The Pennsylvania State University
The Graduate School

GRAVITATIONAL WAVE SEARCHES TOWARDS MULTI-MESSENGER AND MULTI-BAND ASTRONOMY

A Dissertation in
Physics
by
Becca Ewing

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2024

The dissertation of Becca Ewing was reviewed and approved by the following:

Chad Hanna
Professor of Physics & Astronomy and Astrophysics
Dissertation Co-Advisor
Co-Chair of Committee

Bangalore Sathyaprakash
Elsbach Professor of Physics & Professor of Astronomy and Astrophysics
Dissertation Co-Advisor
Co-Chair of Committee

Sarah Shandera
Associate Professor of Physics

Caryl Gronwall
Senior Scientist, Astronomy and Astrophysics

Irina Mocioiu
Associate Professor of Physics, Director of Graduate Studies

Abstract

Einstein's general theory of relativity describes how the curvature of spacetime is determined by the energy and momentum of massive objects. He found that his field equations admit wave solutions which he called *gravitational waves*: plane waves traveling at the speed of light and carrying a perturbation of the spacetime metric as they propagate throughout the universe. The physical existence of such waves was in doubt for many years, until their indirect observation via the slow orbital decay of the binary pulsar system, PSR 1913 + 16, by Russell Hulse and Joseph Taylor. At this point, the endeavor to directly observe gravitational waves was begun.

It took many years to develop experimental techniques capable of observing the tiny magnitude of strain expected from compact binary mergers ($h \sim 10^{-21}$), but finally on September 14th 2015 the Advanced LIGO detectors succeeded in making the first observation of gravitational waves from a pair of merging black holes with the signal known as GW150914. This was a landmark moment in physics, confirming the prediction of Einstein's theory nearly one hundred years later and opening up an entirely new window with which to observe the universe. Since that date, gravitational waves have become regular cosmic messengers carrying with them information about their extreme sources and unexpected new insights in many diverse areas of physics. Without a doubt the single most informative detection has been the joint observation of gravitational waves and electromagnetic radiation from the merging neutron stars and short gamma ray burst, known as GW170817 and GW170817A, respectively. This observation marked the beginning of multi-messenger astronomy with gravitational waves and advanced our knowledge of the universe by leaps and bounds. However, as all great discoveries do, this observation brought even more unanswered questions. While we have seen an unprecedented rate of growth in the field of gravitational wave astronomy in just a decade of observation, we are still in the early years. There is much still to learn from gravitational waves, however to get there we must aim to make as many detections as possible, building up populations of gravitational waves to reveal bulk properties of their sources. This has been the subject of my dissertation.

GstLAL is a data analysis algorithm, known as a *search pipeline* used to identify quiet gravitational wave signals within the noisy LIGO, Virgo, and KAGRA strain data. My work has focused on the development and operation of this pipeline, particularly towards detecting gravitational waves in near real time to enable multi-messenger astronomy. This dissertation will overview the methods and results of the various gravitational wave searches which I have led, culminating in $\mathcal{O}(100)$ new detections in just a few years. I additionally present a study towards the next generation of gravitational wave

observation, showing how detection in ground-based interferometers can be used to identify stellar-mass binary black hole mergers in the low-frequency data of LISA. The methods presented here all seek to advance our detection capabilities and particularly, to improve our immediate prospects for multi-messenger astronomy. The population of new gravitational wave signals which have been identified as a result of my work will certainly have long-lasting implications for astrophysics, cosmology, and fundamental physics.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I want to thank my twin, Rachel, who is my best friend and probably the main reason why I have accomplished anything in life. Since we were young she has been a source of inspiration for me and someone who I wish to emulate in many ways. Without her I would not have begun this endeavor and I certainly would not have finished it. She supported me at every critical juncture, always making me laugh and helping me find levity in the most stressful times, and for that I will always be grateful.

I also want to thank my family, particularly my mom and dad, Lynda and Steve Ewing, for encouraging me from a young age to take my education seriously. They always believed in me and from them I learned about integrity and hard work, a few lessons which I have carried with me through all these years. I want to thank my Grandpa Childers too, who always memorized the admittedly verbose titles of my undergraduate papers and made sure that I knew he was proud of me. His encouragement and love were a constant source of motivation for me. Cassie, Doug, and Matt have each been incredible role models and friends and I'm grateful for their support over the years.

I had the pleasure of spending my years in graduate school alongside an incredible cohort of colleagues and friends. Divya, Bing, Kokki, Dan, and many others made living in State College worth it. I feel lucky to have gone through this experience with them. From the chaos of first year, the holidays and countless meals we've shared, to now going through the process of graduating together, these are all really precious memories of what I now consider to be some of my best years. Divya and Bing were the best roommates anyone could ask for. Every meal we cooked together (and the ones they cooked that I just ate), the parties we hosted, the long and loud gossip sessions in the kitchen, and every time we sang karaoke in the living room all kept me going these past years. I can't imagine having done this without them.

I was lucky to have several great mentors here at Penn State. Firstly, my advisor Chad who provided the perfect environment for me to learn and grow into an independent researcher. Working with and learning from Chad over the years has been an absolute pleasure (even when I had to wake up in the middle of the night for alerts). I can't overstate how much I learned from working with Sathya in my first few years of graduate school. Reading his papers as an undergraduate sparked my curiosity for this field and getting to work with him is something I will always be grateful for. Additionally, I have to give credit to my seniors in Chad's group: Surabhi, Cody, Ryan, Patrick, and Leo who sat with me and helped me write my first Python code, always answered my

questions with patience, and more than anything made working in this group fun. Finally, I couldn't have survived these past years without Divya and Rachael in my corner. I feel incredibly lucky to have had two colleagues as silly, inspiring, and intelligent as them, who never judged me when I asked dumb questions; we truly got through this together.

The material in this dissertation is based on work funded by the National Science Foundation under grants: NSF-1454389, NSF-1642391, NSF-1700765, NSF-1841480, NSF-1841598, NSF-1934752, NSF-2011865, NSF-2018299, NSF-2103662, NSF-2110594, NSF-2201445, NSF-2308881, NSF-2310912, NSF-2311355, NSF-2346596. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The authors of Chapter 5 are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation (NSF) and operates under cooperative agreement PHY-1764464. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Netherlands Organization for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors are grateful for computational resources provided by the Pennsylvania State University's Institute for Computational and Data Sciences (ICDS) and the University of Wisconsin Milwaukee Nemo and support by NSF PHY-2011865, NSF OAC-2103662, NSF PHY-1626190, NSF PHY-1700765, and NSF PHY-2207728. M.W.C. acknowledges support from the National Science Foundation with grant numbers PHY-2010970 and OAC-2117997. R.E. is supported by the Natural Sciences & Engineering Research Council of Canada (NSERC).0 Chapter 5 carries LIGO Document Number LIGO-P2300124.

The authors of Chapter 7 thank Anuradha Gupta for providing access to her simulation of binary black holes that can be observed both in *LISA* and the 3G network. BE is supported by the Eberly Graduate Fellowship of Penn State. SS is supported by the Eberly Postdoctoral Fellowship of Penn State. BSS and SB are supported in part by NSF Grant No. PHY-1836779.

Chapter 1
Introduction

On September 14, 2015 the Advanced LIGO detectors made the first observation of gravitational waves from a pair of merging black holes in a signal known as GW150914. These gravitational waves carry information about their source and travel unimpeded across millions of light years until they pass through the Earth. Gravitational waves are therefore a new kind of cosmic messenger, completely distinct from electromagnetic radiation. By observing gravitational waves we can now "listen" to the universe in addition to viewing it with light. However, we gain a more full picture by combining observations from multiple cosmic messengers in joint detections. This is the field known as multi-messenger astronomy; while this field has existed for a long time, it has only been in the past decade that gravitational waves have joined the cast of viable messengers. The watershed detection of gravitational waves from a pair of merging neutron stars (known as GW170817) and the accompanying short gamma ray burst observed at frequencies all across the electromagnetic spectrum marked the beginning of multi-messenger astronomy with gravitational waves. This single detection brought a wealth of information about its unique source and new insights to astrophysics, cosmology, and fundamental physics. We can now say that the era of gravitational wave astronomy is in full swing.

My work has directly contributed to the now regular detection of gravitational waves from merging black holes and neutron stars. This dissertation will focus on the methods I have developed and led to detect gravitational waves and improve our prospects for multi-messenger astronomy. I will begin in Chapter 1 with background material starting with the prediction of gravitational waves from general relativity, and discussing the sources of these waves, the instrumental and data analysis methods used to detect them, and an outlook for the next generation of observation. In Chapter 2, I overview the general methods of the GstLAL search pipeline used to detect gravitational waves in near real time, as well as the technical software implementation of the pipeline in the LIGO

Scientific, Virgo and KAGRA (LVK) Collaboration's fourth observing run. Chapter 3 overviews a detection strategy aimed at improving our prospects for multi-messenger astronomy by sharing sub-threshold gravitational wave detections with astronomers at the Neil Gehrels Swift Observatory in real time with the goal of enabling joint gamma ray burst observations. A significant portion of my dissertation was spent contributing to the production of the Gravitational Wave Transient Catalog and the addition of 35 new gravitational wave candidates in the second part of the LVK Collaboration's third observing run; this work is described in Chapter 4. Returning to low latency detection, Chapters 5 and 6 detail the preparation of the GstLAL search pipeline ahead of the fourth observing run and the new public results from the run, including 81 new gravitational wave detections. Finally, in Chapter 7, I present a study of our capabilities for joint detections among ground- and space-based gravitational wave detectors of the next generation. I conclude in Chapter 8 with a summary and prospects for future work.

1.1 Gravitational waves in general relativity

The existence of gravitational waves was first postulated by Albert Einstein as a mathematical consequence of his general theory of relativity [8, 9]. Although in the following years, Einstein himself erroneously cast doubt on the physical existence of gravitational waves, viewing them as merely mathematical artifacts of the linearized theory. This led to several decades of stunted progress in the study of gravitational waves, until it was finally proven in the mid-1900's that these waves do in fact exist in the fully non-linearized theory and that they carry energy and momentum [10, 11]. After this, the experimental pursuit to directly detect gravitational waves finally began, eventually culminating in the construction of the LIGO, Virgo, KAGRA interferometric detectors and the observation of gravitational waves from two merging black holes, a signal known as GW150914 [12]. In this section, I will briefly review the aspects of general relativity leading to the existence of gravitational waves [13–15].

In the Newtonian theory, gravity is a force which is transmitted instantaneously between massive objects. However, the theory of special relativity imposes the speed of light in vacuum as the theoretical maximum speed of information transfer. Therefore, in general relativity Einstein reformulated gravity as a consequence of the geometry of spacetime. In other words, it is the spacetime metric which transmits the force of gravity. Massive sources cause the local spacetime to warp and curve and only in a vacuum, far away from any masses can the spacetime be approximated as flat Euclidean

space. Otherwise, no matter the choice of coordinates, the spacetime will always have a curvature. This curvature determines how objects move in the absence of other forces. Further, the acceleration of massive sources will cause spacetime to distort dynamically. It is natural to expect then that the distortion will "ripple" throughout the spacetime as waves moving at a finite speed.

The Einstein field equations formalize how massive objects determine the curvature of spacetime, and in turn how the shape of spacetime determines how massive objects move. The equations are written as follows,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \kappa T_{\mu\nu}. \tag{1.1}$$

The left hand side describes the curvature of spacetime. $R_{\mu\nu}$ is the Ricci curvature tensor which describes how much the local spacetime metric differs from flat Euclidean space, in which case we would have $R_{\mu\nu} = 0$. R is the Ricci curvature scalar, which is defined as the trace of $R_{\mu\nu}$. The metric tensor $g_{\mu\nu}$ defines the proper distance between nearby points, i.e. $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$. The simplest example is for a flat spacetime, for which we have the Minkowski metric,

$$\eta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1.2}$$

The right hand side of Eqn. 1.1 describes the energy and momentum of massive objects, given by the stress-energy tensor $T_{\mu\nu}$. The constant $\kappa = 8\pi G/c^4$ has units of inverse force (N^{-1}) and can be thought of as indicating the magnitude of force required to warp the spacetime.

To predict the existence of gravitational waves, we consider a metric which is described by a small perturbation away from the Minkowski metric, i.e.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu},\ |h_{\mu\nu}| \ll 1. \tag{1.3}$$

Then Eqn. 1.1 can be linearized by keeping terms only to leading order in the perturbation. The vacuum equations (that is, outside the source where $T_{\mu\nu} = 0$) then admit wave solutions:

$$\Box \bar{h}_{\mu\nu} = 0,\ \bar{h}_{\mu\nu} = A_{\mu\nu} e^{ik_\sigma x^\sigma}, \tag{1.4}$$

where $\Box$ indicates the four-dimensional Laplacian or wave operator, $\Box = g^{\mu\nu}\partial_\nu\partial_\mu$. Here,

we have denoted the perturbation as $\bar{h}_{\mu\nu}$ to indicate the choice of Lorenz gauge. From this, it can be shown that metric perturbations are plane waves traveling at the speed of light, with the perturbations perpendicular to the direction of propagation of the wave. By further choosing the transverse-traceless gauge, the metric perturbations are (i) in the spatial dimension only and perpendicular to the propagation of the wave and (ii) symmetric in rotations about the direction of propagation of the wave. We can choose the propagation direction to be $\hat{z}$ so that the metric perturbations are in the $\hat{x}$ and $\hat{y}$ directions with $h_{xx} = -h_{yy}$ and $h_{xy} = h_{yx}$. These quantities represent the two different polarization states of a gravitational wave and are typically referred to as h_+ and $h_\times$, i.e.

$$h = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.5}$$

The magnitude of the perturbation, $|h|$ is the gravitational wave amplitude. The amplitude determines the fractional change in distance between two fixed points (or test masses) in each direction perpendicular to the wave propagation direction. This is referred to as the *strain*. Considering for example, the h_+ polarization as a gravitational wave passes the distances in the $\hat{x}$ direction will be stretched from $L \to L(1+h)$ while distances in the $\hat{y}$ direction are compressed, $L \to L(1-h)$. The gravitational wave strain is then,

$$|h| = \frac{\Delta L}{L}. \tag{1.6}$$

1.2 Sources of gravitational waves

We expect the strain induced by a passing gravitational wave to be extremely small, making their direct detection difficult. To understand the kinds of sources which will emit gravitational waves and the detectability of these signals, we must understand the energy carried by gravitational waves. This is related to the wave amplitude, which depends on the quadrupole mass moment of the source as follows [14]:

$$h_{ij} = \frac{2}{c^4}\frac{G}{D_L}\frac{\mathrm{d}^2 Q_{ij}}{\mathrm{d}t^2}. \tag{1.7}$$

Here, $G = 6.674 \times 10^{-11}\ \mathrm{N \cdot m^2/kg^2}$ is the gravitational constant, $c = 2.998 \times 10^8\ \mathrm{m/s^2}$ is the speed of light in vacuum, and D_L is the luminosity distance to the source. Q_{ij} is the quadrupole mass moment of the source which is defined as:

$$Q_{ij} = \int \mathrm{d}^3\vec{x}\rho(\vec{x}) \left(x_i x_j - \frac{1}{3} r^2 \delta_{ij} \right), \tag{1.8}$$

where the indices $\{i, j\}$ indicate spatial coordinates, r is the distance from the coordinate origin (typically chosen as the center of mass), and δ_{ij} is the Kronecker delta. From this we see that $\ddot{Q}_{ij}$ has units of energy, making h_{ij} a dimensionless quantity, as required. The quadrupole mass moment of a spherically symmetric source will be zero, therefore only sources with non-spherically symmetric acceleration may emit gravitational waves [14].

In Eqn. 1.7, we see again the factor of G/c^4 (whose inverse appeared first in the Einstein field equations, Eqn. 1.1). The tiny magnitude of this factor, $\sim 10^{-44}$ indicates that we require an extremely energetic source to produce even very small strain amplitudes. Therefore, any man-made sources will produce gravitational waves far too weak to ever be detected. On the other hand, the universe provides many astrophysical sources which are powerful enough to produce radiation which is observable. For the purpose of detection, these sources are generally categorized as transient or continuous and modeled or unmodeled. I will briefly give examples of sources in each category, but the rest of this section and indeed this dissertation as a whole will focus on the detection of transient modeled sources only.

The transient sources are those for which the signal duration is short relative to the observation time. Examples include compact binary coalescences (CBCs) – mergers of the most compact objects in the known universe, i.e. black holes and neutron stars – as well-modeled sources (see Sec. 1.4). Famously, in 1975 Richard Alan Hulse and Joseph Hooten Taylor discovered a binary pulsar system later inferring the loss of energy due to gravitational wave emission by observing the orbital decay [16, 17]. This was the first indirect observational evidence for the existence of gravitational waves, many years ahead of the construction of the LIGO, Virgo, and KAGRA interferometers.

Neutron stars and black holes are known to be formed by the gravitational core-collapse of massive stars. These powerful events are potential sources of transient bursts of gravitational emission, assuming that the collapse is sufficiently non-spherical. However, the many complex dynamics at play in such an event are formidable for current simulations to accurately capture and thus the exact strength and form of the emitted gravitational radiation is highly uncertain [18–20]. Still, current estimates suggest that

the LIGO, Virgo, and KAGRA detectors may be sensitive to such sources and searches for these signals have been carried out since the first observing run [21–23]. These searches represent an exciting prospect to gain unexpected new information about the universe.

Sources which produce gravitational radiation lasting much longer than the typical observation duration are known as continuous sources. For example, an isolated neutron star with some asymmetry such as a "bump" of mass on its surface may slowly radiate energy away as weak gravitational waves over many millions of years [24–26]. This radiation is expected to leak away the rotational energy of the neutron star, producing a spindown such that older neutron stars rotate much more slowly. Depending on the size of the mass asymmetry (which would determine the energy loss to gravitational waves), the LIGO, Virgo, and KAGRA detectors may be able to observe these signals, or otherwise place upper limits on the mass asymmetries, providing interesting insights to neutron star physics which may not be obtainable through other astrophysical messengers [27, 28].

In the unmodeled continuous regime we expect stochastic gravitational waves. This is a superposition of weak gravitational waves originating from the various sources throughout the universe which together produce a continuous background [29, 30]. Such a signal cannot be modeled as the individual wave components are not discernible, but could nevertheless be detected by observing cross-correlated data from multiple detectors over very long durations [31, 32]. Particularly, these stochastic gravitational waves could carry information directly from the very early universe, just fractions of a second after the Big Bang. Therefore a measurement of these waves would be hugely influential for cosmology and high energy physics [33].

Returning to the transient, well-modeled case of CBCs – for example a BBH or BNS, I will now give an overview of the characteristic strain amplitude produced by such sources as well as the known signal morphology. For these systems, the motion in the source is purely non-spherical. This fact makes these sources excellent natural laboratories for studying gravitational waves as the amplitude of the emitted waves is maximal in this case. By simple dimensional analysis of Eqn. 1.8 we can see that $\ddot{Q}_{ij}$ will be of the order Mv^2 where $M = m_1 + m_2$ is the total mass of the binary with component masses m_1, m_2 and v is the internal velocity. The strain of a CBC source is then,

$$h = \frac{\Delta L}{L} \sim \frac{2}{c^4}\frac{G}{D_L}(Mv^2) \tag{1.9}$$

Here, the internal velocity of the system is not an observable, but we can use the virial

theorem[1] to obtain an order of magnitude estimate for the strain of a typical BNS system. This allows us to replace the unknown v^2 with the internal gravitational field, $\phi_{\rm int} = GM/r$, where r is the orbital separation. We can obtain a convenient relation by further replacing $\phi_{\rm ext} = GM/D_L$, i.e.

$$h \sim \frac{2}{c^4}\phi_{\rm ext}\phi_{\rm int}. \tag{1.10}$$

Using this relation, we can find the typical strain of a BNS system with component masses $m_i = 1.4M_\odot$, where $M_\odot = 1.988 \times 10^{30}$ kg is the mass of the sun. At a distance of 40 Mpc and an assumed orbital separation of several hundred kilometers, we obtain $h \sim 10^{-21}$. In the next section, I will briefly summarize the methods used by the LIGO, Virgo, and KAGRA interferometers to measure strains of this magnitude.

Conceptually, we understand that since gravitational waves carry energy they will gradually deplete the energy of the orbit,

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\rm orb} = -\frac{\mathrm{d}}{\mathrm{d}t}E_{\rm GW}. \tag{1.11}$$

There are several references which work out this calculation in detail [14, 34, 35]. The end result is a relation between the orbital frequency and a combination of the binary masses known as the *chirp mass*:

$$\dot{\omega} = \left(\frac{96}{5}\right)\frac{\omega^{11/3}}{c^5}(G\mathcal{M}_c)^{5/3}. \tag{1.12}$$

Here, $\omega = 2\pi f_{\rm orb}$ is the angular velocity of the orbit[2]. The chirp mass, $\mathcal{M}_c$, defined as:

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{1.13}$$

is closely related to the amplitude of the gravitational waves, and thus the amount of energy carried away from the system. Eqn. 1.12 shows the well known "chirping" morphology of gravitational radiation from a binary orbit: as the waves carry away energy the orbital frequency increases while, by Kepler's laws the radius of the binary decreases. This results in the eventual merger of the two objects and the formation of a final remnant object.

[1]The virial theorem says that the kinetic energy of a self-gravitating system is approximately equal to half its gravitational potential energy [14].

[2]The orbital frequency is *half* the gravitational wave frequency, so $\omega = \frac{2\pi f_{\rm GW}}{2}$.

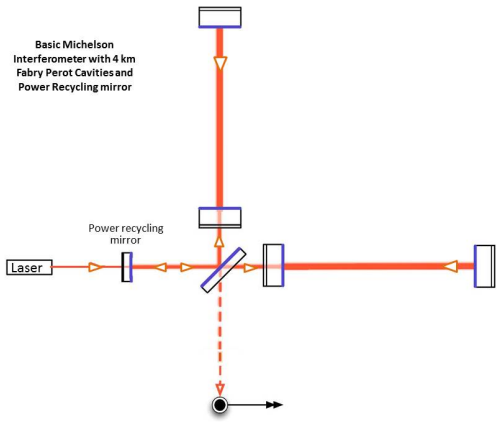

Figure 1.1. Schematic of a basic Michelson interferometer with the addition of Fabry-Perot optical cavities and power recycling as used by the advanced LIGO detectors. Reprinted from [1].

1.3 Current gravitational wave detectors

The problem of gravitational wave detection is to measure the tiny changes in distance induced by a passing wave. As shown in the previous section, for some of the most energetic and efficient sources of gravitational radiation, i.e. merging black holes and neutron stars, the gravitational wave strain can be on the order of $h \sim 10^{-21}$. This is orders of magnitude smaller than the size of a proton. To observe such a small change in length, any detector must be extremely sensitive and employ advanced technologies to reduce sources of noise.

Today, the advanced LIGO, Virgo, and KAGRA interferometers are the most sensitive gravitational wave detectors ever built [36–40]. These detectors are built on the principle of basic Michelson interferometry with a few critical modifications to improve sensitivity. Fig. 1.1 shows a basic illustration of a Michelson interferometer. The detector is constructed with two perpendicular arms where a laser beam is directed through a beam splitter. In LIGO, these arms are 4 km in length, while Virgo and KAGRA have 3 km arms. The mirrors at the end of each arm reflect the laser beams back to the beam splitter where they recombine. The transmitted light is then incident on an output photodiode. Here, the mirrors act as test masses. As a gravitational wave passes through the detector plane the distance between each test mass and the beam splitter will oscillate: as one arm increases in length the other will decrease and vice versa throughout wave periods. This change in distance translates to a difference in the time taken by the laser beam to traverse the arm length, leading to a phase shift in the recombined beams and

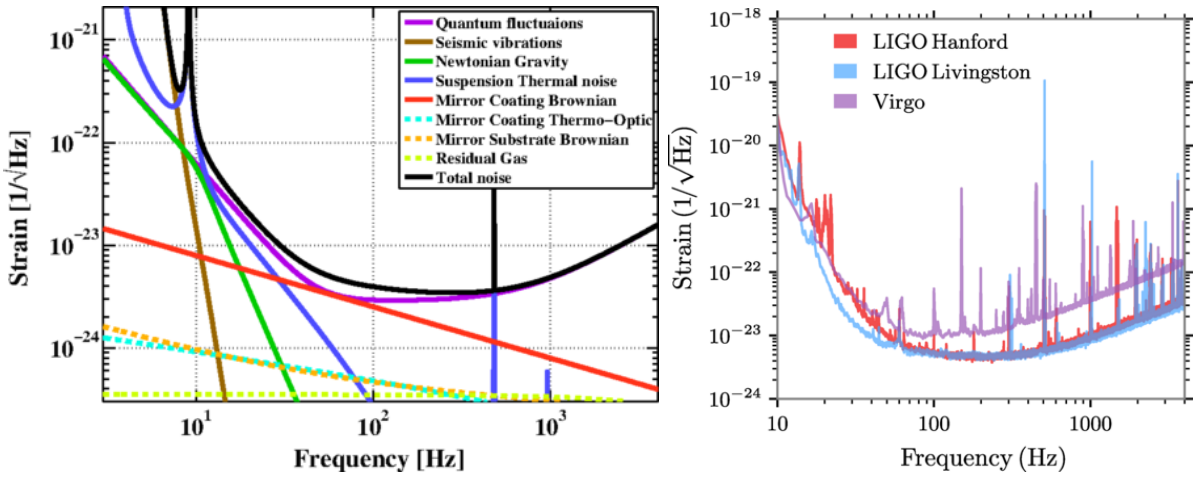

Figure 1.2. **Left**. Noise budget of the advanced LIGO detectors. Reprinted from [2]. Representative amplitude spectral density (ASD) of the advanced LIGO and Virgo detectors as of the third observing run. Reprinted from [3].

an interference pattern at the output. The LIGO, Virgo, and KAGRA interferometers are operated in a "dark fringe mode" meaning that the arm lengths are calibrated such that no light appears at the output in the absence of a gravitational wave. Then, the presence of any interference fringe may be used to infer the presence of a signal.

Fig. 1.1 is a vastly simplified view of the LIGO, Virgo, and KAGRA detectors. Basic Michelson interferometry is not sensitive enough to capture the strain magnitudes produced by gravitational waves. Since their initial construction, each interferometer implemented Fabry-Perot optical cavities, indicated by the mirror between the beam splitter and the end test mass in Fig. 1.1 [41, 42]. A Fabry-Perot cavity consists of highly reflective mirrors placed at such a distance to produce resonance within the cavity. This resonance can increase the circulating power in the detector and the phase shift produced by a change in the arm lengths, effectively amplifying the signal. Power recycling is implemented by placing a mirror between the laser input and the beam splitter, also illustrated simply in Fig. 1.1 [43–45]. With this addition, the input laser power to the detector can be significantly increased which improves the sensitivity as described below.

Even with these modifications the detectors are still noise dominated. This means that the gravitational wave signals will be buried within the detector's inherent noise and we require advanced data analysis techniques to identify their presence. This is discussed in more detail in the next section and at length in Chapter 2 for the GstLAL search algorithm, specifically. The detector noise is characterized by the PSD, typically denoted

as $S_n(f)$. For a stationary random process $n(t)$ the PSD is defined as:

$$\frac{1}{2}S_n(f)\delta(f-f') = \langle \tilde{n}(f), \tilde{n}^*(f')\rangle, \tag{1.14}$$

where $\tilde{n}(f)$ is the Fourier transform of the noise time-series, $n(t)$ and $\langle\cdot,\cdot\rangle$ denotes an ensemble average. Conceptually, this quantifies the amount of noise present in the detector as a function of the frequency.

The left panel of Fig. 1.2 shows the noise budget of the advanced LIGO detectors. This shows the expected noise from various fundamental sources across the frequency band from a few Hz to several thousand Hz. At low frequencies below a few tens of Hz, seismic noise dominates due to the vibrations of the Earth as well as human activity and weather conditions [46, 47]. The seismic noise is mitigated by suspending the test mirrors from pendulum systems which isolates them from much of the ground vibrations. In the middle of the frequency band from a few tens to hundreds of Hz we are limited by thermal noise which arises from light absorption at the end mirrors as well as vibrations in the suspensions holding the mirrors [48, 49]. Additionally quantum noise is a limiting factor across the full bandwidth from ~ 10 to 1000 Hz [37, 39, 50]. The quantum noise originates from the particle nature of light in two ways. Shot noise dominates at high frequencies due to uncertainty in the number of photons hitting the output photodiode at any time. This can be reduced by increasing the laser power. At low frequencies radiation pressure noise dominates originating from momentum imparted to the test masses from the incident photons [51]. The broadband quantum noise can be reduced through frequency dependent squeezing of light [52, 53].

The right panel of Fig. 1.2 shows the representative measured amplitude spectral densities (ASDs) (the square root of the PSD) for the advanced LIGO, Virgo, and KAGRA detectors as of the third observing run [3]. Overall, we see that the detectors are sensitive to gravitational wave signals in the bandwidth from about $10 - 2000$ Hz. At lower or higher frequencies, the power from fundamental noise sources diverges and we are unable to detect any signal. However, from this ASD we can see that the noise floor in the most sensitive band of the detector is lower than 10^{-23}, meaning that a signal from the canonical BNS source considered in the previous section is in fact observable by current detectors.

1.4 Gravitational wave detection

1.4.1 Approximating the waveform

In the previous sections, it was mentioned that gravitational wave signals from CBC sources are well-modeled. Conceptually, we can understand how these signals should look based on the orbital dynamics and energy balance. However, it is not easy to solve these equations analytically so rather than finding exact solutions we must rely on approximations. Namely, we use the Post-Newtonian (PN) approximation. In this approximation, we make the assumptions of (i) slow motion, i.e. (v/c) is small and (ii) a weak-field, i.e. far from the source GM/D_L is small. Of course, as the objects inspiral their velocities will speed up becoming relativistic near the merger causing the PN approximation to break down. But it is accurate enough for the inspiral portion of the waveform. There are many sources which derive the energy, phase evolution, and gravitational waveform polarizations in the PN expansion. I will only provide a summary of the key points here but readers are referred to [14, 34, 54] for the full derivations.

First, we know that the gravitational waveform has two independent polarizations, $h_+(t)$ and $h_\times(t)$ but at a given detector we will observe the combined waveform as:

$$h(t) = F_+ h_+(t) + F_\times h_\times(t). \tag{1.15}$$

Here, F_+ and $F_\times$ are the detector response functions or *antenna patterns*. In general, a gravitational wave detector has varying sensitivity to signals coming from different directions and orientations with the maximal sensitivity being reached for a face-on merger directly above (or below) the plane of the detector [41]. It can be shown that the two polarization states combine to give:

$$h(t) = \frac{G}{c^4}\frac{4\nu M}{D_L} Q(\text{angles}) v^2 \cos(2\varphi(t) + 2\varphi_0), \tag{1.16}$$

where all of the dependence on the source sky position and inclination have been packaged neatly into $Q(\text{angles})$ and $\nu = m_1 m_2/M^2$ is the symmetric mass ratio. From here, under the assumption that the frequency varies slowly (which is true in the early inspiral portion of the waveform, but again will breakdown near merger) we can write the Fourier

transform as,

$$\tilde{h}(f) = \frac{1}{D_L} Q(\text{angles}) \sqrt{\frac{5\pi}{24c^3}} \left(G\mathcal{M}_c\right)^{5/6} (\pi f)^{-7/6} e^{i\psi(f)}, \tag{1.17}$$

where we have also replaced ν and M in favor of the chirp mass, $\mathcal{M}_c$. Here, $\psi(f)$ encodes the phase evolution of the gravitational wave. It is written as a sum of expansion terms,

$$\psi(f) = 2\pi f t_c + \varphi_c + \frac{3}{128\nu} \sum_k \alpha_k (\pi M f)^{\frac{k-5}{3}}. \tag{1.18}$$

To leading order ($k = 0$) we have,

$$\psi_0(f) = \psi(f, \mathcal{M}_c) = 2\pi f t_c + \varphi_c + \frac{3}{128} \left(\frac{c^3}{G}\right)^{5/3} (\pi \mathcal{M}_c f)^{-5/3}, \tag{1.19}$$

where t_c and φ_c are the time and phase at coalescence. The notation $\psi(f, \mathcal{M}_c)$ emphasizes that at leading order the phase is only a function of the frequency and chirp mass; other intrinsic parameters such as the mass ratio and the component spins come in at higher orders and thus leave a weaker imprint on the signal.

1.4.2 Matched filtering

Given our models of the waveforms, we can used matched filtering to search for signals in the data. This is a standard practice in signal processing for detecting a known signal within noisy data. I will give a general summary of the application of matched filtering to detection of GWs from CBC sources in this section following the method of [55]. In the next chapter, I will specialize these methods to the case of the GstLAL search pipeline.

The strain data from GW interferometers such as LIGO and Virgo can be expressed as:

$$d(t) = n(t) + s(t). \tag{1.20}$$

The noise, $n(t)$, is always present while the signal, $s(t)$, is an additive component which may or may not be present at any given time. The task of GW analysis pipelines is to determine when a signal is present and what its form is. We assume that the noise is stationary, meaning that the noise properties do not change over timescales comparable to the signal durations. In this case, the noise is described by the PSD, $S_n(f)$.

To identify signals, we construct complex-valued template waveforms, $h(t)$,[3] where

[3]Until now, I have used $h(t)$ to denote the actual gravitational waveform. Now, we consider that the

the real and imaginary parts of the waveform correspond to the two polarization phases,

$$h(t) = h_+(t) + ih_\times(t). \tag{1.21}$$

The shape of each waveform is described by the intrinsic parameters, $\vec{\lambda} = \{m_1, m_2, s_{1,z}, s_{2,z}\}$[4], while its strength observed at the detector will be determined by the extrinsic parameters $\vec{\mu} = \{\alpha, \delta, \iota, t_c, \varphi_c, \psi\}$, describing the sky position, orientation, polarization, and time and phase at coalescence. Therefore we can denote the templates as $\mathcal{A}\ h(t; \vec{\lambda})$ where all the dependence on extrinsic parameters has been packaged into the amplitude term.

The SNR in a given stretch of data is defined by the inner product between the data and a filter, $u(t)$:

$$\rho^2(t) \equiv \langle d(t), u(t) \rangle = 4 \int_0^\infty \mathrm{d}f \frac{\tilde{u}^*(f)\tilde{d}(f)}{S_n(f)}, \tag{1.22}$$

where $\tilde{u}(f)$ is the Fourier transform of $u(t)$ and $\tilde{u}^*(f)$ indicates a complex conjugate. The theory of matched filtering tells us that the maximal SNR is achieved when the filter $u(t)$ identically matches the signal. In other words, assuming that $h(t; \vec{\lambda}) = s(t)$, we have that the template is the matched filter and the complex-valued matched filter response, $R(t)$, is

$$R(t) = R_+(t) + iR_\times(t) = 4 \int_0^\infty \mathrm{d}f \frac{\tilde{h}^*(f)\tilde{d}(f)}{S_n(f)}. \tag{1.23}$$

Taking the magnitude of this and maximizing over the unknown extrinsic parameters, $\vec{\mu}$, gives the real-valued SNR,

$$\rho = \max_{\vec{\mu}} |R(t)|, \tag{1.24}$$

which can be used as a naive detection statistic for a signal with intrinsic parameters $\vec{\lambda}$.

Of course, in practice we don't know a priori what values the intrinsic parameters will take for any given signal. This motivates the need for *template banks* where we construct and store many templates across a broad parameter space that we are interested in searching. In constructing such a bank, we typically define the *match* between two nearby templates (that is, having similar intrinsic parameters $\vec{\lambda}$ and $\vec{\lambda} + \Delta\vec{\lambda}$) as the inner product or overlap between the templates,

$$M(\vec{\lambda}, \Delta\vec{\lambda}) = \max_{\vec{\mu}, \Delta\vec{\mu}} \left[\langle h(\vec{\mu}, \vec{\lambda}), h(\vec{\mu} + \Delta\vec{\mu}, \vec{\lambda} + \Delta\vec{\lambda}) \rangle \right], \tag{1.25}$$

exact physical waveform is unknown and use $h(t)$ as a model to describe the observed signal.

[4]Here, we make the typical simplifying assumption that the spins are aligned or anti-aligned with the orbital angular momentum which we take to be in the $\hat{z}$ direction. Therefore, the in-plane spins $s_{i,x} = s_{i,y} = 0$ are ignored. We additionally neglect the orbital eccentricity and tidal deformabilities.

where we have maximized over the extrinsic parameters. From this expression, [55] derives the notion of a *minimal match* as the worst-case scenario match between a signal and any template in the bank. This is a common metric for measuring the efficacy of a template bank. A high minimal match ($MM \sim 1$) implies a small mismatch between a potential signal and the most similar template in the bank, which would result in a near maximal SNR. However, constructing a template bank with a high minimal match requires a large number of densely packed templates and filtering all of these presents a significant computational burden. Therefore, practical methods for constructing template banks must balance the desire to maximally cover the space while keeping the number of templates manageable [56–58].

1.4.3 Characterizing detector sensitivity

Finally, given the gravitational waveforms from the CBC sources we expect, the noise present in our detectors, and the detection statistic, ρ, we can define a simple metric for the sensitivity of the detector, referred to as the *horizon distance*. We consider a canonical BNS merger – an equal mass, non-spinning, circular orbit with both binary components $m_i = 1.4M_\odot$. Further we will assume this source is optimally located and oriented, i.e. it is a face-on merger directly above the plane of the detector. The horizon distance is the distance to which we can expect to observe GWs from such an event with some chosen SNR typically taken to be $\rho = 8$ [59]. Using the gravitational waveform of Eqn. 1.17 and the definition of the SNR in Eqn. 1.23 we can solve for the distance as,

$$D_H = \frac{1}{\rho}\mathcal{A}\sqrt{4\int_{f_{\rm low}}^{f_{\rm hi}} \mathrm{d}f \frac{f^{-7/3}}{S_n(f)}}. \tag{1.26}$$

Here I have wrapped all the extraneous factors of $\tilde{h}(f)$ which do not depend on f into the amplitude term $\mathcal{A}$ for simplicity. The horizon distance is by construction an optimistic measure of the detector sensitivity. To get a more realistic estimate, we average this quantity over the sky position and inclination, obtaining a factor of ~ 2.25, giving the detector *range* [59, 60],

$$\mathcal{R} = \frac{D_L}{2.25}. \tag{1.27}$$

This is a common quantity used to estimate sensitivity of gravitational wave detectors[5].

[5]During an observation run of the LIGO, Virgo, and KAGRA Collaboration, estimates of the detector range are updated continuously and made available to the public via `https://online.igwn.org`.

1.5 Next generation gravitational wave detection

The LIGO, Virgo, and KAGRA detectors are to-date the only instruments which have made detections of gravitational waves from merging black holes and neutron stars. These detectors have seen phenomenal success and an unprecedented rate of growth in the field over just ten years of observing. The scientific impact of these detections cannot be understated. Ten years ago, it was not clear whether binary black holes even existed; now there have been over one hundred detections of these mergers giving us insight into a previously unknown population of stellar-mass black holes [12]. We have already been able to begin teasing out structure in the black hole mass distribution from the supposed lower mass gap through to the upper mass gap imposed by pair-instability supernovae [61–63]. These observations have brought advancements in cosmology and fundamental physics with new measurements of the Hubble constant and constraints on alternative theories of gravity [64,65]. Further, the first observation of gravitational waves from a binary neutron star merger confirmed such events as progenitors of short gamma ray bursts [7, 66, 67]. The joint detection of GWs and electromagnetic (EM) radiation from the same merger event proved that GWs move at the speed of light [67]. This observation also elucidated the origin of some of the heaviest elements in the universe as r-process nucleosynthesis in the kilonovae following a BNS merger [68–70]. Thus, the observation of gravitational waves has already brought advances in astrophysics, cosmology, fundamental physics, and high energy physics.

However, even at their design sensitivities, the current ground-based interferometers will only be capable of probing a small fraction of the cosmological volume, corresponding to the nearest or most recently merging binaries. Next Generation (XG) detectors are already being planned which will have sensitivities at least an order of magnitude better than the current network. The Einstein Telescope (ET) is a European proposal for an underground triangular detector with three nested interferometers of arm lengths ~ 10 km (compared to the 4 km of LIGO and 3 km of Virgo) [71]. These specifications will lead to vastly improved sensitivities, lower noise (especially at low frequencies), and (due to the nested configuration) an isotropic antenna pattern free of the blind spots that plague current interferometers. The Cosmic Explorer (CE) is a proposal for two L-shaped detectors in the U.S.: one with 40 km arms and one with 20 km arms [72, 73]. The order-of-magnitude increase in arm length will bring a vast improvement to sensitivity compared to current detectors and the two-detector network configuration is designed to balance cost with science goals, namely the desire for sky localization.

Detectors such as ET and CE, especially when operated simultaneously as a network, will be able to far surpass the capabilities of the current ground-based detector network. These detectors will be sensitive to the high redshift universe, observing stellar mass binary black holes to redshifts $z \sim 20$ or greater [74–76]. This means that we will effectively be sensitive to every BBH merging within a Hubble time, allowing for unprecedented direct observations of the very early Universe. The improvement at low frequencies will bring sensitivity to higher mass binary black holes up to $M \sim 1000 M_{\odot}$. In the low mass regime, the improved low frequency sensitivity means that BNS signals will last much longer, leading to improved prospects for pre-merger detection and early warning for multi-messenger EM follow-up. In addition to the CBC sources, these detectors are expected to bring measurements of continuous gravitational waves such as the stochastic background and galactic pulsars as well as unmodeled gravitational wave sources like supernovae. Further, many of these signals will be extremely loud in XG data having SNRs $\sim 10^2 - 10^3$. This will allow for precise measurements of the waveform and improved tests of general relativity.

