covariance matrix, we use a multivariate normal random generator to generate observed value of the parameters, **pO**, based on the multivariate distribution with the mean equal to the true value of the parameters, **pT** and covariance matrix as $\Sigma$. The error in parameter estimation is then given by

$$\begin{aligned} \mathbf{R} = [\Delta\theta_1, \Delta\theta_2, \Delta\theta_3] &= \mathbf{pO} - \mathbf{pT}, \\ \Delta t_c &= \frac{\Delta\theta_1}{f_0}, \\ \Delta\phi_c &= \Delta\theta_2, \end{aligned}$$

and

$$\Delta\mathcal{M} = \mathcal{M}\Delta\theta_3. \tag{8.23}$$

## 8.4 Population synthesis for multiple detectors

We simulate a population of BBHs and BNSs upto a redshift of 10, and then calculate an estimate of $\Omega_{\text{cbc, rec}}$, and $\Omega_{\text{error}}$ as outlined in Sec. 8.2 and Sec. 8.3. The list of compact binaries (neutron star binaries or black hole binaries) is generated following a Monte Carlo procedure described in [15, 178–180], and using the fiducial model of [142] for the distribution of the parameters (masses, redshift, position in the sky, polarization, and inclination angle of the binary). In particular, we assume a redshift distribution which is derived from the star formation rate of [13] and accounts for a delay between the formation of the progenitors and the merger. We further consider the most recent median rates estimated from the LIGO/Virgo detections. The procedure for generating the list of sources is described below.

1. For BBHs, the intrinsic masses $m_1, m_2$ (in the source frame) are selected from the power-law distribution (Saltpeter initial mass function [181]) considered in [18, 30] of the primary (i.e., the larger mass) companion $p(m_1) \propto m_1^{-2.35}$, and from a uniform distribution of the secondary companion. In addition, we require that the component masses take values in the range 5-100 $M_\odot$.

   For BNSs, the intrinsic masses $m_1, m_2$ (in the source frame) are both drawn from uniform distribution ranging from 1 to 2 $M_\odot$. This is also the distribution that was used to calculate the rates of binary neutron star mergers in [11].

2. The redshift z is drawn from a probability distribution $p(z)$ given by (See Fig. 8.3),

   $$p(z) = \frac{R_z(z)}{\int_0^{10} R_z(z)dz}, \tag{8.24}$$

   obtained by normalizing the merger rate of binaries in the observer frame, $R_z(z)$ per interval of redshift, over the range $z \in [0, 10]$. We choose to cut off the redshift integral at $z_{\text{max}} = 10$, since redshifts larger than 5 contribute little to the background [142]. The merger rate in the observer frame is

   $$R_z(z) = \int \frac{R_m(z)}{1+z}\frac{dV}{dz}(z)dz, \tag{8.25}$$

where $dV/dz$ is the comoving volume element and $R_m(z)$ is the rate per comoving volume in the source frame, given by

$$R_m(z) = \int_{t_{\min}}^{t_{\max}} R_f(z_f) P(t_d) dt_d, \tag{8.26}$$

where $R_f(z_f)$ is the binary formation rate as a function of the redshift at formation time, $t_f$, and $P(t_d)$ is the distribution of the time delay $t_d$ between the formation and merger of the binary. The value of $R_m(z = 0)$ is chosen as the local merger rate estimate from the LIGO-Virgo observations. For the rate of BBH mergers, we use the most recent published result associated with the power-law mass distribution $103^{+110}_{-63}\mathrm{Gpc}^{-3}\mathrm{yr}{-}1$ [30]. For the BNS case, we set $R_m(z = 0)$ to $1540^{+3200}_{-1220}\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ obtained from the GW170817 calculation [11].

We assume that the binary formation rate $R_f(z)$ scales with the star formation rate. We follow the the GRB-based cosmic star formation model of [13]. It uses the Springer Hernquist functional form [182]

$$R_f(z) = \nu \frac{a e^{b(z-z_m)}}{a - b + b e^{(a(z-z_m)}}, \tag{8.27}$$

to fit to the GRB-based high-redshift SFR data of [183] but normalized based on the [184] SFR ( [185, 186]). This fit results in $\nu = 0.146 M_\odot/\mathrm{yr}/\mathrm{Mpc}^3, z_m = 1.72, a = 2.80, \mathrm{and} b = 2.46$ [13]. The resulting model in illustrated in Fig. 8.1 [14] by the solid line.

We consider a time delay distribution $p(t_d) \propto 1/t_d$, for $t_{\min} < t_d < t_{\max}$. For BNS, we set $t_{\min} = 20\mathrm{Myr}$ [142, 187], whereas for BBH, we set $t_{\min} = 50\mathrm{Myr}$ [141, 142, 188]. The maximum time delay, $t_{\max}$ is set to the Hubble time [188, 188–195].

Massive black holes are formed preferentially in low-metallicity environments [44, 141]. For systems where at least one black hole has a mass larger than $30M\odot$, we re-weight the star formation rate $R_f(z)$ by the fraction of stars with metallicities less than half the solar metallicity [142]. Following [141, 142], we follow the mean metallicity-redshift relation of [196], and scale it upwards by a factor of 3 to account for local observations [13, 197]. The mean metallicity-redshift relation is illustrated in Fig. 8.1 [14] by the solid line.

3. The location in the sky, the cosine of orientation, the polarization, and the coalescence phase are drawn from uniform distributions.

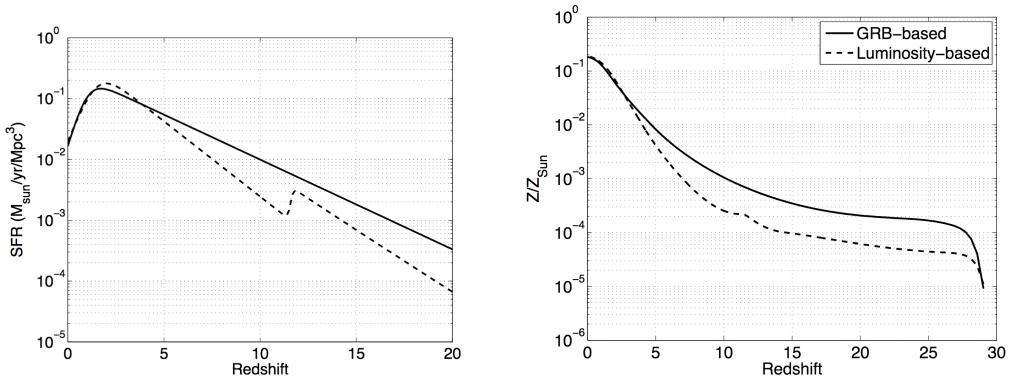



Figure 8.1: The GRB-based star-formation rate from [13] (left, in solid). The mean metallicity-redshift relation based on the GRB-SFR from [13] (right, solid). Figure taken from [14].

### 8.4.1 Detector Network

We consider a network of five third generation detectors (Fig 8.2), specifically a detector with the sensitivity of ET at the location of Virgo, and detectors with the CE sensitivity at locations of LIGO Hanford, LIGO Livingston, LIGO India, and KAGRA. We choose this configuration of detector-network because it was shown in [15], that the astrophysical "confusion" background from unresolved BBH sources is of the order $\Omega_{\rm GW}(10{\rm Hz}) = 10^{-16} - 10^{-14}$, making the network sensitive to detection of other backgrounds (PGWB) at that level.