All ground-based interferometers are fundamentally limited by seismic noise at low frequencies. Therefore, even the XG detectors will not be able to reach below a few Hz. However, the Universe is full of systems producing gravitational waves at lower frequencies which can provide new insights to astrophysics and cosmology. This motivates the need for the Laser Interferometer Space Antenna (LISA), a European proposal for a triangular spacecraft formation in an Earth-like orbit around the sun [77, 78]. LISA would be free of many of the noise sources that limit ground-based interferometers and is proposed to have arm lengths $\sim 10^6$ km, leading to an incredible sensitivity in the frequency band of ~ 10 mHz $-$ 1 Hz. With such a detector, we could observe low-mass compact binaries in the galaxy, such as binary neutron stars or white dwarves, whose gravitational radiation would be far too weak to every detect in ground-based interferometers [77, 79]. LISA will also be sensitive to the inspirals, mergers, and ringdowns of supermassive black hole binaries with component masses $\sim 10^5 - 10^7 M_{\odot}$ allowing for tests of general relativity under the most extreme dynamical conditions [80, 81]. Additionally, LISA will observe a stochastic GW foreground at very low frequencies from millions of unresolvable systems in the universe and could additionally detect a gravitational wave background from the early universe [82].

Despite the incredible success of the LIGO, Virgo, and KAGRA gravitational wave detectors to date, we are still in the early days of gravitational wave astronomy. Next generation detectors such as ET, CE, and LISA will give us a direct window to the

universe all the way back to the Big Bang. The next generation of gravitational wave detection will certainly bring new surprises in astrophysics, cosmology, and fundamental physics, revolutionizing our understanding of the universe for many years to come.

Chapter 2
GstLAL inspiral search

2.1 Introduction

In this chapter, I will review the methods of the GstLAL pipeline in searching for GW signals from modeled CBC sources. GstLAL uses matched filtering to search the detector strain data, $d(t)$, for potential signals, $s(t)$, which are modeled by complex-valued template waveforms, $h(t)$, as described in Sec. 1.4.2. However, GstLAL typically performs matched-filtering in the time-domain, i.e. Eqn. 1.23 becomes a convolution:

$$\rho(t) = 4 \int_{-\infty}^{\infty} \mathrm{d}\tau \hat{h}(\tau) \hat{d}(t+\tau). \tag{2.1}$$

Additionally, here the data and template have been whitened as indicated by the "hat". Whitening ensures that the data is uncorrelated in time and has unit variance [83]. In other words, each sample of whitened data is independent and follows a Gaussian distribution. The whitened data is defined as:

$$\hat{d}(t) = \int_{-\infty}^{\infty} \mathrm{d}f \frac{\tilde{d}(f) e^{2\pi i f t}}{\sqrt{S_n(f)}}. \tag{2.2}$$

In this chapter, I will begin by describing the inputs to this equation – the data conditioning in Sec. 2.2 and the templates in Sec. 2.3. After we obtain GW candidates with matched filtering, we rank them against a background of noise triggers with the likelihood ratio (Sec. 2.4) and then estimate the significance with the false alarm probability (FAP) and FAR (Sec. 2.5). Finally, in Sec. 2.7 I overview the technical implementation of the pipeline in terms of the software infrastructure.

2.2 Data conditioning

As in Sec. 1.4.2, we denote the output strain data from GW interferometers such as LIGO, Virgo, and KAGRA as:

$$d(t) = n(t) + s(t) \tag{2.3}$$

and make the initial assumption of stationary noise. However, over long durations such as days to weeks or even months, we know that this assumption breaks down and indeed the detector noise properties do vary significantly. We also assume that the noise is represented by a Gaussian random variable. The mean of the Gaussian noise may not be zero, however with the assumption of stationarity the mean will be constant and can therefore always simply be shifted to zero.

Before the detector data, $d(t)$, can be used in the matched filtering, we require several conditioning steps. First, it is downsampled. The data from the detectors has sampling frequency 16384 Hz, but typically the GstLAL pipeline will decimate the incoming data to a lower sampling frequency before proceeding to the filtering. As the matched filtering involves an inner product between the data and the template waveform, we must ensure that the two are sampled at the same rate. This is discussed in more detail in Sec. 2.3.

After downsampling, the data is whitened using the PSD according to Eqn. 2.2. As described in Chapter 1, the PSD is defined as

$$\langle \tilde{n}(f)\tilde{n}(f')\rangle = \frac{1}{2}S_n(f)\delta(f-f'), f > 0. \tag{2.4}$$

However, this quantity is not known analytically and instead we must measure it empirically. From here, I will denote the data as a discrete quantity with time index k, $\hat{d}[k]$, as opposed to the continuous quantity $\hat{d}(t)$.

To compute the instantaneous PSD, we first split the incoming data into segments of length T seconds, i.e. $N = f_s T$ samples. We apply a window function and take the fast Fourier transform (FFT),

$$\tilde{d}_j[k] = \sqrt{\frac{N}{\sum_{k=0}^{N-1} w[l]^2}}\Delta t \sum_{l=0}^{N-1} w[l]\cdot d_j[l]\exp\left(-2\pi i k l/N\right). \tag{2.5}$$

The subscript j indexes the data segments and l is the time index. We use a Hann window, $w[k]$, to smoothly zero the data at either end of each segment [4, 84]. For each

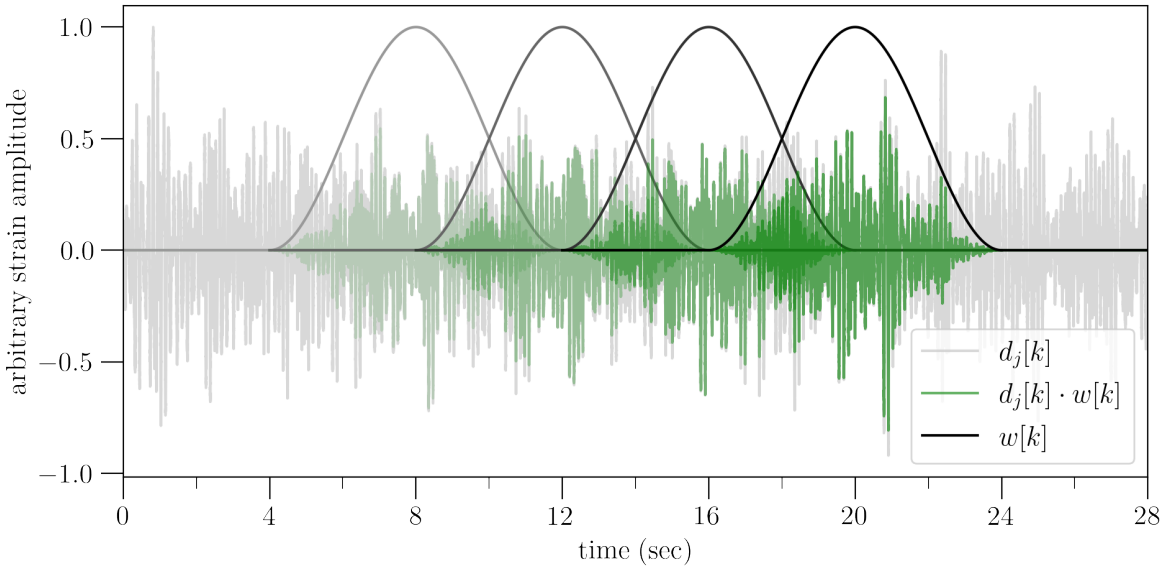

Figure 2.1. Illustrative example of the data segments used for PSD estimation by GstLAL. The normalized strain time-series data, $d[k]$, is shown in light grey. The solid curves show four overlapping Hann windows, $w[k]$, with FFT length $= f_s \times 16$ seconds. The green curves show the convolved $d_j[k] \cdot w[k]$ in each segment with varying opacities.

segment of length N, the first and last $N/4$ samples are zeroed by the window function. Therefore, only the middle $N/2$ samples are used in the PSD calculation. Each segment is then overlapped with the next by $N/4$ samples so that every sample of data is present in 4 segments of the instantaneous PSD. Fig. 2.1 shows an illustrative example of four such overlapping windowed data segments. Historically, the data segment length N has been chosen as $f_s \times 16$ seconds corresponding to a 4-second zero pad on either side of the segment and 8 seconds of data in the middle being used to compute the instantaneous PSD. From Eqn. 2.4, the instantaneous PSD in the j^{th} data segment is:

$$S_j^{\text{inst}}[l] = 2\Delta f \left|\tilde{d}_k[l]\right|^2 \tag{2.6}$$

The instantaneous PSD is a rapidly changing quantity and not robust to short time-scale fluctuations in the noise. We desire an estimate of the PSD which more accurately represents the noise properties over long durations. [4] derives a suitable PSD by first taking the median PSD, $S_j^{med}[l]$, over the recent $n_{med} = 7$ instantaneous PSD samples. The median ensures that any short duration non-Gaussianities or outliers in the statistical properties of the data will be disregarded. The recent $n_{avg} = 64$ median PSDs are then used to construct the arithmetic mean of the PSD, $S_j[l]$, which ensures that the PSD is

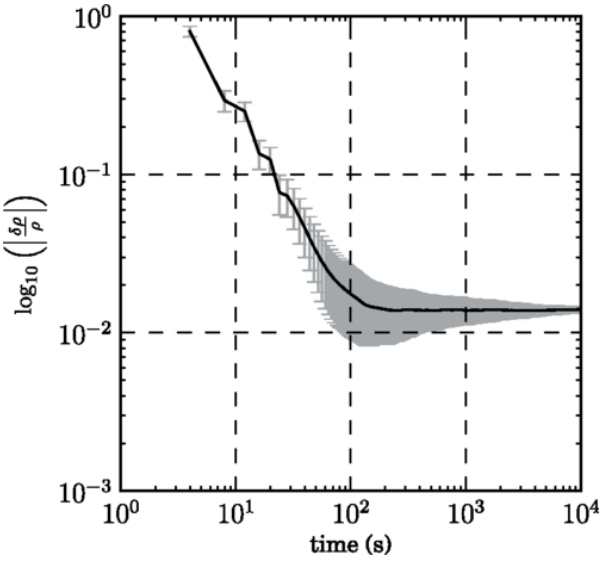

Figure 2.2. Convergence of the PSD as measured by GstLAL. The vertical axis shows the log fractional loss in SNR as a result of filtering with the estimated PSD. Within about 200 seconds, the loss in SNR drops to the level of $\mathcal{O}(1\%)$. Reprinted from [4].

representative of data characteristics over long durations, and will not fluctuate rapidly. To move the median PSD, an excursion in the data would have to last at least $\frac{1}{2}n_{med}N$ seconds and still it would only contribute $\frac{1}{n_{avg}}$ of the mean PSD. $S_j[l]$ has been shown to converge quickly to an accurate estimate of the true PSD, shown in Fig. 2.2, and remain robust to frequent short duration excursions in the data [4, 85]. This is the quantity used by the GstLAL search pipeline for whitening the data.

However, since the PSD is estimated in real-time as the data is streamed, it is not possible to use the same $S_j[l]$ to whiten the templates as is used for the data. The templates must be pre-whitened before an analysis begins. For this purpose, we use a *reference PSD* computed using the same method but on a past segment of data. The effect of whitening the templates with a different PSD is expected to be small as long as the data is sufficiently stationary, that is the reference should be chosen as recently as possible to the data being analyzed.

As mentioned previously, the assumption that LIGO noise is stationary and Gaussian breaks down on long timescales. In fact, even on short timescales the LIGO noise is subject to transient non-Gaussian excursions known as *glitches*. This poses a significant problem for GW detection as glitches can mimic real signals and make it difficult to distinguish terrestrial noise from astrophysical signals. Developing methods to reduce glitches, identify their sources, and distinguish them from signals is an ongoing area of study [86–91]. To mitigate the negative effects of glitches, the GstLAL search pipeline implements a gate

thresholded on the whitened strain amplitude [4]. Under the assumption that no signal is present in the data the whitened strain has a unit variance. We can choose a threshold in units of standard deviations from the mean as a way to identify excursions in the data which have significant deviations from stationarity. Any stretches of the data which surpass the threshold are then set to zero before the filtering proceeds. Since whitening further ensures that neighboring samples of data are uncorrelated, zeroing the data in a window around potential glitches is expected to have a negligible effect on the matched filter response to the surrounding data. We must be careful in choosing an appropriate threshold so that we do not gate real signals. A discussion of the choice of threshold used in the first part of the O4 is given in Chapter 5.

After down sampling, whitening, and gating the data is ready for input to the matched filter equation. In the next section I will discuss the second ingredient to the matched filtering equation which is the templates.

2.3 The LLOID method

In the previous section, I discussed the data conditioning and specified that the templates are whitened with a pre-computed reference PSD before filtering. In fact, the GstLAL pipeline first transforms the templates according to the Low-Latency Online Inspiral Detection (LLOID) method before the matched filtering can proceed [92]. In this section, I will provide an overview of the LLOID method and its motivation.

To enable multi-messenger astronomy and deep EM follow up campaigns it is imperative that we can detect GW candidates as quickly as possible. EM counterparts in the first few seconds or minutes after the merger are expected to carry rich information about the source dynamics [93–99]. These signals may fade very quickly so any delay in the initial detection of a GW candidate reduces our chances of observing the prompt emission. However, GW searches which rely on template banks and matched filtering to identify signals can suffer from several inefficiencies which become critical when the goal is low-latency detection [92]. First, the number of computations required to convolve the template waveform with the data is formidable especially when the waveforms and data are sampled at many thousands of Hertz. Second, these searches require hundreds of thousands to millions of templates to reach an acceptable sensitivity across the full mass spectrum from $\sim 1.0\ M_\odot$ for the smallest neutron stars to $\mathcal{O}(100) M_\odot$ for the heaviest black holes. Especially in the lower mass range, where waveforms can last several minutes in the LIGO and Virgo sensitive frequency band, we require an extremely dense bank.

This is because for these long duration waveforms, a very small change in the masses or spins of the source can lead to a phase shift which builds up considerably over many cycles leading to a large SNR loss. This means that we must convolve millions of templates with every second of data from each detector we wish to analyze. A search such as this would require vast computational resources and still may struggle to keep up with the real time arrival of data.

The LLOID method seeks to optimize the matched filtering to produce a GW search which is highly performant without sacrificing sensitivity [92]. This is a two-pronged method addressing first the issues of over-sampling by treating the waveform as independent *time slices* each down-sampled to an appropriate frequency according to its bandwidth. Secondly, the LLOID method employs a singular value decomposition (SVD) of the full template bank, reducing the total number of templates required to fully cover the physical parameter space. In this section, I will describe the method of time-slicing first, followed by the SVD computation.

The Nyquist-Shannon sampling theorem stipulates that to accurately represent a waveform whose highest frequency component is f_N, it must be sampled with a frequency, $f_s \geq 2f_N$ [100]. That is, the sampling rate must be at least twice as high as the maximum frequency component present in the signal. Otherwise, there can be bias in the sampled waveform known as aliasing. Given this requirement, the problem of the large number of computations required to filter a single GW waveform is exacerbated by the specific waveform morphology of CBC signals. The waveforms monotonically increase in frequency with time, spanning a range from $\sim 10 - 1000$ Hz from the early inspiral through to the merger in our frequency band of interest. This means that the sampling frequency required to accurately represent the later portion of the waveform necessarily over samples the early portion. This poses an obvious problem to the efficiency of the search, however we can leverage the same waveform morphology along with the Nyquist theorem to split up the time-domain waveform into *slices* which can each be safely down-sampled [92].

We use the Nyquist-Shannon theorem to dictate appropriate bounds corresponding to slices of a CBC waveform which may be sampled at a given rate. The raw strain data is sampled at 16384 Hz which dictates the maximum possible sampling frequency for the waveforms. In GstLAL we choose allowed sampling frequencies in decreasing powers of two, i.e. $f_s : \{16384, 8192, 4096, 2048, ... 2\}$ Hz. To construct time slices, we will start at the end of the waveform, i.e. at the coalescence time which we define as $t = 0$. We count time as increasing backwards towards the beginning of the waveform. For each

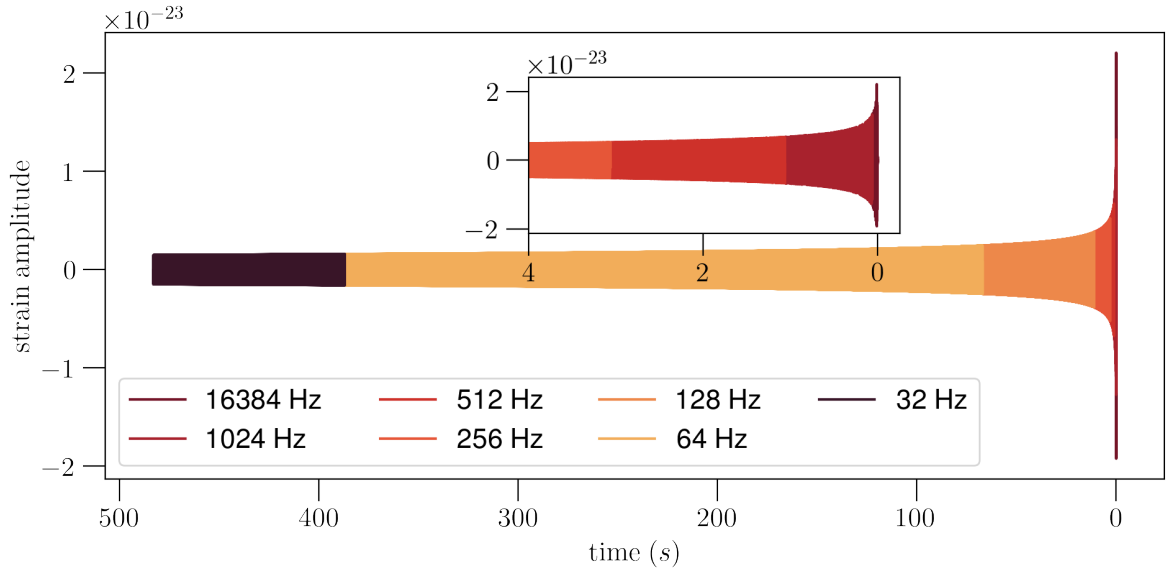

Figure 2.3. Illustration of a time-sliced BNS waveform, with a low frequency cut off 15 Hz and upper frequency cut off 16384 Hz. Padding was not included in the construction of these time slices. The colors indicate the sampling frequency of each band. The final few seconds of the waveform, shown in the inset axis, has time slices that are extremely short in duration.

time slice, defined by its sampling frequency, we need to choose the lower frequency cut off f_{low} to define the bounds of the slice. This is conveniently determined by the Nyquist theorem. The f_{low} of the i^{th} slice, with sampling frequency, $f_{s,i}$, is the Nyquist frequency of the next slice, that is [92]:

$$f_{\text{low}} = f_{N,i+1} = \frac{f_{s,i+1}}{2} = \frac{f_{s,i}}{4}. \tag{2.7}$$

For example, a slice sampled at $f_s = 16384$ Hz will extend to the lower frequency cut off $f_{\text{low}} = 4096$ Hz. The next slice, sampled at $f_s = 8192$ Hz will extend to the lower frequency cut off $f_{\text{low}} = 2048$ Hz, and so on. The time-slicing is illustrated for a BNS waveform with $m_1 = m_2 = 1.2\ M_\odot$ in Fig. 2.3. Eqn. 2.7 is the highest allowable choice of f_{low}, since a higher cut off will necessarily mean that the next slice includes frequency content higher than its own Nyquist, and therefore the waveform will no longer be accurately represented. In practice, the GstLAL pipeline also uses a *padding* factor. With padding, the lower frequency cut off of each slice is

$$f_{\text{low}} = \frac{f_{N,i+1}}{\text{padding}} = \frac{f_{s,i}}{4 \times \text{padding}}. \tag{2.8}$$

This produces time slice boundaries which are more conservative by extending a higher sampling frequency to lower frequency portions of the inspiral (i.e. longer duration slices) than is strictly necessary.

The waveform slicing and down-sampling only needs to be computed once and can be done in advance of an analysis. Therefore, it can lead to a significant reduction in the computations required to filter a single waveform, as the number of samples in the early portions of the waveform can be reduced by several orders of magnitude with respect to the late portions near the merger [92]. The down-sampled waveform slices need to be interpolated and re-combined after filtering to produce the full SNR time-series output. However this operation is cheap compared to the reduction in computation afforded by the down sampling. We will discuss the waveform reconstruction in more detail below with regards to the SVD of the template bank.

The second step in the LLOID method involves computing a SVD on the physical templates in the bank [101, 102]. As mentioned above, in order to achieve a high search sensitivity in the parameter space of interest, the template bank is constructed very densely to ensure a small maximum SNR loss between a real signal and the nearest template. This is known as the maximum mismatch and is usually chosen to be $\sim 3\%$ in gravitational wave searches (or equivalently, a 97% minimum match) [55]. Therefore the denseness of the bank is quantified in terms of the degree of similarity between neighboring templates given by their overlap or inner product, $\langle h_\alpha, h_\beta \rangle$. This high overlap is a redundancy in the matched filtering process. Reducing this redundancy allows a more sensitive search to be carried out using fewer computational resources. The SVD is used to reduce the set of physical templates to a smaller set of orthonormal, albeit non-physical, basis vectors which are used to filter the data.

As mentioned earlier in this chapter, for each physical waveform we would like to search for we must perform the matched filtering with two templates corresponding to each quadrature phase of the waveform, $h_+(t)$ and $h_\times(t)$. That is, for M complex waveforms there are $N = 2M$ real-valued templates to filter. In the rest of this section, I will follow the notation of [101] and denote the real and imaginary parts of the α^{th} waveform as $\vec{h}_{(2\alpha-1)}$ and $\vec{h}_{(2\alpha)}$ respectively, such that $\alpha : \{1, 2, ... M\}$. The matched filter response to the α^{th} template is the inner product of the data with the template, $\rho_\alpha = \vec{h}_\alpha \cdot \vec{d}$, where the inner product is over the time index, $l : \{1, 2, ... L\}$.

We can construct an $N \times L$ dimensional matrix by stacking the templates as rows,

i.e.

$$\mathbf{H} = \begin{bmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \dots \\ \vec{h}_N \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \dots & h_{1L} \\ h_{21} & h_{22} & h_{23} & \dots & H_{2L} \\ & & \dots\dots\dots\dots & & \\ h_{N1} & h_{N2} & h_{N3} & \dots & h_{NL} \end{bmatrix}. \tag{2.9}$$

Then, the SVD theorem states that this matrix $\mathbf{H}$, as is true for all matrices, can be factorized as follows [103]:

$$\mathbf{H} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}. \tag{2.10}$$

Here, $\mathbf{V}$ is an $N \times N$ matrix of reconstruction coefficients, $\mathbf{\Sigma}$ is an $N \times N$ diagonal matrix of the singular values of $\mathbf{H}$, and $\mathbf{U}$ is an $N \times L$ matrix of orthogonal basis vectors. We can consider a specific $H_{\alpha j}$ corresponding to the j^{th} sample of the α^{th} waveform in the bank,

$$H_{\alpha j} = \sum_{\nu=1}^{N} v_{\alpha\nu}\sigma_\nu u_{\nu j}. \tag{2.11}$$

Here, $\vec{v}_\alpha$ are the reconstruction coefficients for the α^{th} template, σ_ν are the singular values of matrix $\mathbf{H}$, $\vec{u}_j$ are the j^{th} samples of each basis vector, and the summation is over the N basis vectors. The $\vec{u}$ being basis vectors means that the matrix $\mathbf{H}$ is reconstructed by a linear combination of the $\vec{u}$ with coefficients given by $\vec{\sigma}$.

Until this point, the factorization has not reduced the computational burden required to filter the N physical waveforms. The crux of this method lies in the fact that the $\mathbf{H}$ may be approximated by a smaller subset of the $\vec{u}$ to a high accuracy. The singular values represent the relative contributions of each of their respective basis vectors to the reconstruction. The elements of $\vec{\sigma}$ are arranged in decreasing order, where the smallest values are least important in accurately reconstructing the $\mathbf{H}$. Given that we do not require a perfect coverage of the parameter space, only up to some $\sim 97\%$, we can therefore discard some of the basis vectors with the smallest values of σ_ν, i.e.

$$H_{\alpha j} \approx \sum_{\nu=1}^{N'} v_{\alpha\nu}\sigma_\nu u_{\nu j}, \tag{2.12}$$

where $N' < N$. That is, a specific waveform, h_α is approximated by the linear combination:

$$h_\alpha \approx (v_{\alpha 1}\sigma_1)\cdot\vec{u}_1 + (v_{\alpha 2}\sigma_2)\cdot\vec{u}_2 + ... \, (v_{\alpha N'}\sigma_{N'})\cdot\vec{u}_{N'}. \tag{2.13}$$

Then instead of filtering all N physical template waveforms, we can filter the reduced set of N' basis waveforms, and reconstruct the output of any physical template by

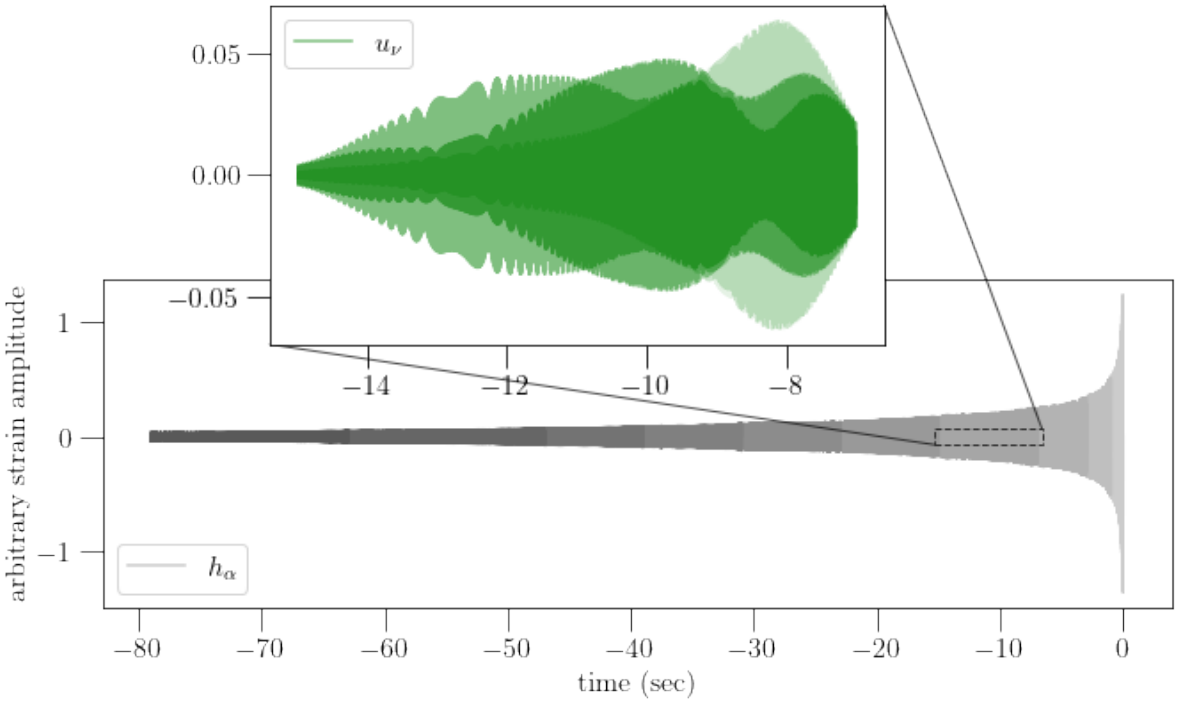

Figure 2.4. A time domain BNS waveform with time slice boundaries indicated by shades of grey. The inset axis shows the first five basis waveforms, u_ν, for the fourth time slice in shades of green.

multiplying the correct reconstruction coefficients [101]. An illustration of a few such basis waveforms for a BNS template is shown in Fig. 2.4, emphasizing the non-physical shape of the $\vec{u}$.

The question of how many basis waveforms we can safely discard, i.e. the choice of N', is determined by the allowable SNR loss. [101] derives the SNR loss in the α^{th} complex template in terms of the reconstruction coefficients,

$$\frac{\delta\rho}{\rho} = \frac{1}{4}\sum_{\nu=N'+1}^{N}\left(v^2_{(2\alpha-1)\nu} + v^2_{(2\alpha)\nu}\right)\sigma^2_\nu. \tag{2.14}$$

And [101] and [92] show that with the number of basis waveforms N' about an order of magnitude less than the original number of templates N, the physical waveforms are reconstructed to an accuracy of about $1 - 10^4$ and the fractional loss in SNR is only about 10^{-3}. This is a powerful method to reduce the computational burden of matched filtering searches for GW signals.

Combining the two steps of the LLOID method, GstLAL first splits the full template bank into smaller *sub-banks* of neighboring templates in the parameter space. The bank

splitting can be computed with respect to several parameters. Historically, the chirp mass $\mathcal{M}_c$ and χ_{eff} have been used [4, 104, 105]. More recently, methods have been developed to group templates based on the PN phase terms [106, 107]. Several nearby sub-banks are grouped together into what we refer to as *background bins* (this choice of naming convention will be motivated in Sec. 2.4). The waveforms in each background bin are first time-sliced and downsampled, such that the redundancy in the time dimension is reduced. Then, the SVD is computed over the groups of waveforms in each time slice independently. This reduces the redundancy in the template dimension, i.e. the overlap between neighboring templates with similar intrinsic parameters. The filtering proceeds with this reduced set of down-sampled orthogonal basis vectors. The physical down-sampled waveforms are reconstructed via the pre-computed reconstruction matrices. Finally, the SNR is computed by up-sampling and combining across time slices.

2.4 Trigger generation and ranking

The matched filtering produces an SNR time-series which we use to identify *triggers*. The GstLAL pipeline defines a trigger as any point in the SNR time-series which passes a threshold of 4.0. Each background bin may include $\mathcal{O}(1000)$ templates and the filtering of each template may produce $\mathcal{O}(1)$ trigger per second. Further, we will have independent SNR time-series from filtering data in each of the detectors participating in an analysis. We must reduce this large set of triggers, which are expected to mostly originate from noise, to a smaller set of viable GW candidates.

A coincidence is formed when a template rings up a trigger in multiple detectors within the light travel time. The network SNR, defined as,

$$\rho_{net} = \sqrt{\sum_{i \in \text{IFOs}} \rho_i^2}, \tag{2.15}$$

is used as an initial threshold in forming GW candidates. Any trigger with $\rho_{net} \geq 7.0$ is kept and considered as a potential GW candidate. GstLAL does not require coincidence among more than one detector to form a candidate [104]. That is, a non-coincident trigger with $\rho \geq 7.0$ in a single detector is still considered as a GW candidate.

To estimate the significance of the candidates we must first understand the distribution of triggers produced by the detector noise. We construct a *background* of triggers which we can confidently assume are only terrestrial in origin. We know that coincidence among independent detectors is a powerful validation test of true GW triggers. Therefore, the

GstLAL pipeline constructs background from the non-coincident triggers [4]. Any trigger found during times when at least two detectors were in operation which was identified in only a single detector is used to construct a background histogram. This choice relies on the assumption that a real GW would ring up a trigger from the same template across at least two detectors. In constructing the background, we conservatively exclude all triggers from times when only a single detector was operational. This is because during these times we lack information about possible coincidence and we cannot confidently exclude the possibility that any trigger is astrophysical in origin. It is imperative for accurate significance estimation that the background is "clean", i.e. that it includes only triggers produced by noise and no signals. On the other hand the candidate list, defined only by the threshold of $\rho_{net} \geq 7.0$ certainly includes noise triggers. The task is to rank these candidate in order of their likelihood of being produced by real GWs.

As mentioned in Sec. 1.4.2, the SNR can be used as a naive ranking statistic for GW candidates. However, in non-Gaussian data this is not ideal as transient instrumental artifacts may ring up high SNRs mimicking true signals. Instead, we use the *likelihood ratio* [108, 109], defined for a template with intrinsic parameters $\vec{\theta} = \{m_1, m_2, s_{1,z}, s_{2,z}\}$ as: [110, 111]

$$\mathcal{L} = \frac{p(\vec{O}, \vec{\theta}|\vec{s})}{p(\vec{O}, \vec{\theta}|\vec{n})}. \tag{2.16}$$

Here, $\vec{O}$ is a set of observable properties of the trigger and the data. The numerator is the PDF under the signal hypothesis, i.e. the probability of finding $\vec{O}$ given that a signal $\vec{s}$ is present in the data. The denominator is the PDF under the noise hypothesis, i.e. the probability of finding $\vec{O}$ given that no signal is present and the data only includes noise. In GstLAL, we take $\vec{O} = \left\{\vec{\rho}, \vec{t}, \vec{\phi}, \vec{\xi^2}, \vec{D}\right\}$. The vectors $\vec{\rho}$, $\vec{t}$, and $\vec{\phi}$ are the SNRs and coalescence times and phases in each of the participating detectors, $\vec{D}$. The $\vec{\xi^2}$ is a signal consistency test statistic based on the autocorrelation function of the complex SNR timeseries.

To estimate $\mathcal{L}$, we consider the numerator and denominator separately and factor the PDFs into simpler terms. Here, I will give a brief overview of the terms beginning with the signal model in Sec. 2.4.1 followed by the noise model in Sec. 2.4.2, but readers are referred to [110, 111] for more details.

2.4.1 Signal model

The numerator of Eqn. 2.16 factors as follows:

$$\begin{aligned} p(\vec{O},\vec{\theta}|\vec{s}) =& P(t_{\rm ref},\phi_{\rm ref}|\vec{\theta},\vec{s}) \\ & \times P(\vec{D}|t_{\rm ref},\vec{s}) \\ & \times P(\vec{\rho},\vec{\Delta t},\vec{\Delta\phi}|\vec{D},t_{\rm ref},\vec{s}) \\ & \times P(\vec{\xi^2}|\vec{\rho},\vec{\theta},\vec{s}) \\ & \times P(\vec{\theta}|\vec{s}) \end{aligned} \tag{2.17}$$

The first term, $P(t_{\rm ref},\phi_{\rm ref}|\vec{\theta},\vec{s})$ is the probability of finding a trigger at the coalescence time and phase $t_{\rm ref},\phi_{\rm ref}$ measured in a reference detector[1] under the assumption that a signal, $\vec{s}$, described by a template with the parameters $\vec{\theta}$ is present in the data [110, 111]. This term depends on the sensitivity of the detector network as a function of time which can be quantified using the reference detector's horizon distance, $D_H(t_{\rm ref},\vec{\theta})$, as defined in Sec. 1.4.3. This is the distance to which the detector is expected to observe an optimally oriented source with SNR $\rho = 8$. The horizon distance is computed at any time by using the estimated PSD which is tracked continuously by GstLAL as described in Sec. 2.2. The probability that a trigger is observed in the reference detector at $t_{\rm ref}$ is then proportional to the detector's sensitive volume which is simply D_H^3.

The second term, $P(\vec{D}|t_{\rm ref},\vec{s})$ is the probability of the trigger being found in the exact subset $\vec{D}$ of detectors operating at $t_{\rm ref}$ [110, 111]. $\vec{D}$ is defined as the set of detectors which observed the trigger with $\rho > 4.0$ and is not necessarily the same as the set of detectors operating at the time. This term depends on the horizon distances of each detector. It is modeled by simulating triggers uniformly oriented and distributed over the sky and calculating the number of detectors in which each signal is detectable at a given horizon distance.

Given a time and phase in the reference detector we have an expectation of what the times and phases should be in each other detector based on the distance between the detectors and the assumption that gravitational waves travel at the speed of light. $P(\vec{\rho},\vec{\Delta t},\vec{\Delta\phi}|\vec{D},t_{\rm ref},\vec{s})$ represents the consistency of the relative times and phases of the signals observed in each of the participating detectors, $\vec{D}$ [104, 111]. The time (phase) difference is computed relative to the reference, $\Delta t = t - t_{\rm ref}$ ($\Delta\phi = \phi - \phi_{\rm ref}$). This term is also modeled by simulation. We construct the PDF by calculating the Δt and $\Delta\phi$ for

[1] In GstLAL we arbitrarily choose the first participating detector in alphabetical order as the reference.

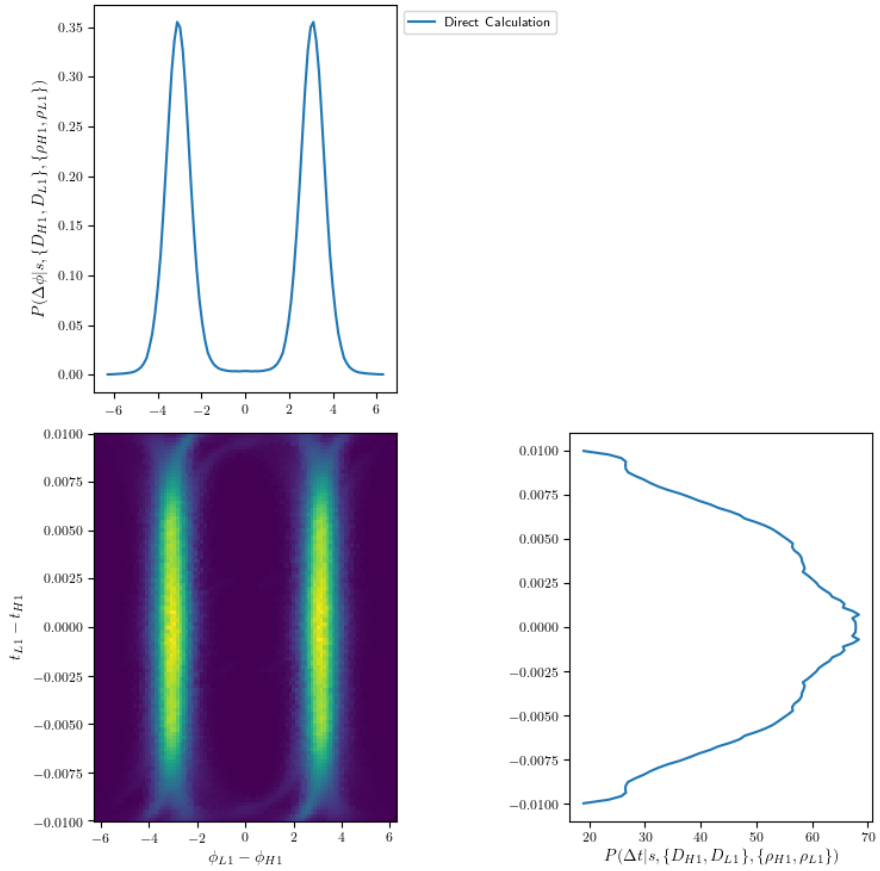

Figure 2.5. The simulated probability distributions of Δt (lower right) and $\Delta\phi$ (upper left) for a trigger found in LIGO Livingston and LIGO Hanford. The joint PDF over Δt and $\Delta\phi$ is shown in the lower left. We expect real GW signals to fall somewhere on the bright regions of this distribution.

simulated sources uniformly distributed on the sky. We discard any triggers observed with $|\Delta t|$ greater than some conservative cut off which is taken to be larger than the light travel time between detectors. Finally, we smooth the resulting distribution with a Gaussian kernel density estimate (KDE). An example of the resulting PDF for LIGO Hanford and LIGO Livingston is shown in Fig. 2.5.

The next term, $P(\vec{\xi^2}|\vec{\rho},\vec{\theta},\vec{s})$ is also a signal consistency test [4, 111]. The detectors are noise dominated and include glitches which can ring up triggers with high SNR even when the template waveform is not a good match for the glitch morphology. If we assume that the template, h_α, with parameters $\vec{\theta}$ exactly represents the signal in the data, then we know what the matched filter response time-series should be. This is given by the

autocorrelation time-series,

$$R_\alpha(t) = \int_{-\infty}^{\infty} \mathrm{d}f \frac{|\tilde{h}_{2\alpha}(f)|^2 + |\tilde{h}_{2\alpha+1}(f)|^2}{S_n(f)} \exp\left(2\pi i f t\right). \tag{2.18}$$

We can compare this to the measured SNR time-series to understand how closely the signal shape matches the template, i.e.

$$\xi^2_\alpha(t) = \left|\rho_\alpha(t) - \rho_\alpha(0) R_\alpha(t)\right|^2, \tag{2.19}$$

where the autocorrelation time-series is scaled by the value of the complex SNR time-series, $\rho_\alpha(t)$, at $t = 0$ corresponding to the peak SNR at coalescence. The ξ^2 is computed by integrating this time-series over a small window surrounding the trigger time, i.e. $[-\delta t, \delta t]$, and the vector $\vec{\xi^2}$ is the ξ^2 value in each participating detector. Because the noise in each detector is assumed to be independent, we can model the PDF as a product over the detectors,

$$P(\vec{\xi^2}|\vec{\rho}, \vec{\theta}, \vec{s}) \approx \prod_{i \in \mathrm{IFOs}} P(\xi_i^2|\rho_i, \vec{\theta}, \vec{s}). \tag{2.20}$$

Historically, the PDFs for each detector had been estimated empirically by tuning a number of free parameters. However recently, [111] re-derived the term as a semi-analytic function.