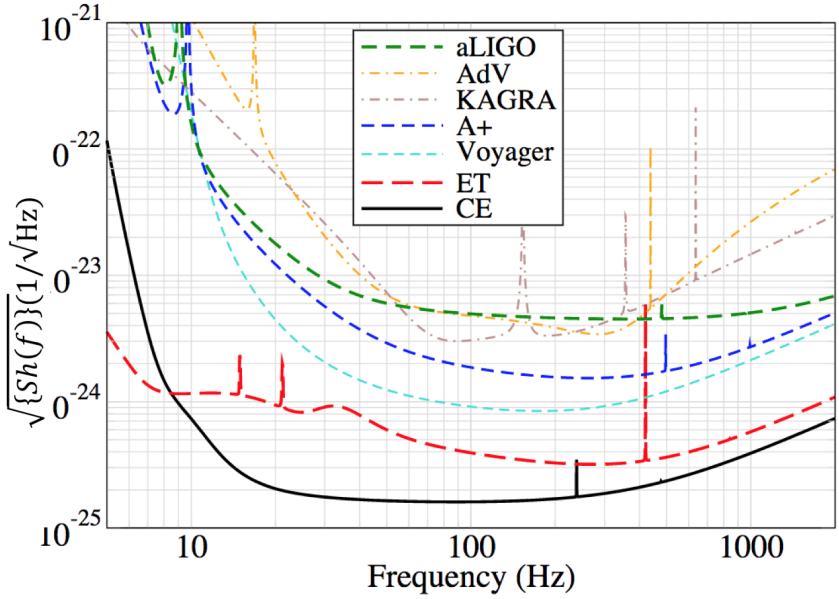



Figure 8.2: Design amplitude spectral density of second and third generation detectors. Expected intermediate sensitivities such as Advanced LIGO Plus (A+) and Voyager are also shown. Figure taken from [15].

## 8.5 Results

We simulate a population of BBH and BNS mergers according to the procedure described in Sec. 8.4 for a year's worth of data. There are 139,879 BBH and 2,408,262 BNS signals in our simulation. The redshift distribution for the generated sources is shown in Fig. 8.3.

For each source, we calculate the expected network SNR assuming perfect template match, given by

$$\rho_i^{\text{net}} = \sqrt{\sum_{\text{det}} (\rho_i^{\text{det}})^2}, \tag{8.28}$$

where index $i$ runs over all the sources, and $\rho_i^{\text{det}}$

$$(\rho_i^{\text{det}})^2 = 4 \int_0^\infty \text{df} \frac{\left|\tilde{h}_i^{\text{det}}(f)\right|^2}{S_h^{\text{det}}(f)} \tag{8.29}$$

is the SNR for each source and detector pair $(i, \text{det})$, and $\tilde{h}_i^{\text{det}}(f) = F_+^{\text{det}} \tilde{h}_{i,+} + F_\times^{\text{det}} h_{i,\times}$ is the Fourier domain waveform projected on the detector.

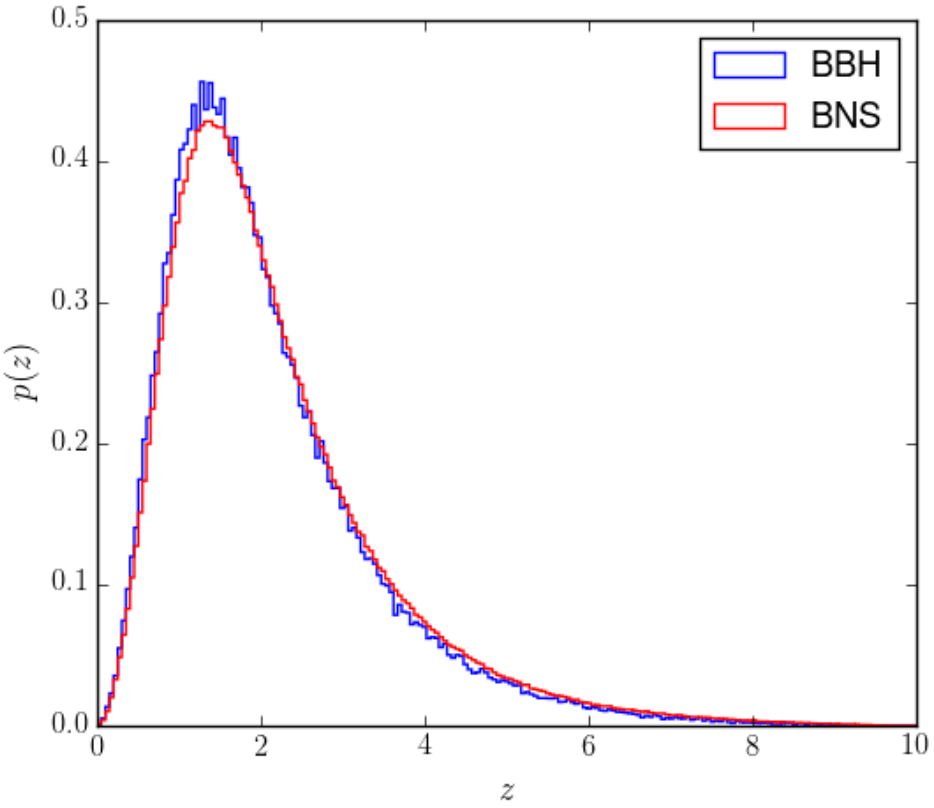



Figure 8.3: Redshift distribution of the simulated BBH (blue) and BNS (red) sources.

We considered a source as resolvable and a part of the "foreground", whenever $\rho_i^{\rm net} \geq \rho_{\rm thresh} = 12.0$. To begin with we calculate the quantities, $\Omega_{\rm BBH,total}$, $\Omega_{\rm BBH,unresolved}$, $\Omega_{\rm BNS,total}$, and $\Omega_{\rm BNS,unresolved}$. We use two waveform models, a full IMR model `IMRPhenomD`, and a model using $0^{th}$ order PN approximation (Eq. 8.12). We show the results in Fig. 8.4.

We can see that the results from the two models agree to a great extent below 100 Hz. It has been shown for various detector combinations that frequencies below 100 Hz account for more than 99% of the SNR for the stochastic search [179, 187]. Therefore for calculating $\Omega_{\rm error}$, we only consider the 0 PN model to compute the Fisher matrix for each source in our simulation.

Now we calculate the Fisher matrices (and the variance-covariance matrices) for all the sources in our simulation, and recover a set of parameters in order to calculate $\Omega_{\rm error,BNS}$ and $\Omega_{\rm BBH}$.

We find that ~20 % of all the BNS sources are unresolved, whereas only ~0.1 % of

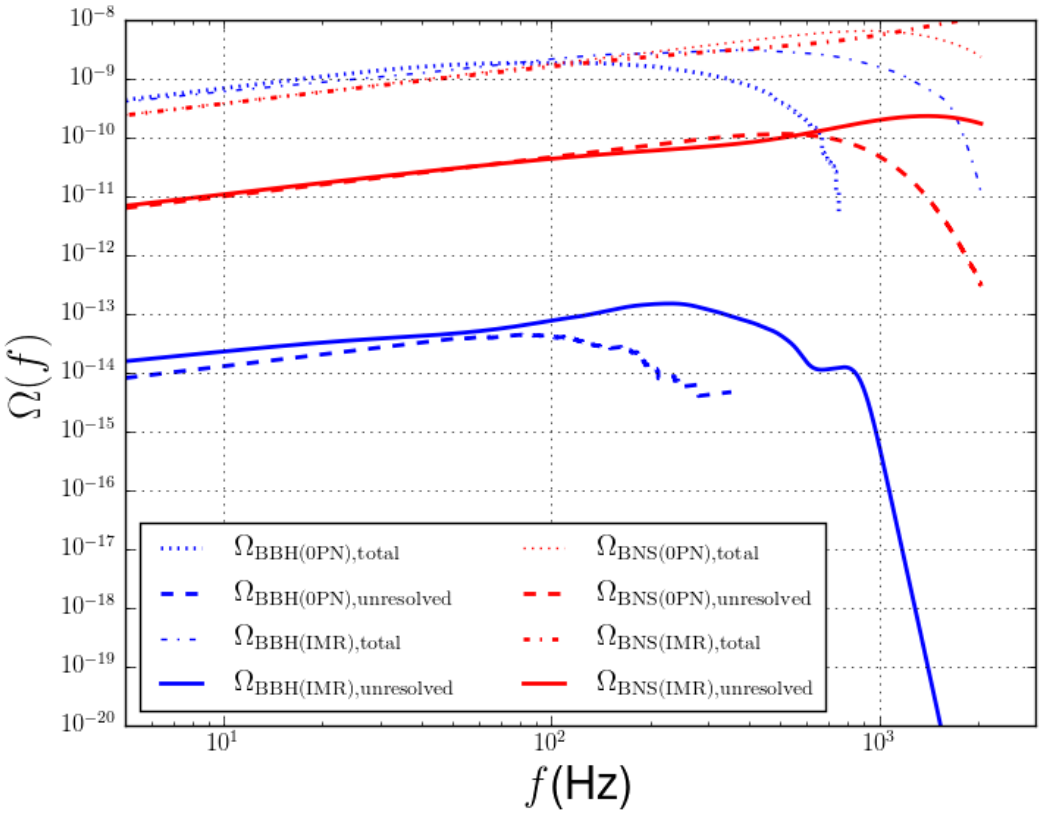



Figure 8.4: Gravitational-wave energy spectrum from total and unresolved BBH (blue) and total and unresolved BNS (red) sources using an IMR model and a 0 PN waveform model. We can see that the results from the two models agree to a great extent below 100 Hz.

all the BBH sources fall in the unresolved category.

The results are shown in Fig. 8.5.

Residual background from BBH sources (solid, blue: $\Omega_{\mathrm{BBH,\ residue}}$) is dominated by the deviation from true value of recovered parameters ($\Omega_{\mathrm{BBH,\ error}}$), while the residual background from BNS sources (solid, red:$\Omega_{\mathrm{BNS,\ residue}}$) is dominated by the unresolved sources ($\Omega_{\mathrm{BNS,\ unresolved}}$). We notice that for the BNS sources, $\Omega_{\mathrm{BNS,\ unresolved}}$ is very small as compared to $\Omega_{\mathrm{BNS,\ unresolved}}$. Therefore, we lower the SNR threshold for what we call something as resolved for the BNS sources to 8. The results are shown in green in Fig. 8.5. The residual background from the BNS sources goes down as expected when lowering the SNR threshold. It is still, however, dominated by the unresolved sources. So we could keep lowering the threshold in our framework to get a lower residual from the BNS sources. But the Fisher matrix method to calculate the errors on parameters breaks down for low SNR values. So we

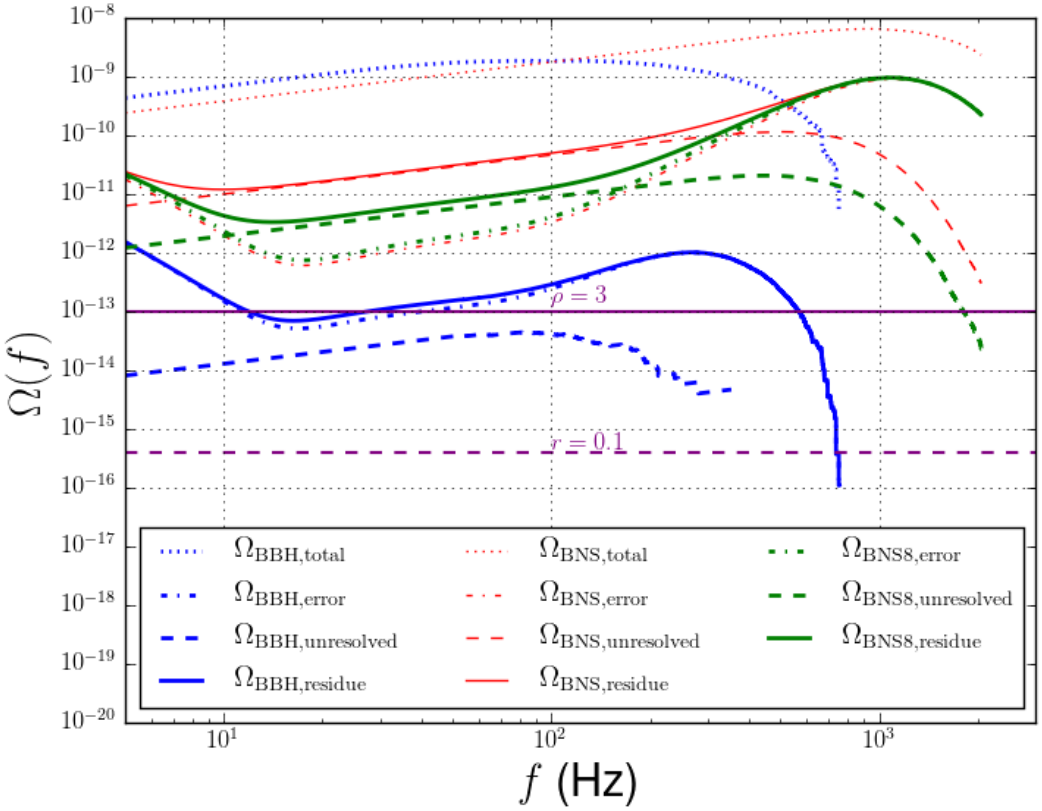



Figure 8.5: Gravitational-wave energy spectrums from the total unresolved error, and residue from BBH (blue) and the total, unresolved, error and residue from BNS (red) sources. We see that the residue from BBH sources is dominated from the error due to background subtraction, whereas for the BNS case it is dominated by the unresolved sources. We therefore perform another simulation for the BNS case, where we change the $\rho_{\rm thresh}$ to 8.0 instead of 12.0. The results are shown in green. Overall the residual background is dominated by the BNS sources. The cosmological background from inflation assuming a tensor-to-scalar ratio $r = 0.1$ is shown for comparison, the horizontal solid purple line is the minimal flat spectrum that can be detected with $\rho = 3$ with the HLVIK network after 5 years.

will stop here and conclude that the residual background after removing the resolved CBC sources will be dominated by the unresolved BNS sources. It will be at the level of $\Omega_{\text{BNS, unresolved}}(10\text{Hz}) = 10^{-12} - 10^{-11}$. The residual background from the BBH sources will be at a level of $\Omega_{\text{BBH, residue}}(10\text{Hz}) \sim 10^{-13}$. For reference, we also show the cosmological background from inflation assuming a tensor-to-scalar ratio of 0.1 (in dashed, purple). But there is a lot of uncertainty in the BNS rates, and it will be interesting to see how that will change with more detections. In a future study, we will also repeat these calculations with a full IMR waveform model, and eventually run them through a pipeline on mock data.