The final term in Eqn. 2.16, $P(\vec{\theta})|\vec{s})$ is a prior on our astrophysical expectations of the distribution of GW source properties in the parameter space [110, 111]. This term is often referred to as a *population model* or *mass model*. It represents our expectation of how likely it is to find a GW signal with intrinsic parameters $\vec{\theta} = \{m_1, m_2, s_{1,z}, s_{2,z}\}$ corresponding to each template in the bank. The choice of distribution is made in advance and often informed by previous detections.

2.4.2 Noise model

The denominator of Eqn. 2.16 factors as follows:

$$\begin{aligned} p(\vec{O}, \vec{\theta}|\vec{n}) =& P(\vec{D}|t_{\mathrm{ref}}, \vec{n}) \\ & \times P(\vec{\Delta t}, \vec{\Delta\phi}|\vec{D}, \vec{n}) \\ & \times P(\vec{\rho}, \vec{\xi^2}|t_{\mathrm{ref}}, \vec{\theta}, \vec{n}) \end{aligned} \tag{2.21}$$

The first term, $P(\vec{D}|t_{\rm ref},\vec{n})$, is the probability that a random noise event would produce a trigger in the exact combination of detectors, $\vec{D}$, at $t_{\rm ref}$ [110, 111]. We first consider the rate of noise events in a single detector which can be approximated as a Poisson process and measured simply by counting triggers over a long period of time. The mean rate of noise events in a single detector is denoted by μ_1. The rate of coincident events among a pair of detectors, D_1 and D_2, is $\mu_{1\wedge 2} = \mu_1\mu_2\Delta t_{12}$, where Δt_{12} is a coincidence window chosen based on the light travel time between the two detectors. For a two detector analysis these are the only quantities we need. For more detectors, we would need to compute triple (or more) coincident rates which is non-trivial but a method for doing so is derived in [110]. To estimate the PDF, we need to compute the coincident rate among the exact subset of detectors $D = \{D_1, D_2, ... D_N\}$, which we denote as $\mu_{\not> \vec{D}}$. This is estimated as the sum of the coincident event rates among each combination, $\vec{N}$, of detectors in $\vec{D}$:

$$\mu_{\vec{D}} = \sum_{\vec{N}\in\vec{D}} \mu_{N_1\wedge N_2\wedge ... N_n}. \tag{2.22}$$

But, any of the detector combinations in this sum may occur as part of a coincidence among a larger set of detectors, $\vec{O}$ (the total set of detectors considered in the analysis), which must be subtracted from the overall rate, i.e.

$$\mu_{\not> \vec{D}} = \mu_{\vec{D}} - \mu_{\vec{O}}. \tag{2.23}$$

Finally, the probability is given by the ratio of the mean coincidence rate for the subset of detectors $\vec{D}$ to the sum of mean coincidence rates of all possible detector combinations (including non-coincidents, i.e. single detectors) in the total set, $\vec{O}$,

$$P(\vec{D}|t_{\rm ref},\vec{n}) = \frac{\mu_{\not> \vec{D}}}{\mu_{\vec{O}}} = \frac{\mu_{\not> \vec{D}}}{\mu_{\not>1} + \mu_{\not>2} + ... \mu_{\not>1\wedge 2} + \mu_{\not>1\wedge 3} + ... \mu_{1\wedge 2\wedge ... 3} + ...}. \tag{2.24}$$

The second term, $P(\vec{\Delta t}, \vec{\Delta\phi}|\vec{D},\vec{n})$, is similar to the signal consistency test over times and phases in the numerator. In the noise hypothesis we assume that any spurious coincidences formed by noise in the independent detectors are not coherent and therefore their relative times and phases will be uniformly distributed [104, 111]. This makes the PDF a constant.

Finally, we have the probability of finding noise triggers with the given ρ and ξ^2 values in each detector, $P(\vec{\rho},\vec{\xi^2}|t_{\rm ref},\vec{\theta},\vec{n})$ [4, 111]. For this term, we compute the ξ^2 from Eqn. 2.19 for all of the triggers in the background, i.e. non-coincident triggers found during times when at least two detectors were in operation. For each detector, we

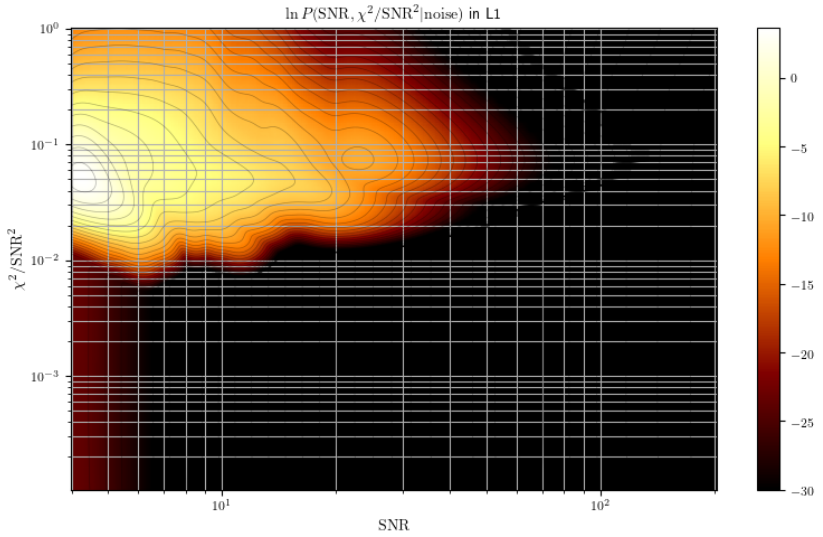

Figure 2.6. An example of a $P(\rho, \xi^2 | \{\theta\}, \vec{n})$ constructed for a background bin of templates with $\mathcal{M}_c \sim 4.39 - 4.75 M_\odot$ in LIGO Livingston data. The χ^2 values referenced in the title and axis label are a holdover of a historical naming convention, but refer to the ξ^2 values of Eqn. 2.19.

compute a two dimensional background histogram over the $\rho - \xi^2$ space. We smooth this distribution with a Gaussian KDE to obtain the PDF, and similarly to Eqn. 2.20 we have:

$$P(\vec{\rho}, \vec{\xi^2} | t_{\rm ref}, \vec{\theta}, \vec{n}) \approx \prod_{i \in \rm IFOs} P(\rho_i, \xi_i^2 | \vec{\theta}, \vec{n}). \tag{2.25}$$

This equation references a particular template defined by the parameters $\vec{\theta}$, however in practice we construct a single PDF over each background bin, $\{\theta\}$, assuming that the noise properties do not vary significantly over the neighboring templates in the bin. One such histogram is shown in Fig. 2.6. The bright region in the upper left corner of the plot, for low SNRs and high ξ^2 values, is dominated by noise. Any GW candidates falling in this region of the plot would be difficult to distinguish from noise. However, candidates falling in the dark regions of the plot, where SNR is high and ξ^2 is low, are considered more likely to be real GW signals.

While GstLAL does not require coincidence among detectors to form candidates, we do choose to down rank the candidates found in only a single detector by a tunable *penalty*. This penalty is applied as an extra term in the denominator of the natural logarithm of the likelihood ratio, i.e. for single detector candidates:

$$\ln \mathcal{L} \to \ln \mathcal{L} - \ln(P). \tag{2.26}$$

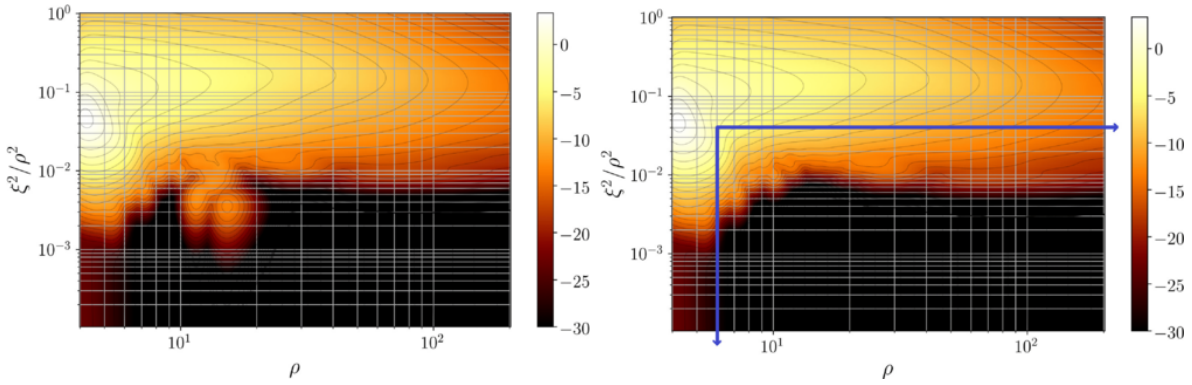

Figure 2.7. An example of an $\rho - \xi^2$ PDF with significant background contamination (left). The contamination is shown by the "islands" in the distribution around SNR $\sim 10 - 20$ and $\xi^2/\rho^2 \lesssim 10^{-2}$. The right panel shows the same panel after background count removal, where the blue arrows indicate the region in $\rho - \xi^2$ space where counts are removed. The disappearance of the "island" features shows that the unwanted contributions from signal have been removed and the PDF now more accurately represents the noise distribution. Reprinted from [5].

Given all the terms discussed in Eqn. 2.21, we can construct histograms of the $\mathcal{L}$ distributions in each background bin, $P(\mathcal{L}|\{\theta\}, \vec{n})$. By adding these histograms together, i.e. *marginalizing* across the background bins, we produce an overall distribution of the likelihoods in the noise hypothesis, $P(\mathcal{L}|\vec{n})$. This quantity is used to estimate the significance of GW candidates in Sec. 2.5.

2.4.3 Background contamination

As mentioned previously, it is imperative to the sensitivity of the search that we do not accidentally include real signals in the construction of our background histograms. Any contribution in the background from a trigger originating from a real GW signal (referred to as *background contamination*) can negatively effect our ability to distinguish signal from noise. In low-latency operation, the background histograms are built up over time. The background used to assess the likelihood of a given trigger cumulatively includes all the noise triggers throughout the analysis up to some short period (less than four hours) before the detection time. In other words, we use our previous observations to inform the likelihood of future observations and so the likelihood estimation is causal. Even if there is background contamination, a trigger will never be ranked against a background which includes itself. However, contamination will effect future triggers, making their estimated likelihoods lower than if there was no contamination and effectively lowering the sensitivity of the search [5].

Our construction of the background hinges on the assumptions that real signals will

ring up triggers in coincidence among more than one detector and that such signals are rare. For a real signal, especially a signal with high SNR, many nearby templates may ring up a trigger. Although we expect the "best" matching template (the one with the highest SNR) to be found in coincidence among detectors, it is possible that nearby templates could ring up non-coincident triggers. These triggers may be placed in the background even though they originate from real signals. Another possible, albeit rarer, avenue for background contamination is a trigger with a GW source located in the blind spot of one operating detector. In this case, a real signal may only ring up a trigger in a single detector and therefore be added to the background. If signals are rare then even if we have background contamination, the effect on the overall distribution can be assumed to be negligible. However, as the detectors have become more sensitive, our detection rates increased beyond what was expected when algorithms such as GstLAL were initially designed and background contamination has become a significant problem. Therefore, recent developments have focused on methods of mitigating the negative effects of background contamination [5].

A recent development has allowed us to track time stamps associated with triggers added to the background histograms. This enables us to ex post facto remove any counts from the background which are near in time to triggers determined to be real signals. In [5] the background contamination was studied and a region of the $\rho - \xi^2$ space was identified where background contamination is particularly likely. Fig. 2.7 shows an example of a background histogram before and after signal contamination was removed from this region. Any counts in a 10 second window around a "true signal" which fall in this region are subtracted, effectively "cleaning" the background. Since we of course do not a priori know which triggers are true GW signals, in practice we must make a choice for how to determine when to remove counts from the background. This choice must balance the desire to aggressively remove all signals while being careful not to remove actual noise signatures from the background.

2.5 Significance estimation

In the previous section I described how the GstLAL pipeline chooses to categorize triggers as background (likely to originate from noise) and candidates (potentially astrophysical in origin) and uses the *likelihood ratio* to rank the candidates such that those with the highest $\mathcal{L}$ are thought most likely to originate from an astrophysical GW source. In order to claim a detection we must estimate the significance of these candidates. To do this,

we use the false alarm probability (FAP) and the false alarm rate (FAR).

At a likelihood threshold of $\mathcal{L}^*$[2], the FAP is the probability that an analysis yielding M noise triggers (again, meaning non-coincident triggers during times when at least two detectors were operating) will produce at least $N = 1$ noise trigger with $\mathcal{L} \geq \mathcal{L}^*$. First, consider the toy example where $M = 1$. The probability that this single noise trigger has likelihood above the threshold is:

$$P(N = 1) = P(\mathcal{L} \geq \mathcal{L}^*). \tag{2.27}$$

This is the complement of the probability that the trigger has $\mathcal{L} < \mathcal{L}^*$, i.e.

$$P(\mathcal{L} \geq \mathcal{L}^*) = 1 - P(\mathcal{L} < \mathcal{L}^*). \tag{2.28}$$

We can approximate the rate of noise triggers with $\mathcal{L} \geq \mathcal{L}^*$ as a Poisson process so that the probability of finding k triggers with $\mathcal{L} \geq \mathcal{L}^*$ is given by:

$$P(N = k) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{2.29}$$

Here, λ is the expected number of events with $\mathcal{L} \geq \mathcal{L}^*$ in a given interval (which we take to be the analysis duration). We can find λ by integrating over the distribution of observed $\mathcal{L}$ for the entire analysis, $P(\mathcal{L}|\vec{n})$ [4]. Therefore, we have that

$$\lambda = C(\mathcal{L}^*|\vec{n}) = \int_{\mathcal{L}^*}^{\infty} \mathrm{d}\mathcal{L} P(\mathcal{L}|\vec{n}), \tag{2.30}$$

is given by the complementary CDF over $\mathcal{L}$. Taking $k = 0$ and substituting into Eqns. 2.29 and 2.28, we have that for a single noise trigger:

$$P(\mathcal{L} \geq \mathcal{L}^*) = 1 - e^{-C(\mathcal{L}^*|\vec{n})}. \tag{2.31}$$

Now, for an analysis producing M noise triggers, we need the probability that *at least one* noise trigger has $\mathcal{L} \geq \mathcal{L}^*$ which is the complement of the probability that *all* of the triggers have $\mathcal{L} < \mathcal{L}^*$ [4]:

$$P(N \geq 1) = 1 - (P(N = 0))^M = 1 - e^{-MC(\mathcal{L}^*|\vec{n})}. \tag{2.32}$$

[2]In practice, we typically refer to the natural logarithm of the likelihood ratio, $\ln \mathcal{L}$, however here I will use $\mathcal{L}$ to simplify the notation.

This is the FAP.

The FAR at a threshold of $\mathcal{L}^*$ is the number of noise triggers with $\mathcal{L} \geq \mathcal{L}^*$ produced over the observation time, T [4]:

$$\text{FAR} = \frac{C(\mathcal{L}^*|\vec{n})}{T}. \tag{2.33}$$

In the low-latency mode of operation, T is tracked as the live time of the analysis, essentially the running wall clock time.

2.6 Online injection campaign

Typically, GW searches are validated by analyzing *injections* which are simulated signals added to the strain data. Since the exact signal is known, we can then assess the performance of the search – its sensitivity to the known signal and how accurately it identifies the signal parameters. Having this functionality, especially in low-latency, is critical for monitoring a GW search. If we hope to observe rare merger events with multi-messenger counterparts, we need GW searches which are highly performant and stable. This is not possible to achieve if the health of an analysis is not easily available to users. In the past, before injection campaigns were implemented in low-latency, analysts struggled from a frustrating inability to identify weaknesses in the analysis until a real GW candidate was observed, or potentially missed. A low-latency injection campaign allows an analysis to continuously demonstrate its sensitivity to potential signals and analysts to quickly identify and diagnose failures when they occur. More than this, an injection campaign allows to measure the sensitive *volume-time*, $\langle VT \rangle$, which is an important science output of GW searches used to place limits on astrophysical merger rates [3, 62]. Historically, the sensitive $\langle VT \rangle$ of GstLAL has been estimated with time consuming re-analysis of the data, long after its initial availability.

Injection campaigns in low-latency are achieved by duplicating the data source, so that we can analyze the original non-injected data stream alongside the injected data stream. However, as discussed in Sec. 2.4.3, our estimation of the background noise histograms relies upon the assumption that signals in the data are rare. Therefore, if the rate that signals are injected in the data is much larger than the expected astrophysical rate, it poses a problem for the significance estimation. On the other hand, an injection campaign where signals are only injected at the astrophysical rate will not be very effective. In GstLAL this problem is solved by using the non-injected data stream to

collect noise statistics in each background bin. Then, the injection candidates are ranked against the background collected in the corresponding non-injected bin. The technical implementation of this method is described in the next section.

2.7 Software infrastructure

GstLAL, which gets its name from "LIGO Algorithms Library[3] for GStreamer", fundamentally depends upon the use of GStreamer[4] for data streaming and signal processing. GStreamer is an open-source framework for real-time audio and video processing which is built on *graphs* representing the flow and manipulation of data through sources, filters, and sinks. These elements can be combined together in complex ways to form a data analysis *pipeline*. GstLAL makes use of many standard GStreamer elements as well as custom elements written specifically for the problem of gravitational wave data analysis [112].

The GstLAL software package is written in Python and C and comprises four subpackages (in order of dependence):

- `gstlal` is the base package. It includes fundamental data analysis and signal processing tools.
- `gstlal-ugly` is built upon `gstlal`. It is a temporary holding place for new modules and applications in the development stage.
- `gstlal-inspiral` and `gstlal-burst` include post-processing layers which are necessary for the specific use case of analyzing LIGO and Virgo strain data to identify CBC GW signals and unmodeled *burst* signals respectively. These packages have dependencies in `gstlal` and `gstlal-ugly`.

The software is public and maintained via the LIGO Scientific Collaboration's GitLab instance at: `https://git.ligo.org/lscsoft/gstlal`. The code is primarily distributed as Docker Images built by the GitLab continuous integration (CI) and stored at the LIGO containers registry: `containers.ligo.org`. The images include all the necessary dependencies to run the GstLAL software and are currently built upon the Rocky Linux 8 operating system (OS), the default OS on LIGO Data Grid (LDG) clusters at the time of writing.

[3] `https://git.ligo.org/lscsoft/lalsuite`
[4] `https://gstreamer.freedesktop.org/`

The GstLAL software suite is vast and written with modularity in mind so that GStreamer elements can easily be linked together through a simple Python API to construct arbitrary pipelines [112]. In the rest of this section, I will focus on the technical implementation of the specific GstLAL low-latency inspiral pipeline and its deployment workflow developed for the LVK Collaboration's fourth observing run (O4).

The pipeline is built as a simple directed acyclic graph (DAG) and deployed via HTCondor[5] in two stages: an initial *set up* stage where pre-computed data products are prepared and an *inspiral* stage which continuously analyzes the incoming strain data and produces the GW candidates. However, since the inspiral analysis is designed to run continuously and essentially perpetually until user intervention, there is actually no "direction" in the directed acyclic graph, i.e. there are no dependencies between the applications at run time.

Any information which applications require from each other is shared asynchronously and largely without the use of file handling. We rely on the micro web-framework Bottle[6] for autonomous information sharing among applications as well as to command dynamic configurational changes by users throughout the runtime. Additionally, we leverage the power of the Apache Kafka messaging system via the Confluent Kafka[7] Python API. Every application in the GstLAL low-latency inspiral DAG acts as a Kafka consumer, producer, or both by ingesting messages from or sending messages to various Kafka topics. This messaging is used to trigger critical event-based actions in the pipeline as well as to communicate a high volume of monitoring data outputs. We store the monitoring data as time-series measurements in InfluxDB[8]. This allows us to easily visualize the live health and performance of the analysis via custom web-based dashboards powered by Grafana[9]. Additionally, the longtime storage of monitoring data enables archival look-backs on the state of the analysis at any given time.

2.7.1 Set up stage

In this section, I will outline each of the applications shown in the set up DAG in Fig. 2.8. These applications prepare data products needed to begin a low-latency GstLAL analysis. In contrast to the inspiral DAG of Sec. 2.7.2, these applications are only run once per experiment.

[5] `https://htcondor.readthedocs.io/`
[6] https://bottlepy.org/docs/dev/
[7] `https://docs.confluent.io/home/overview.html`
[8] `https://www.influxdata.com/`
[9] `https://grafana.com/`

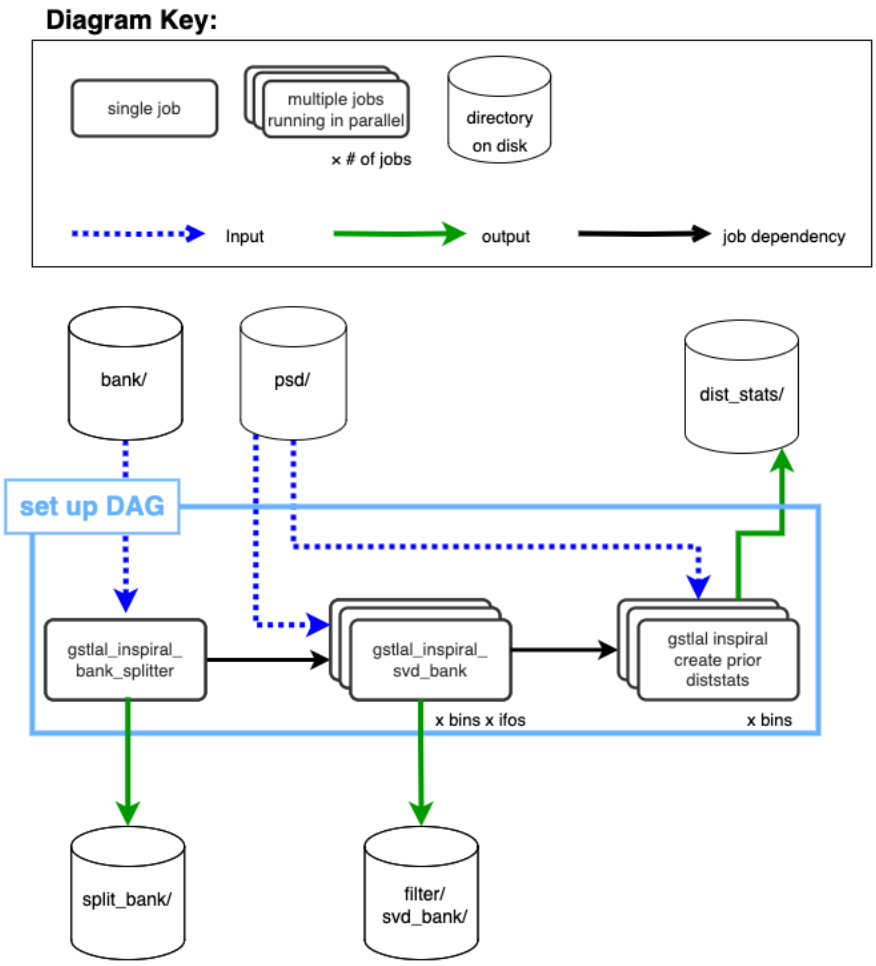

Figure 2.8. Schematic of the GstLAL set up DAG for the inspiral pipeline. The black arrows indicate dependencies among applications (or *jobs*), e.g. the `gstlal_inspiral_bank_splitter` stage must complete before the `gstlal_inspiral_svd_bank` stage may proceed. Job inputs and outputs are indicated by the blue-dashed and green arrows, respectively.

`gstlal_inspiral_bank_splitter`. As the name suggests, this application splits a full template bank into small sub-groups, sometimes referred to as *split banks*, of neighboring templates. The user must make a choice of which parameters to use for grouping templates. Ideally, the choice would produce split banks with maximally overlapping templates so that we can achieve the maximum SVD compression, i.e. the sampling frequency weighted ratio of the number of basis waveforms to the number of physical templates required to achieve the given reconstruction accuracy. A greater compression means we can filter a set of templates more efficiently, leading to a lower-latency search. Further, the similarity between templates in a split bank allows for the assumption that the background statistics will not vary appreciably over the set of

templates, $\{\theta\}$. The choice of parameters for template bank splitting in O4 is discussed in Chapter 5 [113].

`gstlal_inspiral_svd_bank`. This application combines neighboring split banks into background bins of templates which are time-sliced and decomposed into basis waveforms for filtering according to the LLOID method [92, 101]. This application also pre-whitens the templates according to the reference PSD.

`gstlal_inspiral_create_prior_diststats`. This application initializes the ranking statistic terms in Eqns. 2.17 and 2.21. The PDF is written to a file and will later be populated by noise triggers and candidates throughout the analysis.

2.7.2 Inspiral stage

Here, I will outline each of the applications shown in the inspiral DAG in Fig. 2.9 and Fig. 2.10.

`gstlal_inspiral`. The GstLAL low-latency inspiral pipeline is centered around this application. It encompasses the main GStreamer graph – from data retrieval and conditioning to filtering, reconstruction of the SNR timeseries and identification and ranking of triggers. A heavily simplified illustration of this GStreamer graph is shown in Fig. 2.11. The DAG includes $2N$ instances (or in the language of HTCondor, *jobs*) of this application, where N is the number of background bins and the factor 2 accounts for duplicated jobs to support the parallel injection campaign.

Fig. 2.9 illustrates the required inputs and the outputs produced, including files and Kafka messaging. The application requires three input files: one which specifies the SVD basis waveforms to filter, one which defines a reference PSD used to mitigate the initial loss in SNR in the several minutes while the live PSD is converging, and one which stores the background histogram mentioned in Sec. 2.4.2. The background histogram is then periodically updated, or *snapshotted*, with the cumulative background triggers collected throughout the runtime. The strain data is processed in one second packets of data, referred to as *buffers*. Upon processing of each buffer, the job outputs a significant amount of monitoring data via Kafka, thereby acting as a Kafka producer. This data includes scientific metrics such as the SNR timeseries, likelihood and FAR history of triggers, as well as performance metrics such as the latency of each element in the GStreamer pipeline and the random access memory (RAM) usage. Information about any GW candidates, i.e. triggers with $\rho_{net} \geq 7$, identified during processing of each buffer is output to the so-called *coinc*[10] Kafka topic. Next, the candidates produced by each job

[10]This topic name is poorly chosen as it incorrectly suggests that we require coincidence among

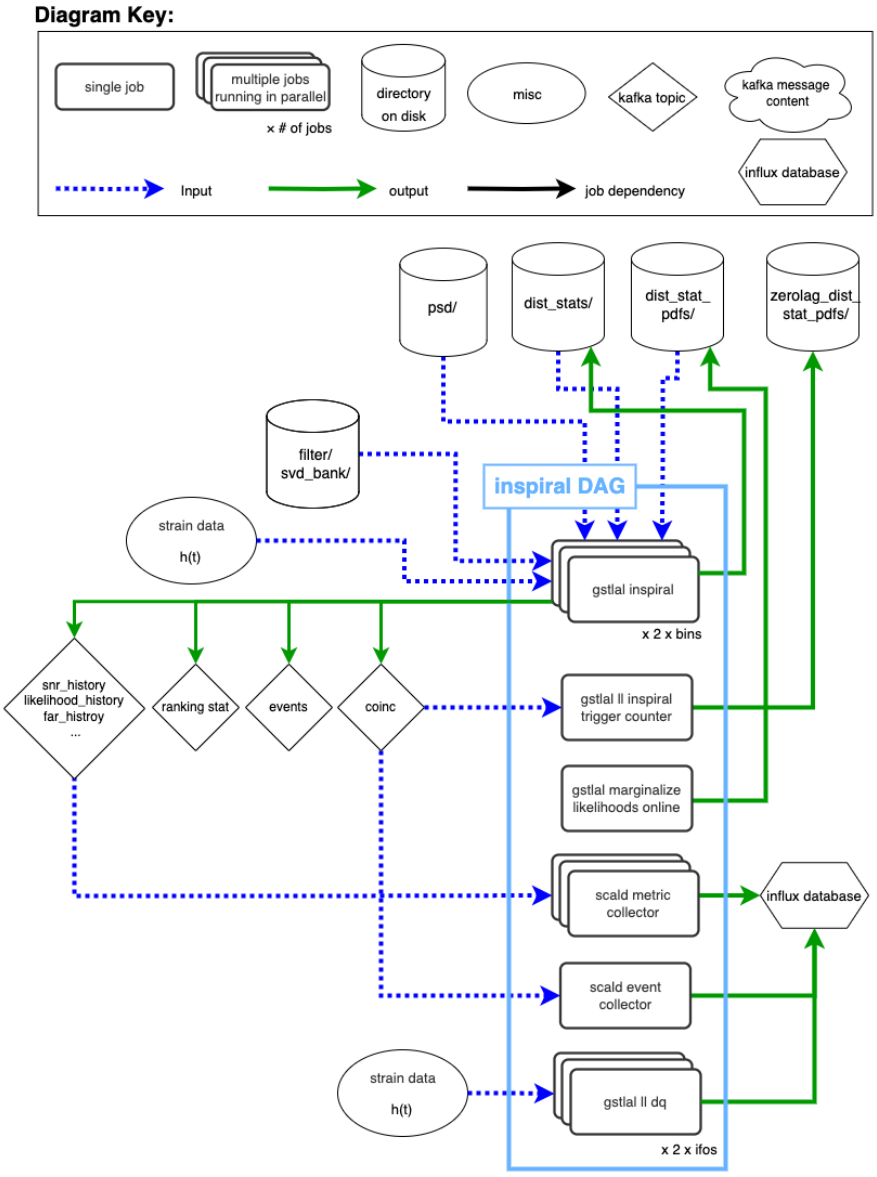

Figure 2.9. Schematic of the GstLAL low-latency inspiral pipeline DAG. This figure illustrates the analysis portion of the DAG with the uploading and post-processing portion being continued in Fig. 2.10. There is no distinction between the two parts in practice. The diagram key is the same as in Fig. 2.8 with the addition of Kafka topics and message content indicated by diamonds and cloud shapes, respectively and InfluxDB indicated by the hexagonal shape. Note that there are no job dependencies in the inspiral DAG. The factors of $\times 2$ on the number of parallel jobs account for the duplicated jobs needed to support the parallel injection campaign.

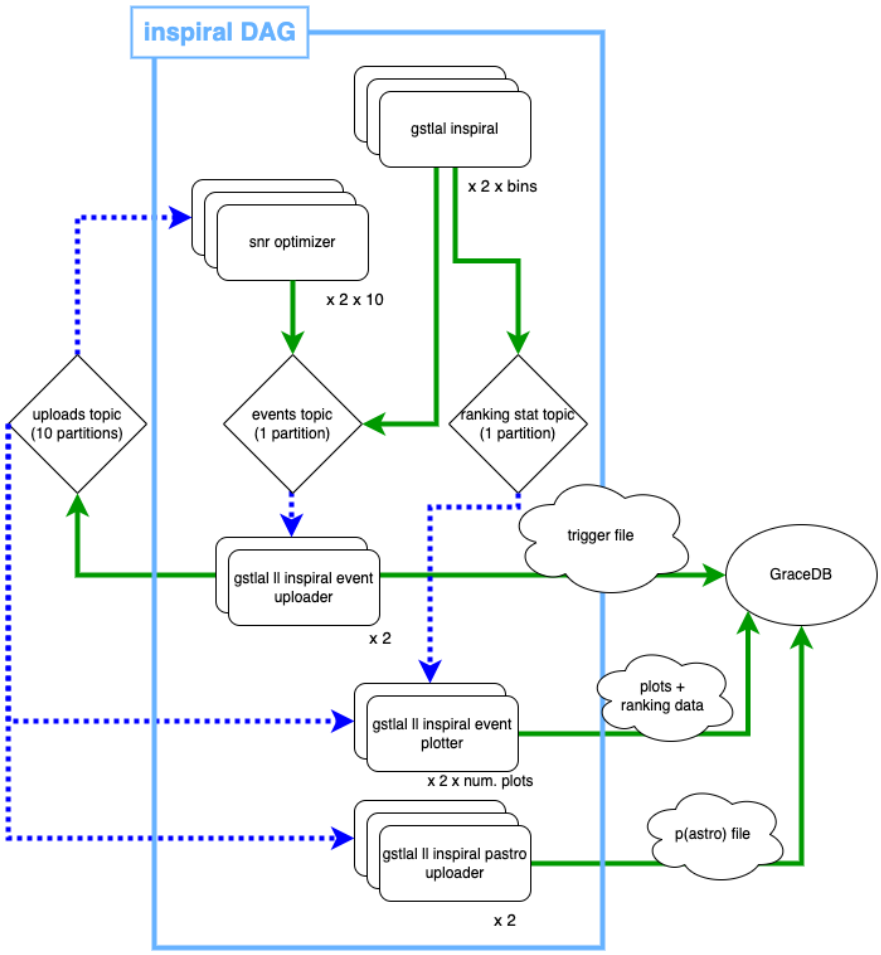

Figure 2.10. Continuation of the schematic of the GstLAL low-latency inspiral pipeline DAG shown in Fig. 2.9. This figure illustrates the uploading and post-processing portion of the DAG. There is no distinction between the two parts in practice. The diagram key is the same as in Fig. 2.9.

(i.e. in a single background bin) are aggregated by keeping only the maximum likelihood over ten second windows. Finally, any candidates which pass a pre-determined FAR threshold are communicated via the so-called *events* Kafka topic to be used down-stream.

`gstlal_ll_inspiral_trigger_counter`. This application is primarily a Kafka consumer which ingests information about GW candidates from `gstlal_inspiral` via the *coinc* Kafka topic. It aggregates the incoming candidates across all background bins by taking the maximum likelihood in ten second windows and, as the name suggests, produces the likelihood counts histogram. The job periodically writes the updated histogram as a file throughout the runtime. As shown in Fig. 2.9, there is a single

detectors to form GW candidates.

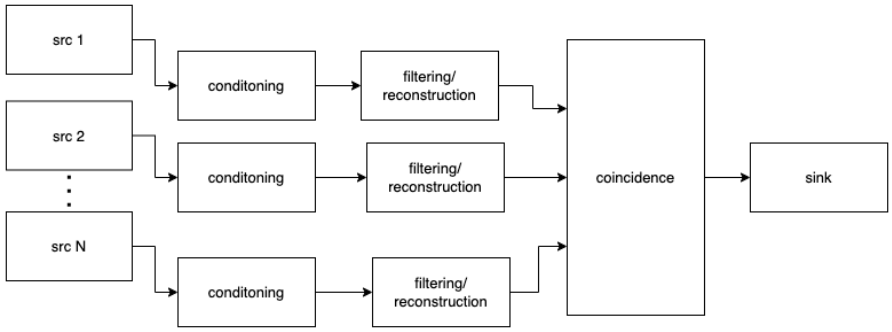

Figure 2.11. A simplified illustration of the GStreamer graph run by the `gstlal_inspiral` application. The N sources correspond to data retrieval from each detector to be analyzed. Data is passed through several GStreamer elements in one second *buffers* where it is conditioned and filtered. All N data streams are combined by a single GStreamer element which is the coincidence engine of the pipeline. Finally, every valid GStreamer pipeline ends with a sink.

instance of this job in the DAG, as we only want to collect non-injection candidates to inform the significance estimation, therefore it is not necessary to have a duplicate job to count injection candidates.

`gstlal_marginalize_likelihoods_online`. This application collects noise likelihood PDFs from each background bin via HTTP requests managed by Bottle. Each running `gstlal_inspiral` job is queried for its background histogram consecutively in a running loop and the counts are cumulatively added across bins. After the noise PDF has been marginalized, the signal PDF is queried from the `gstlal_ll_inspiral_trigger_counter`. It can take several hours to marginalize the PDFs across all the background bins, at which point the final PDF is written as a file and the process starts over. The `gstlal_inspiral` jobs periodically re-load this file so that the background used for FAR assignment remains current up to some lag, typically $4-8$ hours. Again, since the PDFs are only collected from the non-injection background bins, we only require one instance of this job in the DAG and the output PDF it produces is used to rank both injection and non-injection candidates.

The upload stage of the inspiral DAG, illustrated in Fig. 2.10, encompasses all of the post-processing steps after GW candidates have been identified and assigned FARs. As indicated in Fig. 2.10, each of the applications in the upload stage of the inspiral DAG are duplicated. This is so that the injection campaign is supported in parallel but kept completely distinct from the default analysis, i.e. the injection data is never mixed with non-injection data on the same Kafka topics or processed by the same application.

`gstlal_ll_inspiral_event_uploader`. This application acts as a Kafka consumer, ingesting candidates passing a FAR threshold from `gstlal_inspiral` on the *events* topic.

These are aggregated across background bins by keeping either the maximum SNR or the minimum FAR candidates in some small time window. The aggregation scheme is a configurational choice and discussed in more detail with regards to the O4 configuration in Chapter 5 [113]. Candidates surviving this second round of aggregation are uploaded via an HTTP post request to the GraceDB[11] and communicated to the *uploads* Kafka topic.

`gstlal_ll_inspiral_pastro_uploader` and `gstlal_ll_inspiral_event_plotter`. These applications are Kafka consumers triggered by messages sent to the *uploads* topic to produce metadata for each candidate uploaded to GraceDB.

`manifold_cbc_bank_snr_optimizer_online`. This application is also a Kafka consumer triggered by messages on the *uploads* Kafka topic. It uses the methods described in [58, 114] to perform a targeted search over the candidate time window with the goal of achieving a higher SNR. This application then also acts as a Kafka producer, communicating its candidates back to the *events* topic for upload to GraceDB. The SNR optimization for a single candidate can take several minutes so in the inspiral DAG we divide the incoming candidates among ten parallel processors by partitioning the Kafka topic. In this way, each individual optimizer job is assigned messages from a single partition and is only responsible to process every tenth candidate.

The remaining applications in the inspiral DAG are used to store the significant amount of monitoring data produced by the analysis. We use the SCalable Analytics for LIGO/Virgo/KAGRA Data (`scald`)[12] software which provides GstLAL with bespoke wrappings for the Python Confluent Kafka and InfluxDB APIs.

`gstlal_ll_dq`. This application tracks the live data quality of LIGO and Virgo strain data including the detector *states*, indicating whether the detector is in observation mode, as well as a live estimate of the PSD. These metrics are directly stored to InfluxDB for via the `scald` utilities.

`scald_metric_collector` and `scald_event_collector`. These applications ingest Kafka messages from the `gstlal_inspiral` jobs which make up the bulk of the monitoring metrics of the analysis. The data is aggregated in various time windows and written to InfluxDB for live visualization on Grafana and archival storage.

[11] `https://gracedb.ligo.org/`

[12] `https://git.ligo.org/gstlal-visualisation/ligo-scald`

2.8 Conclusion

The potential for scientific gain from gravitational wave observations is vast. Already, in just a few years of observing, we have used GW detections to confirm longstanding theories as well as uncover new and unexpected information in areas such as astrophysics and fundamental physics. These exciting advances depend on sensitive, performant, and reliable search algorithms to dig out the quiet signals from detector strain data and accurately estimate their significance. GstLAL is one such search pipeline which has demonstrated state-of-the-art sensitivity to gravitational waves from CBC sources. It has been used to search LIGO and Virgo's strain data in a low-latency configuration since the first observing run and has participated in many of the most groundbreaking detections, including the first detection of GWs from a pair of merging black holes, GW150914 [12], the first direct detection of GWs from a pair of merging neutron stars, GW170817 [66], and the first detections of a neutron star merging with a black hole, GW200115 and GW200105 [115] to name a few. Since its initial development, the pipeline has been in a constant state of flux being updated and improved incrementally over time for expanded functionality and better performance [4, 104, 111, 113]. The methods summarized here are the most up to date at the time of writing.

The remainder of this dissertation will cover methods and results from several searches for gravitational waves which I have led using the GstLAL search pipeline. First, in Chapter 3, a sub-threshold search for BNS systems coincident with gamma ray bursts. In Chapter 4, I provide an overview of the new GW detections added to the Gravitational Wave Transient Catalog (GWTC-3) after the LVK Collaboration's third observing run. Finally, in Chapter 5 - 6 I detail the latest configuration and performance of the GstLAL search and demonstrate its extremely successful contribution to public results from the first part of the fourth observing run.

Chapter 3
Joint GW-GRB sub-threshold search

3.1 Introduction

One of the most exciting scientific prospects of GW detection is to enable multi-messenger astronomy. The field has existed for decades, combining observations of neutrinos, cosmic rays, and EM radiation from such sources as solar flares, supernovae, and blazars [116–121]. However, it has only been in recent years that gravitational waves have been added to the list of viable cosmic messengers.

BNS mergers had been an expected source of gravitational waves long before the development of the LIGO and Virgo interferometers. In 1975, Hulse and Taylor measured the systematic variation in the period of the known pulsar, PSR $1913 + 16$, to determine that it was captured in a binary orbit [16]. After observing the system for many more years, an energy loss was inferred via the slow monotonic decay of the orbital period [17]. This observation was consistent with general relativity predictions of the energy loss due to gravitational radiation and represented the first indirect detection of gravitational waves.