*Chapter 9*

# CONCLUSIONS AND FUTURE-WORK

## 9.1 Summary of results

A century after their prediction, the stretching and squeezing effect of gravitational waves was measured [6] on September 14, 2015 by the Advanced LIGO detectors [19], thereby marking the start of the era of gravitational-wave astronomy. The signal observed by the two LIGO detectors was consistent with that produced during the final fraction of a second of the merger of two black holes. This observation came after more than five decades of careful planning and building of the LIGO detectors [19, 26] and development of relevant technical and scientific research and data analysis techniques. Since the first detection in 2015, Advanced LIGO and Virgo have detected several other signals from binary black hole mergers and a signal from a binary neutron star inspiral, which was extensively followed in the elctromagnetic spectrum. Gravitational waves carry signatures of their sources, and therefore play a unique role in shaping our understanding of the extremely violent astrophysical events they originate from. We have tested general relativity in the strong regime [27], and studied matter at densities [28] beyond reach of the most powerful laboratories on our planet through these observations. They also help us gain insight into the evolution of stars, galaxies, and even the Universe as a whole by studying their merger rates [11, 18, 29, 30]. Joint electromagnetic and gravitational-wave observations help develop our understanding of the physical processes that occur in such systems [12, 31], and also provide a gravitational-wave standard siren method to measure the Hubble constant [32, 33].

The main focus of this thesis is the detection of gravitational waves from compact binary coalescences. We started with a brief description of gravitational waves (Chapter 2), coalescing binaries (Chapter 3), and methods to detect them. `GstLAL` is a matched-filtering search pipeline which used to detect compact binary mergers in low-latency. Low-latency detection is crucial for sending out timely alerts to the broader electromagnetic community to facilitate multi-messenger observations. We go over the methods employed in the search (Chapter 4), and thereafter, we discuss the development of a BBH `GstLAL` pipeline using full IMR aligned-spin waveforms (Chapter 5). During O1, it was deployed as part of the uberbank search (Fig. 6.5), and

as part of the hyperbank search (Fig. 7.2) in O2 (Chapters6, and 7). We discussed not only the successful results but also the many challenges faced to achieve them. We have shown that the GstLAL pipeline is not only capable of detecting gravitational waves from compact binaries, but is able to do so in near real time. All the loudest candidates found by the offline search, were also found online (except in the cases of mislabelled data, or a detector not being in science mode), Table 7.1. The main goal of the GstLAL pipeline was to provide prompt alerts to our partner electromagnetic observatories to facilitate multi-messenger observations of electromagnetic transients associated with the gravitational-wave phenomena [50]. With GW170817, we did exactly that [11]. The timely alert sent out by the pipeline made possible the extensive followup carried out across the electromagnetic spectrum which led to the detection of the associated optical transient (SSS17a/AT 2017gfo) in the galaxy NGC4993. The joint gravitational-wave and electromagnetic observations of GW170817 support the hypothesis that it was produced by the merger of two neutron stars in the galaxy NGC 4993, followed by a GRB with a time-delay of ~1.7s and a kilonova/macronova powered by the radioactive decay of r-process nuclei synthesized in the ejecta [12]. The time-delay between the gravitational-wave signal and the burst of gamma rays of 1.7 s also told us that the speed of GWs is equal to the speed of light to a part in $10^{15}$ [45].

## 9.2 Future work

It sounds like we have achieved what we set out to do. Yes, partially but there is still a long way to go.

From the recent joint gravitational-wave and gamma-ray detection of GW170817 [11], we now know that the time delay between gravitational-wave emission and the onset of the following burst of gamma rays is approximately 2 seconds, motivating achieving alert latencies below 2 seconds. The median latency of the GstLAL pipeline in O2 was about 22 s (Fig. 7.7). This is already significantly lower than what it was in O1 (~ 67 s), but we need to work towards bringing it down even lower. One of the places where the pipeline could benefit from a speed-up in in the assignment of the likelihood-ratios. We have already started to look into the application of machine learning to do this. The initial results look favourable, but it will need a lot of work before making it to an end-to-end pipeline test.

The detection of GW170817 by the online pipeline using a single detector proves that the GstLAL pipeline is capable of detecting signals from a single detector's data.

We would like to search over all of O1 and O2 using this search.

We have also started to develop a search to look for lensed counterparts of our detection candidates among the sub-threshold candidates of the pipeline[1]. Any lensed counterparts of the detection candidates should have the same intrinsic parameters, but be louder or quieter in magnitude. In our case we are looking for quieter, sub threshold counterparts as the louder ones would already have been detected. To accomplish this, we make use of a restricted template bank which contains only those templates that get rung up by the candidate events used as injections in a search. Reducing the search space decreases the amount of false triggers from noise, and therefore has the potential of elevating sub-threshold lensed signals.

In Chapter 8, we conducted a study about removing the foreground astrophysical CBC signals from our data to detect buried cosmological backgrounds with third generation detectors. We found that after removal of the signals, we are left with a residual background at the level of $\Omega_{\rm residual}(10Hz) = 10^{-12} - 10^{-11}$, most of which is dominated by the background from unresolved BNS sources. Therefore, our study concludes that we would only be sensitive to cosmological backgrounds at the level of $\Omega_{\rm cosmo} \sim 10^{-11}$. But our study had a number of drawbacks. We used waveforms using 0 order in post-Newtonian approximation, and we calculated the errors from deviation of the resolved parameters from true values using the Fisher matrix approach. The Fisher matrix approach breaks down for low SNRs. Therefore, we would like to repeat our study using full IMR waveforms and run them through a full CBC pipeline. If indeed some of the signals we have observed have been magnified through lensing, we have under-estimated the true luminosity distance to the source, and have therefore under-estimated the source redshift, and over-estimated the redshift-corrected source masses. Therefore, although unlikely, it is important to pursue this possibility.

There is lots of exciting work ahead of us. The Advanced LIGO detectors will start taking data for the third observing run, in early 2019 and will continue taking data for about a year. The sensitivities of the detectors is expected to be the best ever recorded. No doubt we will detect many more CBC signals and use them to enhance our knowledge of the Universe by leaps and bounds.

[1]This search is being developed in collaboration with Alvin Li, Tjonnie Li, and Alan Weinstein.

# BIBLIOGRAPHY


[1] Sean M. Carroll. Lecture notes on general relativity. 1997.

[2] Frank Ohme. Analytical meets numerical relativity: status of complete gravitational waveform models for binary black holes. *Classical and Quantum Gravity*, 29(12):124002, 2012.

[3] Cody Messick, Kent Blackburn, Patrick Brady, et al. Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data. *Physical Review D*, 95(4), 2017.

[4] Kipp Cannon, Adrian Chapman, Chad Hanna, et al. Singular value decomposition applied to compact binary coalescence gravitational-wave signals. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 82(4), 2010.

[5] Lisa Barsotti and Peter Fritschel. Early aLIGO Configurations: example scenarios toward design sensitivity. *LIGO Document T1200307v4*.

[6] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116(6), 2016.

[7] Abdul H. Mroué, Mark A. Scheel, Béla Szilágyi, et al. Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy. *Phys. Rev. Lett.*, 111:241104, Dec 2013.

[8] D. V. Martynov, E. D. Hall, B. P. Abbott, et al. Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy. *Phys. Rev. D*, 93:112004, Jun 2016.

[9] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Physical Review Letters*, 116(24), 2016.

[10] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Physical Review Letters*, 119(14), 2017.

[11] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Physical Review Letters*, 119(16):30–33, 2017.

[12] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Multi-messenger Observations of a Binary Neutron Star Merger. *The Astrophysical Journal Letters*, 848(2):L12, 2017.

[13] Elisabeth Vangioni, Keith A. Olive, Tanner Prestegard, et al. The impact of star formation and gamma-ray burst rates at high redshift on cosmic chemical evolution and reionization. *Monthly Notices of the Royal Astronomical Society*, 447(3):2575–2587, 2015.

[14] K. Crocker, V. Mandic, T. Regimbau, et al. Model of the stochastic gravitational-wave background due to core collapse to black holes. *Phys. Rev. D*, 92:063005, Sep 2015.

[15] T. Regimbau, M. Evans, N. Christensen, et al. Digging Deeper: Observing Primordial Gravitational Waves below the Binary-Black-Hole-Produced Stochastic Background. *Phys. Rev. Lett.*, 118:151105, Apr 2017.

[16] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Physical Review Letters*, 118(22), 2017.

[17] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence. *The Astrophysical Journal Letters*, 851(2):L35, dec 2017.

[18] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Binary Black Hole Mergers in the First Advanced LIGO Observing Run. *Phys. Rev. X*, 6:041015, Oct 2016.

[19] J. Aasi, B. P. Abbott, R. Abbott, et al. Advanced LIGO. *Classical and Quantum Gravity*, 32(7), 2015.

[20] The Virgo Collaboration. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, 2015.

[21] Henri Poincaré. http://www.academie-sciences.fr/pdf/dossiers/Poincare/Poincare_pdf/Poincare_CR1905.pdf.

[22] Albert Einstein. Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 688-696.*, 1916.

[23] A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.*, 1915.

[24] Joel M. Weisberg, Joseph H. Taylor, and Lee A. Fowler. Gravitational Waves from an Orbiting Pulsar. *Scientific American*, 245(4):74–82, 1981.

[25] 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.

[26] B. P. Abbott et al. LIGO: The Laser interferometer gravitational-wave observatory. *Rept. Prog. Phys.*, 72:076901, 2009.

[27] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of General Relativity with GW150914. *Physical Review Letters*, 116(22):221101, June 2016.

[28] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, et al. GW170817: Measurements of neutron star radii and equation of state. *ArXiv e-prints*, May 2018.

[29] B. P. Abbott, R. Abbott, T. D. Abbott, et al. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. *Astrophys. J. Lett.*, 833:L1, December 2016.

[30] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118:221101, Jun 2017.

[31] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Estimating the Contribution of Dynamical Ejecta in the Kilonova Associated with GW170817. *The Astrophysical Journal Letters*, 850(2):L39, 2017.

[32] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551:85–88, November 2017.

[33] B. F. Schutz. Determining the Hubble constant from gravitational wave observations. *Nature*, 323:310, September 1986.

[34] J. Cervantes-Cota, S. Galindo-Uribarri, and G. Smoot. A Brief History of Gravitational Waves. *Universe*, 2:22, September 2016.

[35] David Kennefick. *Traveling at the Speed of Thought Einstein and the Quest for Gravitational Waves*. 2016.

[36] B. S. Sathyaprakash and B. F. Schutz. Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Reviews in Relativity*, 12:2, March 2009.

[37] J. Creighton and W. Anderson. *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis.* November 2011.

[38] Bernard Schutz. *A First Course in General Relativity*. 2009.

[39] Jolien D. E. Creighton and Warren G. Anderson. *Gravitational-wave physics and astronomy: An introduction to theory, experiment and data analysis*. 2011.

[40] Alessandra Buonanno. Gravitational waves. In *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2007.

[41] Éanna É Flanagan and Scott A Hughes. The basics of gravitational wave theory. *New Journal of Physics*, 7(1):204, 2005.

[42] B. P. Abbott, R. Abbott, T. D. Abbott, et al. First narrow-band search for continuous gravitational waves from known pulsars in advanced detector data. *Phys. Rev. D*, 96:122006, Dec 2017.

[43] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Observing gravitational-wave transient GW150914 with minimal assumptions. *Phys. Rev. D*, 93:122004, Jun 2016.

[44] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Astrophysical Implications of the Binary Black-hole Merger GW150914. *The Astrophysical Journal Letters*, 818(2):L22, 2016.

[45] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. *The Astrophysical Journal Letters*, 848(2):L13, 2017.

[46] N. Cornish, D. Blas, and G. Nardini. Bounding the Speed of Gravity with Gravitational Wave Observations. *Physical Review Letters*, 119(16):161102, October 2017.

[47] M. Isi and A. J. Weinstein. Probing gravitational wave polarizations with signals from compact binary coalescences. *ArXiv e-prints*, October 2017.

[48] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Astrophysical Implications of the Binary Black-hole Merger GW150914. *Astrophys. J. Lett.*, 818:L22, February 2016.

[49] Surabhi Sachdev, Sarah Caudill, Heather Fong, et al. The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO's Second and Advanced Virgo's First Observing Runs. *arXiv e-prints*, page arXiv:1901.08580, Jan 2019.

[50] Kipp Cannon, Romain Cariou, Adrian Chapman, et al. Toward early-warning detection of gravitational waves from compact binary coalescence. *Astrophysical Journal*, 748(2), 2012.

[51] Alex Nitz, Ian Harry, Duncan Brown, et al. gwastro/pycbc: 1.12.3 Release, September 2018.

[52] 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. *Astrophys. J.*, 849(2):118, 2017.

[53] Samantha A. Usman et al. The PyCBC search for gravitational waves from compact binary coalescence. *Class. Quant. Grav.*, 33(21):215004, 2016.

[54] LA Waistein and VD Zubakov. Extraction of signals from noise. *European Science Notes*, 1, 1970.

[55] Kip S Thorne. Gravitational radiation. *Three hundred years of gravitation*, 1:330–458, 1987.

[56] B. S. Sathyaprakash and S. V. Dhurandhar. Choice of filters for the detection of gravitational waves from coalescing binaries. *Phys. Rev. D*, 44:3819–3834, 1991.

[57] Curt Cutler et al. The Last three minutes: issues in gravitational wave measurements of coalescing compact binaries. *Phys. Rev. L*, 70:2984–2987, 1993.

[58] Lee S Finn. Detection, measurement, and gravitational radiation. *Physical Review D*, 46(12):5236, 1992.

[59] Lee Samuel Finn and David F Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. *Physical Review D*, 47(6):2198, 1993.

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

[61] R. Balasubramanian, B. S. Sathyaprakash, and S. V. Dhurandhar. Gravitational waves from coalescing binaries: Detection strategies and Monte Carlo estimation of parameters. *Phys. Rev. D*, 53:3033–3055, 1996.

[62] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 1. Signal-to-noise for inspiral, merger, and ringdown. *Phys. Rev. D*, 57:4535–4565, 1998.

[63] Luc Blanchet, Thibault Damour, Bala R Iyer, Clifford M Will, and Alan G Wiseman. Gravitational-radiation damping of compact binary systems to second post-Newtonian order. *Physical Review Letters*, 74(18):3515, 1995.

[64] Luc Blanchet, Bala R Iyer, Clifford M Will, and Alan G Wiseman. Gravitational waveforms from inspiralling compact binaries to second-post-Newtonian order. *Classical and Quantum Gravity*, 13(4):575, 1996.

[65] Luc Blanchet, Guillaume Faye, Bala R Iyer, and Benoit Joguet. Gravitational-wave inspiral of compact binary systems to 7/2 post-Newtonian order. *Physical Review D*, 65(6):061501, 2002.

[66] Luc Blanchet, Thibault Damour, Gilles Esposito-Farese, and Bala R Iyer. Gravitational radiation from inspiralling compact binaries completed at the third post-Newtonian order. *Physical review letters*, 93(9):091101, 2004.

[67] Luc Blanchet. Gravitational radiation from post-Newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity*, 9:4, 2006.

[68] Alessandra Buonanno and Thibault Damour. Effective one-body approach to general relativistic two-body dynamics. *Physical Review D*, 59(8):084006, 1999.