At the same time, the bright transient bursts of photons known as gamma ray bursts (GRBs) have been a popular focus of astronomical observation and study since their first detection in 1969 [122]. Since then, the observed population of GRBs have been empirically grouped into two classes deemed "short" and "long" bursts [123]. The distinction is based upon a strong bi-modality in the burst duration, specifically the amount of time it takes for 90% of the isotropic energy of the burst to be emitted, shown in Fig. 3.1. This quantity is typically referred to as T_{90}. Short GRBs are those for which $T_{90} < 2$ seconds, and long GRBs have $T_{90} > 2$ seconds. Since 1998, it had been suspected that long GRBs originate from the death of large stars in core-collapse supernovae and this claim has been supported with conclusive evidence over many years

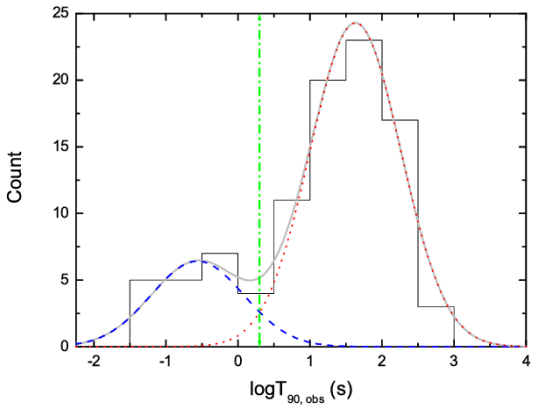

Figure 3.1. The observed bi-modality in the distribution of T_{90} for GRBs. The cut off between "short" and "long" GRBs is indicated by the vertical line at 2 seconds. Reprinted from [6].

of observations [124, 125]. On the other hand, BNS mergers had been proposed as progenitors for short GRBs for many years, but the claim remained unproven. However, this conjecture presented an exciting possibility for the dual detection of gravitational waves and gamma-ray emission from the same source. Due to the complimentary nature of the information carried by GWs and EM radiation, observing both from the same source provides a profound potential for insight on the properties and dynamics of the progenitor system. A detection such as this was sought after for many years until finally achieved in 2017 with the observation of GWs from the BNS merger known as GW170817 along with the subsequent short GRB, GRB170817A, and kilonova AT2017gfo [7, 66, 126–128].

The extreme and violent nature of such events make them spectacular natural laboratories for fundamental and nuclear physics as well as astrophysics and cosmology. In addition to the multi-messenger detection of GW170817 and GRB170817A, to date hundreds of short GRBs have been identified by EM telescopes such as the Neil Gehrels Swift Observatory (*Swift*) and the Fermi Gamma-ray Burst Monitor (Fermi-GBM) [129, 130]. However, the physical mechanisms governing these processes remain relatively opaque. There are still many open questions related to the central engine that drives the burst, exactly what happens in the first few moments of prompt emission, the composition and structure of the ejecta material, and so on [131–137]. To answer such questions, we must prepare to make more joint detections.

However, over the past years we have learned that detections like GW170817 are rare. This motivates the need for highly performant day-to-day operation of GW searches

and the development of novel strategies to expand our opportunities for detection. The former ensures that we meet the basic detection standards. Since the LIGO and Virgo interferometers already have a relatively low coincident duty cycle compared to EM telescopes, it is imperative that our search algorithms operate stably at high sensitivity in low latency so that we are prepared for potential signals any time the detectors are in operation. The latter will allow us to dig even deeper into the limited data we have and extend our detection horizons. The LIGO interferometers regularly go through extended commissioning periods, meaning that for long durations of time we are completely blind to any potential GW counterparts to EM observations. Therefore, we must develop strategies to make full use of any coincident data we have, ensuring that no potential discoveries are left on the table.

In this chapter, I will describe a search for coincident sub-threshold GWs and GRBs from the LIGO and Virgo detectors and *Swift*, respectively. We will begin in Sec. 3.2 with a case study of GW170817, where we aim to outline the detection scenario, what we have learned, and the ways in which we can even further optimize the detection scenario for similar events in the future. We then give an overview of necessary background information on the *Swift* Burst Alert Telescope (BAT) in Sec. 3.3. In Sec. 3.4 we describe the joint search and conclude with current and potential science outcomes in Sec. 3.5.

3.2 GW170817 case study

The detection of gravitational waves from GW170817 was the first direct detection of merging neutron stars and provided exciting new insights into GRB physics [66, 67]. However, in many ways this detection was serendipitous, and as we have seen over the recent years of observation, a detection scenario such as this is very rare. As shown in the spectrogram of Fig. 3.2, GW170817 was an extremely loud signal with a network SNR in LIGO's strain data of 32.4 [66]. The signal was determined to have originated from a pair of merging neutron stars at an estimated luminosity distance of only 40 Mpc [66]. This distance corresponds to a cosmological redshift, $z \sim 0.008$ [66]. To date, this is the closest merger event which has been observed by LIGO and Virgo. To put the numbers in context, the next closest source in the GWTC-3 has a luminosity distance ~ 0.16 Gpc and redshift $z \sim 0.03$, about four times further away [3, 138]. Compared to the known short GRB population, GRB170817A is also an outlier, being several orders of magnitude closer and dimmer [67, 139].

Despite its high SNR and all three detectors (LIGO Hanford, LIGO Livingston,

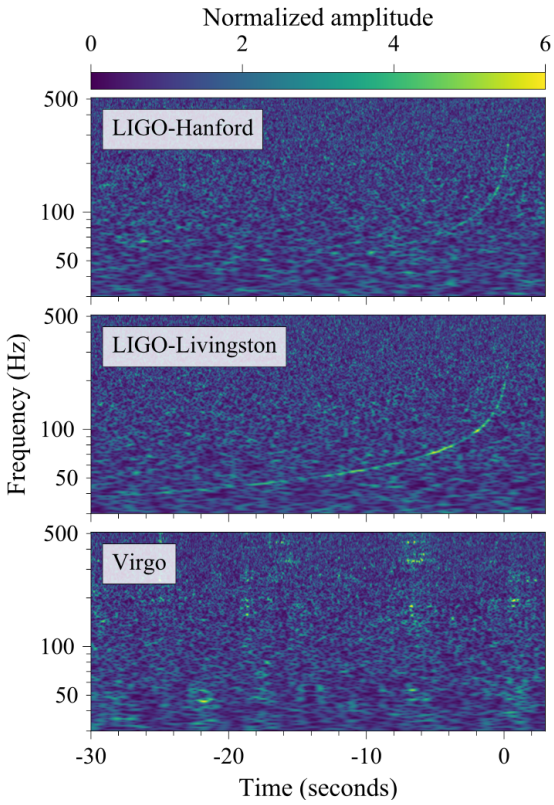

Figure 3.2. Time frequency spectrogram of the GW170817 signal. The chirping inspiral track is clearly visible in LIGO Hanford and LIGO Livingston. Reprinted from [7].

and Virgo) being operational at the time, GW170817 was originally detected in only a single interferometer, LIGO Hanford. This was due to the presence of a loud glitch ~ 1 second before the coalescence time in LIGO Livingston. However, since the high SNR made the signal chirp clearly visible in the spectrogram data, the LIGO Livingston data was manually re-filtered after excising the glitch with a zeroing window function, allowing a coincidence to be formed among both LIGO detectors [66]. At the time, the LVK Collaboration had not yet begun sending automated public alerts. The first preliminary GCN notice [140], which included the potential detection as a single detector candidate, was sent about 40 minutes after the trigger time and the initial Fermi-GBM detection notice [141]. Subsequent re-analysis of the trigger using cleaned data from LIGO Livingston came after this initial alert.

Another peculiarity of the detection scenario for GW170817 was its relatively small sky localization. The best localized sources in GWTC-3 are those for which all three detectors participated in the observation. These candidates have 90% credible sky areas of tens of square degrees [3, 138]. GW170817 was localized to only 28 deg^2 [66]. While

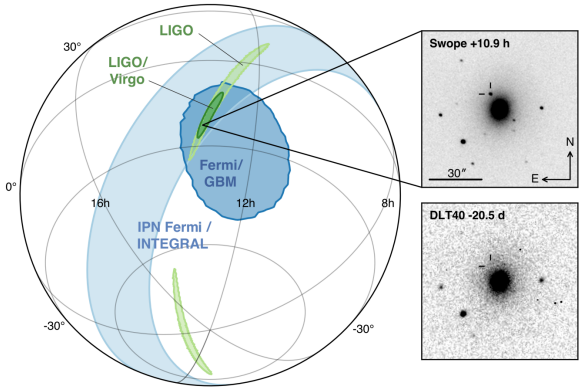

Figure 3.3. The source localization contours for GW170817. The LIGO-only localization is shown in light green and the improved LIGO-Virgo localization in dark green. Reprinted from [7].

this is still much larger than the field of view of many EM telescopes, it is clearly one of the best localized gravitational wave candidates to date. As mentioned in Chapter 1, gravitational wave detectors have an antenna pattern of sensitivity varying across the sky. The precise sky localization, despite the trigger only appearing in the LIGO detector data, is attributed to the event being located in Virgo's blind spot. Virgo was operating stably at the time of the trigger with a BNS horizon distance of 58 Mpc [66]. Therefore, the lack of signal in Virgo data indicated that the source was in Virgo's blind spot. Using data from both LIGO detectors and the Virgo detector produced a well-constrained sky localization, as shown in Fig. 3.3.

At the same time, the GRB was bright enough to be independently observed by Fermi-GBM at high significance [67]. The time and spatial coincidence with the GW trigger provided immediate overwhelming evidence that the two indeed originated from the same source. The relatively small sky localization allowed astronomers with optical telescopes to tile the region quickly and identify the host galaxy as NGC 4993 [7]. With the host galaxy identified, the source was essentially pinpointed and deep follow up observations of the accompanying kilonova could proceed in frequency bands across the EM spectrum [7, 95, 142–146].

GW170817 was truly extraordinary but we can't count on the same observation scenario for the next multi-messenger event. In some ways the detection of GW170817 was also sub-optimal, as it took more than half an hour to produce a preliminary GCN notice with identification of the gravitational wave trigger and more than an hour to produce an initial detection statement [140]. Therefore, although we were able to observe the event across the EM spectrum over long timescales, we completely missed the early

	Energy band	Field of view	Source localization
BAT	$15 - 150$ kev	1.4 sr	$\sim 1' - 4'$
UVOT	$170 - 600$ nm	$17' \times 17'$	0.3"
XRT	$0.2 - 10$ kev	$23' \times 23'$	5"

Table 3.1. Various technical specifications of each instrument onboard *Swift*.

post-merger stage of the evolution which is of extreme interest for elucidating exactly what happens after two compact objects merge. Further, we cannot expect the next BNS merger to be localized as well by the GW observation alone. If the signal is found in only one or two detectors, the localization can be hundreds of square degrees or as large as the entire hemisphere. In this case, simply tiling the sky area will not be a feasible way to identify the host galaxy of the source.

There is still much work to be done towards preparing for the next multi-messenger event and increasing our chances of observing its early evolution. We must prepare for the possibility of detecting a quieter gravitational wave signal or a source which is further away. This kind of event may not be highly significant in gravitational waves or gamma rays independently. Additionally, we must develop a more robust infrastructure that will allow us to make the detection, localize its source, and alert the community as quickly as possible so that we can observe all stages of the EM emission. The detection strategy presented in this chapter addresses each of these goals.

3.3 Swift Burst Alert Telescope

Swift is a satellite telescope designed to observe GRBs across multi-wavelength EM emission and localize their positions on the sky to the level of arcsecond precision within minutes [147]. *Swift* encompasses three telescopes: BAT [148], the X-ray Telescope (XRT) [149], and the UV/Optical Telescope (UVOT) [150]. The relevant technical capabilities of each onboard telescope are outlined in Table 3.1. Since its launch in 2004, *Swift* has regularly detected GRBs at a rate of about 100 yr^{-1} [129, 151, 152].

In its typical mode of operation, BAT identifies burst triggers by observing increased counts above the expected background [153]. If the source of the excess is determined to be a point source, the sky position is automatically located to a precision of $\sim 1' - 4'$. This initial localization can take up to about 10 sec to produce. Then, the XRT and UVOT are autonomously commanded to slew to the position, where they begin observing the afterglow emission. The delay time between the initial burst and the first UV/Optical

and X-ray observations will include the 10 sec for BAT to locate the source plus the time it takes to slew the XRT and UVOT, which is typically about ~ 10 sec $- 1$ min. The XRT is capable of further localizing the source to within 5" after about 10 sec of the slew. Finally, the source may then be even better localized by the UVOT to within only 0.3". These triggers are communicated to the ground immediately and distributed to the EM follow-up community via GCN. This *event data* includes timing, source position, and energy information for each detected count. Following the detection, the afterglow is periodically observed for days to weeks by the XRT and UVOT.

Due to the large field of view (FOV) of the BAT instrument, the typical data dumps for onboard triggered bursts comprise a large volume of data which requires significant bandwidth to communicate to ground [154]. For this reason, *Swift* was not designed to communicate or archivally store event data for times without significant bursts triggered onboard. This means that *Swift* astronomers have suffered an inability to detect weaker GRBs that may not trigger the onboard detection algorithm and historical targeted searches have additionally been limited by the lack of complete data. This limitation motivated the development of the Gamma-Ray Urgent Archiver for Novel Opportunities (GUANO) infrastructure for downlinking event data on demand, surrounding astrophysically interesting times, as triggered by external searches [154]. This notably includes GW triggers from the LIGO and Virgo detectors but could additionally include weak short GRBs detected by e.g., Fermi-GBM, neutrinos from IceCube, etc.

BAT event data is stored on a ring buffer such that data is overwritten on timescales of $\sim 20 - 30$ min [154]. This means that if an external command does not reach *Swift* within this time frame, the data of interest will be lost forever. GUANO manages a queue of commands based on alerts received from external sources. The commands specify the start time and duration of data to be saved and are queued in order of priority based on the external trigger's FAR. The *Swift* analysis team utilizes the Tracking and Data Relay Satellite System (TDRSS) to schedule commands to *Swift* with a minimum latency of 14 min. The success of the GUANO command depends upon several factors, most importantly, the latency with which the command reaches the instrument. In [154] the GUANO performance was characterized as reaching a 99% success rate for commands received by *Swift* with a total latency less than 16 min. The development of this method has been crucial in increasing the detection rate of *Swift* GRBs, especially for weak bursts and targeted searches.

The event data saved by GUANO is then analyzed on the ground by the Non-Imaging Transient Reconstruction and Temporal Search (NITRATES), a detection

algorithm designed for identifying sub-threshold GRBs in BAT data. Details of the NITRATES method are given by [155]. In contrast to the BAT onboard detection algorithm, NITRATES relies upon a significantly more computationally heavy maximum likelihood algorithm to achieve a vast improvement in sensitivity. In combination with GUANO, the use of NITRATES has further extended the *Swift* detection horizon and led to additional detections at a rate of $\sim 13\ \mathrm{yr}^{-1}$ since its operation began [155]. The joint search described in this chapter is only feasible due to the development of the GUANO commanding infrastructure and NITRATES detection algorithm.

3.4 The joint search

In the O3, the LVK Collaboration released automated public alerts for candidate gravitational waves reaching a FAR threshold of 1.2 per year [3]. This threshold was chosen as a balance so as not to inundate astronomers with a high number of false alarms while maximizing the number of true astrophysical candidates that were released. The joint GW-GRB search with *Swift* used a much higher FAR threshold of two per day. This high threshold was chosen to allow the joint search to probe deeply into the available data, with a higher priority placed on identifying all "true alarms" rather than avoiding "false alarms". For each gravitational wave trigger found below this threshold, astronomers at *Swift* were alerted via automated notices. In O3, these alerts were sent by private email directly to *Swift* astronomers under a special Memorandum of Understanding (MOU). In O4, the LVK issued *low-significance* public alerts for all triggers reaching a FAR threshold of two per day [156]. Therefore, *Swift* astronomers could receive these alerts via GCN rather than the bespoke email alert infrastructure of O3.

In O3, the alert contents included the GPS time associated with the GW coalescence time and a measure of the trigger's significance in the form of both the network SNR and the FAR. Additionally, information about the search producing the trigger was included. The LVK GW search pipelines producing public alerts are grouped into two categories: "CBC" (compact binary coalescence) and "Burst". These groups indicate the types of sources each pipeline searches for. CBC searches (e.g., GstLAL) use matched filtering to search for modeled waveforms from merging black holes and neutron stars. Burst searches perform coherent searches over the time-frequency data in at least two detectors to look for unmodelled transient signals from a variety of potential searches. The alert carries information about the expected probability of astrophysical origin ($p(\mathrm{astro})$) of the signal source, including the probability that the source is a BNS, NSBH, or BBH.

The search group and p(astro) information allow *Swift* astronomers to prioritize the follow up of each alert based on an expected likelihood that the source will have an associated EM counterpart. For example, the followup of a BNS alert may be prioritized over a BBH alert, since BNS mergers are known progenitors of short GRBs. Finally, the alert included an initial skymap Flexible Image Transport System (FITS) file produced by BAYESTAR [157]. The sky localization information was used in the estimation of joint significance, described below.

Our bespoke email alert system in O3 was built using the `LVAlert` (now International Gravitational Wave Network Alert System (`igwn-alert`)) notification system for receiving state changes from GraceDB[1]. The system achieved latencies of only a few seconds, measured from the trigger's time of upload on GraceDB to the time the email was sent. This extremely low latency was imperative to the success of the search as the GUANO data dump success rate was highly dependent upon the time of arrival of alerts relative to the trigger time. Given the low latency of alerts reaching the *Swift* ground team, the overall latency of the alerts reaching the instrument was typically limited by the latency of sending the GUANO command via the TDRSS, as described in [154]. Beginning midway through the O3 run in September 2019, we sent 636 sub-threshold alerts to the *Swift* team, reaching a 90% data dump success rate by the end of the run in March 2020 [154, 158].

As described in Sec. 3.3, upon successful downlinking of the *Swift* BAT sub-threshold event data, the NITRATES pipeline identifies the maximum likelihood GRB trigger. If a coincident trigger is identified, the joint significance is estimated by computing the combined FAR [159]:

$$\mathrm{FAR} = \frac{Z}{I_\Omega}\left[1 - \ln\left(\frac{Z}{Z_{max}}\right)\right]. \tag{3.1}$$

Here, Z indicates the joint ranking statistic,

$$Z = \mathrm{FAR}_{\mathrm{GW}} \times \mathrm{FAR}_{\mathrm{GRB}} \times \Delta t. \tag{3.2}$$

The independently computed FARs of the GW trigger and the GRB trigger are $\mathrm{FAR}_{\mathrm{GW}}$ and $\mathrm{FAR}_{\mathrm{GRB}}$, respectively. The coincidence search window is $\Delta t = 30$ sec, corresponding to $[t_c - 10, t_c + 20]$, with t_c being the GW coalescence time. Z_{max} is computed using upper FAR cut offs, $\mathrm{FAR}_{\mathrm{GW},max} = 2.315 \times 10^{-5}$ Hz (two per day) and $\mathrm{FAR}_{\mathrm{GRB},max} = 1000$ Hz, giving $Z_{max} = 6.945 \times 10^{-7}$ Hz. I_Ω is the sky area overlap integral as derived in [160].

Given this formula for computing the joint GW-GRB FAR, we can see that a sub-

[1] `https://gracedb.ligo.org/`

threshold gravitational wave trigger which is found in coincidence with a sub-threshold GRB trigger will be substantially increased in significance. As an example, consider a GW trigger with $\mathrm{FAR}_{\mathrm{GW}} = 2 \times 10^{-5}$ Hz (just under two per day) and a GRB trigger with $\mathrm{FAR}_{\mathrm{GRB}} = 1 \times 10^{-5}$ Hz. Neither event could constitute a significant detection alone. The joint ranking statistic is $Z = 6 \times 10^{-9}$ Hz. If we conservatively choose an uninformative $I_\Omega \gtrapprox 1$ [160], then we find that the combined FAR is 3.45×10^{-8} Hz, or about one per year. This example further motivates the choice of the high two per day threshold for sending alerts.

3.5 Conclusion

None of the 636 alerts sent to *Swift* in O3 produced significant coincidences in low-latency. In [158] a detailed offline analyses was performed for the triggers associated with each alert. Neither did this offline analysis produce detection quality candidates. However, the null result still allows constraints to be placed on the photon flux and luminosity upper limits of GW candidates [158].

Despite the null result, the joint search described in this chapter represents one of the most promising avenues for detecting an EM counterpart to a GW signal and a kilonova. With this method, we can leverage the superior localization capabilities of *Swift* to promptly identify where on the sky a GW signal has originated from. For bright GRBs which trigger the onboard BAT detection algorithm automatically, this localization can be determined and communicated to ground observers on a timescale of seconds to minutes. Such an observation scenario presents extraordinary potential for scientific discovery, especially in observing the very early gamma-ray and kilonova emission. However, the recent development of the GUANO and NITRATES infrastructure means that *Swift* can further be utilized to detect sub-threshold GRBs in targeted searches coincident with GW triggers, albeit at a much higher latency due to the computational burden of the analysis. Still, this method pushes the joint detection horizon to the limits, allowing us to fully probe the limited GW-GRB data. Even a coincident detection with *Swift* which is delayed by $\mathcal{O}$(hours), presents an extraordinary potential for new scientific discoveries.

Chapter 4
The third GW transient catalog

This chapter is adapted from [3].

4.1 Introduction

In the previous chapter, I demonstrated a strategy for detection of gravitational waves relying on search pipelines operating in a low-latency configuration to enable multi-messenger astronomy. This kind of strategy focuses on enabling a deep study of a single GW and coincident EM transient event. However, as the LIGO and Virgo detectors get more sensitive we seek to uncover populations of GW sources and study their bulk properties. In addition to digging deeply into a single observation, we aim to detect broad populations of GW sources which can inform our understanding in such areas as black hole and neutron star mass distributions, binary formation channels and merger rates, the early universe, and star formation [62, 161–163].

With data from the LIGO and Virgo detectors we have been regularly observing gravitational wave candidates since 2015. These candidates are collected cumulatively into the Gravitational Wave Transient Catalog (GWTC-3) following re-analysis of the strain data on long timescales relative to its initial availability. This is referred to as an *offline analysis* in contrast to the low-latency *online analysis*. The first instance of this catalog, GWTC-1, added candidates detected during the first observing run (O1) (12 September 2015 – 19 January 2016) and the second observing run (O2) (30 November 2016 – 25 August 2017) [164]. In this analysis, 11 significant GW candidates were identified which all passed a threshold on the probability of astrophysical origin, $p(\text{astro}) > 0.50$. These notably included the first detection of gravitational waves from a pair of merging black holes, GW150914 [12], and the first detection of gravitational waves from a pair of merging neutron stars, GW170817 [66]. GWTC-2 added 39 candidates

from the first part of the third observing run (O3a), which took place from 1 April 2019 – 1 October 2019 [138]. These 39 candidates were identified based on a FAR threshold of 2.0 per year. The data from O3a was later re-analyzed with improved data calibration, resulting in GWTC-2.1 which reported on 44 GW candidates with $p(\text{astro})$ above the threshold of 0.50 [165]. These 44 candidates included 36 of the same candidates from GWTC-2 and 8 new candidates. Therefore, using the $p(\text{astro})$ threshold for consistency, by the end of O3a the catalog cumulatively included 55 candidates.

In this chapter I will discuss the latest addition to the catalog, GWTC-3, which adds 35 new significant GW candidates from the second part of the third observing run (O3b) which took place from 1 November 2019 – 27 March 2020 [3]. In this analysis, we followed the precedent of GWTC-1 and GWTC-2.1 by using the threshold $p(\text{astro}) > 0.50$ to identify highly significant candidates. These candidates are all consistent with originating from compact binary coalescences: binary black holes and for the first time in O3b the confident detection of a neutron star–black hole binary, GW200115_042309 [115]. We additionally find candidates which have significant probabilities of having at least one component mass in either of the upper or lower mass gaps [166–169] and several sources which have confidently unequal masses and high effective inspiral spin. This addition brings the total number of highly significant candidates reported by the LVK Collaboration throughout each of the first three observing runs, O1 – O3, to 90.

The catalog additionally includes 7 *marginal* candidates which had $p(\text{astro}) \leq 0.50$ but FAR passing a threshold of 2.0 per year [3]. The significant and marginal candidates are discussed in Sec. 4.4. Additionally, GWTC-3 included a deep sub-threshold list of 1048 gravitational wave triggers extending to a FAR threshold of 2.0 per day. These triggers are mostly thought to be consistent with noise, and therefore we forego detailed follow up studies of them. However, these triggers were made available in a public data release via the Gravitational Wave Open Science Center (GWOSC) [170].

The success of the third observing run was due in large part to the improvement in detector sensitivity. This can be quantified in terms of the BNS *time-volume* which is roughly proportional to the observed detection rate, as shown in Fig. 4.1. The time-volume is given in units of Gpc^3 yrs. The observing time in O1 – O2 was taken to be any time when at least two detectors were observing. In O3, since the GstLAL search pipeline allowed for candidates identified in only a single detector we take the observing time to be any time when at least one detector was observing. The volume is the Euclidean sensitive volume, which is computed as the volume of a sphere with radius the BNS inspiral range (as defined in Sec. 1.4.3) of the second most sensitive detector in the network at

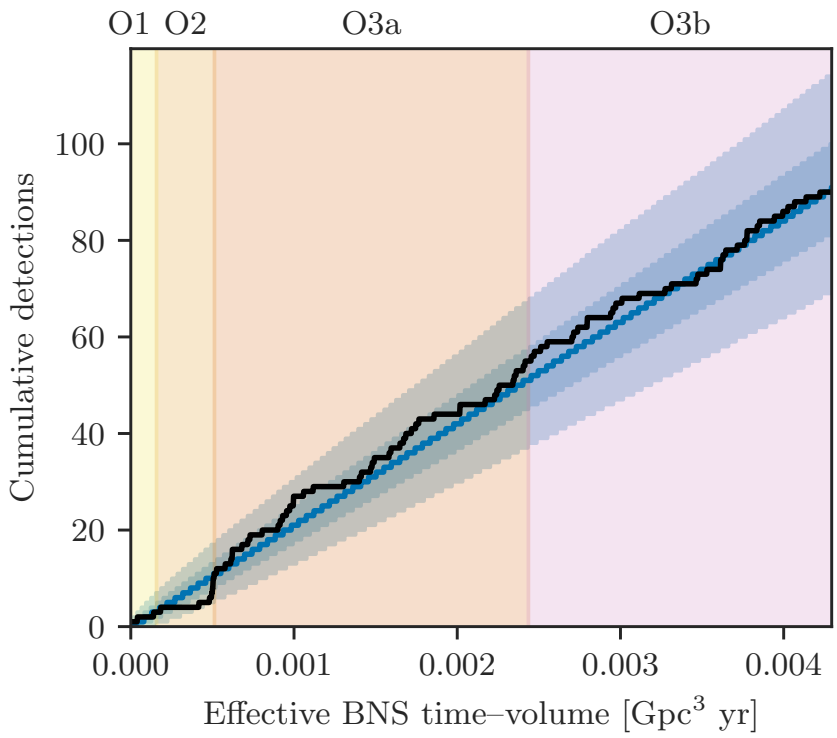

Figure 4.1. The horizontal axis shows the sensitive BNS time-volume of the detector network in each observing run, indicated by the colored bands. The vertical axis is the number of GW detections with $p(\text{astro}) > 0.50$. The cumulative observed candidates are shown by the black line, while the blue line and dark and light blue shaded regions indicate the median, 50% confidence interval, and 90% confidence interval on the expected number of detections for a Poisson distribution fit to the total number of candidates at the end of O3b. While O2 and O3b had roughly the same observing duration, the sensitive time-volume covered by O3b is several times that of O2, indicating the drastic improvement in detector sensitivity. Reprinted from [3].

each time [171, 172]. Fig. 4.1 shows that our number of detections has grown rapidly throughout each observing run, correlating with the increase in sensitive time-volume.

As a member of the LVK Collaboration, I am expected to dedicate 90% of my research time to collaboration activities. Towards the production of GWTC-3, my contributions are as follows. I operated the GstLAL search pipeline to produce many of the candidate events in the catalog, listed here in Sec. 4.4 [3]. The search sensitivities presented in Sec. 4.5 were estimated by *injection campaigns* run over the full duration of O3. For the GstLAL analysis, I carried out these injection analyses amounting to months of work and millions of CPU hours, essentially doubling the computation time of the initial offline analysis. As an editorial team member, I was a lead author for [3] and saw the paper

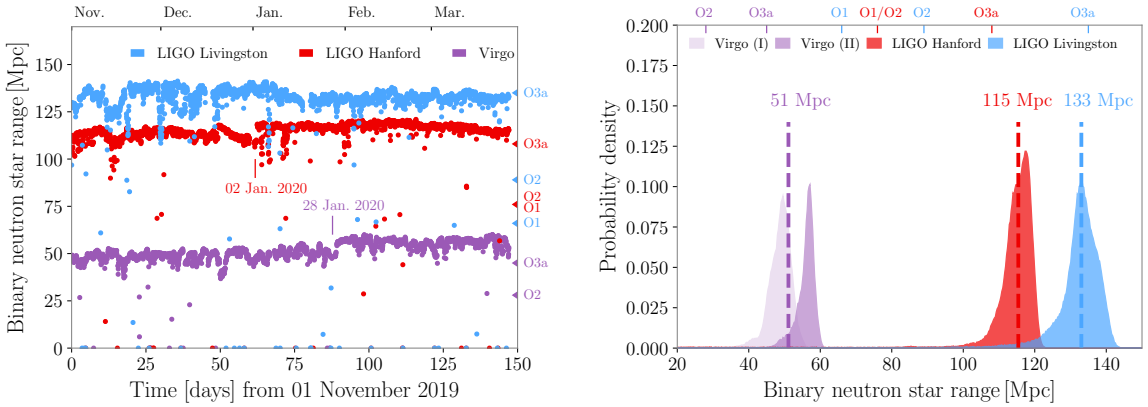

Figure 4.2. Left. The BNS inspiral range of both LIGO detectors and Virgo as a function of time in O3b. Each data point indicates the median inspiral range taken over hour-long segments of data. The right axis indicates the median BNS inspiral ranges over the entire duration of the O1 - O3a runs for comparison. **Right.** Distribution of the BNS inspiral range in each detector with the median values over the full duration of O3b indicated by the dashed lines. Reprinted from [3].

through from data collection, writing, and publication. In the rest of this chapter, I will begin by briefly summarizing the state of the LIGO and Virgo detectors during O3b, the strain data calibration, and the noise mitigation strategies in Sec. 4.2. In Sec. 4.3 I will overview the methods used by the four participating search pipelines and describe the detection statistic, i.e. the $p(\text{astro})$, used to determine inclusion as a high significance candidate in the catalog. In Sec. 4.5, I will describe the methods used to assess search sensitivity. Finally, in Sec. 4.6 I will give an overview of the inference of some of the source parameters before concluding with Sec. 4.7.

4.2 Detectors and data

The LIGO and Virgo detectors have always operated in planned periods alternating between observing and commissioning [173]. Before the start of O3 there was a long commissioning period where the detectors underwent significant upgrades. Then there was a short one month commissioning break between O3a and O3b from 1 October – 1 November 2019. During this time the LIGO and Virgo detectors each had certain goals for their improvement, which are briefly summarized below along with their corresponding impacts on the BNS inspiral range.

LIGO Hanford and Livingston. The goals between O3a and O3b included: improving the BNS inspiral range, improving data quality by mitigating noise sources that

produce glitches, and improving the stability and joint duty cycle of the detectors [3, 38]. In both detectors, the main improvement to the range came from improvements to the light squeezer between O3a and O3b [174–176]. In LIGO Livingston, the squeezed light improvements led to a gain in the BNS inspiral range of about 3 Mpc and in LIGO Hanford the improvements led to a gain of about 7 Mpc. The laser power in both detectors was increased significantly between O2 and O3 leading to a reduction in photon shot noise. LIGO Hanford's laser power was not changed between O3a and O3b. In LIGO Livingston, new point absorbers in the test masses were found between O3a and O3b [49] which led to a reduction in the laser power and corresponding reduction in the BNS inspiral range of about 5 Mpc. Therefore, the range of the LIGO Livingston detector was relatively unchanged from O3a to O3b, with median value 135 Mpc and 133 Mpc, respectively. The LIGO Hanford median BNS inspiral range increased from 108 Mpc in O3a to 115 Mpc in O3b. The comparison of BNS inspiral range between O3a and O3b is shown in Fig. 4.2.

One of the main known glitch types in LIGO Hanford and Livingston are caused by scattered light in the detector recombining with the main beam incoherently [177, 178]. The scattered light produces two known glitch types, referred to as *slow scattering* and *fast scattering* glitches, as shown in the upper and lower panels of Fig. 4.3, respectively [3]. Slow scattering glitches are common in both LIGO detectors and correlate strongly with microseismic activity from ocean waves and earthquakes. Fast scattering glitches, which are more common in LIGO Livingston, are caused by bad weather conditions and human activity, particularly nearby logging trucks. In both detectors, upgrades were implemented during the commissioning break between O3a and O3b to reduce the effects of these noise sources [46]. New baffles were installed to absorb scattered light and keep it from recombining with the main beam and reaction-chain tracking which reduces the relative motion between test masses. While these upgrades do not lead to increased BNS inspiral range, they do significantly improve the data quality by reducing the glitch rate [3].

Finally, the joint duty cycle of the LIGO detectors was improved from O3a to O3b [3, 138]. The duty cycles of each detector are summarized in Table 4.1. LIGO Hanford and LIGO Livingston each increased their independent duty cycles from 71% and 76% in O3a to 79% and 79% in O3b, respectively. This led to a significant increase in the percent of time when at least two detectors in the network were observing, i.e. 81.9% in O3a to 85.3% in O3b.

Virgo. The main goal between O3a and O3b was to use the time to gain understanding

	H	L	V	1+ IFO	2+ IFO	3 IFO
O3a	71%	76%	76%	96.9%	81.9%	44.5%
	(130.3 days)	(138.5 days)	(139.5 days)	(177.3 days)	(149.9 days)	(81.4 days)
O3b	79%	79%	76%	96.6%	85.3%	51.0%
	(115.7 days)	(115.5 days)	(111.3 days)	(142.0 days)	(125.5 days)	(75.0 days)

Table 4.1. The duty cycles of LIGO Hanford (H), LIGO Livingston (L), and Virgo (V) in O3a and O3b given as the percentage of the total observing time as well as the duration in days. The final three columns in the table show the joint duty cycles where at least one, two, and three detectors were observing together. The increase in LIGO Hanford and LIGO Livingston duty cycles led to an improved 2+ IFO duty cycle in O3b compared to O3a [3].

of the various noise sources and sensitivity of the detector [3]. Then over the course of O3b several minor upgrades were made leading to incremental improvements in the detector range from 49 Mpc to 56 Mpc. This improvement is indicated in Fig. 4.2 with the bump in median BNS inspiral range after 28 January 2020. The main upgrade between O3a and O3b was the increase of laser power from 18 to 26 W. The various minor upgrades throughout O3b included reduction in electronic noise, software improvements, and improvements to the Fabry-Perot cavities [179]. Additionally, there was mitigation of scattered light noise [47, 180] and adjustments to the light squeezer which had been initially installed prior to O3a [181, 182]. Each of these upgrades led to approximately $1 - 3$ Mpc gain in the BNS inspiral range.

The detector strain data which is analyzed by searches is reconstructed from various detector output channels using complex models of the detector responses [3]. This includes calibrating the strain to these known models and estimating systematic errors in the reconstruction process as well as removing known noise sources or lines [183–186]. The strain data is initially calibrated in low-latency but often it is re-calibrated before use by the offline searches. This re-calibration can benefit from improved understanding of the detector noise and response models. In O3b, the LIGO Hanford and Livingston data was re-calibrated offline, however the Virgo data used to produce GWTC-3 was the same as the low-latency data.

Before the searches obtain the calibrated strain data it is pre-processed by data quality experts [187–189]. Segments of data determined to be particularly noisy, or containing excess glitches, are removed so that searches only process the clean data [190, 191]. This is done to reduce the possibility of search algorithms mistaking glitches for real GW candidates. After the search algorithms produce a list of candidates, data is further validated in windows around each significant candidate. Known glitches in these

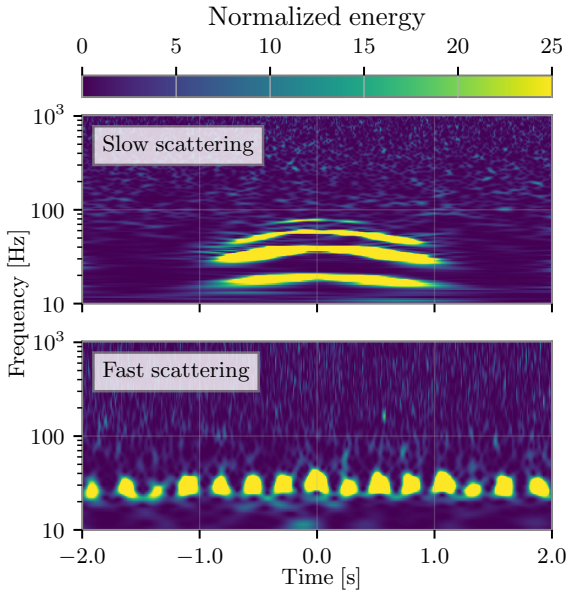

Figure 4.3. Time-frequency spectrograms showing examples of slow scattering (top panel) and fast scattering (bottom panel) glitches in O3b. Reprinted from [3].

segments may be subtracted if they are thought likely to bias the estimation of source parameters [192].

4.3 Search methods

In O3b, there were five search pipelines operating in low-latency: cWB [193, 194], GstLAL [4, 92, 104], MBTAOnline [195, 196], PyCBC Live [197, 198], and SPIIR [199, 200]. For the offline analyses, SPIIR did not contribute results so the new candidates added in GWTC-3 are only produced by cWB, GstLAL, the Multi-Band Template Analysis (MBTA) and PyCBC. I will briefly summarize the methods used by these search pipelines in this section.

Modeled searches. The *modeled* searches, GstLAL, MBTA, and PyCBC, use matched-filtering to search for CBC signals, i.e. merging black holes and neutron stars whose waveforms are well understood. Each search employs a template bank to cover the parameter space of interest. PyCBC used two independent configurations: the PyCBC-broad configuration which covers a broad region of the parameter space with the goal of being sensitive to as many different source types as possible and the PyCBC-BBH configuration which restricts its parameter space to that of previously observed BBH systems. The template bank boundaries of each search are summarized as follows [3].

The GstLAL template bank was the same as that used for the GWTC-2 analysis [138]. The template bank covers waveforms with detector frame total masses ranging from $2M_\odot$ to $758M_\odot$. The spins are assumed to be parallel (aligned or anti-aligned) to the orbital angular momentum. A conservative upper mass cutoff for neutron stars is taken to be $3M_\odot$. Components with masses below $3M_\odot$ have spin magnitudes $|s_{i,z}| < 0.05$. For components with masses above $3M_\odot$, the spin magnitudes are $|s_{i,z}| < 0.999$. The `SEOBNRv4_ROM` waveform approximant [201] is used for templates with chirp mass $\geq 1.73M_\odot$. This is a frequency-domain reduced-order model [202] of the time-domain inspiral–merger–ringdown model `SEOBNRv4` which models quasicircular, nonprecessing BBHs based upon the effective-one-body (EOB) equations of motion [201]. The `TaylorF2` waveform approximant [203] is used for lower-mass systems. This is a frequency-domain, inspiral-only model of aligned-spin CBC systems built from closed-form PN approximations.

The MBTA template bank covers waveforms with detector frame component masses from $1M_\odot$to $195M_\odot$ [3]. MBTA takes $2M_\odot$ as the cutoff between NS components and black hole (BH) components. Templates where the smaller mass is below $2M_\odot$ are constrained to have the larger mass $< 100M_\odot$. The spins are similarly assumed to be parallel to the orbital angular momentum. Components with masses below $2M_\odot$ have spin magnitudes $|s_{i,z}| < 0.05$ and otherwise $|s_{i,z}| < 0.997$. Templates for which both components are below $2M_\odot$ are constructed with the `SpinTaylorT4` waveform approximant [203]. This is an inspiral-only, time-domain model for CBC systems based on the PN equations of motion. If either component has mass above $2M_\odot$, the waveform is generated with the `SEOBNRv4` approximant which is a full inspiral–merger–ringdown waveform [201].

The PyCBC-broad template bank covers templates with total masses ranging from $2M_\odot$ to $500M_\odot$ and the PyCBC-BBH bank restricts to BBH templates with the larger component mass ranging from $5M_\odot$ to $350M_\odot$ and the smaller component mass ranging from $5M_\odot$ to m_1 [3] . The total mass of templates in the PyCBC-BBH bank ranges from $10M_\odot$ to $500M_\odot$. Additionally, the PyCBC-BBH bank places a constraint on the mass ratio, $q = m_2/m_1$, of $1/3 < q \leq 1$. All templates in the PyCBC-BBH bank assume parallel spins with respect to the orbital angular momentum and spin magnitudes $|s_{i,z}| < 0.998$. Both PyCBC bank configurations use the `SEOBNRv4_ROM` [201] waveform approximant for templates with total mass above $4M_\odot$, and `TaylorF2` [203] otherwise.

After matched filtering, these searches identify GW triggers based on an SNR threshold. For GstLAL and PyCBC this threshold is taken as 4.0 while for MBTA the threshold varies across the parameter space but is always in the range 4.5 to 4.8. While PyCBC

and MBTA require triggers to be found in coincidence among at least two detectors in order to qualify as a candidate, GstLAL allows for candidates identified in only a single detector. Some of the most significant differences between the modeled searches are in the construction of the ranking statistic and estimation of the background [3]. Each search uses the likelihood ratio [108, 109] as the ranking statistic, however its specific construction and implementation vary across the searches. MBTA and PyCBC each estimate the background using *time slides*. This means that the data from independent detectors is translated in time with respect to each other by more than the light travel time between detectors [204]. The time-sliding ensures that no coincident signals can remain in the background and therefore any triggers can be safely assumed to be due to noise. As discussed in Chapter 2, GstLAL does not use time-slides to estimate the background but rather uses non-coincident triggers observed during times when at least two detectors were observing. Further, MBTA and PyCBC each rank their candidates against a local background spanning one to a few weeks prior to the detection time while GstLAL ranks candidates against a global background collected over the full observation duration. In other words, the GstLAL candidates are ranked asynchronously; a candidate identified in the first week of the observing run will still be ranked against all of the "future" noise.

Unmodelled search. cWB is referred to as an *unmodelled* search because it can in principle identify signals from any exotic sources with unknown gravitational waveforms [193,194]. While cWB is sensitive to BBH signals, especially those with high masses and short durations, it does not restrict to CBC signals and does not use any template bank or waveforms to filter the data. Instead, it uses the time-frequency spectrograms among pairs of detectors to search for coherent excess power. That is, the analysis used in GWTC-3 restricts to candidates detected among the HL, HV, and LV detector pairs. cWB operated with two configurations in the GWTC-3 analysis – one which is optimized for the detection of high mass BBHs and one for low mass BBHs [3]. The two configurations both operate in the frequency range 16 Hz to 512 Hz, however the high mass BBH search targets signals with central frequency, $f_c < 80$ Hz while the low mass search targets $f_c > 80$ Hz. After a signal is identified, the SNR is estimated by reconstructing the waveforms. Similarly to the modeled searches MBTA and PyCBC, cWB constructs background using time-slides of detector data and uses this background to rank the significance of its candidates.

For inclusion in GWTC-3 we use a FAR threshold of 2.0 per day which represents a deep extended list of triggers, most of which are likely to originate from noise. We

require a significance threshold to identify the triggers in this list which are most likely to be real GW candidates [3]. The FAR indicates how often we expect the detector network to produce a noise trigger at or above the given ranking statistic of a GW candidate in question, however it does not incorporate our astrophysical knowledge about merger rates or populations of compact objects. In other words, the FAR only depends on the estimated background distribution but does not include the foreground estimate. This is not an ideal significance estimator for a collection of events which originate from a broad set of populations. At the same FAR, a GW candidate originating from a well known population will be more significant than one from an uncertain population. For example, a BBH with a FAR of 2.0 per year is more significant than say, an NSBH at the same FAR. This is because we know that merging BBHs are very common, whereas we have much more uncertainty on the rate of merging NSBHs and have detected far fewer of them. A FAR of 2.0 per year may not be enough to claim an NSBH detection given the foreground expectation. If we used FAR as our detection threshold, we would need to use a more stringent FAR threshold to claim BNS and NSBH detections, thereby leaving out BBHs which may likely be real.

The probability of astrophysical origin, $p(\text{astro})$, is a better detection threshold since it incorporates our knowledge of both the background and foreground distributions [205–210]. We follow the precedent of GWTC-1 [164] and GWTC-2.1 [165] and use a threshold on the $p(\text{astro})$of > 0.50 to produce a higher purity set of GW candidates which are followed up with detailed source parameter estimation and careful data quality checks. For cWB candidates, we make the additional constraint that candidates be found in at least one of the modeled searches with $p(\text{astro}) > 0.1$, sine cWB can potentially detect signals from other astrophysical sources, while the $p(\text{astro})$ calculation assumes a CBC source. These candidates are discussed in the next section. By taking the threshold as $p(\text{astro}) > 0.50$ we ensure that all of the significant candidates included have a higher probability of being astrophysical in origin than terrestrial since,

$$p(\text{astro}) = 1 - p(\text{terrestrial}). \tag{4.1}$$

The $p(\text{astro})$ is estimated by comparing the number of events in the background distribution to the foreground distribution at a given ranking statistic [205–207]. The ranking statistic and background distributions are unique to each search pipeline and are the same used to estimate FARs. The foreground distribution represents the likelihood of observing a GW candidate at a given ranking statistic with given source intrinsic

parameters, $\{m_1, m_2, s_{1,z}, s_{2,z}\}$. Construction of the foreground distribution depends on a *population model* – our expectation of the black hole and neutron star mass and spin distributions and binary merger rates. We can further estimate the probability that each candidate originated from the three source classes: BNS, NSBH, and BBH, i.e.:

$$p(\text{astro}) = p(\text{BNS}) + p(\text{NSBH}) + p(\text{BBH}). \tag{4.2}$$

These classes are defined based on a conservative upper limit of the neutron star maximum mass which we take to be $3.0M_\odot$: p(BNS) is the probability that a candidate has both component masses $< 3.0M_\odot$, p(NSBH) is the probability that a candidate has the larger mass $> 3.0M_\odot$, and p(BBH) is the probability that a candidate has both masses $> 3.0M_\odot$. Since it is infeasible to provide detailed source parameter estimation on all the triggers in the catalog (i.e. those which pass the FAR threshold of 2.0 per day) the GWTC-3 provides source classifications based on Eqn. 4.2 for the candidates with $p(\text{astro}) \leq 0.50$ [170].

4.4 Candidates

GWTC-3 includes 35 candidates with $p(\text{astro}) > 0.50$ [3]. Table 4.2 lists each of these candidates and indicates the instruments observing at the time, along with the FAR, network SNR, and p(astro) as found by each of the four search pipelines. Each of these significant candidates are given a GW name following the convention GWYYMMDD_HHMMSS according to the coalescence date and time. Table 4.3 shows the 7 marginal candidates which did not meet the p(astro) threshold but still had FAR less than 2.0 per year. These candidates are also given the GW identifier, except for three candidates which were determined likely to have an instrumental origin. The marginal candidates are discussed in more detail in Sec. 4.4.3. Overall, GWTC-3 includes a total of 1048 candidate events in the deep sub-threshold list extending to 2.0 per day. These events are not discussed individually here, but the data release includes their p(astro) source classifications and estimated sky localizations [170].

During low-latency operation, GW triggers from all search pipelines are uploaded to GraceDB[1] and clustered into *superevents* based on the coalescence time, including all triggers nominally within a window of $[-1, +1]$ seconds. The superevents are automatically assigned an identifier with naming convention SYYMMDD indicating the date of upload.

[1] `https://gracedb.ligo.org/`

In O3, GW candidates originating from CBC searches (i.e. GstLAL, MBTAOnline, PyCBC Live, or SPIIR) with FAR below a public alert threshold of one per two months (6 yr^{-1}) were distributed to the public via GCN alert. There were 39 GCN alerts during O3b, of which 16 were subsequently determined to be instrumental in origin and retracted [3]. None of these 16 candidates were found above the thresholds for inclusion in the GWTC-3 analysis.

Of the 39 public candidates from O3b, 5 were not retracted but also not found above the p(astro) threshold for inclusion as a high significance candidate in GWTC-3 [3]. Of these, S191205ah [211], S200213t [212], and S191213g [213] were GstLAL candidates with modest significance. S191205ah and S200213t were both identified in only a single detector, LIGO Livingston and LIGO Hanford, respectively. S191213g was identified by GstLAL as a quiet signal in both LIGO Hanford and Livingston. These three candidates were either not found by the GstLAL analysis for GWTC-3 or found with FAR > 2.0 per day. S200105ae [214] was an NSBH candidate identified in low-latency in LIGO Livingston and published in [115]. However, in the GWTC-3 analyses this candidate was found with p(astro) below the threshold and is therefore included only as a marginal candidate [3]. This candidate is discussed further in Sec. 4.4.3. Lastly, S200114f [215] was identified in low-latency by the cWB search pipeline. However in the re-analysis for GWTC-3 it was found with a low p(astro) below the threshold for inclusion and was not found by any of the three modeled searches.

The remaining 18 public GW candidates in O3b were found in the GWTC-3 analyses and are included as high significance candidates in Table 4.2. The 17 candidates listed in bold in Table 4.2 are new candidates published in [3] for the first time. These candidates are discussed in the following sections: candidates consistent with a BBH source in Sec. 4.4.1 and candidates consistent with an NSBH source in Sec. 4.4.2. There were no candidates from O3b consistent with a BNS source.

The differences among the candidates identified in low-latency and the GWTC-3 re-analysis are not unexpected. In general, the offline analyses benefit from a better understanding of data quality and the ability to rank candidates against more background. Especially candidates which have modest significance are likely to fluctuate above or below the threshold for inclusion due to slight changes in the data analyzed or the search methods used from low-latency to the offline analysis [3]. The p(astro) threshold used offline is expected to elucidate additional signals in the regions of parameter space where astrophysical rates are high, i.e. BBHs. This expectation is confirmed as we find most of the 17 new candidates to be consistent with BBH sources.

In addition to differences compared to the candidates identified in low-latency, it is expected that the candidate list will differ slightly among the search pipelines. These differences can be attributed to (i) different configurational and tuning choices made by the searches, (ii) different methods in constructing the ranking statistic, (iii) different responses by each search to the detector noise and differences in the background estimation by each search, and (iv) varying sensitivity across the parameter space due to the different template banks used by each search [3]. Despite these differences, the searches are still highly redundant and identify many of the same signals, especially those for which $p(\text{astro}) \sim 1$. Using multiple searches serves as an internal consistency check and makes our results as a collaboration more robust.

We expect the 35 candidates with $p(\text{astro}) > 0.50$ to be mostly consistent with signals, however we can't rule out the possibility that some of these in fact originated from noise. We can estimate the purity of this list, i.e. the fraction of signals which are astrophysical in origin, by summing the $p(\text{astro})$ across all candidates and comparing to the total number of candidates. Although we use a threshold on $p(\text{astro})$ of only 0.50, many of our candidates are identified with $p(\text{astro}) \sim 1$, so the naive guess that 50% of the candidates originate from noise is not true. We find a purity of $\sim 85 - 90\%$ – or rather, we expect a contamination of triggers with terrestrial origin of $\sim 10 - 15\%$ (approximately four to six candidates) [3]. Of course we cannot a priori know which candidates these are, but the signals closest to the $p(\text{astro})$ threshold are the most likely to be terrestrial in origin.

4.4.1 Binary black hole candidates

Most of the 35 candidates in Table 4.2 are consistent with a BBH source, i.e. they were identified as having both component masses $> 3.0 M_{\odot}$. In this section I will give an overview of a few of the notable BBH candidates [3].

The GstLAL search pipeline was the only pipeline in O3 which allowed for the detection of candidates identified in a single detector. In O3b there were 2 single detector candidates identified only by GstLAL – GW200112_155838 and GW200302_015811, indicated with a † on the FAR in Table 4.2 [3]. These candidates were both initially identified in low-latency and had relatively loud SNRs > 10.0. In the offline re-analysis they were found with high $p(\text{astro}) > 0.9$.

There were 4 candidates identified only by the MBTA search pipeline above the $p(\text{astro})$ threshold. All of these candidates were identified with high FARs which would not have met the public alert threshold in O3b and thus were published in GWTC-3 for the first time [3].

Name	Inst.	cWB			GstLAL			MBTA			PyCBC-broad			PyCBC-BBH		
		FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}
GW191103_012549	HL	–	–	–	–	–	–	*27*	*9.0*	*0.13*	4.8	9.3	0.77	0.46	9.3	0.94
GW191105_143521	HLV	–	–	–	*24*	*10.0*	*0.07*	0.14	10.7	> 0.99	0.012	9.8	> 0.99	0.036	9.8	> 0.99
GW191109_010717	HL	< 0.0011	15.6	> 0.99	0.0010	15.8	> 0.99	1.8×10^{-4}	15.2	> 0.99	0.096	13.2	> 0.99	0.047	14.4	> 0.99
GW191113_071753	HLV	–	–	–	–	–	–	26	9.2	0.68	*1.1×10^{4}*	*8.3*	*< 0.01*	*1.2×10^{3}*	*8.5*	*< 0.01*
GW191126_115259	HL	–	–	–	*80*	*8.7*	*0.02*	*59*	*8.5*	*0.30*	*22*	*8.5*	*0.39*	3.2	8.5	0.70
GW191127_050227	HLV	–	–	–	*0.25*	*10.3*	*0.49*	1.2	9.8	0.73	*20*	*9.5*	*0.47*	4.1	8.7	0.74
GW191129_134029	HL	–	–	–	$< 1.0 \times 10^{-5}$	13.3	> 0.99	0.013	12.7	> 0.99	$< 2.6 \times 10^{-5}$	12.9	> 0.99	$< 2.4 \times 10^{-5}$	12.9	> 0.99
GW191204_110529	HL	–	–	–	*21*	*9.0*	*0.07*	*1.3×10^{4}*	*8.1*	*< 0.01*	*980*	*8.9*	*< 0.01*	3.3	8.9	0.74
GW191204_171526	HL	$< 8.7 \times 10^{-4}$	17.1	> 0.99	$< 1.0 \times 10^{-5}$	15.6	> 0.99	$< 1.0 \times 10^{-5}$	17.1	> 0.99	$< 1.4 \times 10^{-5}$	16.9	> 0.99	$< 1.2 \times 10^{-5}$	16.9	> 0.99
GW191215_223052	HLV	0.12	9.8	0.95	$< 1.0 \times 10^{-5}$	10.9	> 0.99	0.22	10.8	> 0.99	0.0016	10.3	> 0.99	0.28	10.2	> 0.99
GW191216_213338	HV	–	–	–	$< 1.0 \times 10^{-5}$	18.6	> 0.99	9.3×10^{-4}	17.9	> 0.99	0.0019	18.3	> 0.99	7.6×10^{-4}	18.3	> 0.99
GW191219_163120	HLV	–	–	–	–	–	–	–	–	–	4.0	8.9	0.82	–	–	–
GW191222_033537	HL	$< 8.9 \times 10^{-4}$	11.1	> 0.99	$< 1.0 \times 10^{-5}$	12.0	> 0.99	0.0099	10.8	> 0.99	0.0021	11.5	> 0.99	9.8×10^{-5}	11.5	> 0.99
GW191230_180458	HLV	0.050	10.3	0.95	0.13	10.3	0.87	*8.1*	*9.8*	*0.40*	*52*	*9.6*	*0.29*	0.42	9.9	0.96
GW200112_155838	LV	–	–	–	$< 1.0 \times 10^{-5}$†	17.6	> 0.99	–	–	–	–	–	–	–	–	–
GW200115_042309	HLV	–	–	–	$< 1.0 \times 10^{-5}$	11.5	> 0.99	0.0055	11.2	> 0.99	$< 1.2 \times 10^{-4}$	10.8	> 0.99	–	–	–
GW200128_022011	HL	1.3	8.8	0.63	0.022	10.1	0.97	3.3	9.4	0.98	0.63	9.8	0.95	0.0043	9.9	> 0.99
GW200129_065458	HLV	–	–	–	$< 1.0 \times 10^{-5}$	26.5	> 0.99	–	–	–	$< 2.3 \times 10^{-5}$	16.3	> 0.99	$< 1.7 \times 10^{-5}$	16.2	> 0.99
GW200202_154313	HLV	–	–	–	$< 1.0 \times 10^{-5}$	11.3	> 0.99	–	–	–	–	–	–	0.025	10.8	> 0.99
GW200208_130117	HLV	–	–	–	0.0096	10.7	0.99	0.46	10.4	> 0.99	0.18	9.6	0.98	3.1×10^{-4}	10.8	> 0.99
GW200208_222617	HLV	–	–	–	*160*	*8.2*	*< 0.01*	*420*	*8.9*	*0.02*	–	–	–	4.8	7.9	0.70
GW200209_085452	HLV	–	–	–	0.046	10.0	0.95	12	9.7	0.97	*550*	*9.2*	*0.04*	1.2	9.2	0.89
GW200210_092254	HLV	–	–	–	*1.2*	*9.5*	*0.42*	–	–	–	17	8.9	0.53	7.7	8.9	0.54
GW200216_220804	HLV	–	–	–	0.35	9.4	0.77	*2.4×10^{3}*	*8.8*	*0.02*	*970*	*9.0*	*< 0.01*	7.8	8.7	0.54
GW200219_094415	HLV	0.77	9.7	0.85	9.9×10^{-4}	10.7	> 0.99	0.18	10.6	> 0.99	1.7	9.9	0.89	0.016	10.0	> 0.99
GW200220_061928	HLV	–	–	–	–	–	–	–	–	–	–	–	–	6.8	7.5	0.62
GW200220_124850	HL	–	–	–	*150*	*8.2*	*< 0.01*	1.8×10^{3}	8.2	0.83	–	–	–	*30*	*7.8*	*0.20*
GW200224_222234	HLV	$< 8.8 \times 10^{-4}$	18.8	> 0.99	$< 1.0 \times 10^{-5}$	18.9	> 0.99	$< 1.0 \times 10^{-5}$	19.0	> 0.99	$< 8.2 \times 10^{-5}$	19.2	> 0.99	$< 7.7 \times 10^{-5}$	18.6	> 0.99
GW200225_060421	HL	$< 8.8 \times 10^{-4}$	13.1	> 0.99	0.079	12.9	0.93	0.0049	12.5	> 0.99	$< 1.1 \times 10^{-5}$	12.3	> 0.99	4.1×10^{-5}	12.3	> 0.99
GW200302_015811	HV	–	–	–	0.11†	10.6	0.91	–	–	–	–	–	–	–	–	–
GW200306_093714	HL	–	–	–	–	–	–	410	8.5	0.81	*3.4×10^{3}*	*7.8*	*< 0.01*	*24*	*8.0*	*0.24*
GW200308_173609	HLV	–	–	–	*680*	*8.1*	*< 0.01*	*6.9×10^{4}*	*8.3*	*0.24*	*770*	*7.9*	*< 0.01*	2.4	8.0	0.86
GW200311_115853	HLV	$< 8.2 \times 10^{-4}$	16.2	> 0.99	$< 1.0 \times 10^{-5}$	17.7	> 0.99	$< 1.0 \times 10^{-5}$	16.5	> 0.99	$< 6.9 \times 10^{-5}$	17.0	> 0.99	$< 7.7 \times 10^{-5}$	17.4	> 0.99
GW200316_215756	HLV	–	–	–	$< 1.0 \times 10^{-5}$	10.1	> 0.99	*12*	*9.5*	*0.30*	0.20	9.3	0.98	0.58	9.3	0.98
GW200322_091133	HLV	–	–	–	–	–	–	450	9.0	0.62	*1.4×10^{3}*	*8.0*	*< 0.01*	*140*	*7.7*	*0.08*

Table 4.2. Candidate GW signals. The time (UTC) of the signal is encoded in the name as GWYYMMDD_hhmmss (e.g., GW200112_155838 occurred on 2020-01-12 at 15:58:38). The names of candidates not previously reported are given in **bold**. The detectors observing at the merger time of the candidate are indicated using single-letter identifiers (e.g., H for LIGO Hanford); these are not necessarily the same detectors that contributed triggers associated with the candidate. Where a candidate was found with p(astro) above the threshold value of 0.50 by at least one analysis but below the threshold by others, we include in *italics* the results from the other analyses, where available. A dash (–) indicates that a candidate was not found by an analysis. The 2 candidates labeled with a dagger (†) were found only above threshold in a single detector with the GstLAL analysis, and the FAR estimates were made using significant extrapolation of the background data, meaning that single-detector candidates have higher uncertainty than coincident candidates. A conservative estimate of the FAR for these single-detector candidates is one per live time of the analysis; this is $\sim 3.16\ \mathrm{yr}^{-1}$ for both LIGO Hanford and LIGO Livingston. Reprinted from [3].

There were 5 BBH candidates identified only by the PyCBC-BBH analysis as well as 2 identified in both PyCBC analysis configurations. Among these, GW191103_012549 was the most significant, identified with a FAR of 0.46 per year. None of these candidates were identified in low-latency and thus all are reported in GWTC-3 for the first time [3].

GW200208_222617 and GW200220_061928 were two candidates identified above the p(astro) threshold only by the PyCBC-BBH analysis. These each were identified with very high masses, making them relevant in the search for intermediate mass black hole (IMBH) candidates [216–218]. However, both of these candidates were identified with low SNRs ~ 8.0, high FARs of $\sim$ several per year, and modest p(astro) of 0.70 and 0.62, respectively. Neither candidate was identified as significant in the dedicated IMBH search [216].

The candidate GW191109_010717 was found with high significance in all search pipelines in the GWTC-3 analyses as well as the dedicated IMBH search [216]. While it was identified with high component masses, we cannot confirm that its masses lie in the range above which black holes are expected to form from stellar collapse [168, 217]

4.4.2 Neutron star–black hole binary candidates

The new candidates in O3b included the first confident detections of NSBH binary mergers. After the first detection of a BBH merger in O1 and a BNS merger in O2 the detection of an NSBH completed observations of the three expected CBC source classes. GW200115_042309 was found with high significance in all three of the modeled searches: GstLAL, MBTA, and PyCBC-broad. This candidate was first initially published in [115] along with GW200105_162426 which is included in GWTC-3 as a marginal candidate [3]. GW200105_162426 will be discussed in more detail in Sec. 4.4.3. Although GW200115_042309 was identified in low-latency, no EM counterparts were identified. This is not unexpected for NSBH signals as the presence of a counterpart will depend upon the source intrinsic parameters, namely the mass ratio and component spins [219–222]. The particular source parameters of this candidate made it unlikely to have remnant mass outside of the final compact object to produce an EM counterpart.

GW191219_163120 was a candidate identified only by the PyCBC-broad analysis [3]. The template which produced the trigger had secondary mass $1.84 M_\odot$ which makes it consistent with an NSBH origin. However, due to technical details of the PyCBC p(astro) calculation, its significance was estimated with respect to the BBH population. This resulted in the candidate having an artificially higher p(astro) $= 0.82$. If it had been ranked as part of the NSBH population its p(astro) would have dropped to only 0.085,

Name	Inst.	cWB			GstLAL			MBTA			PyCBC-broad		
		FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}	FAR (yr^{-1})	SNR	p_{astro}
GW191118_212859	LV	–	–	–	–	–	–	*7.4×10^5*	*$8.0 < 0.01$*		1.3	9.1	0.05
GW200105_162426	LV	–	–	–	0.20†	13.9	0.36	–	–	–	–	–	–
200121_031748*	HV	–	–	–	*58*	*9.1*	*0.02*	1.1	10.7	0.23	–	–	–
GW200201_203549	HLV	–	–	–	1.4	9.0	0.12	*850*	*$8.9 < 0.01$*	*1.0×10^3*	*$8.3 < 0.01$*		
200214_224526*	HLV	0.13	13.1	0.91	–	–	–	–	–	–	–	–	–
200219_201407*	HLV	–	–	–	–	–	–	0.22	13.6	0.48	–	–	–
GW200311_103121	HL	–	–	–	*110*	*$9.0 < 0.01$*		1.3	9.0	0.03	1.3	9.2	0.19

Table 4.3. Marginal candidates found by the various analyses. The candidates in this table have a FAR below a threshold of 2.0 yr^{-1} in at least one analysis, but were not found with p(astro) that meets our threshold for Table 4.2 (p(astro) > 0.50 from a search analysis, with the additional requirement that cWB candidates have a counterpart from a matched-filter analysis). Detector-identifying letters are the same as given in Table 4.2. The instruments for each candidate are the ones which were operating at the time of the trigger, and are not necessarily the same as those which participated in the detection. The candidates are named according to the same convention as in Table 4.2 except that here we omit the GW prefix for the candidates found to be likely caused by instrumental artifacts, indicated with an asterisk ($*$). Where a candidate was seen below the FAR threshold in at least one analysis but above threshold in others, we include in *italics* the information on that trigger from the other analyses as well where available. As in Table 4.2, the dagger (†) indicates a candidate found by a single detector with the GstLAL analysis. Reprinted from [3].

excluding it from Table 4.2. Future studies may shed more light on the true nature of this event, but the example highlights the sensitivity of the p(astro) value to the population model used to estimate the foreground distribution.

Finally, GW200210_092254 was a possible NSBH candidate identified above the p(astro) threshold only by the PyCBC analyses. This candidate had a modest significance with FAR nearly 10 per year in both PyCBC analysis configurations as well as p(astro) close to the threshold for inclusion [3]. The templates which produced this trigger in the PyCBC-broad and PyCBC-BBH analyses both had secondary mass $5.61M_{\odot}$. If astrophysical, we cannot confidently determine whether the signal originated from an NSBH binary with a heavy neutron star or a BBH with a light secondary in the supposed lower mass gap, extending from $3 - 5M_{\odot}$ [166, 167].

4.4.3 Marginal candidates

There were 7 candidates identified in the GWTC-3 analyses which had p(astro) below the threshold for inclusion in Table 4.2 but still had a significant FAR below 2.0 per year [3].

These candidates are referred to as *marginal* and listed in Table 4.3.

Most notably, the marginal candidates in GWTC-3 include GW200105_162426 which is an NSBH candidate identified by GstLAL in LIGO Livingston data and previously published in [115]. In the offline analysis, this candidate was identified with a high SNR of 13.9 making its signal clearly visible in the LIGO Livingston data. The SNR peak was similarly identified in low-latency by MBTAOnline, PyCBC Live, and SPIIR however none of these searches were configured to assign FARs to single detector candidates. Similarly, the MBTA and PyCBC offline analyses identified SNR triggers but do not allow candidates without coincidence, hence they are not included in Table 4.3.

For single detector candidates, we can place a conservative upper bound on the FAR of one per observing time. However, in practice we obtain a stronger bound, albeit with added uncertainty, by extrapolating the background population to a longer observing duration [115]. Despite this uncertainty, the trigger clearly stands out from the background compared to noise triggers collected over the duration of O3b. This is demonstrated by the right panel of Fig. 4.4, which shows the $\rho - \xi^2$ background histogram of noise triggers with similar intrinsic parameters to the template producing GW200105_162426, as described in Chapter 2. The SNR and ξ^2 of GW200115_042309 are also shown, indicating that without coincidence this trigger would not have been separable from the background. In contrast, the SNR and ξ^2 of GW200105_162426 are unlike any noise trigger obtained throughout the observing duration. For this reason, [115] concludes that GW200105_162426 is very likely astrophysical.

The uncertainty in the FAR estimation due to extrapolation propagates into the $p(\text{astro})$ calculation as uncertainty in the background distribution. There is further uncertainty in the foreground distribution due to the relative lack of NSBH signals observed to date [3]. Therefore, the low $p(\text{astro})$ estimated for GW200105_162426 in GWTC-3 is not necessarily unexpected. In the future, once we have made more observations of NSBHs we can better constrain the population model and merger rate of similar events. At this time, the significance of GW200105_162426 may be re-evaluated and its true nature determined with less uncertainty.

The three candidates in Table 4.3 indicated with an asterisk, $(*)$, were determined by data quality checks to be likely instrumental in origin. These triggers were therefore not assigned a GW name to reflect our doubt that they are astrophysical in origin [3]. 200121_031748 occurred at a time when LIGO Hanford data was found to contain excess power consistent with a blip glitch [188]. 200214_224526 was found by the cWB analysis in both LIGO detectors, however both detectors were found to have excess power at

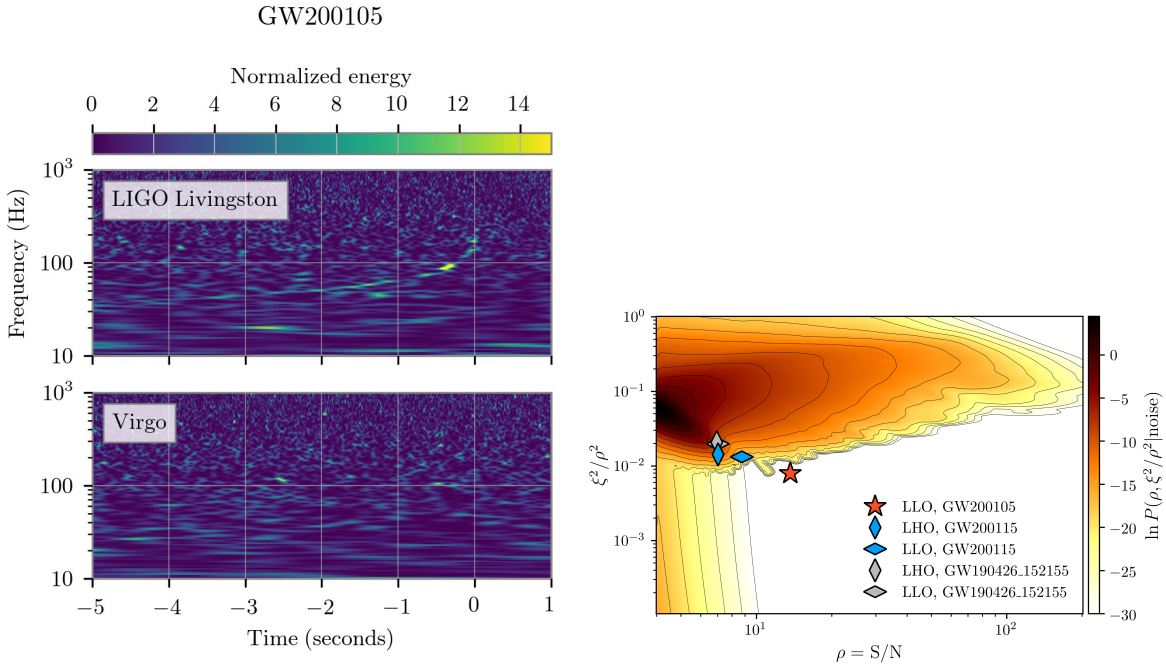

Figure 4.4. **Left**. Time-frequency representation of the LIGO Livingston and Virgo data in a six second window around the coalescence time of GW200105_162426. The chirping signal track is clearly visible in the LIGO Livingston data due to its high SNR of 13.9. **Right**. The probability density function of noise triggers collected by GstLAL from templates with $\mathcal{M}_c < 4M_\odot$ collected over the full duration of O3, combined over the LIGO Hanford, Livingston, and Virgo detectors. The SNR and ξ^2 values of GW200115 and GW190426 are shown for comparison by the blue and grey diamonds respectively. The single SNR and ξ^2 value of GW200105 in LIGO Livingston data is indicated by the red star, where it is clear that it stands out from the rest of the background distribution. Reprinted from [115].

the time consistent with a weak scattering arch and fast scattering in LIGO Hanford and LIGO Livingston, respectively [178]. Additionally, even without the identification of these instrumental artifacts, this candidate would not have met the criteria for inclusion in Table 4.2 as it was not found in any of the three modeled searches. Finally, 200219K was determined to be instrumental in origin as there were several loud glitches in the LIGO Hanford data near the time of the signal.

4.5 Search sensitivity

To quantify the search sensitivity, we use the sensitive time-volume, $\langle VT \rangle$, over the duration of O3b computed using *injection campaigns* [3]. We simulate signals, referred to as injections, in each of the CBC source classes we are interested in: BNS, NSBH,

	Mass distribution	Mass range ($M_\odot$)	Spin range	Redshift evolution	Maximum redshift
BBH	$p(m_1) \propto m_1^{-2.35}$	$2 < m_1 < 100$	$\lvert\chi_{1,2}\rvert < 0.998$	$\kappa = 1$	1.9
	$p(m_2\lvert m_1) \propto m_2$	$2 < m_2 < 100$			
NSBH	$p(m_1) \propto m_1^{-2.35}$	$2.5 < m_1 < 60$	$\lvert\chi_1\rvert < 0.998$	$\kappa = 0$	0.25
	Uniform	$1 < m_2 < 2.5$	$\lvert\chi_2\rvert < 0.4$		
BNS	Uniform	$1 < m_1 < 2.5$	$\lvert\chi_{1,2}\rvert < 0.4$	$\kappa = 0$	0.15
	Uniform	$1 < m_2 < 2.5$			

Table 4.4. Parameter distributions used to generate injections for estimating search sensitivity. We always use the convention that $m_1 \geq m_2$ with masses defined in the source frame. The redshift-evolution parameter κ controls the injected distribution as described in Eqn. (4.3).

and BBH mergers. The injections are added to the strain data and distributed uniformly in time throughout the duration of the run. The injections are drawn from a uniform distribution in comoving volume, V_c, according to a flat Λ-cold-dark-matter (ΛCDM) cosmology. The probability distribution of the injection redshifts goes as,

$$p(z) \propto \frac{dV_c}{dz}(1+z)^{\kappa-1}, \tag{4.3}$$

where the value of κ is specified in Table 4.4 for each source class. The injection mass and spin distributions were constructed in each of the known CBC source classes. The injection spins are drawn from isotropic distributions with $|s_{i,z}| < 0.998$ for components with mass $> 2.0M_\odot$ and 0.4 otherwise. These distributions are outlined in Table 4.4, with masses always referring to the source frame and assuming $m_1 \geq m_2$. The BNS injections are generated using the `SpinTaylorT4` waveform model [203], while the BBH and NSBH injections are generated using the `SEOBNRv4PHM` model [223], or the `SEOBNRv4P` model [201, 223] for injections corresponding to binaries with redshifted total mass below $9M_\odot$.

By counting the number of injections, N, recovered by each search pipeline above the p(astro) threshold of 0.50, we can compute the sensitive time-volume as:

$$\langle VT \rangle = \frac{N}{R}, \tag{4.4}$$

where R is the rate of injected signals per unit volume and time. The time is taken as the livetime of each analysis. For cWB, the livetime was 94.9 days – this is a subset

Binary masses ($M_\odot$)			Sensitive hypervolume (Gpc^3 yr)					
m_1	m_2	$\mathcal{M}$	cWB	GstLAL	MBTA	PyCBC-broad	PyCBC-BBH	Any
35.0	35.0	30.5	$2.6^{+0.1}_{-0.1}$	$4.1^{+0.1}_{-0.1}$	$3.3^{+0.2}_{-0.1}$	$3.3^{+0.1}_{-0.1}$	$4.3^{+0.2}_{-0.1}$	$5.3^{+0.1}_{-0.2}$
35.0	20.0	22.9	$1.35^{+0.09}_{-0.10}$	$2.3^{+0.2}_{-0.1}$	$1.8^{+0.1}_{-0.1}$	$1.9^{+0.1}_{-0.1}$	$2.5^{+0.1}_{-0.1}$	$3.1^{+0.1}_{-0.2}$
35.0	1.5	5.2	–	$1.8^{+0.2}_{-0.3} \times 10^{-2}$	$1.9^{+0.3}_{-0.3} \times 10^{-2}$	$3.1^{+0.3}_{-0.3} \times 10^{-2}$	–	$3.3^{+0.4}_{-0.3} \times 10^{-2}$
20.0	20.0	17.4	$0.56^{+0.04}_{-0.04}$	$1.34^{+0.06}_{-0.05}$	$1.10^{+0.05}_{-0.05}$	$1.14^{+0.05}_{-0.05}$	$1.42^{+0.06}_{-0.05}$	$1.71^{+0.06}_{-0.07}$
20.0	10.0	12.2	$0.24^{+0.03}_{-0.04}$	$0.60^{+0.05}_{-0.05}$	$0.51^{+0.05}_{-0.04}$	$0.56^{+0.05}_{-0.05}$	$0.65^{+0.05}_{-0.05}$	$0.77^{+0.06}_{-0.06}$
20.0	1.5	4.2	–	$1.9^{+0.2}_{-0.2} \times 10^{-2}$	$1.9^{+0.2}_{-0.2} \times 10^{-2}$	$2.7^{+0.2}_{-0.2} \times 10^{-2}$	–	$2.9^{+0.3}_{-0.2} \times 10^{-2}$
10.0	10.0	8.7	$6.8^{+0.8}_{-0.9} \times 10^{-2}$	$0.26^{+0.01}_{-0.02}$	$0.26^{+0.01}_{-0.02}$	$0.27^{+0.01}_{-0.02}$	$0.28^{+0.02}_{-0.02}$	$0.32^{+0.02}_{-0.01}$
10.0	5.0	6.1	$1.3^{+0.5}_{-0.4} \times 10^{-2}$	$0.10^{+0.02}_{-0.01}$	$0.10^{+0.02}_{-0.01}$	$0.12^{+0.01}_{-0.02}$	$0.11^{+0.02}_{-0.01}$	$0.13^{+0.02}_{-0.01}$
10.0	1.5	3.1	–	$1.6^{+0.1}_{-0.1} \times 10^{-2}$	$1.5^{+0.2}_{-0.1} \times 10^{-2}$	$1.8^{+0.1}_{-0.1} \times 10^{-2}$	–	$2.1^{+0.1}_{-0.1} \times 10^{-2}$
5.0	5.0	4.4	$5^{+1}_{-2} \times 10^{-3}$	$5.8^{+0.5}_{-0.4} \times 10^{-2}$	$4.5^{+0.4}_{-0.4} \times 10^{-2}$	$6.5^{+0.5}_{-0.4} \times 10^{-2}$	$5.0^{+0.5}_{-0.4} \times 10^{-2}$	$7.4^{+0.5}_{-0.5} \times 10^{-2}$
5.0	1.5	2.3	–	$1.12^{+0.05}_{-0.06} \times 10^{-2}$	$1.19^{+0.06}_{-0.05} \times 10^{-2}$	$1.21^{+0.06}_{-0.06} \times 10^{-2}$	–	$1.43^{+0.06}_{-0.06} \times 10^{-2}$
1.5	1.5	1.3	–	$2.7^{+0.1}_{-0.1} \times 10^{-3}$	$3.4^{+0.1}_{-0.1} \times 10^{-3}$	$3.5^{+0.1}_{-0.2} \times 10^{-3}$	–	$3.9^{+0.1}_{-0.2} \times 10^{-3}$

Table 4.5. Sensitive hypervolume from O3b for the various search analyses with $p_{\text{astro}} > 0.50$ at the assessed points in the mass parameter space. The *Any* results come from calculating the sensitive hypervolume for injections found by at least one search analysis. For each set of binary masses, the given values are the central points of a log-normal distribution with width 0.1. For some regions and analyses, few injections were recovered such that the sensitive hypervolume cannot be accurately estimated; these cases are indicated by a dash (–). As an example of this, the PyCBC-BBH and cWB analyses analyzed only injections in the designated BBH set, and so no injections were found in the BNS or NSBH regions. Reprinted from [3].

of the times when strictly two detectors were observing [3]. The MBTA and PyCBC livetimes were 124.5 days and 124.2 days, respectively – these are each subsets of the amount of time when at least two detectors were observing [3]. The GstLAL livetime was 142.0 days, which is the amount of time when at least one detector was observing [3].

The injected distributions indicated in Table 4.4 do not represent the same astrophysical assumptions chosen for construction of the foreground in the p(astro) calculation. Further, each search pipeline computes p(astro) independently and the foreground models used are not consistent among pipelines. For the precise population models used by each search, see Appendix D7 of [3]. For the cWB and PyCBC-BBH analyses, we exclude injections outside of the BBH region of parameter space, as indicated by the (–) entries in Table 4.5. Further, to count an injection as found by cWB we require that it also be found with $p(\text{astro}) \geq 0.1$ in at least one of the modeled searches. This requirement has a negligible effect on the reported sensitivity of cWB.

Rather than computing the $\langle VT \rangle$ across the entire parameter space of injected signals in Table 4.4, we compute it independently in several sub-regions across the space. The sensitivity estimates in Table 4.5 are computed using injections drawn from log-normal distributions with central masses $m_i \in \{1.5, 5.0, 10, 20, 35\} M_\odot$, indicated by the first two

columns in the table. Each distribution has a width of 0.1. These distributions are chosen to be representative of the signals we have observed in the past. The results in Table 4.5 demonstrate the variation in sensitivity across the parameter space by each search [3]. As is expected based on previous observations, we find that each search is most sensitive in the $35-35M_{\odot}$ mass bin, representing GW150914-like signals [12]. The *Any* column in the table combines results by counting injections which are found in at least one search pipeline. This demonstrates how our overall sensitivity as a collaboration improves by including multiple search pipelines.

4.6 Source parameters

After the initial detection of GW candidates by search pipelines, there are detailed parameter estimation analyses based on full Bayesian inference to identify the source properties and estimate their uncertainties [224–227]. The Bayesian inference is founded on the well-known Bayes theorem [228, 229],

$$p(\vec{\theta}|\vec{d}) = \frac{p(\vec{d}|\vec{\theta})p(\vec{\theta})}{p(\vec{d})}. \tag{4.5}$$

Here, $\vec{\theta}$ is the vector of unknown source parameters and $\vec{d}$ are the time-series data observed from each participating detector. $p(\vec{\theta}|\vec{d})$ are the posterior distributions on the unknown source parameters (for example masses, spins, sky position and distance). Therefore, these analyses estimate uncertainty on the inferred source parameters as opposed to the point estimates produced by searches. $p(\vec{d}|\vec{\theta})$ is the likelihood of observing the data given that a signal with parameters $\vec{\theta}$ is present. Finally, $p(\vec{\theta})$ is the prior distribution on the source parameters and $p(\vec{d})$ is obtained by marginalizing the likelihood over the prior and is typically regarded as a normalization constant. These analyses can be complicated by uncertainty on systematic errors in the strain data calibration and waveform models used, as well as the high dimension of the parameter space to search. In this section, I will only focus on the masses, spins, and sky localizations, but see Section V of [3] and the associated data release for the full results [170].

In parameter estimation, the choice of prior distributions can have a significant effect on the inferred posterior distributions. Especially for parameters which have sub-dominant effects on the waveform evolution, the results can rely heavily upon the prior. For this reason, the GWTC-3 analysis used broad or uniform priors which are meant to be uninformative and reduce assumptions about the underlying astrophysics [3].