[69] Alessandra Buonanno and Thibault Damour. Transition from inspiral to plunge in binary black hole coalescences. *Phys. Rev. D*, 62:064015, Aug 2000.

[70] J. Blackman, S. E. Field, M. A. Scheel, et al. Numerical relativity waveform surrogate model for generically precessing binary black hole mergers. *Phys. Rev. D*, 96(2):024058, July 2017.

[71] P Ajith, S Babak, Y Chen, et al. A phenomenological template family for black-hole coalescence waveforms. *Classical and Quantum Gravity*, 24(19):S689, 2007.

[72] P. Ajith, M. Hannam, S. Husa, et al. Inspiral-Merger-Ringdown Waveforms for Black-Hole Binaries with Nonprecessing Spins. *Physical Review Letters*, 106(24):241101, June 2011.

[73] S. Husa, S. Khan, M. Hannam, et al. Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal. *Phys. Rev. D*, 93(4):044006, February 2016.

[74] S. Khan, S. Husa, M. Hannam, et al. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev. D*, 93(4):044007, February 2016.

[75] M. Hannam, P. Schmidt, A. Bohé, et al. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. *Physical Review Letters*, 113(15):151101, October 2014.

[76] Andrea Taracchini, Alessandra Buonanno, Yi Pan, et al. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys. Rev. D*, 89:061502, Mar 2014.

[77] Alejandro Bohé, Lijing Shao, Andrea Taracchini, et al. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. *Phys. Rev. D*, 95:044028, Feb 2017.

[78] Michael Pürrer. Frequency-domain reduced order models for gravitational waves from aligned-spin compact binaries. *Classical and Quantum Gravity*, 31(19):195010, 2014.

[79] Curt Cutler and Éanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658–2697, Mar 1994.

[80] Ian W. Harry, Alexander H. Nitz, Duncan A. Brown, et al. Investigating the effect of precession on searches for neutron-star–black-hole binaries with Advanced LIGO. *Phys. Rev. D*, 89:024010, Jan 2014.

[81] I. Harry, S. Privitera, A. Bohé, and A. Buonanno. Searching for gravitational waves from compact binaries with precessing spins. *Phys. Rev. D*, 94(2):024012, July 2016.

[82] Lee Samuel Finn and David F. Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. *Phys. Rev. D*, 47:2198–2219, Mar 1993.

[83] S. Babak, R. Biswas, P. R. Brady, et al. Searching for gravitational waves from binary coalescence. *Phys. Rev. D*, 87:024033, Jan 2013.

[84] 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. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 85(12), 2012.

[85] Matteo Frigo and Steven G. Johnson. The Design and Implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[86] Benjamin J Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Physical Review D*, 53(12):6749, 1996.

[87] Benjamin J Owen and BS Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Physical Review D*, 60(2):022002, 1999.

[88] Ian W Harry, Bruce Allen, and BS Sathyaprakash. Stochastic template placement algorithm for gravitational wave data analysis. *Physical Review D*, 80(10):104014, 2009.

[89] P Ajith, N Fotopoulos, S Privitera, et al. Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins. *Physical Review D*, 89(8):084041, 2014.

[90] Stephen M. Privitera. *The importance of spin for observing gravitational waves from coalescing compact binaries with LIGO and Virgo*. PhD thesis, California Institute of Technology, 2014.

[91] Theocharis A. Apostolatos. Construction of a template family for the detection of gravitational waves from coalescing binaries. *Phys. Rev. D*, 54:2421–2437, Aug 1996.

[92] B. S. Sathyaprakash. Filtering post-Newtonian gravitational waves from coalescing binaries. *Phys. Rev. D*, 50:R7111–R7115, Dec 1994.

[93] Duncan A Brown, Ian Harry, Andrew Lundgren, and Alexander H Nitz. Detecting binary neutron star systems with spin in advanced gravitational-wave detectors. *Physical Review D*, 86(8):084017, 2012.

[94] Stanislav Babak. Building a stochastic template bank for detecting massive black hole binaries. *Classical and Quantum Gravity*, 25(19):195011, 2008.

[95] Stephen Privitera, Satyanarayan RP Mohapatra, Parameswaran Ajith, et al. Improving the sensitivity of a search for coalescing binary black holes with nonprecessing spins in gravitational wave data. *Physical Review D*, 89(2):024003, 2014.

[96] GStreamer. `https://gstreamer.freedesktop.org`.

[97] LALSuite. `https://www.lsc-group.phys.uwm.edu/daswg/projects/lalsuite.html`. Accessed: 2015-07-01.

[98] Kipp Cannon, Chad Hanna, and Jacob Peoples. Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation. apr 2015.

[99] Kipp Cannon, J. D. Emberson, Chad Hanna, Drew Keppel, and Harald P. Pfeiffer. Interpolation in waveform space: Enhancing the accuracy of gravitational waveform families using numerical relativity. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 87(4), 2013.

[100] B P Abbott, R Abbott, T D Abbott, et al. Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. *Classical and Quantum Gravity*, 33(13):134001, 2016.

[101] Gravitational Wave Candidate Event Database.

[102] Kipp Cannon, Chad Hanna, and Drew Keppel. Efficiently enclosing the compact binary parameter space by singular-value decomposition. *Phys. Rev. D*, 84:084003, Oct 2011.

[103] Kipp Cannon, Chad Hanna, and Drew Keppel. Method to estimate the significance of coincident gravitational-wave observations from compact binary coalescence. *Phys. Rev. D*, 88:024025, Jul 2013.

[104] Bruce Allen. $\chi$ 2 time-frequency discriminator for gravitational wave detection. *Physical Review D*, 71(6):062001, 2005.

[105] Collin D Capano. *Searching for gravitational waves from compact binary coalescence using LIGO and virgo data*. PhD thesis, Syracuse University, 2011.

[106] Drew Garvin Keppel. *Signatures and dynamics of compact binary coalescences and a search in LIGOâs S5 data*. PhD thesis, California Institute of Technology, 2009.

[107] HTCondor DAGMan.

[108] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. *Phys. Rev. D*, 93:122003, Jun 2016.

[109] S Chatterji, L Blackburn, G Martin, and E Katsavounidis. Multiresolution techniques for the detection of gravitational-wave bursts. *Classical and Quantum Gravity*, 21(20):S1809, 2004.

[110] D. G. Blair. *The Detection of Gravitational Waves*. October 2005.

[111] S. Karki, D. Tuyenbayev, S. Kandhasamy, et al. The Advanced LIGO photon calibrators. *Review of Scientific Instruments*, 87(11):114503, November 2016.

[112] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914. *Phys. Rev. D*, 95(6):062003, March 2017.

[113] A. D. Viets, M. Wade, A. L. Urban, et al. Reconstructing the calibrated strain signal in the Advanced LIGO detectors. *Classical and Quantum Gravity*, 35(9):095015, May 2018.

[114] C. Cahillane, J. Betzwieser, D. A. Brown, et al. Calibration uncertainty for Advanced LIGO's first and second observing runs. *Phys. Rev. D*, 96(10):102001, November 2017.

[115] M. Coleman Miller and Jon M. Miller. The masses and spins of neutron stars and stellar-mass black holes. *Physics Reports*, 548:1 – 34, 2015. The masses and spins of neutron stars and stellar-mass black holes.

[116] Jeffrey E. McClintock, Ramesh Narayan, and James F. Steiner. Black Hole Spin via Continuum Fitting and the Role of Spin in Powering Transient Jets. *Space Sci. Rev.*, pages 1–28, July 2013.

[117] J Aasi, J Abadie, B P Abbott, et al. Characterization of the LIGO detectors during their sixth science run. *Classical and Quantum Gravity*, 32(11):115012, 2015.

[118] Nelson Christensen, the LIGO Scientific Collaboration, and the Virgo Collaboration. LIGO S6 detector characterization studies. *Classical and Quantum Gravity*, 27(19):194010, 2010.

[119] Jessica McIver. Data quality studies of enhanced interferometric gravitational wave detectors. *Classical and Quantum Gravity*, 29(12):124010, 2012.

[120] D M Macleod, S Fairhurst, B Hughey, et al. Reducing the effect of seismic noise in LIGO searches by targeted veto generation. *Classical and Quantum Gravity*, 29(5):055006, 2012.

[121] Joshua R Smith, Thomas Abbott, Eiichi Hirose, et al. A hierarchical method for vetoing noise transients in gravitational-wave detectors. *Classical and Quantum Gravity*, 28(23):235005, 2011.

[122] Tomoki Isogai, the Ligo Scientific Collaboration, and the Virgo Collaboration. Used percentage veto for LIGO and virgo binary inspiral searches. *Journal of Physics: Conference Series*, 243(1):012005, 2010.

[123] R Essick, L Blackburn, and E Katsavounidis. Optimizing vetoes for gravitational-wave transient searches. *Classical and Quantum Gravity*, 30(15):155010, 2013.

[124] Rahul Biswas, Lindy Blackburn, Junwei Cao, et al. Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. *Phys. Rev. D*, 88:062003, Sep 2013.

[125] T A Callister, J B Kanner, T J Massinger, S Dhurandhar, and A J Weinstein. Observing gravitational waves with a single detector. *Classical and Quantum Gravity*, 34(15):155007, 2017.

[126] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Properties of the Binary Black Hole Merger GW150914. *Phys. Rev. Lett.*, 116:241102, Jun 2016.

[127] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Improved Analysis of GW150914 Using a Fully Spin-Precessing Waveform Model. *Phys. Rev. X*, 6:041014, Oct 2016.

[128] B. P. Abbott, R. Abbott, T. D. 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.

[129] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Search for sub-solar mass ultracompact binaries in Advanced LIGO's first observing run. *ArXiv e-prints*, August 2018.

[130] R. Magee, A.-S. Deutsch, P. McClincy, et al. Methods for the detection of gravitational waves from sub-solar mass ultracompact binaries. *ArXiv e-prints*, August 2018.

[131] Alessandra Buonanno, Bala R Iyer, Evan Ochsner, Yi Pan, and Bangalore Suryanarayana Sathyaprakash. Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Physical Review D*, 80(8):084043, 2009.

[132] Debnandini Mukherjee, Sarah Caudill, Ryan Magee, et al. The GstLAL template bank for spinning compact binary mergers in the second observation run of Advanced LIGO and Virgo. *arXiv e-prints*, page arXiv:1812.05121, Dec 2018.

[133] Leo Tsukada, Kipp Cannon, Chad Hanna, et al. Application of a zero-latency whitening filter to compact binary coalescence gravitational-wave searches. *Phys. Rev. D*, 97:103009, May 2018.

[134] N. Damera-Venkata, B. L. Evans, and S. R. McCaslin. Design of optimal minimum-phase digital FIR filters using discrete Hilbert transforms. *IEEE Transactions on Signal Processing*, 48(5):1491–1495, May 2000.

[135] 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.

[136] X. J. Zhu, E. Howell, T. Regimbau, D. Blair, and Z. H. Zhu. Stochastic Gravitational Wave Background from Coalescing Binary Black Holes. *Astrophys. J.*, 739:86, 2011.

[137] P. A. Rosado. Gravitational wave background from binary systems. *Phys. Rev. D*, 84:084004, 2011.

[138] S. Marassi, R. Schneider, G. Corvino, V. Ferrari, and S. P. Zwart. Imprint of the merger and ring-down on the gravitational wave background from black hole binaries coalescence. *Phys. Rev. D*, 84:124037, 2011.

[139] C. Wu, V. Mandic, and T. Regimbau. Accessibility of the gravitational-wave background due to binary coalescences to second and third generation gravitational-wave detectors. *Phys. Rev. D*, 85:104024, 2012.

[140] Xing-Jiang Zhu, Eric J. Howell, David G. Blair, and Zong-Hong Zhu. On the gravitational wave background from compact binary coalescences in the band of ground-based interferometers. *MNRAS*, 431:882–899, 2013.

[141] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW150914: Implications for the Stochastic Gravitational-Wave Background from Binary Black Holes. *Phys. Rev. Lett.*, 116:131102, Mar 2016.

[142] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences. *Phys. Rev. Lett.*, 120:091101, Feb 2018.

[143] L. P. Grishchuk. Amplification of gravitational waves in an isotropic universe. *Sov. Phys. JETP*, 40:409, 1975.

[144] A. A. Starobinskiĭ. Spectrum of relict gravitational radiation and the early state of the universe. *JETP Lett.*, 30:682, 1979.

[145] L. P. Grishchuk. Relic gravitational waves and limits on inflation. *Phys. Rev. D*, 48:3513–3516, 1993.

[146] M. Gasperini and G. Veneziano. Pre-big-bang in string cosmology. *Astropart. Phys.*, 1:317–339, 1993.

[147] A. Buonanno, M. Maggiore, and C. Ungarelli. Spectrum of relic gravitational waves in string cosmology. *Phys. Rev. D*, 55:3330–3336, 1997.

[148] J. F. Dufaux, D. G. Figueroa, and J. García-Bellido. Gravitational waves from Abelian gauge fields and cosmic strings at preheating. *Phys. Rev. D*, 82:083518, 2010.

[149] T. Damour and A. Vilenkin. Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows. *Phys. Rev. D*, 71:063510, 2005.

[150] X. Siemens, V. Mandic, and J. Creighton. Gravitational-Wave Stochastic Background from Cosmic Strings. *Phys. Rev. Lett.*, 98:111101, 2007.

[151] S. Ölmez, V. Mandic, and X. Siemens. Gravitational-wave stochastic background from kinks and cusps on cosmic strings. *Phys. Rev. D*, 81:104028, 2010.

[152] T. Regimbau, S. Giampanis, X. Siemens, and V. Mandic. Stochastic background from cosmic (super)strings: Popcorn-like and (Gaussian) continuous regimes. *Phys. Rev. D*, 85:066001, 2012.

[153] C. Caprini, R. Durrer, and G. Servant. Gravitational wave generation from bubble collisions in first-order phase transitions: An analytic approach. *Phys. Rev. D*, 77:124015, 2008.

[154] C. Caprini, R. Durrer, T. Konstandin, and G. Servant. General properties of the gravitational wave spectrum from phase transitions. *Phys. Rev. D*, 79:083519, 2009.

[155] C. Caprini, R. Durrer, and G. Servant. The stochastic gravitational wave background from turbulence and magnetic fields generated by a first-order phase transition. *Journal of Cosmology and Astroparticle Physics*, 12:024, 2009.

[156] M. Maggiore. Gravitational wave experiments and early universe cosmology. *Physics Reports*, 331:283–367, 2000.

[157] P. Binétruy, A. Bohé, C. Caprini, and J.-F. Dufaux. Cosmological backgrounds of gravitational waves and eLISA/NGO: phase transitions, cosmic strings and other sources. *Journal of Cosmology and Astroparticle Physics*, 6:27, 2012.

[158] T. Regimbau. The astrophysical gravitational wave stochastic background. *Res. Astron. Astrophys.*, 11:369, 2011.