Using priors which more closely resemble our previous observations could produce a bias against any unexpected or diverse results.

These analyses typically use different waveform models than those used by the search pipelines to initially identify signals. For BBH signals, i.e. those for which both components have mass $> 3M_{\odot}$, the `IMRPhenomXPHM` and `SEOBNRv4PHM` waveform models are used [223, 230]. These models allow for the effects of spin precession and higher modes. For potential NSBH signals, the analyses use waveform models which allow for matter effects – `IMRPhenomNSBH` and `SEOBNRv4_ROM_NRTidalv2_NSBH` [231, 232].

4.6.1 Masses

The source masses are typically the best constrained parameters from GW observations due to their strong imprint on the waveform evolution. Particularly, the combination of component masses, m_1 and m_2, known as the chirp mass,

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m1 + m2)^{1/5}}, \tag{4.6}$$

has a strong effect on the frequency evolution in the inspiral phase of the waveform since it enters at leading PN order [233, 234]. Therefore, especially for low mass systems where the inspiral portion of the waveform lasts for many cycles in the LIGO sensitive frequency band, we can infer the $\mathcal{M}_c$ with high accuracy. The total mass, $M = m_1 + m_2$, mainly effects the merger and ringdown phase of the waveform, making it better measured for high mass, short duration signals where the merger makes up a larger proportion of the overall waveform in LIGO's bandwidth. Finally, the mass ratio,

$$q = \frac{m_1}{m_2},\ m_1 \geq m_2 \tag{4.7}$$

effects the phase of the inspiral, but only enters at the first PN order, that is after the $\mathcal{M}_c$ [233, 234]. Therefore the mass ratio has a sub-dominant effect and is usually not constrained as well as the chirp mass. The credible region contours for the inferred component masses, m_1 and m_2 as well as total mass M and mass ratio, $q = m_2/m_1$ are shown in Fig. 4.5.

I will highlight a few of the candidates with interesting source masses below, starting with unambiguous BBHs (i.e. candidates for which both component masses were inferred to be $> 3M_{\odot}$ at 99% credibility) followed by signals which may be consistent with an NSBH origin. The threshold $3M_{\odot}$ is chosen as a conservative divider between black hole

and neutron star component masses. This does not necessarily represent any statement on the assumed neutron star maximum mass, which is currently unknown but typically thought to lie in the range $2.1 - 2.7 M_\odot$ [96, 235–237].

Unambiguous BBHs. Overall, the BH mass distributions inferred for the new candidates added to GWTC-3 from O3b are consistent with previous observations from GWTC-2.1. Unless stated otherwise, the masses quoted here refer to the 90% symmetric credible region (see Table IV of [3]).

The highest mass binary was probably GW200220_061928, with total mass $M = 148^{+55}_{-33} M_\odot$ and $\mathcal{M}_c = 62^{+23}_{-15} M_\odot$ [3]. Each of its component masses have significant probability of lying within the so-called *upper mass gap*, a region of the stellar mass parameter space where pair-instability supernovae results in the death of a star leaving no remnant compact object [168]. This corresponds to an inferred gap in the BH mass spectrum from about $65 - 120 M_\odot$. Observation of black holes in this mass region would indicate formation via alternative channels, for example hierarchical mergers [238–240]. The primary mass of GW200220_061928's source is inferred to be $87^{+40}_{-23} M_\odot$ and the secondary mass, $61^{+26}_{-25} M_\odot$. The parameter estimation analyses identified probabilities of 94% and 39% that the primary and secondary mass components were $> 65 M_\odot$, respectively. There was a 7% probability that GW200220_061928's source had its primary mass above the mass gap, i.e. $m_1 > 120 M_\odot$. The remnant BH formed by the merger of GW200220_061928's source objects was inferred as $141^{+51}_{-31} M_\odot$, with a 99% probability of lying above $100 M_\odot$, the threshold for inclusion as an IMBH.

At the other end of the black hole mass spectrum, there is a so-called *lower mass gap* from $3 - 5 M_\odot$ inferred from a lack of x-ray binary observations [166, 167]. Among the unambiguous BBH candidates in O3b, GW191113_071753's source has the highest probability of having a component in this region with secondary mass inferred to be $5.9^{+4.4}_{-1.3} M_\odot$ [3]. There is a 13% probability that GW191113_071753's source had its secondary mass $< 5 M_\odot$.

Similarly to past observations, most of the BBH candidates identified in O3b are consistent with having equal masses, i.e. $q \sim 1$. Probably the candidates with the most extreme mass ratios among these are GW191113_071753 and GW200208_222617 with $q = 0.202^{+0.490}_{-0.087}$ and $q = 0.21^{+0.67}_{-0.16}$, respectively [3]. This is consistent with the O3a candidate GW190412 which was the first detection of a BBH with confidently unequal masses at a mass ratio of $q = 0.28^{+0.12}_{-0.06}$ [138, 165, 241].

Potential NSBHs. GWTC-3 includes the first confident detections of gravitational waves from NSBH sources. These candidates are all consistent with having the secondary

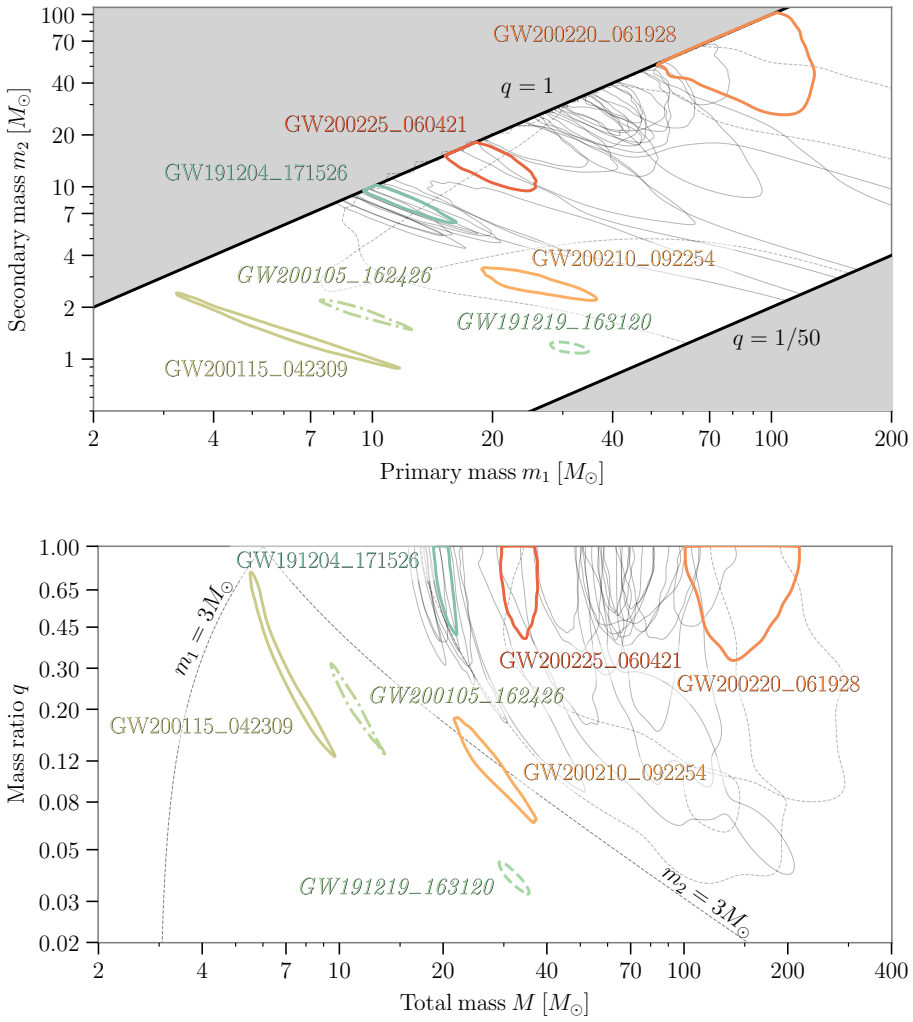

Figure 4.5. Credible-region contours for the inferred masses of the O3b candidates with $p(\text{astro}) > 0.50$ plus GW200105_162426. **Top**: Results for the primary and secondary component masses m_1 and m_2. The shaded areas indicate regions excluded by the convention $m_1 \geq m_2$, and by the most extreme mass ratio considered in our analyses, i.e. $q = 0.05$. **Bottom**: Results for total mass M and mass ratio q. Each contour represents the 90% credible region for a different candidate. Highlighted contours are for the NSBH candidates GW191219_163120, GW200105_162426 and GW200115_042309; the NSBH or low-mass BBH candidate GW200210_092254; GW191204_171526, which has inferred $\chi_{\text{eff}} > 0$; GW200225_060421, which has 85% probability that $\chi_{\text{eff}} < 0$, and GW200220_061928, which probably has the most massive source of the O3b candidates. We highlight with italics GW200105_162426 as it has $p(\text{astro}) < 0.50$, as well as GW191219_163120 because of significant uncertainty in its $p(\text{astro})$ and because it has significant posterior support outside of mass ratios where the waveform models have been calibrated. Results for GW200308_173609 and GW200322_091133 are indicated with dashed lines to highlight that these include a low-likelihood mode at large distances and high masses, and are particularly prior sensitive. The dotted lines delineate regions where the primary and secondary can have a mass below $3M_\odot$. For the region above the $m_2 = 3M_\odot$ line, both objects in the binary have masses above $3M_\odot$. Reprinted from [3].

mass $m_2 < 3M_\odot$.

Among these, GW200115_042309 was the most significant. GW200115_042309's source was probably the lowest mass binary among the O3b candidates. The primary mass, inferred as $5.9^{+2.0}_{-2.5}M_\odot$ has a 29% probability of lying in the lower mass gap region, i.e. $m_1 < 5M_\odot$ [3]. The secondary component mass, inferred as $1.44^{+0.85}_{-0.28}$, is consistent with the Galactic NS mass distribution [242, 243].

GW200210_092254's source is similar to the previously published GW190814 from GWTC-2 [138,165]. This candidate was inferred to have a secondary mass of $2.83^{+0.47}_{-0.42}M_\odot$ with a 76% probability of having $m_2 < 3M_\odot$ [3]. This means that the secondary mass could be consistent with being a heavy neutron star or a light black hole lying in the lower mass gap. This candidate also has one of the most extreme mass ratios from O3b, with $q \leq 0.150$ at 90% credibility.

If astrophysical, GW191219_163120's source would be the most extreme mass ratio binary in O3b with $q = 0.038^{+0.005}_{-0.004}$ and $q \leq 0.041$ at 90% probability [3]. However, this candidate's p(astro) is subject to significant uncertainty as described in Sec. 4.4. Further, this very extreme mass ratio is outside the region where the waveforms are calibrated and thus may be subject to significant systematic uncertainty.

4.6.2 Spins

The binary component spins are typically not well constrained by GW measurements, due to the high PN order at which they enter the waveform [233, 234]. The spins only effect the inspiral evolution in the mass-weighted combination of spins known as the effective inspiral spin,

$$\chi_{\rm eff} = \frac{m_1\vec{s_1} + m_2\vec{s_2}}{m_1 + m_2}\hat{L}. \tag{4.8}$$

We typically assume that the component spin vectors, $\vec{s_i}$, are aligned or anti-aligned with the orbital angular momentum. Then, defining $\hat{L}$ to be in the $\hat{z}$-direction the definition can be simplified as,

$$\chi_{\rm eff} = \frac{m_1 s_{1,z} + m_2 s_{2,z}}{m_1 + m_2}. \tag{4.9}$$

The $\chi_{\rm eff}$ mainly has an effect on the inspiral duration of the GW waveforms. The secondary component spin, $s_{2,z}$ is typically not well-constrained, however the primary spin, $s_{1,z}$ may be constrained for the most extreme mass ratio binaries as $s_{1,z}$ will dominate the $\chi_{\rm eff}$ in this case.

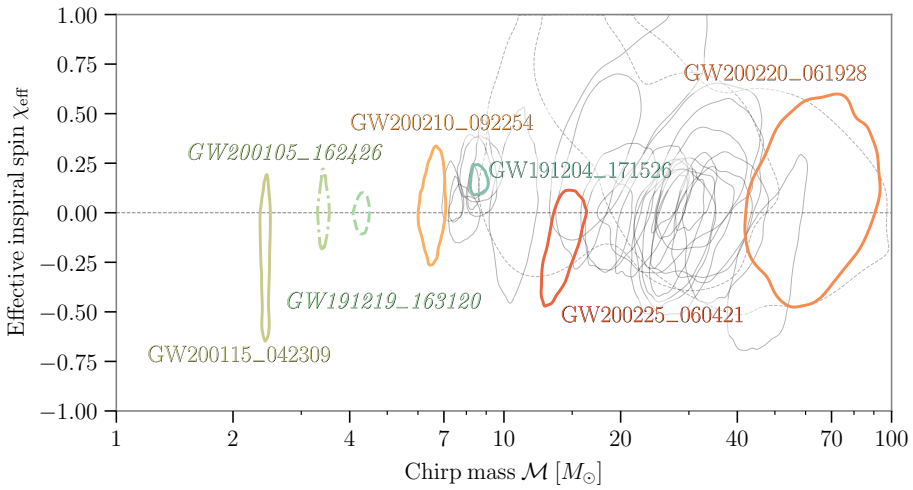

Figure 4.6. Credible-region contours in the plane of chirp mass $\mathcal{M}_c$ and effective inspiral spin $\chi_{\rm eff}$ for O3b candidates with $p(\text{astro}) > 0.50$ plus GW200105_162426. Each contour represents the 90% credible region for a different candidate. Highlighted contours are for the NSBH candidates GW191219_163120, GW200105_162426 and GW200115_042309; the NSBH or low-mass BBH candidate GW200210_092254; GW191204_171526, which has inferred $\chi_{\rm eff} > 0$; GW200225_060421, which has 85% probability that $\chi_{\rm eff} < 0$, and GW200220_061928, which probably has the most massive source of the O3b candidates. We highlight with italics GW200105_162426 as it has $p(\text{astro}) < 0.50$, as well as GW191219_163120 because of significant uncertainty in its $p(\text{astro})$ and because it has significant posterior support outside of mass ratios where the waveform models have been calibrated. Results for GW200308_173609 and GW200322_091133 are indicated with dashed lines to highlight that these include a low-likelihood mode at large distances and high masses, and are particularly prior sensitive.

The in-plane spin components are characterized by the precession spin [244–246],

$$\chi_p = \max\left\{\chi_{1,\perp}, \frac{q(4q+3)}{4+3q}\chi_{2,\perp}\right\}, \tag{4.10}$$

where $\chi_{i,\perp}$ indicates the spin component which is perpendicular to the orbital angular momentum, i.e. the $\hat{x}$ and $\hat{y}$ components of the spin vector. Measurement of the precession spin is strongly dominated by the choice of prior and also correlated to the $\chi_{\rm eff}$ measurement in that the total spin magnitude is limited to $|\vec{s_i}| \leq 1$, thus we only expect to place weak constraints on the χ_p.

Constraining the effective inspiral spin can give insight to the various channels through which binaries formed [161, 247, 248]. The isolated field formation scenario would tend to produce mostly aligned spins [249, 250], while dynamically formed binaries are expected to have isotropically distributed spin orientations [251–253]. The 90% credible region

contours for chirp mass and effective inspiral spin are shown in Fig. 4.6. Most of the spin posteriors have support for $\chi_{\rm eff} = 0$ and, in agreement with GWTC-2.1, there are more binaries with inferred $\chi_{\rm eff} > 0$ than $\chi_{\rm eff} < 0$ [3, 138, 165]. This indicates a possible preference for the isolated field formation channel, as binaries formed dynamically in say, dense cluster environments may be more likely to have mis-aligned spins. Below, I will highlight a few of the candidates with notable spin values, quoting the inferred 90% credible symmetric intervals unless otherwise stated.

The highest $\chi_{\rm eff}$ source is most likely from GW200208_222617, with $\chi_{\rm eff} = 0.48^{+0.42}_{-0.46}$ [3]. This is similar to the highest $\chi_{\rm eff}$ source from GWTC-2.1 [165]. GW191109_010717's source probably has the lowest $\chi_{\rm eff}$ value among O3b candidats at $-0.29^{+0.42}_{-0.31}$ [3].

The inferred χ_p posteriors are largely uninformative, however the inferred values for the sources of GW200129_065458 and GW191219_163120 are worth noting. GW200129_065458 was the highest SNR signal observed in O3b, with $\rho = 26.5$ as measured by the GstLAL search pipeline [3]. This places the signal in a region where waveform systematics become more important relative to statistical uncertainties [254,255]. Therefore, the measurement of χ_p and the mass ratio (which χ_p depends on) are highly dependent on the waveform used – `SEOBNRv4PHM` or `IMRPhenomXPHM`. With the former, the inferred χ_p is $0.36^{+0.31}_{-0.25}$ and with the latter, $0.77^{+0.19}_{-0.44}$ [3]. If astrophysical, GW191219_163120's source would have probably the lowest χ_p, inferred to be ≤ 0.14 at 90% probability [3]. However, again, the extreme mass ratio of this candidate means that it may be subject to additional systematic uncertainty, the extent of which is not currently well understood.

4.6.3 Localization

Our ability to localize a source on the sky, as well as identify its luminosity distance and redshift, is highly dependent on the number of detectors which participated in the observation. The best localized source in O3b was that of GW200311_115853 which was observed in all three detectors – LIGO Hanford, Livingston, and Virgo with SNRs ~ 12, ~ 10, and ~ 7, respectively [3]. Among the candidates detected in all three detectors, this one had the highest SNR in Virgo which may have contributed to its precise localization. The 90% credible sky area was 35deg^2, which is comparable to the best localized sources in GWTC-2.1 [3, 165].

The localization volume depends upon the ability to constrain both the sky area and the luminosity distance. In O3b, the closest observed source was probably the marginal NSBH GW200105_162426 with an inferred luminosity distance of $D_L = 0.27^{+0.12}_{-0.11}$Gpc

and redshift $z = 0.06^{+0.02}_{-0.02}$ [3]. This is similar to the closest event observed in O3a, but still nearly an order of magnitude further away than the closest source ever detected, the BNS merger GW170817 [66, 165]. The furthest observed source in O3b is probably that of GW200308_173609, with an inferred luminosity distance $D_L = 7.1^{+13.9}_{-4.4}$Gpc and redshift $z = 1.04^{+1.47}_{-0.57}$ [3]. The candidates with the smallest localization volume are the NSBH candidate GW200115_042309 and the BBH candidate GW200202_154313 with inferred volumes of 0.0023Gpc^3 and 0.0063Gpc^3, respectively [3]. According to galaxy catalogs there are thousands of galaxies in these volumes [256, 257], so it is not possible to identify the host galaxies of these sources.

4.7 Conclusion

At the time of writing, GWTC-3 represents the most comprehensive catalog of GW candidates in existence. The 35 new candidates added to the catalog in O3b bring the total number of highly significant detections to 90. This is an incredible result in just ten years of observing. The rate of growth in the field of gravitational wave astronomy is truly unprecedented, but even still we are only beginning to scrape the surface of what these observations can teach us about the universe. Continuing to build up populations of GW events across the full CBC parameter space of BNS, NSBH, and BBH sources will allow us to probe more deeply into astrophysical population models, binary formation scenarios, and merger rates. With GW200115_042309, we have now confidently detected GWs from each of the three CBC source classes and unambiguously proven that neutron star-black hole systems do exist [115]. However, significant uncertainty remains in the NSBH mass distribution and merger rates until we make more detections of similar systems. We have already seen that gravitational wave observations allow for advances in fundamental physics such as constraining alternative theories of gravity [258, 259], providing independent measures of the Hubble constant [64], and studying the high redshift universe [162, 260]. These analyses will continue to be improved by including more GW observations. In addition to the high significance candidates presented here, the deep sub-threshold data release will be a boon for broad analysis efforts seeking to glean as much information from the LIGO, Virgo, and KAGRA strain data as possible [170].

The GWTC-3 paper is published with several companions which I will briefly mention here. The new candidates published in GWTC-3 are used to estimate merger rates and source populations in [62] and to perform tests of general relativity in [65]. While the GWTC-3 analysis presented here focused on the so-called *all sky* parameter space,

broadly searching for CBC signals from BNS, NSBH, and BBH systems, dedicated offline searches in specific parameter regions of interest were also conducted. [216] presents a dedicated search for IMBH binaries with component masses $\sim 50-600M_{\odot}$ and [261,262] present searches for sub-solar mass black hole binaries in O3a and O3b, respectively. A search for lensed gravitational wave signals throughout all of O3 was conducted by [263]. Finally [23] and [264] present searches for short- and long-duration unmodeled GW bursts in O3. None of these searches identify significant candidates, however the null results are used to place upper bounds on merger rates in their respective parameter regions.

Chapter 5
Performance of the low-latency GstLAL inspiral pipeline towards O4

This chapter is a reprint of [113].

GstLAL is a stream-based matched-filtering search pipeline aiming at the prompt discovery of gravitational waves from compact binary coalescences such as the mergers of black holes and neutron stars. Over the past three observation runs by the LIGO, Virgo, and KAGRA (LVK) collaboration, the GstLAL search pipeline has participated in several tens of gravitational wave discoveries. The fourth observing run (O4) is set to begin in May 2023 and is expected to see the discovery of many new and interesting gravitational wave signals which will inform our understanding of astrophysics and cosmology. We describe the current configuration of the GstLAL low-latency search and show its readiness for the upcoming observation run by presenting its performance on a mock data challenge. The mock data challenge includes 40 days of LIGO Hanford, LIGO Livingston, and Virgo strain data along with an injection campaign in order to fully characterize the performance of the search. We find an improved performance in terms of detection rate and significance estimation as compared to that observed in the O3 online analysis. The improvements are attributed to several incremental advances in the likelihood ratio ranking statistic computation and the method of background estimation.

5.1 Introduction

Since the first observing run (O1) of the LVK Collaboration, GstLAL, a matched filtering based gravitational wave search pipeline [4], has participated in the discovery

of groundbreaking gravitational wave events. GstLAL was among the search pipelines that made the first direct detection of gravitational waves from a merging binary black hole (BBH), known as GW150914 [12]. In the second observing run (O2), GstLAL was the first pipeline to observe the binary neutron star (BNS) merger known as GW170817, whose discovery kickstarted the field of multi-messenger astronomy [66, 140]. In the third observing run (O3), GstLAL detected $\mathcal{O}(10)$s of gravitational wave signals including the first ever neutron star–black hole binary (NSBH) mergers [115] and the very heavy BBH merger, GW190521, which resulted in a remnant object in the IMBH mass region [265].

The GstLAL pipeline can be operated in one of two configurations: a low-latency or "online" mode and an "offline" mode. The online configuration of the GstLAL analysis proceeds in near real time as strain data becomes available from the interferometers (currently, H, L and V). The online analysis enables the prompt detection of gravitational wave events, allowing for rapid communication to the external community for electromagnetic follow-up. In order to provide the best opportunities for multi-messenger astronomy, it is imperative that the low-latency analyses perform optimally. This includes reliable signal recovery, accuracy of source property estimation, and the ability of the search to keep up with real-time data and provide results as quickly as possible.

In contrast, the offline analysis proceeds on long timescales relative to the low-latency distribution of strain data. The offline analysis can benefit from a fuller understanding of the detector noise and the ability to re-rank the significance of candidates against the full asynchronous background estimate collected over the entire run duration. Since the likelihood ratios and FARs of the candidates are re-computed relative to the full background, it is also possible to make adjustments to the signal model and mass model compared to what is used in the online analysis [111]. All of these factors can contribute to higher sensitivity, as quantified by the sensitive volume-time $\langle VT \rangle$, in the offline analysis.

In this paper, we will focus on the online configuration and aim to characterize the GstLAL pipeline's performance toward the fourth observing run (O4). In Section 5.2 we will describe the current configuration of the GstLAL online analysis. Additionally, we describe the Gravitational Wave Low-latency Test Suite (`gw-lts`) software package as a tool for monitoring the performance of gravitational wave search pipelines in low-latency. Then, in Section 5.3 we demonstrate the performance of the pipeline by presenting results from a Mock Data Challenge (MDC). We will conclude in Section 5.4 with a description of on-going development towards O4.

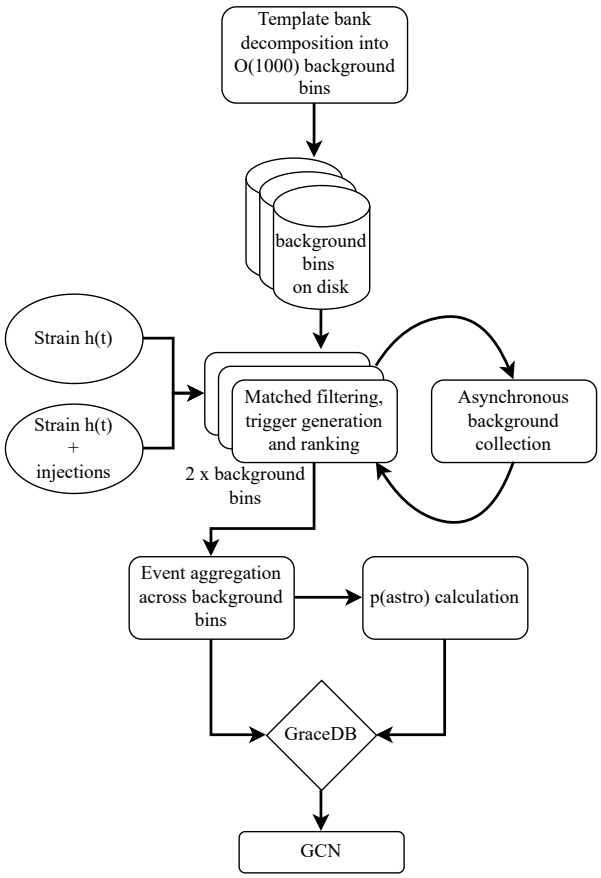

Figure 5.1. The low-latency GstLAL inspiral analysis workflow. The full template bank must first be decomposed into $\mathcal{O}(1000)$ independent background bins via the LLOID method of singular value decomposition and time slicing [92]. The strain data is transferred from the interferometer sites at LIGO Livingston, LIGO Hanford, and Virgo to the computing clusters where it will be read from disk by the GstLAL pipeline. Filtering, trigger generation, and candidate ranking proceeds in parallel for each background bin independently. These filtering and ranking jobs are duplicated to process strain channels which include simulated signals injected into the data. Background statistics are collected independently in each background bin and asynchronously marginalized over the full parameter space in order to inform the FAR estimation. Candidate events are aggregated in time across all background bins, using the maximum SNR or minimum FAR as a metric for determining which candidates will be uploaded to GraceDB. Finally, candidates passing the public alert FAR threshold will be disseminated via GCN.

5.2 Software description

5.2.1 GstLAL

The low-latency GstLAL inspiral workflow consists of two broad stages: a set up stage where pre-computed data products are generated and stored on disk and a persistent

analysis stage where strain data is filtered in near real time and candidate events are identified. We will give a brief description of the current workflow and configuration choices to be used in operating the GstLAL analysis during O4. A diagram of the low-latency workflow is shown in Fig. 5.1. For a more detailed description of the GstLAL analysis methods as of the end of O1 and O2 see [4] and [104], respectively. The GstLAL software package is described in [112].

Before the GstLAL analysis is launched, the template bank is first split into two halves in a process referred to as "checkerboarding". Each checkerboarded bank is constructed by taking alternating neighboring templates from the full bank. The checkerboarded banks are redundant as they cover the same parameter space while having unique individual templates. The full O4 template bank includes 1.8×10^6 templates, meaning that each checkerboarded bank has about 9×10^5 templates. The effectualness of the checkerboarded banks is validated in [107]. With this configuration the overall analysis can be split across two independent computing sites which improves the robustness of the analysis to upstream failures. According to the Low-Latency Online Inspiral Detection (LLOID) method, the checkerboarded template banks must then be split into independent bins of waveforms, hereafter referred to as background bins, as shown in Fig. 5.1. A full derivation and motivation of the LLOID method, including the singular value decomposition (SVD) and template time-slicing, is given by [92]. We first sort the template bank by the orthogonal Post-Newtonian (PN) phase terms μ_1 and μ_2. These are linear combinations of the PN coefficients, ψ^0, ψ^2, and ψ^3, defined as follows [106]:

$$\begin{aligned} \mu_1 &= 0.974\psi^0 + 0.209\psi^2 + 0.0840\psi^3 \\ \mu_2 &= -0.221\psi^0 + 0.823\psi^2 + 0.524\psi^3. \end{aligned} \tag{5.1}$$

Using these parameters to sort, we split the template bank into sub-banks each with ~ 500 templates. Each background bin is then constructed by grouping 2 sub-banks together. When computing the decomposition, we require a 99.999% match between the re-constructed template waveforms and the initial physical waveforms. This value is chosen by balancing the need for computational efficiency with the need for accurately reconstructed waveforms. For the checkerboarded O4 template bank in [107], this produces ~ 1000 background bins.

As part of the background bank construction during the set up stage of the analysis, the SVD waveforms are also whitened. For the initial whitening before filtering has begun, we use a reference power spectral density (PSD) generated from several hours of O3 data. As the analysis stage proceeds we re-whiten the SVD waveforms on a

weekly timescale using recent PSDs in order to account for any long term changes to the detector characteristic noise. As the analysis runs, the PSD is continuously tracked using a fast Fourier transform (FFT) length of 4 seconds. Such a short length of FFT in the whitening stage of the pipeline reduces latency at the cost of a less accurate PSD measurement which could potentially bring a loss in sensitivity while filtering.

The low-latency analysis ingests strain data, as well as data quality and interferometer state information from frame files. Each frame includes 1 second of data. The frames are distributed from the detectors via Apache Kafka, an open source event streaming platform. After streaming from the detector sites, frames are stored in shared memory partitions, where they are accessed by the GstLAL analysis. The frames are then processed in buffers 4096 bytes at a time by each filtering job in the GstLAL pipeline as shown in Fig. 5.1. In O4, the GstLAL pipeline will also ingest a parallel stream of strain data including simulated compact binary coalescence (CBC) signals injected into the data. These injections will be based on the inferred astrophysical distribution of sources based on the Gravitational Wave Transient Catalog (GWTC-3) [62].

It is known that the LIGO and Virgo data are not "well-behaved" and include transient and non-Gaussian noise components known as glitches. These glitches can be mistaken for astrophysical signals, especially high mass BBH templates which are short in duration within the LIGO-Virgo frequency band. To mitigate the negative effects of non-Gaussian data, the GstLAL pipeline gates particularly glitchy whitened $h(t)$ strain data using a threshold on the amplitude of the data in units of standard deviations. In gating the strain data, we must be careful to balance the desire to reduce false positives (i.e. mistaking a glitch for an astrophysical signal) with the desire to avoid false negatives (i.e. mistaking an astrophysical signal for a glitch). Since we know that signals from heavy mass CBC systems (for example, IMBH binaries) with high amplitudes and short durations tend to resemble glitches, we want to be conservative with gating data while filtering templates in this region of the parameter space. However, for low mass systems we can more aggressively gate with a lower threshold since these signals typically have a smaller amplitude and much longer duration. For this reason, we choose a gate threshold which is linear in the chirp mass,

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \tag{5.2}$$

and calculate it for each background bin as follows [266]. We first compute a gate ratio

defined as:

$$R_{\text{gate}} = \frac{\sigma_{\max} - \sigma_{\min}}{\mathcal{M}_{c,\max} - \mathcal{M}_{c,\min}}, \tag{5.3}$$

where we choose the minimum and maximum gate thresholds, $\sigma_{\min}$ and $\sigma_{\max}$ as 15.0 and 100.0, respectively. The $\mathcal{M}_c$ minimum and maximum are taken to be $\mathcal{M}_{c,\min} = 0.8$ and $\mathcal{M}_{c,\max} = 45.0$. Then, the $h(t)$ gate threshold is calculated for each background bin as:

$$\sigma_{thr} = R_{\text{gate}} \times \Delta\mathcal{M}_c + \sigma_{\min}. \tag{5.4}$$

Here, $\Delta\mathcal{M}_c$ is the difference between the maximum $\mathcal{M}_c$ in the given background bin and the $\mathcal{M}_{c,\min}$. This produces gate thresholds ranging from ~ 15 for the smallest mass bins to ~ 325 for the largest mass bins.

In the O4 configuration, we choose a filtering stride of 0.25 seconds, meaning that the matched filter output is computed in stretches of 0.25 seconds at a time. The small stride is chosen to reduce latency in the filtering stage. Triggers are defined by peaks in the SNR time-series output by the matched filtering which pass a threshold of 4.0. The GstLAL analysis has allowed for single detector candidates since O2 and will continue to do so in O4. However, when calculating the likelihood ratio of single detector candidates we apply a penalty to down rank their significance. This is a tunable parameter and the value to be used in O4 will be discussed in more detail in Sec. 5.3.2. Coincident candidates include triggers from the same template in at least two detectors. We require that the end times of coincident triggers be within 0.005 seconds of each other after accounting for the light travel time between detectors. Together, the coincidence threshold and the requirement that triggers across detectors ring up the same template provide a strong signal consistency test for candidate events.

The GstLAL pipeline uses the likelihood ratio as a ranking statistic to assign significance to gravitational wave candidates [110, 267]. Recent improvements to the likelihood ratio computation towards O4 are given in [111]. These include an upgraded analytic SNR-ξ^2 signal model and a method for removing signal contamination from the background which is also described in [5]. The background noise in each detector is estimated by collecting ranking statistic data from single-detector triggers observed in coincident time. We exclude triggers from times when only one detector is operating since these triggers may be astrophysical signals. These background estimations are cumulative and "snapshotted" to disk every 4 hours. The filtering jobs which process injection strain data do not collect their own background estimations. This is because the high rate of injected signals in the data would contaminate the background and corrupt the statistics

used for the FAR estimation. Instead, these injection filtering jobs use a copy of the background statistics collected by the corresponding non-injection filtering job which processes the same background bin.

While the pipeline is designed to run persistently, there is need to take the analysis down periodically. We remove each of our analyses for a short period of time on a weekly timescale with a staggered schedule so that at least one of the checkerboarded analyses is always observing. When an analysis is re-launched after this weekly downtime, we compress the background ranking statistic data by removing any values in the horizon distance history that differ fractionally from their neighbors by less than 0.003. This compression reduces the file size and memory use of the pipeline, which would otherwise grow without bounds over the duration of the observing period.

For FAR estimation, the ranking statistic data is marginalized by adding counts from the SNR-ξ^2 background distributions collected in each background bin. The histograms are marginalized over in a continuous loop, taking several hours to complete each iteration. The marginalization is cumulative in time so that as the run proceeds, we collect more and more background counts. To account for the two redundant checkerboarded analyses, we apply a FAR trials factor of 2 to each trigger.

Gravitational wave events passing a FAR threshold of one per hour will be uploaded to the GraceDB [268]. Because the GstLAL pipeline filters the strain data in ~ 1000 independent background bins, it is not only possible but highly probable that there will be multiple triggers associated with each physical gravitational wave candidate. The number of triggers per candidate could range from a few for quieter signals to several tens for louder signals. In order to reduce API calls to GraceDB we aggregate these triggers in time across background bins by the maximum SNR and only upload the current best candidate. In this aggregation stage, triggers from different background bins are grouped into candidates using a coincidence window defined by rounding $t_{\mathrm{end}} - dt$ down to the nearest half second and rounding $t_{\mathrm{end}} + dt$ up to the nearest half second. Here, t_{end} is the end time of the trigger and $dt = 0.2$ seconds. The first trigger received by the aggregator for a given candidate is uploaded to GraceDB immediately. Any subsequent triggers for the same candidate which are found with higher SNR are uploaded with a 4 second geometric wait time. That is, after the first upload, the second upload will not be made until 4 seconds later, the third upload until 4^2 seconds later, and so on. The aggregation stage of the pipeline is illustrated in Fig. 5.1.

Finally, the GstLAL pipeline calculates a probability of astrophysical origin, or $p(\mathrm{astro})$, for each event uploaded to GraceDB. The $p(\mathrm{astro})$ is a measure of the event's

significance, and as we also compute the probability that the event originates from each CBC source class (BNS, NSBH, or BBH) it gives an indication of the likelihood that an event will have an electromagnetic counterpart. Therefore, the p(astro) is an important quantity to help astronomers determine when to follow up gravitational wave candidates. More information about the GstLAL pipeline's computation of p(astro) can be found in [269].

5.2.2 GW Low-latency Test Suite

The `gw-lts` software is designed to provide consistency checks and real-time feedback on the reliability of science outputs of gravitational wave search pipelines. By using simulated signals injected in the strain data, we can compare the pipeline performance to what is expected.

The Test Suite requires a source of truth for the signals that are present in the data. For this, we rely on an injection set on disk which defines all of the injections, including all intrinsic and extrinsic parameters, and the GPS times at which they appear in the strain data. Using a live estimate of the PSD and the injected signal's sky location we can compute the expected SNR. For information about recovered events, GraceDB is taken as the source of truth. The `igwn-alert` software package is a messaging system built on Apache Kafka which sends notifications of GraceDB state changes to subscribed users. The Test Suite subscribes to notifications from `igwn-alert` which are sent for any new or updated event on GraceDB. The injections are then matched with recovered alerts in low-latency by finding coincidences within a small Δt which we take to be ± 1 seconds. This time window was chosen to be very small compared to a typical injection rate to avoid erroneous coincidences.

Once an injected signal is matched with a recovered event, the information is passed to an arbitrary number of independent jobs via Apache Kafka. The jobs compute metrics associated with the injection recovery such as the $\langle VT \rangle$, accuracy of source classification and sky localization, and accuracy of point estimates of the source intrinsic parameters. The `gw-lts` capabilities are described in further detail in Section 5.3.

All of the metrics computed by the `gw-lts` are stored with InfluxDB, which is an open source platform for storing and querying time series data. We use the data visualization tool Grafana to display the data in real time in online dashboards. With this infrastructure, we are able to track changes in the performance of the analysis on the timescale of seconds. Additionally, from the Influx database we are able to keep an archival record of the performance metrics.

5.3 Mock data challenge results

To demonstrate the performance of the GstLAL analysis and our readiness for O4, we participated in a MDC consisting of a forty day stretch of HLV O3 strain data taken from 5 Jan. 2020 15:59:42 to 14 Feb. 2020 15:59:42 UTC and replayed so as to be analyzed in a low-latency configuration. The MDC also provided a set of identical strain channels with injected BNS, NSBH, and BBH signals. Details of the injection distributions used in the MDC can be found in [270]. Injected signals were placed in the strain data at a rate of one per ~ 40 seconds, leading to a total of 5×10^4 total injections throughout the MDC duration.

In this section we seek to quantify the performance of the GstLAL pipeline in its latest configuration. We will first show the recovery of known gravitational wave events in the MDC data, as well as highlight any potential retraction level events. We will then detail the results of the MDC injection campaign. Finally we present the stability and performance of the pipeline in terms of its uptime and latency.

5.3.1 Gravitational wave events

There are 9 gravitational wave events in the duration of the MDC replay data which were previously published as significant candidates in GWTC-3 [3]. These are described throughout the remainder of this section and summarized in Table 5.1, comparing the GstLAL pipeline's recovery of the signal in O3 to that in the MDC. We recover all of the 9 candidates at the 1 per hour FAR threshold for uploading to GraceDB. Of these, three were found with high significance by GstLAL in the O3 online analysis. Two were found with marginal or sub-threshold significance online but with high significance offline. The remaining four candidates were found by GstLAL only in the offline analysis. The recovery of all previously published candidates shows that the pipeline is performing with at least the same capability as in O3.

Fig. 5.2 shows the count of observed candidates versus IFAR. The expected background counts are calculated using an estimated livetime which is equal to the wall clock time from the first to the last candidate. Additionally, we apply a trials factor of 2 to the FARs since we only include candidates from one of the checkerboarded analyses. Fig. 5.2 shows the 9 known gravitational wave events recovered in the MDC. The candidate IFAR statistics agree with the expected counts from noise at lower IFAR and diverge due to the presence of signals at higher IFAR.

GW200112_155838. This event was a BBH candidate observed in LIGO Livingston

	O3 Online				MDC			
Name	**Inst**	**SNR**	**FAR** (yrs^{-1})	***p*(astro)**	**Inst**	**SNR**	**FAR** (yrs^{-1})	***p*(astro)**
GW200112_155838	L1	18.79	4.05×10^{-4}	> 0.99	L1	18.46	1.01×10^{-7}	> 0.99
GW200115_042309	H1L1	11.42	6.61×10^{-4}	> 0.99	H1L1	11.48	2.55×10^{-4}	> 0.99
GW200128_022011	–	–	> 1.2	–	H1L1	9.98	1.44×10^{-4}	> 0.99
GW200129_065458	H1L1V1	26.61	2.11×10^{-24}	> 0.99	H1L1V1	26.30	1.78×10^{-17}	> 0.99
GW200202_154313	–	–	> 1.2	–	H1L1	11.09	1.69×10^{-2}	> 0.99
GW200208_130117	–	–	–	–	H1L1	10.56	4.92×10^{-5}	> 0.99
GW200208_222617	–	–	–	–	H1L1	8.00	2.02×10^{3}	0.48
GW200209_085452	–	–	–	–	H1L1	9.96	1.20	> 0.99
GW200210_092254	–	–	–	–	H1L1	9.28	3.64×10^{3}	0.27

Table 5.1. Gravitational wave candidates from 5 Jan. 2020 15:59:42 to 14 Feb. 2020 15:59:42 UTC as recovered by the GstLAL pipeline during the O3 online analysis and during the MDC. The instruments provided are those which participated in the event, that is, contributed a trigger with SNR > 4.0. Here, the SNR is the recovered network SNR, FAR is the false alarm rate in inverse years, and p(astro) is the probability of astrophysical origin. The two low-significance candidates identified with FAR above the public alert threshold in the O3 online analysis are indicated with FAR > 1.2 per year. We note that this FAR threshold is after a trials factor corresponding to the number of operating pipelines has been applied. The last four candidates in the table were not recovered by GstLAL in the O3 online analysis.