[159] A. Buonanno, G. Sigl, G. G. Raffelt, H. T. Janka, and E. Müller. Stochastic gravitational-wave background from cosmological supernovae. *Phys. Rev. D*, 72:084001, 2005.

[160] P. Sandick, K. A. Olive, F. Daigne, and E. Vangioni. Gravitational Waves from the First Stars. *Phys. Rev. D*, 73:104024, 2006.

[161] S Marassi, R Schneider, and V Ferrari. Gravitational wave backgrounds and the cosmic transition from Population III to Population II stars. *MNRAS*, 398:293, 2009.

[162] X. J. Zhu, E. Howell, and D. Blair. Observational upper limits on the gravitational wave production of core collapse supernovae. *MNRASL*, 409:L132–L136, 2010.

[163] T. Regimbau and J. A. de Freitas Pacheco. Cosmic background of gravitational waves from rotating neutron stars. *Astron. & Astrophys.*, 376:381, 2001.

[164] P. A. Rosado. Gravitational wave background from rotating neutron stars. *Phys. Rev. D*, 86:104007, 2012.

[165] T. Regimbau and J. A. de Freitas Pacheco. Gravitational wave background from magnetars. *Astron. & Astrophys.*, 447:1–7, 2006.

[166] E. Howell, T. Regimbau, A. Corsi, D. Coward, and R. Burman. Gravitational wave background from sub-luminous GRBs: prospects for second- and third-generation detectors. *MNRAS*, 410:2123, 2011.

[167] S. Marassi, R. Ciolfi, R. Schneider, L. Stella, and V. Ferrari. Stochastic background of gravitational waves emitted by magnetars. *MNRAS*, 411:2549, 2011.

[168] C. J. Wu, V. Mandic, and T. Regimbau. Accessibility of the stochastic gravitational wave background from magnetars to the interferometric gravitational wave detectors. *Phys. Rev. D*, 87:042002, 2013.

[169] J. C. N. de Araujo and G. F. Marranghello. Gravitational wave background from neutron star phase transition. *General Relativity and Gravitation*, 41:1389–1406, 2009.

[170] V. Ferrari, S. Matarrese, and R. Schneider. Stochastic background of gravitational waves generated by a cosmological population of young, rapidly rotating neutron stars. *MNRAS*, 303:258–264, 1999.

[171] E. Howell, D. Coward, R. Burman, D. Blair, and J. Gilmore. The gravitational wave background from neutron star birth throughout the cosmos. *MNRAS*, 351:1237–1246, 2004.

[172] X. J. Zhu, X. L. Fan, and Z. H. Zhu. Stochastic Gravitational Wave Background from Neutron Star r-mode Instability Revisited. *Astrophys. J.*, 729:59, 2011.

[173] M Punturo, M Abernathy, F Acernese, et al. The Einstein Telescope: a third-generation gravitational wave observatory. *Classical and Quantum Gravity*, 27(19):194002, 2010.

[174] B P Abbott, R Abbott, T D Abbott, et al. Exploring the sensitivity of next generation gravitational wave detectors. *Classical and Quantum Gravity*, 34(4):044001, 2017.

[175] B. Allen and J. D. Romano. Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities. *Phys. Rev. D*, 59:102001, 1999.

[176] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. Planck 2015 results - XIII. Cosmological parameters. *A&A*, 594:A13, 2016.

[177] K. G. Arun, Bala R Iyer, B. S. Sathyaprakash, and Pranesh A Sundararajan. Parameter estimation of inspiralling compact binaries using 3.5 post-Newtonian gravitational wave phasing: The nonspinning case. *Phys. Rev. D*, 71:084008, Apr 2005.

[178] T. Regimbau, T. Dent, W. Del Pozzo, et al. Mock data challenge for the Einstein Gravitational-Wave Telescope. *Phys. Rev. D*, 86(12):122001, December 2012.

[179] T. Regimbau, D. Meacher, and M. Coughlin. Second Einstein Telescope mock science challenge: Detection of the gravitational-wave stochastic background from compact binary coalescences. *Phys. Rev. D*, 89(8):084046, April 2014.

[180] D. Meacher, M. Coughlin, S. Morris, et al. Mock data and science challenge for detecting an astrophysical stochastic gravitational-wave background with Advanced LIGO and Advanced Virgo. *Phys. Rev. D*, 92(6):063002, September 2015.

[181] E. E. Salpeter. The Luminosity Function and Stellar Evolution. *Astrophys. J.*, 121:161, January 1955.

[182] Lars Hernquist and Volker Springel. An analytical model for the history of cosmic star formation. *Monthly Notices of the Royal Astronomical Society*, 341(4):1253–1267, 2003.

[183] Matthew D. Kistler, Hasan Yuksel, and Andrew M. Hopkins. The Cosmic Star Formation Rate from the Faintest Galaxies in the Unobservable Universe. 2013.

[184] P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. *Astrophys. J.*, 770:57, June 2013.

[185] Michele Trenti, Rosalba Perna, and Sandro Tacchella. Gamma-Ray Burst and Star Formation Rates: The Physical Origin for the Redshift Evolution of Their Ratio. *The Astrophysical Journal Letters*, 773(2):L22, 2013.

[186] Peter S. Behroozi and Joseph Silk. A Simple Technique for Predicting High-redshift Galaxy Evolution. *The Astrophysical Journal*, 799(1):32, 2015.

[187] D. Meacher, M. Coughlin, S. Morris, et al. Mock data and science challenge for detecting an astrophysical stochastic gravitational-wave background with Advanced LIGO and Advanced Virgo. *Phys. Rev. D*, 92:063002, 2015.

[188] Michal Dominik, Krzysztof Belczynski, Christopher Fryer, et al. Double Compact Objects. II. Cosmological Merger Rates. *The Astrophysical Journal*, 779(1):72, 2013.

[189] K. Belczynski, V. Kalogera, and T. Bulik. A Comprehensive Study of Binary Compact Objects as Gravitational Wave Sources: Evolutionary Channels, Rates, and Physical Properties. *Astrophys. J.*, 572:407, 2002.

[190] Shin'ichiro Ando. Short gamma-ray bursts as a possible probe of binary neutron star mergers. *Journal of Cosmology and Astroparticle Physics*, 2004(06):007, 2004.

[191] K. Belczynski, R. Perna, T. Bulik, et al. A Study of Compact Object Mergers as Short Gamma-Ray Burst Progenitors. *Astrophys. J.*, 648:1110, 2006.

[192] E. Berger, D. B. Fox, P. A. Price, et al. A New Population of High-Redshift Short-Duration Gamma-Ray Bursts. *The Astrophysical Journal*, 664(2):1000, 2007.

[193] Ehud Nakar. Short-hard gamma-ray bursts. *Physics Reports*, 442(1):166 – 236, 2007. The Hans Bethe Centennial Volume 1906-2006.

[194] R. O'Shaughnessy, K. Belczynski, and V. Kalogera. Short Gamma-Ray Bursts and Binary Mergers in Spiral and Elliptical Galaxies: Redshift Distribution and Hosts. *The Astrophysical Journal*, 675(1):566, 2008.

[195] Michal Dominik, Krzysztof Belczynski, Christopher Fryer, et al. Double Compact Objects. I. The Significance of the Common Envelope on Merger Rates. *The Astrophysical Journal*, 759(1):52, 2012.

[196] Piero Madau and Mark Dickinson. Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics*, 52(1):415–486, 2014.

[197] Krzysztof Belczynski, Daniel E. Holz, Tomasz Bulik, and Richard O'Shaughnessy. The first gravitational-wave source from the isolated evolution of two stars in the 40–100 solar mass range. *Nature*, 534(7608):512–515, jun 2016.