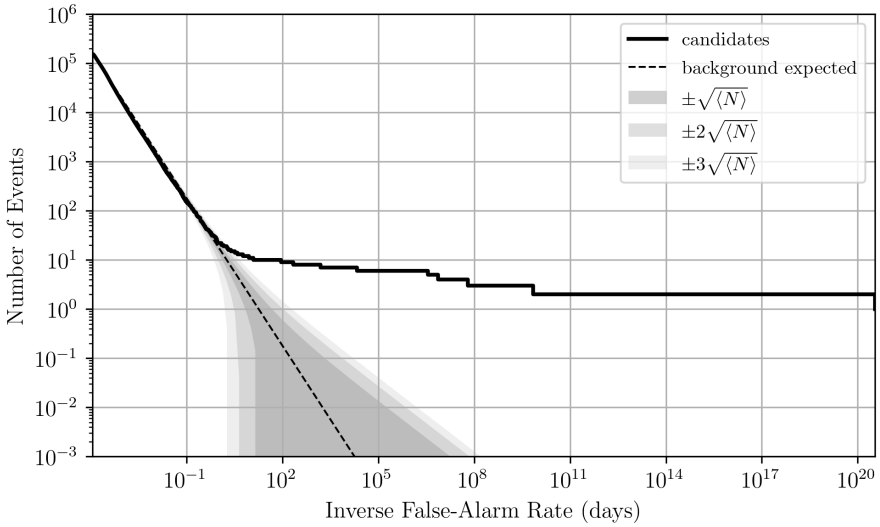

Figure 5.2. Count of observed candidates vs IFAR in days. The dashed line is an estimation of the expected number of background counts, assuming a FAR threshold of 1×10^4 per day. The one, two, and three σ error regions are indicated by the shaded regions.

data as a single detector candidate with chirp mass 35.37 $M_\odot$ in O3 and 33.37 $M_\odot$ in the MDC. In the MDC, we recover the event with a comparable SNR as that observed in O3, however in the MDC the FAR is significantly lower.

GW200115_042309 was the first confident NSBH detection found in O3. The event was found as a coincident trigger in LIGO Hanford and LIGO Livingston data. The SNR recovery, FAR estimation, and chirp mass estimation are all equivalent in the MDC to what was observed in O3.

GW200129_065458 was a BBH and the loudest gravitational wave signal in the duration of the MDC with O3 SNR = 26.61. The event was recovered well below the public alert FAR threshold in both O3 and the MDC.

GW200128_022011 and **GW200202_154313** are both BBH candidates found by GstLAL in the O3 online analysis with low significance. Both candidates were found with FAR above the O3 public alert threshold of 1.2 per year, where a trials factor corresponding to the number of operating pipelines has been applied. Later, during the offline analysis they were recovered as significant candidates and included in GWTC-3 [3]. In the MDC we recover both candidates with significantly lower FARs, both well below the public alert threshold. Therefore, if similar events occur during O4, we can expect to recover them as significant public alerts.

GW200208_130117 was not recovered by GstLAL in the O3 online analysis, however it was found in the offline analysis by GstLAL as a highly significant candidate [3]. As recovered in the MDC, this event is a BBH candidate with chirp mass 34.50 $M_\odot$ and a much lower FAR than what was found in either the O3 online or offline analyses.

GW200208_222617 was only recovered by GstLAL as a sub-threshold candidate in the O3 offline analysis, and its inclusion as a significant candidate in GWTC-3 was due to its recovery by other CBC pipelines [3]. The GstLAL pipeline did not recover this event in O3 online. In the MDC the event was recovered with low significance at SNR = 8.00 and FAR = 2.02×10^3 per year.

GW200209_085452 and **GW200210_092254** were not recovered by GstLAL in the O3 online analysis, however they were found in the offline analysis by GstLAL. GW200209_085452 was recovered in the MDC as a significant candidate, although with a FAR of 1.20 per year it is less significant than in the O3 offline analysis. This event is a BBH candidate with chirp mass = 39.45 $M_\odot$ found in LIGO Hanford and LIGO Livingston data. GW200210_092254 was found as a sub-threshold candidate in the MDC with FAR = 3.64×10^3 per year. If astrophysical, the event would be an NSBH candidate with chirp mass = 7.89 $M_\odot$. GW200210_092254 was considered as a highly

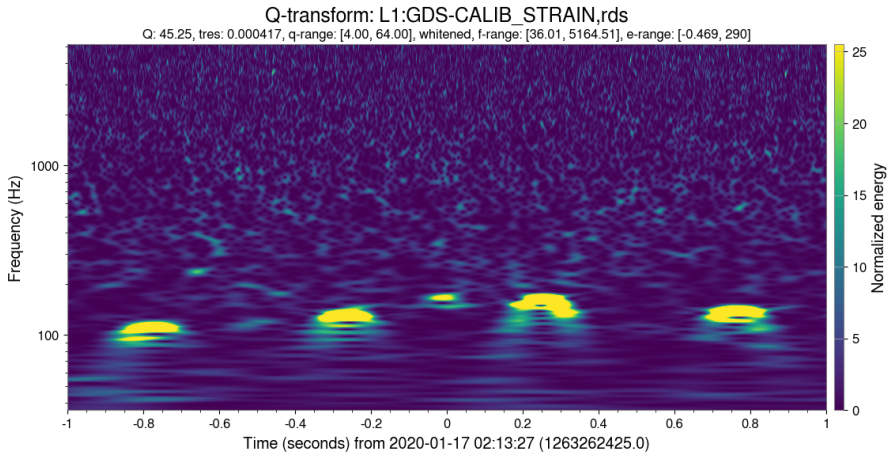

Figure 5.3. Spectrogram of L1 $h(t)$ data for ± 1 second around the time of the retraction level candidate recovered in the MDC. This candidate is expected to be terrestrial in origin due to the clear presence of glitches in this data.

significant candidate in GWTC-3 due to its recovery by other pipelines, however the GstLAL trigger for this candidate had a $p(\text{astro})$ below the threshold of 0.50 and on it's own would be considered marginal. We note that the candidate was recovered with a higher FAR and a lower $p(\text{astro})$ in the MDC than in the O3 offline analysis. Still, the recovery of these candidates in the MDC, even at sub-threshold significance which is the case for GW200210_092254, demonstrates an improvement over the O3 online sensitivity.

The improved performance of the GstLAL pipeline in the MDC as compared to the O3 online analysis can be attributed to a number of incremental improvements made to the likelihood ratio ranking statistic and background estimation. [111] describes an improved signal model and [5] introduces a new method for a time-dependent background wherein contamination is reduced by removing signals counts from the background SNR $- \xi^2$ histograms. Each of these changes have introduced a small improvement to the $\langle VT \rangle$ which, when combined, leads to a noticeable increase in sensitivity and corresponding number of detected events.

5.3.2 Retractions

In O3, there were 23 public gravitational wave candidates which were subsequently determined to be terrestrial in origin and thus retracted. Of these, GstLAL contributed to 15. In O4, we hope to significantly reduce the number of retractions produced by the GstLAL pipeline. Four of the 23 retractions took place during the stretch of data covered

by the MDC. These are: S200106au, S200106av, S200108v, and S200116ah [271–273].

GstLAL did not upload triggers for S200106au and S200106av during O3. In the MDC, these events would have occurred at a time before the pipeline had collected sufficient background to begin ranking candidates, and thus we did not upload triggers for these events. The retractions S200108v and S200116ah were GstLAL-only candidates in O3, both being found as single detector candidates in LIGO Livingston. Again, the time corresponding to S200108v would have been early enough in the MDC cycle that the pipeline was not ranking or uploading triggers yet, so we cannot make any comparison to our performance in O3 for this retraction. Finally, we did not produce any trigger below the 1 per hour FAR threshold for uploading to GraceDB corresponding to S200116ah in the MDC, despite the analysis being fully burned in and operating in a stable state. This means that a similar noise event in O4 may not lead to a spurious candidate and subsequent retraction. However, there is insufficient data within the MDC to infer any changes in performance in terms of the rate of retractions.

In addition to the retracted candidates uploaded by GstLAL in O3, for the purpose of the MDC we define a "retraction level candidate" as any gravitational wave candidate uploaded with a FAR less than one per year which is not in the list of previously published candidates discussed earlier in this section. Over the duration of the MDC, we find one such retraction level candidate. This was a single-detector candidate found in LIGO Livingston with an SNR of 14.5 and a FAR of 1.67 per year which was low enough to be counted as significant. The LIGO Livingston data around the event time shows the clear presence of scattering glitches, as shown in Fig. 5.3. Further evidence of terrestrial origin for this candidate is that no coincident triggers were recovered in LIGO Hanford or Virgo despite both of these detectors operating normally at the time. Candidates recovered in only a single detector are more susceptible to uncertainty since they lack the strong signal consistency test of coincidence in multiple detectors. For this reason, there has been a penalty applied to the ranking statistic of single detector candidates which down weights their significance. In the O3 offline analysis and in the MDC we used a singles penalty of 12 in log likelihood ratio, however in order to reduce the number of similar retraction level events in O4, we plan to increase the singles penalty to 13. With this penalty applied, the retraction event in the MDC would be down weighted and expected to be recovered with a FAR greater than two per year.

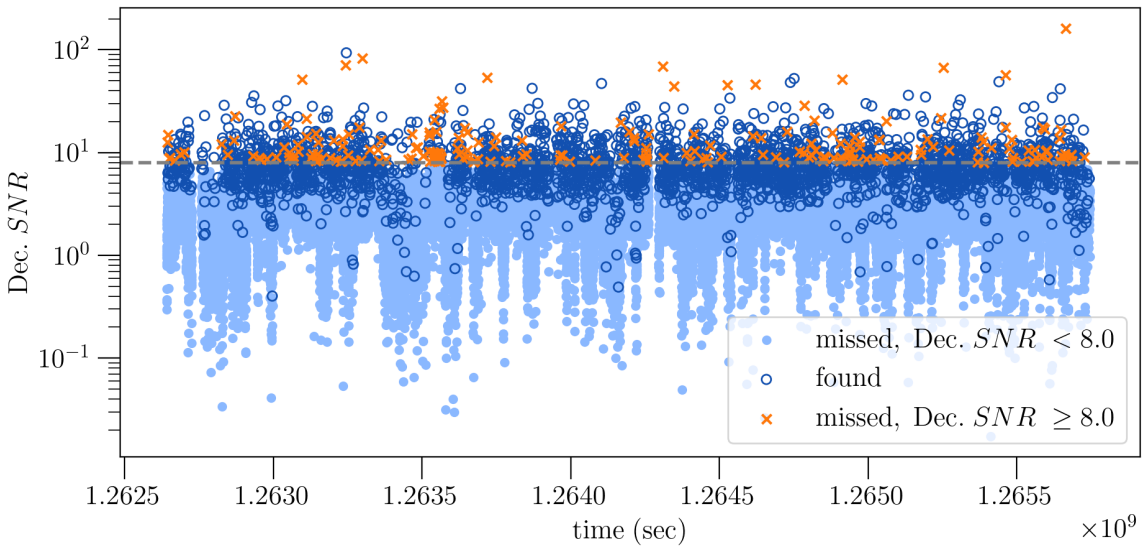

Figure 5.4. Time-series of injected decisive SNR for injections with component masses and spins within the O4 template bank. Dark blue circles indicate injections that were recovered below a 2 per day FAR threshold. Orange crosses and light blue markers indicate injections not recovered below this FAR threshold, where orange crosses are injections with decisive SNR > 8.0. Times on the horizontal axis are GPS times shifted to the original O3 epoch.

5.3.3 Recovered injections

There were 5×10^4 simulated signals injected into the five week duration of the MDC strain data. Of these, many had component masses and spins outside the region of parameter space covered by our template bank. In addition to the injection parameters, the expected recovery of each injection is dependent on the set of interferometers producing science quality data at the time of the injection. During times when no interferometers are operating we of course do not expect to recover any injections. We define the decisive SNR as the SNR in the second most sensitive interferometer during times when multiple interferometers were observing, and the only available SNR otherwise. The decisive SNR is a more informative measure of the loudness of an injection than the network SNR since it wraps in information about the set of operating interferometers. While all injections have network SNR ≥ 4.0, we find that many injections have decisive SNR < 4.0.

Fig. 5.4 shows the time-series of decisive SNR for all injections throughout the MDC. For the purpose of this paper, we focus on injections whose parameters fall inside our bank, that is injections with component masses between 1.0 $M_\odot$ and 200 $M_\odot$, with total masses $m_1 + m_2 < 400.0\ M_\odot$ and mass ratios $q = m_1/m_2 < 20$. For objects with mass $< 3.0\ M_\odot$ the template bank restricts spins perpendicular to the orbital plane

$|s_{i,z}| < 0.05$ and for objects with mass $> 3.0\ M_\odot$ allows $|s_{i,z}| < 0.99$. We use the effective precession spin, χ_p, defined in [245] as:

$$\chi_p = \frac{\max(a_1 \cdot s_1, a_2 \cdot s_2)}{a_1 \cdot m_1^2} \tag{5.5}$$

to quantify the in-plane spin of injections. Here, $a_1 = 2 + 3/2q$, $a_2 = 2 + 3q/2$, $s_i = \sqrt{s_{i,x}^2 + s_{i,y}^2}$ and the mass ratio q assumes $m_1 \geq m_2$. Since the template bank does not include any in-plane spins, we focus on injections with $\chi_p < 1 \times 10^{-3}$. However, we find that all injections with one component mass $< 3.0\ M_\odot$ have spins outside the range of the bank, therefore we relax the spin conditions on these components. The mass and spin restrictions that we use are summarized in Table 5.2. Finally, to account for the fact that not all interferometers were providing science quality data at all times, we highlight injections with an estimated decisive SNR ≥ 8.0. These cuts leave a total of 1457 injections during the five week MDC. Of these, there are 597 BBH, 482 BNS, and 378 NSBH injections.

The injected SNRs are not known in advance of the MDC, but we estimate them using `gw-lts`. We calculate the injected strain time series using the injection end time, sky position, and other source intrinsic parameters assuming an `IMRPhenomPv2_NRTidalv2` waveform [274]. We use a running estimate of the detector PSDs and estimate the SNR with a lower (upper) frequency cut off of 10.0 (1600.0) Hz. Fig. 5.5 shows the recovered and estimated injected SNR for each detector. If the template bank did not have a sufficient minimal match, we may expect to see systematically lower recovered SNRs than the expected values. However, we find that the recovered SNR generally aligns with the expected SNR. We note that the figure shows a wider spread in the recovered SNRs for LIGO Hanford than LIGO Livingston. This is likely due to the greater sensitivity in LIGO Livingston such that injections with low recovered SNRs in LIGO Livingston may be expected to have had even lower SNRs in LIGO Hanford and thus not be recovered at the FAR threshold for inclusion in the figure.

An injection is considered "found" if it is recovered by the pipeline with a FAR passing some pre-determined threshold, and "missed" otherwise. We will quote most results in the following sections with respect to a 2 per day FAR threshold. At the time of writing, this is the threshold expected to be used in O4 for sending public alerts [275]. However, the FARs of CBC signals will be subject to a trials factor corresponding to the number of operating pipelines, so the effective alert threshold will be lower. We define the injection

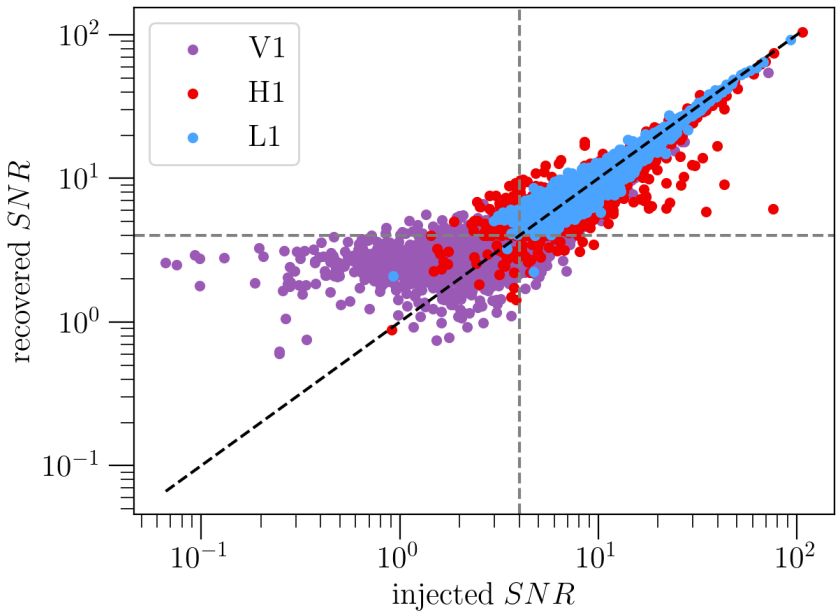

Figure 5.5. Recovered and estimated injected SNR in each interferometer: LIGO Hanford (H), LIGO Livingston (L), and Virgo (V). The black dashed line shows the diagonal and the grey dashed lines show injected and recovered SNRs = 4.0, the threshold which defines a trigger.

m_1	m_2	M	q	$\lvert s_{1,z}\rvert$	$\lvert s_{2,z}\rvert$	$\lvert \chi_p\rvert$
$1, 3$	$1, 3$	< 6	< 0.33	–	–	–
$3, 200$	$1, 3$	< 203	< 20	–	–	–
$3, 200$	$3, 200$	< 400	< 20	< 0.99	< 0.99	< 0.001

Table 5.2. Restrictions on the masses (m_1, m_2, M, and q), spins perpendicular to the orbital plane ($s_{1,z}$ and $s_{2,z}$), and spins parallel to the orbital plane (χ_p) of injections according to the O4 template bank boundaries. The "–" in the first two rows indicates that we make no restrictions on the spins for injections with $m_i < 3.0 M_\odot$. This relaxation is done because the template bank restricts NS spins $|s_{i,z}| < 0.05$, which would effectively remove most BNS and NSBH injections from consideration.

recovery efficiency as,

$$\frac{\text{found injections}}{(\text{found injections} + \text{missed injections})}. \tag{5.6}$$

At the 2 per day FAR threshold the efficiency was = 0.84 for all injections in the template bank. The recovered injection efficiencies for each source class are shown in Table 5.3 at four typical FAR thresholds. As is expected, the efficiencies are better at more conservative FAR thresholds. The analysis has the highest efficiency for injections

FAR	BNS	NSBH	BBH	ALL
2.78×10^{-4} Hz	0.95	0.77	0.87	0.87
2.31×10^{-5} Hz	0.95	0.71	0.84	0.84
3.85×10^{-7} Hz	0.89	0.65	0.77	0.78
3.16×10^{-8} Hz	0.86	0.62	0.71	0.74

Table 5.3. Injection efficiencies as defined in Eq. 5.6 computed using four FAR thresholds to count "found" injections: one per hour (the GraceDB upload threshold), two per day (the public alert threshold), one per month, and two per year. Source categories are defined in Table 5.2 and "ALL" combines injections from the three source categories.

consistent with BNS sources, and the lowest efficiency for NSBH sources. The injections in the MDC included precession effects while the O4 GstLAL template bank was constructed assuming only spins aligned with the orbital angular momentum. Therefore, the relatively lower NSBH recovery efficiency is expected as the precession will have a more significant effect on the gravitational waveform for binaries with more extreme mass ratios [276]. In future work, we may seek to quantify the efficacy of the template bank in the precessing parameter space.

We would expect the pipeline to recover all injections above some decisive SNR or network SNR threshold. However, Fig. 5.4 shows that there are several very high SNR missed injections throughout the duration of the MDC. We find that most of the missed injections with decisive SNR > 20.0 are high mass BBH injections and a few are high mass ratio NSBH injections. This results in a decrease of the BBH recovery efficiency as the injections increase in SNR, which is contradictory to our expectations. These injections are missed due to falling outside of the SNR-ξ^2 signal region used in the likelihood ratio calculation. The signal region is an analytic model which depends on the allowed mismatch[1] between recovered SNR time-series and the template waveform as part of the autocorrelation ξ^2 test. If the allowed mismatch range is too strict, it will result in a narrow signal model which can exclude real signals. This effect is exaggerated at high SNR where we expect larger mismatches due to the discreteness of the template bank as well as waveform systematics. In the MDC, we used a mismatch range of $0.1 - 10\%$. The optimal mismatch range in the signal model is an open area of study. See [111] for a more detailed discussion.

[1]The "mismatch" can also mean the fractional loss in SNR due to differences between the template parameters and the true waveform. However, here we refer to the mismatch as defined in [111] which is an unnormalized quantity, therefore retaining a dependence on the SNR.

5.3.3.1 Injection parameter recovery

In this section we will quantify the accuracy of point estimates of the source intrinsic parameters made by the GstLAL pipeline. These estimates simply come from the template parameters of the trigger which rang up the maximum SNR across background bins. An understanding of the parameter accuracy obtained by search pipelines can be useful to full parameter estimation efforts. For example, the Bayesian inference library Bilby [226, 277] relies on the choice of prior probability distributions for intrinsic parameters. When parameters are well determined by the searches, Bilby can use narrow distributions around those values, otherwise more broad prior distributions must be used. We present parameter accuracy results for the chirp mass $\mathcal{M}_c$, effective inspiral spin χ_{eff}, mass ratio q, and the coalescence end time, t_{end}. The χ_{eff} is a mass-weighted combination of the component spins parallel to the orbital angular momentum $\hat{L}$, defined as:

$$\chi_{\text{eff}} = \frac{(m_1\vec{s_1} + m_2\vec{s_2}) \cdot \hat{L}}{m_1 + m_2}, \tag{5.7}$$

where we take $\hat{L}$ to be in the z-direction. Both the injected and recovered masses quoted in this paper are in the detector frame. The error on a recovered parameter, λ is defined as:

$$\text{error} = \frac{\text{recovered } \lambda - \text{injected } \lambda}{\text{injected } \lambda}, \tag{5.8}$$

for all parameters except for the end time, where we simply take the error as the difference between the recovered and injected end times in milliseconds.

It is well known that the chirp mass, $\mathcal{M}_c$, is one of the best measured parameters in gravitational wave detections, however the recovered accuracy is highly dependent upon the mass of the system. Injections with small $\mathcal{M}_c < 10 - 20M_\odot$ are recovered with very accurate $\mathcal{M}_c$, but above this level, the accuracy starts to fall off, as shown in Fig. 5.6. For BNS injections, we find a mean $\mathcal{M}_c$ error 2.06×10^{-4} with a standard deviation of 8.33×10^{-4}. Similarly, the $\mathcal{M}_c$ in the NSBH region is recovered very well with mean -2.14×10^{-4} and standard deviation 6.26×10^{-3}. The BBH region has a higher $\mathcal{M}_c$ error over all, and additionally a much larger spread in the error with mean 1.54×10^{-1} and standard deviation 4.53×10^{-1}. Histograms of the recovered $\mathcal{M}_c$ error for each source class are shown in Fig. 5.7. The figure shows a skew in the recovered $\mathcal{M}_c$ error for BBH injections toward more positive values, i.e. we are more likely to overestimate the $\mathcal{M}_c$ in this region than to underestimate it. This skew is likely due to differences in the injected waveforms compared to the waveforms used to generate the template bank, however we

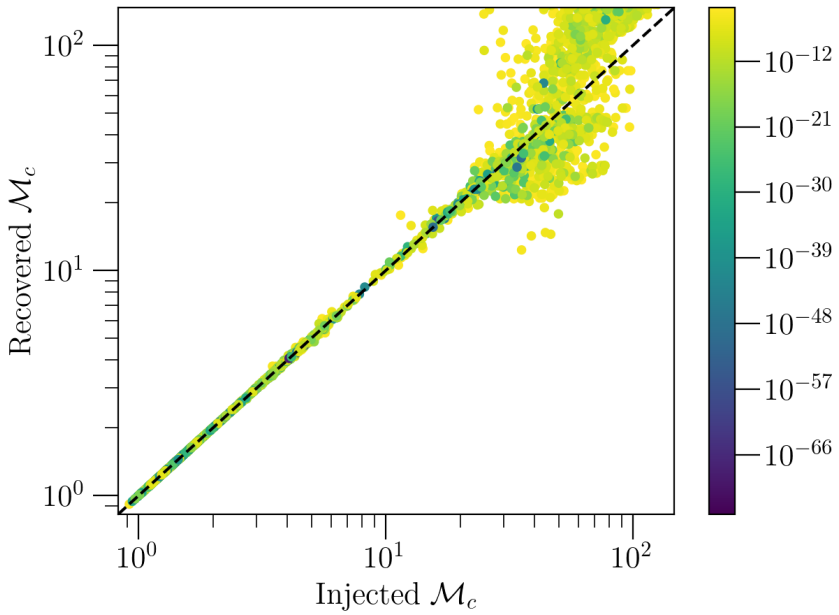

Figure 5.6. Injected $\mathcal{M}_c$ for injections found with FAR $< 2.31 \times 10^{-5}$ Hz is shown on the horizontal axis. The vertical axis shows the recovered $\mathcal{M}_c$. The color bar is FAR.

leave a more in depth investigation of this issue to future work.

Fig. 5.8 and Fig. 5.9 are scatter plots of the injected and recovered χ_{eff} and q respectively. These plots show that there is very little correlation between the injected and recovered values of these parameters. The mean and standard deviation on the recovered error for these parameters are given in Table 5.4.

A histogram of the difference between the injected and recovered injection end times is given in Fig. 5.10. Table 5.4 shows the mean t_{end} difference across all source classes and detectors is 6.23 milliseconds with a standard deviation of 30.22 milliseconds. The 90^{th} percentile on $|t_{\text{end}}|$ is 25.6 milliseconds and we recover every injection with a recovered t_{end} less than a second away from the injected value. Fig. 5.10 shows that for BNS and BBH injections, most of the recovered end times fall within ± 50 milliseconds of the true injected end time. For BNS injections, the mean t_{end} difference is -0.90 milliseconds with a standard deviation 18.0. And for BBH injections, the mean t_{end} difference is 6.03 milliseconds with a standard deviation 11.3. However, for NSBH templates, there is a wider distribution, skewing towards higher positive values of the end time difference, with mean 18.7 milliseconds and standard deviation 59.3. This is expected to be due to waveform systematics.

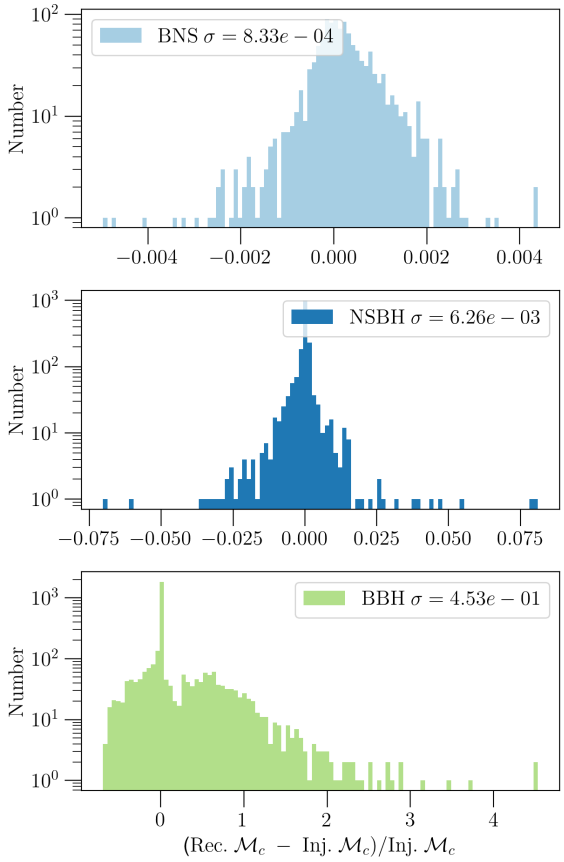

Figure 5.7. Recovered $\mathcal{M}_c$ error for injections found with FAR $< 2.31 \times 10^{-5}$ Hz. The top panel shows injections in the BNS range of the parameter space, the middle panel shows NSBH detections, and the bottom panel shows BBH injections. The σ value in each panel indicates the standard deviation on the recovered $\mathcal{M}_c$ error.

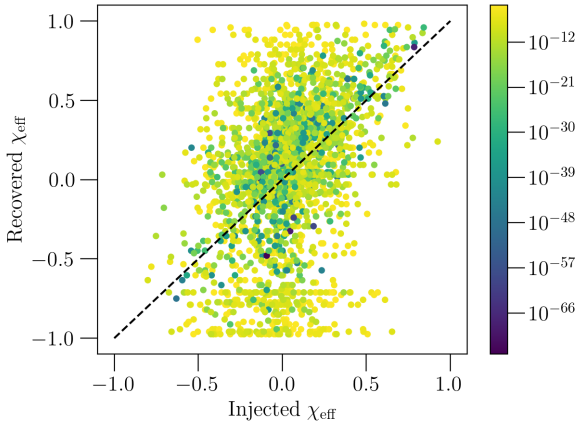

Figure 5.8. Injected $\chi_{\rm eff}$ for injections found with FAR $< 2.31 \times 10^{-5}$ Hz is shown on the horizontal axis. The vertical axis shows the recovered $\chi_{\rm eff}$. The color bar is FAR.

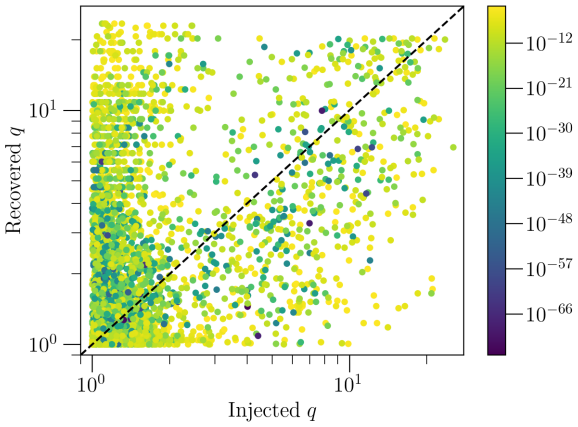

Figure 5.9. Injected mass ratio, $q = m_1/m_2$, for injections found with FAR $< 2.31 \times 10^{-5}$ Hz is shown on the horizontal axis. The vertical axis shows the recovered mass ratio. The color bar is FAR.

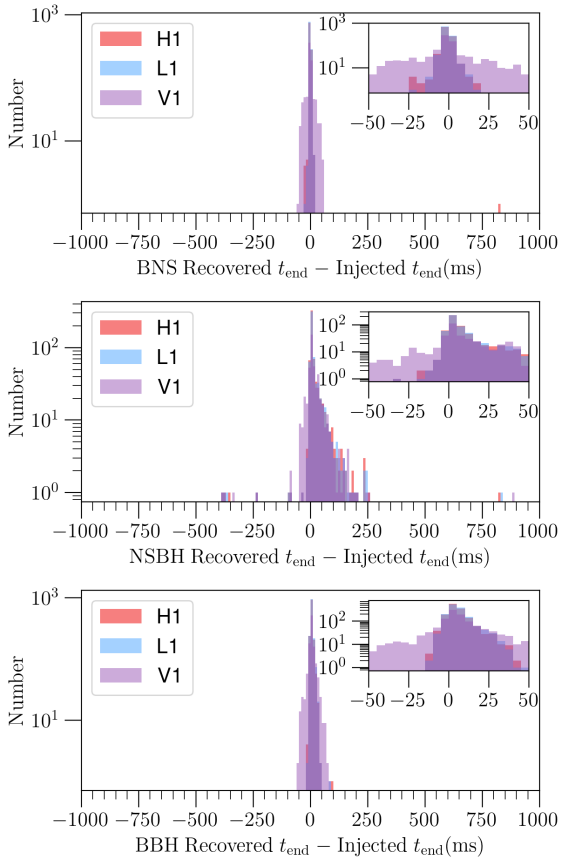

Figure 5.10. Recovered end time accuracy in milliseconds of injections recovered with FAR $< 2.31 \times 10^{-5}$ Hz. Results are shown for each interferometer: LIGO Hanford (red), LIGO Livingston (blue), and Virgo (purple). BNS injections are shown in the upper panel, NSBH in the center, and BBH in the lower panel.

	$\bar{X}$	σ	P_{50}	P_{75}	P_{90}
$\mathcal{M}_c$	0.15	0.45	0.007	0.33	0.73
$\chi_{\rm eff}$	5.77	252	1.34	3.71	10.8
q	1.39	2.86	0.45	1.67	4.97
$t_{\rm end}$	6.23	30.22	3.8	9.78	25.6

Table 5.4. Mean, $\bar{X}$, standard deviation, σ, and the fiftieth, seventy-fifth, and ninetieth percentiles on the recovered parameter error. The error is defined as in Eqn. 5.8 for all parameters except for the end time, where we simply take the difference in milliseconds between the recovered and injected values as the error. Results are computed only including injections which were recovered below a FAR threshold of two per day. The percentiles are computed for the absolute value of each distribution.

5.3.3.2 Search sensitivity

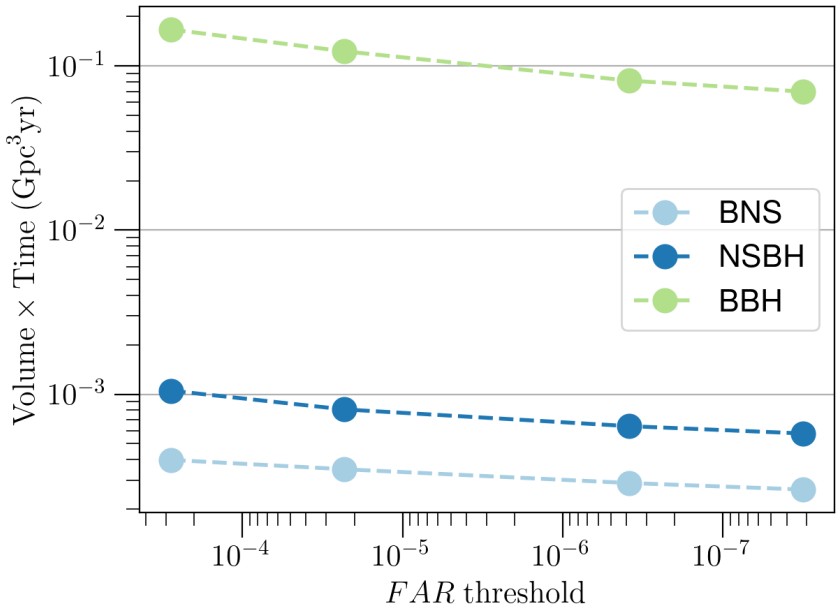

Figure 5.11. $\langle VT \rangle$ in each source class at the end of the MDC at four different FAR thresholds: 2 per year, 1 per month, 2 per day, and 1 per hour. BNS $\langle VT \rangle$ is shown in light blue, NSBH $\langle VT \rangle$ in dark blue, and BBH $\langle VT \rangle$ in green.

The comoving volume is defined as [171, 278]:

$$V_c = 4\pi D_H \int_0^z \mathrm{d}z \frac{(1+z)D_A^2}{E(z)} \tag{5.9}$$

with z as the redshift. D_H is the Hubble distance, D_A is angular distance, and $E(z)$

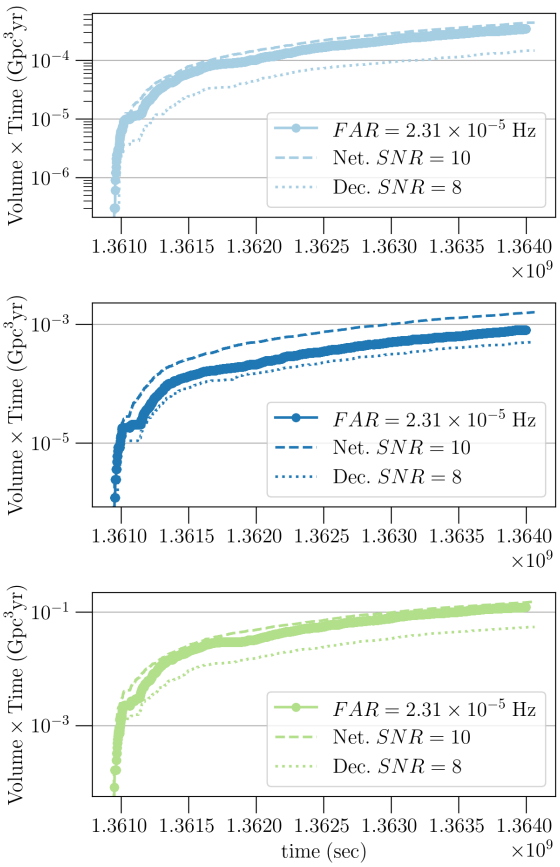

Figure 5.12. The plot shows the cumulative $\langle VT \rangle$ time-series over the duration of the MDC for each source class: BNS in light blue (top panel), NSBH in dark blue (middle panel), and BBH in green (bottom panel). The $\langle VT \rangle$ is calculated using three different thresholds for counting "found" injections: FAR $< 2.31 \times 10^{-5}$ Hz (dot markers), network SNR > 10.0 (dashed line), and decisive SNR > 8.0 (dotted line).

is the Hubble parameter. Multiplying this quantity by an observation time gives the surveyed spacetime-volume. We compute the total injected volume-time, $\langle VT \rangle_{\rm inj}$, using the max redshift to which injections were distributed and the time range over which injections were placed. For the MDC this time range is 3.456×10^6 seconds. With this quantity we can then estimate the online $\langle VT \rangle$ in the MDC as:

$$\langle VT \rangle = N_f \times \langle VT \rangle_{\rm inj}. \tag{5.10}$$

Here, N_f is the fraction of "found" injections out of the total number of injections in the data. We independently compute the $\langle VT \rangle$ for each source population. The max redshifts of the injection distributions are $z =$ 0.15, 0.25, and 1.9 for BNS, NSBH and

$\langle VT \rangle$ Gpc^{3}yrs	BNS	NSBH	BBH
$\langle VT \rangle_{\text{inj}}$	1.08×10^{-1}	4.34×10^{-1}	29.1
$\langle VT \rangle$	3.49×10^{-4}	8.08×10^{-4}	1.23×10^{-1}

Table 5.5. Values of the $\langle VT \rangle$ in cubic gigaparsec-years measured at the end of the MDC using a FAR threshold of 2.31×10^{-5} Hz compared to the injected $\langle VT \rangle$ in each source class.

BBH respectively. Table 5.5 gives the injected $\langle VT \rangle$ for each source class.

Fig. 5.11 shows the final $\langle VT \rangle$ in the MDC for each source class using four different FAR thresholds to determine whether injections count as found. In Fig. 5.12, we show the cumulative $\langle VT \rangle$ over the duration of the MDC using different thresholds to count found injections. For each source class, we find the highest $\langle VT \rangle$ by using a threshold of network SNR = 10.0 and the lowest $\langle VT \rangle$ with a threshold of decisive SNR = 8.0.

5.3.3.3 Sky localization and source classification

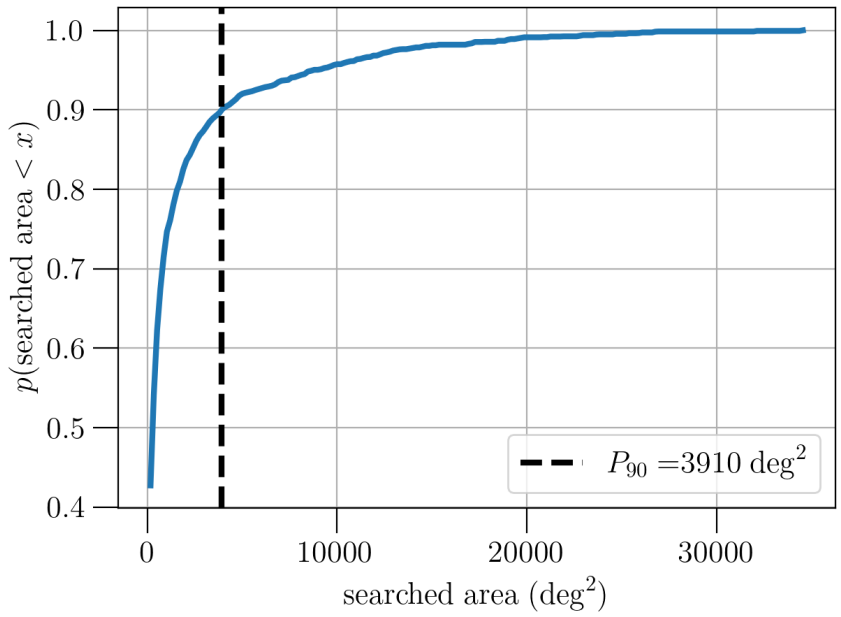

Figure 5.13. Cumulative distribution function of skymap searched area (deg^2) for injections recovered with FAR less than two per day.

For efficient electromagnetic follow-up, it is vital that accurate sky localizations and source classifications are provided to the public in low latency. The sky localization information informs where electromagnetic observers should search on the sky to find coincident events. The accuracy and precision of sky localization information can have a

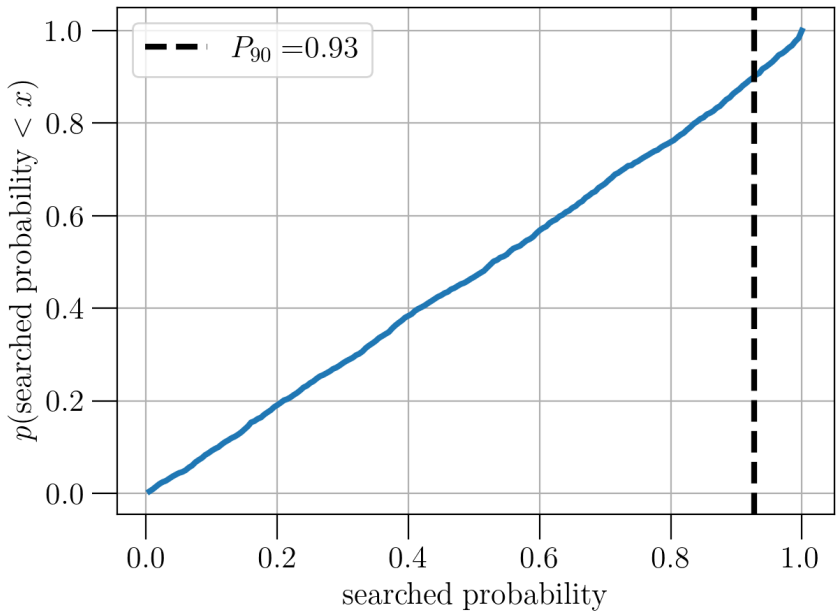

Figure 5.14. Cumulative distribution function of skymap searched probability for injections recovered with FAR less than two per day.

direct impact on the time it takes to identify a counterpart, especially for narrow field of view telescopes. Sky localizations are produced in low latency for all events on GraceDB using Bayestar [157,279]. The sky localization calculation depends on the SNR time-series around the coalescence time of the event. These are uploaded to GraceDB by the search pipelines as part of the event metadata. Since the potential for bright electromagnetic counterparts is highly dependent on the nature of the binary source, it is also important to provide accurate source classification so that observers may make informed decisions about when to follow up gravitational wave events. The probability that a gravitational wave candidate is astrophysical in origin is the p(astro), computed by GstLAL using the multi-component FGMC formalism [205, 206, 269]. We use a population model with a Salpeter distribution for the source component masses m_1, m_2 given by [280]:

$$p(m_1, m_2) \propto \frac{m_1^{-2.35}}{m_1 - m_{\rm min}}, \tag{5.11}$$

with $m_{\rm min} = 0.8 M_\odot$ and a uniform distribution in component spins, $s_{1,z}$, $s_{2,z}$. The p(astro) is further divided into the probability that a gravitational wave candidate originates from

	P_{50}	P_{75}	P_{90}
ALL	271	1080	3910
1 IFO	3150	10,400	18,400
2 IFO	301	893	2470
3 IFO	31.9	140	357

Table 5.6. Fiftieth, seventy-fifth, and ninetieth percentiles on the searched area of injections of each coincidence type recovered with FAR less than two per day. Values are given in deg^2.

a BNS, NSBH, or BBH as:

$$\begin{aligned} p(\text{astro}) &= 1 - p(\text{Terrestrial}) \\ &= p(\text{BNS}) + p(\text{NSBH}) + p(\text{BBH}). \end{aligned} \tag{5.12}$$

For this purpose, we use a cutoff of $3.0 M_\odot$ as the maximum neutron star mass to define the BNS, NSBH, and BBH regions. In this section, we will briefly summarize the accuracy of the Bayestar skymaps as well as the FGMC p(astro) for injections recovered by GstLAL during the MDC. Detailed information about recent developments to the FGMC p(astro) calculation and a comparison between offline and online p(astro) results for GstLAL events is given in [269].

To quantify the sky localization performance we will consider the skymap searched area and searched probability. The searched area is given in square degrees and represents the area within the credible region containing the injection's true source location. This gives an indication of the accuracy of the skymaps. The searched area can be interpreted as the sky area an electromagnetic observer would have to tile before reaching the true injection sky location assuming they start at the highest probability sky position given by the skymap. Therefore, a smaller searched area is desirable as it means an observer could find the injection with minimal telescope pointings. The searched probability is similarly the probability within the credible region containing the injection's true sky location. We aim to find 90% of injections within the 90% credible region of the true source location. If the searched probability P-P plot lies off the diagonal, we can make the interpretation that there is an inconsistency in the sky localizations, i.e. they may be accurate but lack precision, or vice versa.

Fig. 5.13 shows the cumulative distribution function of the searched area for all injections recovered by GstLAL with FAR less than two per day. We find that the 90^{th} percentile of searched area is 3910 deg^2. This is about 19% of one hemisphere of the sky. Since we know that injections recovered in coincidence by two or three interferometers

will have more accurate sky localizations, we give statistics on the searched areas by coincidence type in Table 5.6. Fig. 5.14 shows the cumulative distribution function of the searched probabilities. We find that the fiftieth, seventy-fifth, and ninetieth percentiles on the searched probability are 0.53 , 0.79 , and 0.93 respectively. The CDF lies very near to the diagonal, indicating that Bayestar is producing well-calibrated sky localizations.

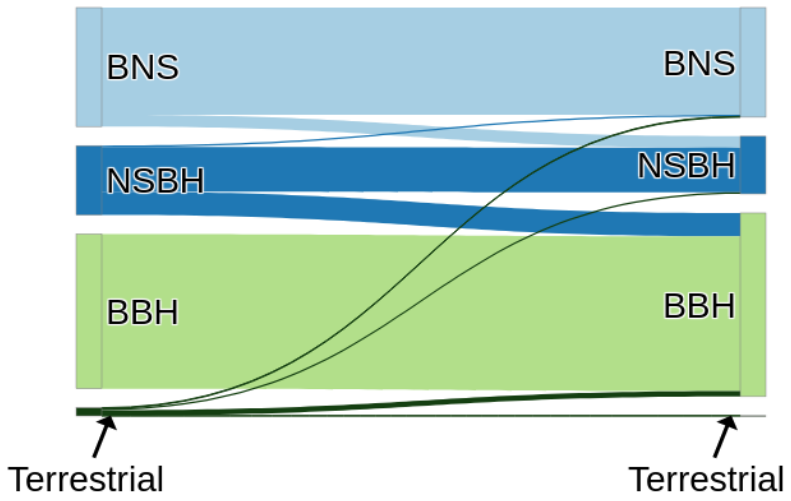

Figure 5.15. Sankey diagram showing p(astro) classification of events uploaded to GraceDB during the MDC with FAR less than two per day. Events with an end time within ± 1 second of an injection are classified as either BNS, NSBH, or BBH using a neutron star mass boundary of $3.0 M_{\odot}$. Events that do not correspond in time with an injection are all classified as terrestrial.

We will briefly discuss the performance of the p(astro) in the MDC using the Sankey diagram in Fig. 5.15. The diagram is read from left to right. The width of each source band on the left corresponds to N_{source}, the number of recovered events in each source class: BNS, NSBH, BBH, and terrestrial. The terrestrial events are any candidates uploaded to GraceDB which do not coincide in time with an injection. The width of each band on the right is the sum of recovered p(source). This diagram gives an indication of the relative misclassification between sources.

For BNS injections, we find that p(BNS) accounts for 90.3% of the recovered probabilities, while p(NSBH) accounts for 9.7%. The recovered BBH and terrestrial probabilities of BNS injections are negligible. This indicates that BNS signals were most commonly mistaken as NSBH. For NSBH injections, we find p(NSBH) is 64.1% of the recovered p(astro), p(BBH) makes up 33.8%, and p(BNS) makes up just 2.10%. There is a sig-

nificant amount of misclassification between BNS and NSBH signals. However, the misclassification is asymmetrical with very few NSBH injections being assigned high p(BNS) whereas many more BNS injections are assigned a high p(NSBH). The large proportion of NSBH signals assigned high p(BBH) is even more concerning as these represent potentially electromagnetically bright signals which might not be followed up by astronomers due to the apparent high probability of originating from a BBH merger. This misclassification is an ongoing area of study and corrections to the p(astro) calculation which mitigate this effect are discussed in [269]. The p(astro) calculation performs very well for BBH classification, with 100% recovered p(BBH).

5.3.3.4 Latency performance

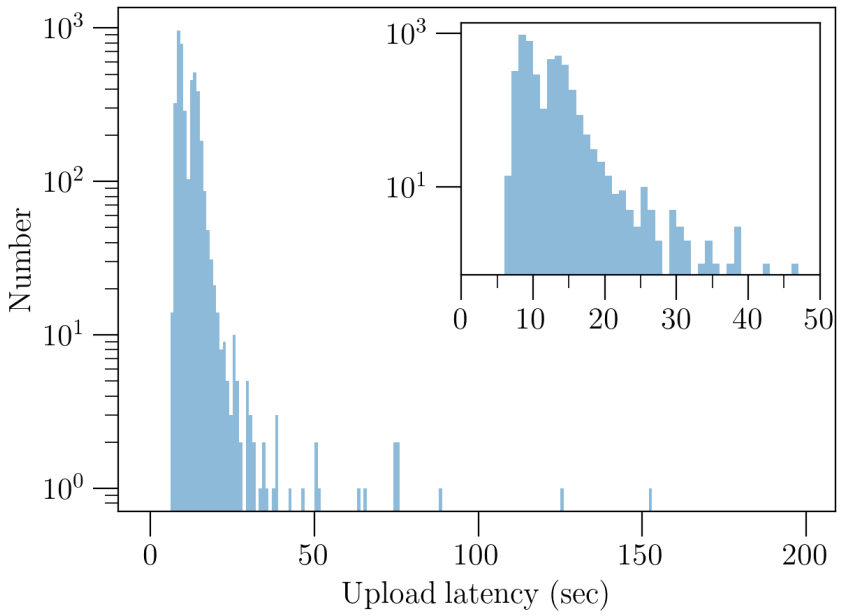

Figure 5.16. Upload latency, defined as the difference between the GPS time of upload to GraceDB and the event coalescence time, in seconds.

Fig. 5.16 shows a histogram of the upload latency for all recovered injections on GraceDB. The upload latency is defined as the difference between the time the event appears on GraceDB and the event coalescence time. The distribution is bimodal as a result of the built-in 4 second geometric wait time between uploads, discussed in Sec. 5.2.1. The first peak in the upload latency distribution is at $8-9$ seconds and the second peak is at $13-14$ seconds. This shows that the GstLAL pipeline is able to keep

up with filtering data in real time and regularly produce gravitational wave candidates within $\mathcal{O}(10)$ seconds of the coalescence time.

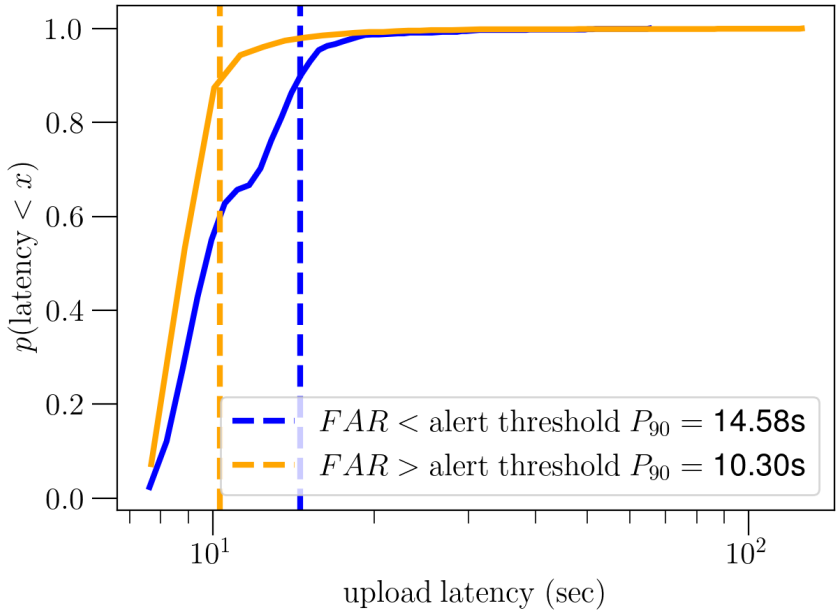

Figure 5.17. Cumulative distribution of event upload latencies. The orange (blue) curve shows the distribution of latencies for the first event uploaded with FAR higher (lower) than the public alert threshold. The dashed lines show the location of the 90^{th} percentile for each distribution.

The 4 second wait time in the event aggregation is implemented in order to reduce the total number of uploads by waiting a long enough time to collect many triggers across background bins before making an upload. Even though there is no built-in wait time for the first event upload, subsequent uploads are necessarily delayed by this method. If an event with FAR below the alert threshold comes after the first upload it can be delayed by at least 4 seconds. Fig. 5.17 shows the cumulative distribution of upload latency for the first events above and below the alert threshold for each superevent in the MDC. We find that the low FAR event uploads are significantly delayed by the event aggregation process. After identifying this issue in the MDC we plan to make changes to the aggregation scheme in order to reduce the latency of alert quality uploads before the start of O4.

5.4 Conclusion

GstLAL is a matched-filtering based gravitational wave search pipeline, which is operated in a low-latency configuration in order to identify signals within seconds of their arrival. We have introduced the `gw-lts` software as a useful auxiliary tool for characterizing the performance of such an analysis in real time. We have presented the performance of the GstLAL pipeline on a mock data challenge consisting of 40 days HLV data from O3 along with an injection campaign of simulated BNS, NSBH, and BBH signals.

Within the MDC data we recover 9 previously published gravitational wave candidates at the one per hour FAR threshold. As only five of these were identified by the GstLAL pipeline in low-latency in O3, we have demonstrated an improvement in the pipeline's signal recovery. We attribute this improvement to several incremental updates to the likelihood ratio computation [111] and the new method of removing signals from the background as introduced in [5]. During the MDC, we find only one candidate which, if uploaded during O4, would likely be retracted. The candidate is identified in only a single detector and we expect that increasing the penalty applied to single detector candidates in the likelihood ratio would reduce our recovery of such spurious signals in the future. We have detailed the results of the injection campaign including efficiency of signal recovery across the parameter space, accuracy of estimated parameters, search sensitivity, sky localization and source classification accuracy, and typical latencies.

The configuration and performance of the GstLAL pipeline as described in Sec. 5.3 is a close approximation to what will be used in the fourth observing run of the LVK Collaboration. However, since the conclusion of the MDC used in this paper, work has been ongoing and several areas for possible improvements have been identified. These changes in configuration and the corresponding improvements in performance are given in Appendix 5.4.

Acknowledgments

The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation (NSF) and operates under cooperative agreement

PHY-1764464. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Netherlands Organization for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors are grateful for computational resources provided by the Pennsylvania State University's Institute for Computational and Data Sciences (ICDS) and the University of Wisconsin Milwaukee Nemo and support by NSF PHY-2011865, NSF OAC-2103662, NSF PHY-1626190, NSF PHY-1700765, and NSF PHY-2207728. This paper carries LIGO Document Number LIGO-P2300124.
M.W.C. acknowledges support from the National Science Foundation with grant numbers PHY-2010970 and OAC-2117997. R.E. is supported by the Natural Sciences & Engineering Research Council of Canada (NSERC).

Software: Confluent Kafka (`https://developer.confluent.io/get-started/python/`), Grafana (`https://grafana.com/docs/`), GraceDB [268], GstLAL [281], `GWCelery` [282], `gw-lts` [283], `igwn-alert` [284], InfluxDB (`https://docs.influxdata.com/influxdb/v2.7/`), `ligo-scald` [285]

Appendix A: GstLAL performance with updated O4 configuration

In this section, we demonstrate the performance of the GstLAL pipeline using an updated configuration. The improved performance in this updated run are attributed to the following areas of development, which proceeded after the conclusion of the MDC presented in Sec. 5.3.

The analytic SNR-ξ^2 signal model used in the likelihood ratio ranking statistic has now been tuned to use a more optimal allowed mismatch region, which is wider for the high SNR region of parameter space. This change is expected to improve the recovery efficiency for very loud signals.

The presence of non-Gaussian noise transients, known as glitches, in the strain data has long been a problem for gravitational wave searches. integrated Data Quality (iDQ) is a machine-learning based algorithm used to assign probabilities of the presence of a glitch in a segment of strain data [88]. In the O3 offline analysis, iDQ was incorporated into the GstLAL ranking statistic as a means of reducing the significance of candidates found during particularly glitchy stretches of data. Although it is no longer used in the

ranking statistic, it is now possible to use iDQ state information as a gate on the strain data, so that segments of data with a high glitch probability will be excluded from the filtering. This should mitigate the negative effects of non-Gaussian data and is expected to result in fewer retraction level candidates and an improved $\langle VT \rangle$. Although the iDQ gate was used in the MDC analysis presented here, the feature is still under development and is not planned for use in the O4 production configuration.

As mentioned in Sec. 5.3.3.3, the p(astro) performance in the BNS and NSBH region of the parameter space was sub-optimal in the MDC, with a significant amount of misclassification between the two source types, as well as NSBH misclassification as BBH. It is imperative that sources including a neutron star are not falsely classified as BBHs since this may discourage astronomers from following up these potentially electromagnetically bright signals. Since the conclusion of the MDC analysis presented in this paper, work has been ongoing toward improving the p(astro) source classification. This effort is discussed in more detail in [269].

Finally, as discussed in Sec. 5.3.3.4, we have made improvements to the event aggregation method to ensure that events with FAR below the public alert threshold will be uploaded with as little latency as possible. By adjusting the geometric cadence factor in the aggregator and removing the wait time for these low FAR events we expect a reduction in the latency of these uploads by up to several seconds.

We re-analyze the final two weeks of the MDC data using the updated pipeline configuration in order to demonstrate the effect on the pipeline performance. The "initial run" presented in Sec. 5.3 was analyzed from 03-15-2023 to 03-28-2023 and the "updated run" was analyzed from 04-24-2023 to 05-07-2023. These runs correspond to identical stretches of O3 replay data with the same injections present in each.

Fig. 5.18 shows the decisive SNR time-series for all injections within the two week span. There are several high SNR injections missed by the initial analysis (upper panel) but found in the updated analysis (lower panel). This indicates an improvement as a result of the widened mismatch region of the SNR-ξ^2 signal model. Table 5.7 compares the injection efficiencies between the two runs. For the BNS and NSBH regions the efficiency is comparable, while there is a 6% improvement in the BBH efficiency from 0.84 in the initial run to 0.90 in the updated run.

Fig. 5.19 and Fig. 5.20 show the improvements in event upload latency between the two runs. The histogram in Fig. 5.19 shows that upload latencies for the updated run are shifted slightly lower overall with respect to the initial run, while the lower edge remains the same. This is expected as the improvements focused on lowering the latency

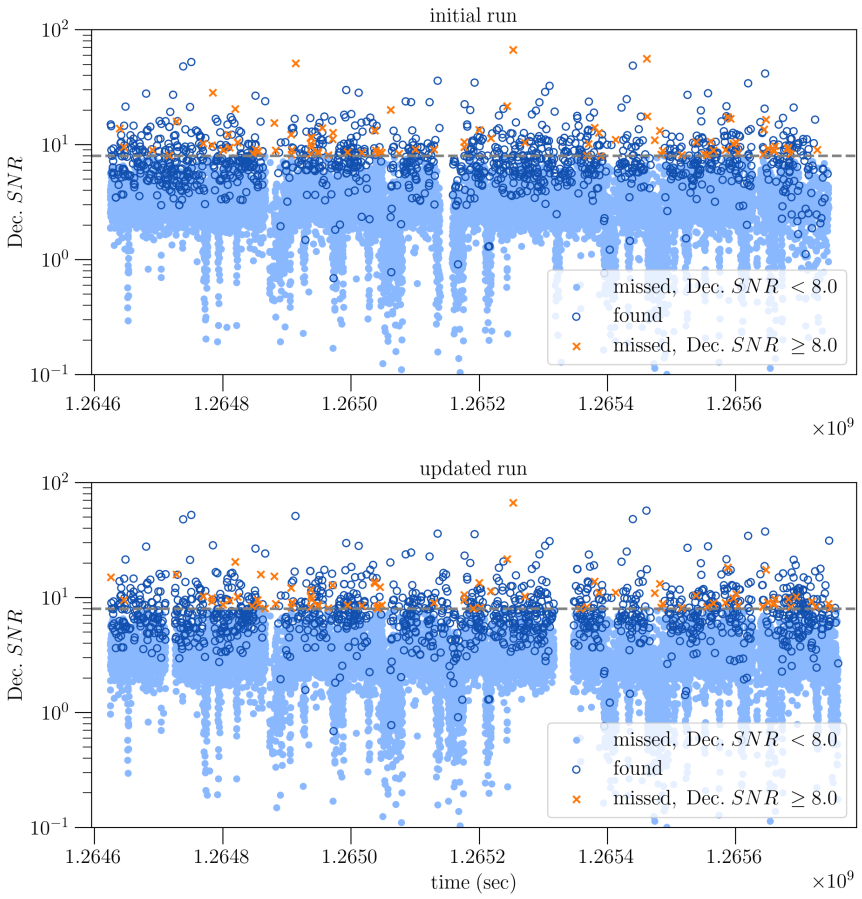

Figure 5.18. Time-series of injected decisive SNR for injections with component masses and spins within the O4 template bank in the initial run, 03-15-2023 to 03-28-2023 (upper panel) and updated run 04-24-2023 to 05-07-2023 (lower panel). Dark blue circles indicate injections that were recovered below a 2 per day FAR threshold. Orange crosses and light blue points indicate injections not recovered below this FAR threshold, where orange crosses are injections with decisive SNR > 8.0. Times on the horizontal axis are GPS times shifted to the original O3 epoch.

of subsequent uploads by reducing the geometric cadence factor from 4 seconds to 2 seconds. Fig. 5.20 shows a similar improvement in upload latency for events below the public alert threshold. We find a 2.54 second reduction in the 90^{th} percentile on upload latency for events with FAR below the public alert threshold, from 14.58 seconds in the initial run to 12.04 seconds in the updated run. This indicates a significant improvement in the latency of alert-quality candidates.

Fig. 5.21 is a Sankey diagram showing the p(astro) classification performance in the updated run. The initial results showed a significant amount of misclassification between BNS and NSBH as well as between NSBH and BBH injections. Here we find

	BNS	NSBH	BBH	ALL
initial run	0.95	0.71	0.84	0.84
updated run	0.93	0.72	0.90	0.86

Table 5.7. Injection efficiencies as defined in Eq. 5.6 computed using a FAR threshold of 2.31×10^{-5} Hz to count "found" injections. The "initial run" is the same as presented in Sec. 5.3 and the "updated run" uses the configuration improvements described in this section. Source categories are defined in Table 5.2 and "ALL" combines injections from the three source categories.

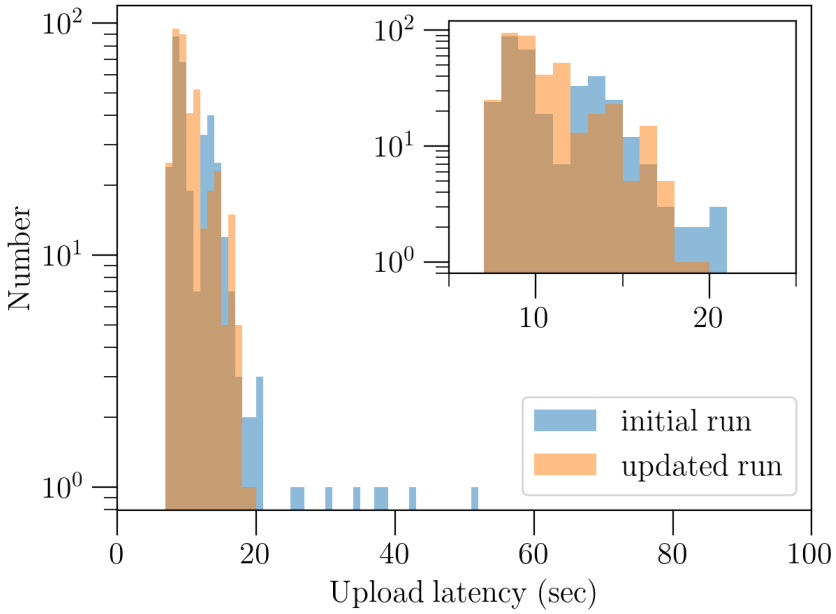

Figure 5.19. Histogram of event upload latencies for the initial run (orange) and updated run (blue).

20.2% of recovered BNS being classified as NSBH and only 6.83% of recovered NSBH being classified as BBH. While the updated run shows more confusion between BNS and NSBH, we see that the confusion between NSBH and BBH is significantly reduced. This is considered as an overall improvement, since the classification more accurately indicates whether a source may be electromagnetically bright, which is more likely the case for BNS and NSBH mergers as opposed to BBH. The reason for this improvement is as follows.

Throughout the p(astro) calculation, misclassification of sources are assumed to be the result of Gaussian noise fluctuations causing a gravitational wave signal to match better with templates that are further away in parameter space than the one that would

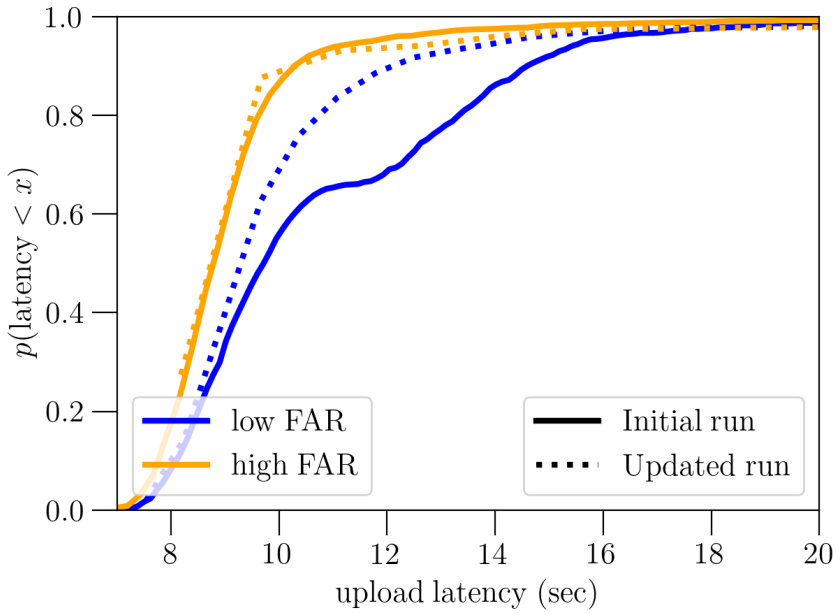

Figure 5.20. Cumulative distribution of event upload latencies for the initial run (solid lines) and updated run (dashed lines). The orange (blue) curves show the distribution of latencies for the first event uploaded with FAR higher (lower) than the public alert threshold. The 90^{th} percentile on the upload latency of high FAR candidates in the initial (updated) run is 10.30 (9.86) seconds. The 90^{th} percentile on the upload latency of low FAR candidates in the initial (updated) run is 14.58 (12.04) seconds.

recover it in the absence of noise [208, 269]. In the initial run, this probability of template migration was estimated while making several simplistic assumptions some of which were later found to be inaccurate. In particular, the geometry of the template bank was modeled as a 4-dimensional unit sphere and the observed SNR assumed to be a single Gaussian distributed real number [208]. In reality however, the geometry of the bank is sufficiently more complex than that of a spherical surface. Furthermore, due to the two-phase matched filtering process, the observed SNR is actually a complex number whose real and imaginary parts can be thought of as independent Gaussians. Marginalizing over the unknown signal phase leads to the absolute value of the SNR having a distribution that is broader than the Gaussian of [208].

Both of these ambiguities can cause an under-estimation of misclassification by the model of ref. [208], leading to the $p(\text{astro})$ calculation confidently classifying, e.g. an NSBH as a BBH even at low SNR. A detailed re-modeling of the misclassification with realistic noise distributions and accurate representations of template bank geometry is beyond the scope of this paper and is part of an ongoing investigation.

Hence, to avoid strongly misclassifying potential electromagnetically bright sources, we instead construct a pessimistic model of noise induced template migration by assuming that the noise degrees of freedom are entirely contained in the signal manifold [269]. Implementing this model in the updated run leads to a significant improvement in the classification of simulated sources that have the potential to be electromagnetically bright. Further improvements in classification are expected to result upon changing the population model in Eq. (5.11) with the true distribution of source parameters [269].

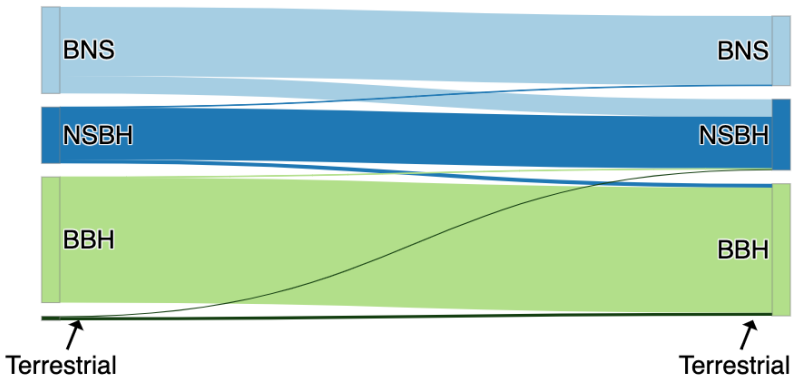

Figure 5.21. Sankey diagram showing p(astro) classification of events uploaded to GraceDB during the updated run from 04-24-2023 to 05-07-2023. Events with an end time within ± 1 second of an injection are classified as either BNS, NSBH, or BBH using a neutron star mass boundary of $3.0 M_{\odot}$. Events that do not correspond in time with an injection are all classified as terrestrial.

Chapter 6
The fourth observing run

6.1 Introduction

The first three observing runs of the Advanced LIGO and Virgo detectors saw the successful beginning of the era of gravitational wave astronomy. This period included the first detections of gravitational waves from a merging binary black hole (BBH) [12], binary neutron star (BNS) [66], and neutron star–black hole binary (NSBH) [115]. As the detectors continue to undergo upgrades, they reach closer to the design sensitivity and the rate of gravitational wave detection is increased substantially. Therefore, with the onset of the fourth observing run we enter a new era of gravitational wave astronomy where we begin to uncover populations of GW sources and study their bulk properties. While there is still much to learn by studying gravitational waves from individual events, we can gain even more insight by observing diverse populations of signals from many different sources. This type of population study will continue beyond O4 into future observing runs and eventually culminate with the operation of next generation detectors such as the Einstein Telescope [75] and Cosmic Explorer [76] which are expected to be sensitive to every merging binary in the observable universe.

The first part of LIGO's fourth observing run (O4a) took place from 24 May 2023 15:00:00 UTC to 16 January 2024 15:00:00 UTC [286]. These eight months of observation saw the detection of 81 new highly significant gravitational wave candidates, each consistent with mergers of neutron star–black hole binaries and binary black holes. These candidates were disseminated to the public via automated GCN notices with typical latencies of $25 - 30$ seconds, measured with respect to the GW merger time. In addition, O4a included over 1500 low significance candidates which were communicated to the public in low-latency to enable broad sub-threshold EM follow up campaigns. We do not identify any significant EM counterparts of the GW candidates in O4a.

This run followed a period of about three years of detector commissioning since the end of the third observing run in March 2020. This commissioning break included several critical upgrades to the LIGO Hanford and Livingston detectors resulting in a significantly improved sensitivity. This is discussed in Sec. 6.2. During O4a, the Virgo detector was undergoing continued commissioning and did not participate in the observing run, however it is planned to join the run in the second part of the fourth observing run (O4b) [286]. The KAGRA detector joined O4 for a short period in the beginning of the run, however since its sensitivity only reached a BNS inspiral range of ~ 1 Mpc, the data was not analyzed by the participating search pipelines [59, 60, 286].

In addition to detector upgrades, the break between O3 and O4 allowed for continued development of the search pipelines and other analyses which produce metadata associated with GW candidates in low-latency. Particularly, the developments of the GstLAL search pipeline towards O4a, which was a major portion of my dissertation work, is discussed in Chapter 5 [5, 107, 111, 113]. While GstLAL has allowed for candidates identified in only a single detector since O2 [104], several of the other modeled search pipelines expanded functionality to allow for single detector candidates for the first time in O4. This development was crucial for O4a, since with only the two LIGO detectors in operation the coincident duty cycle was naturally less than in previous observing runs when Virgo also participated.

Methods for early warning detection of gravitational waves from BNS systems have been under development for many years [92, 287–289]. Several of the modeled search pipelines, including GstLAL, operated an early warning configuration for the first time in O4a. These analyses aim to enable rapid EM follow-up campaigns towards multi-messenger detection of gravitational waves and short GRBs. The prospects for early warning detection and the localization of such candidates on the sky depends heavily on the detector network in operation: the number of detectors participating and their sensitivities as well as the joint duty cycle. Given these factors, the number of early warning detections expected in O4 may be $\lesssim 1$ per year [290, 291] and indeed these searches did not produce any significant detections in O4a. However the development of this infrastructure is still crucial to enable multi-messenger astronomy and operating these searches will ensure that we are prepared for the next BNS merger event.

In this chapter I will focus on results from the *all sky* searches for gravitational waves in O4a, i.e. searches for gravitational waves consistent with CBC signals from merging neutron stars and black holes in the range of $\sim 1 - 200 M_{\odot}$. As a member of the LVK Collaboration, much of my research time was dedicated to the operation of the GstLAL

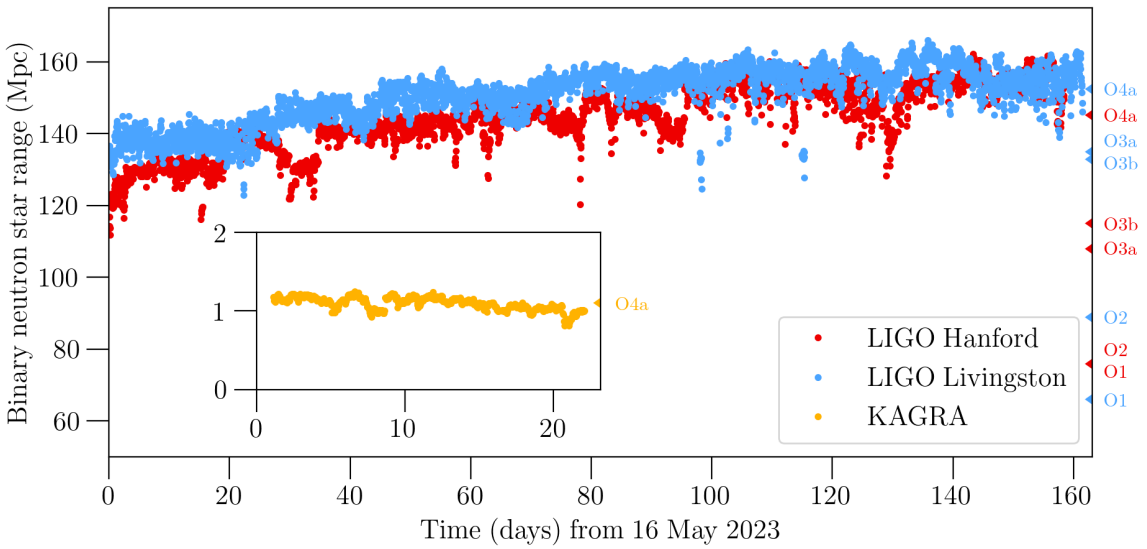

Figure 6.1. Estimated median BNS inspiral range of LIGO Hanford (red), LIGO Livingston (blue), and KAGRA over one hour time segments in O4a. The KAGRA data is shown in the inset axis as it was only observing for the first ~ 20 days of the run and its sensitivity is much less than that of the LIGO detectors. O4a data is taken from the public GstLAL inspiral range estimated by the GW Interferometer Status (GWIStat) [292]. The markers on the right axis indicates median values for each of the previous observing runs, O1 – O3b, as measured by the SenseMonitor [293] BNS inspiral range. Although the GstLAL inspiral range and the SenseMonitor range differ in their technical implementations, they are expected to agree to within a few percent.

low-latency search pipeline and my work directly contributed to the many gravitational wave candidates identified to date. In Sec. 6.3 I will give an overview of the public results from O4a, focusing on the performance of the GstLAL search pipeline in Sec. 6.4. I will conclude in Sec. 6.5 with an outlook towards the second part of the fourth observing run, O4b.

6.2 Detectors

In O4a, both the LIGO Hanford and Livingston detectors were observing, while the Virgo detector was commissioning [286]. KAGRA joined the run for a period of about 3 weeks, followed by a period of extended commissioning. There were several key upgrades made to the LIGO detectors between O3 and O4 which led to significantly improved sensitivity and data quality. The sensitivity throughout O4a, as measured by the BNS inspiral range is shown in Fig. 6.1 for LIGO Hanford and Livingston [59, 60, 171, 172].

The median inspiral range over the duration of the run was ~ 145 Mpc and ~ 155 Mpc for LIGO Hanford and Livingston, respectively. Further, the increasing trend in the range demonstrates continual improvements over the duration of the run.

The sensitivity improvements are owed to increased input power and the implementation of frequency dependent squeezing [294]. Increasing the laser power leads to a reduction in high frequency shot noise [37]. In O3b, the LIGO Hanford and Livingston detectors operated with 34 W and 40 W, respectively [3]. Ahead of O4a, the input power in the detectors was nearly doubled to 60 W and 68 W for LIGO Hanford and Livingston, respectively [294]. Further, since any defects in the test masses can cause diminishing returns on sensitivity gains from the increased power, test masses at both detectors were replaced ahead of the observing run.

Quantum noise presents a fundamental limit for the LIGO and Virgo detector sensitivity in a broad frequency band from $\mathcal{O}(10 - 10^3)$ Hz [37, 39, 50]. This manifests as shot noise at high frequencies (above about 80 Hz) and radiation pressure noise at low frequencies (below about 80 Hz). Prior to O3, frequency *independent* squeezing was implemented which reduced the quantum noise at high frequencies (due to uncertainty in the phase quadrature of the electromagnetic field) at the cost of increased quantum noise at low frequencies (due to uncertainty in the amplitude quadrature) [175, 181, 182, 295]. The development and implementation of frequency *dependent* squeezing in the break between O3 and O4a allowed for the reduction of quantum noise across the bandwidth. With this improvement, the low frequency radiation pressure noise is significantly reduced. This led to an improvement in the BNS inspiral range of $15 - 18\%$ [53].

Reduction of technical noise was another goal for O4a. Particularly, further improvements were made to reduce noise artifacts caused by light scattering [294]. This is one of the main sources of glitches in LIGO data, therefore these upgrades led to improved data quality.

6.3 Public results

In O4a, there were six search pipelines operating in low-latency. These included the four modeled searches: GstLAL [111, 113], MBTAOnline [196], PyCBC Live [296] and SPIIR [200] as well as two unmodeled searches: cWB [194, 297] and oLIB [298]. The modeled searches broadly look for gravitational wave signals consistent with mergers of neutron stars and black holes, covering component mass ranges of $\sim 1 M_\odot$ for the lightest neutron stars to $\mathcal{O}(100) M_\odot$ for heavy black holes. The unmodeled searches do not make

	Low Significance	High Significance	
		modeled searches	unmodeled searches
FAR Threshold	2 per day	1 per 5 months	1 per 4 years
Num. Candidates	1624	88	1

Table 6.1. Summary of alerts in O4a from all sky searches. There were three FAR thresholds: 2 per day for low significance alerts, 1 per 5 months for high significance alerts from modeled searches, and 1 per 4 years for high significance alerts from unmodeled searches. Throughout O4a there were 1624 low significance alerts, 88 high significance alerts from modeled searches, and 1 high significance alert from an unmodeled search.

assumptions about the waveform morphology and in principle may be sensitive to a wide array of potential short-duration transient GW sources. All searches upload gravitational wave candidates to GraceDB[1], which are grouped together by their coalescence times into *superevents*. As mentioned in Sec. 4.4, the superevents are assigned an identifier with naming convention SYYMMDD indicating the date of upload. There were multiple thresholds used for sending public alerts in O4. These are outlined in Table 6.1 [156].

For the first time in O4a, the LVK Collaboration released *low significance* alerts for candidates which passed a FAR threshold of 2 per day. For each of these candidates, an automated GCN notice was produced. The notices include trigger timing and FAR estimates, sky localization estimates from Bayestar [157], and for triggers originating from modeled searches, source classification and EM counterpart probabilities [270, 299]. There were 1624 low significance alerts produced throughout the duration of O4a.

The FAR threshold for *high significance* alerts depends on the search which produced the candidate. For candidates originating from modeled searches, i.e. candidates expected to be consistent with CBC sources, the public alert FAR threshold was 1 per 5 months. This FAR threshold includes a trials factor on the number of searches. For CBC candidates, the trials factor was $4 + 1$ [156] which accounts for the four modeled search pipelines plus The Rapid, on-source VOEvent Coincident Monitor (Raven)[2], which searches for coincidences between GW candidates and external EM triggers. FARs estimated by the individual search pipelines are multiplied by the trials factor before being compared to the threshold of 1 per 5 months. This means that in order to achieve the desired rate of false alarms, i.e. one per five months, the effective false alarm rate per pipeline is one per ten months. There were 88 high significance alerts corresponding to CBC candidates in O4a.

[1] `https://gracedb.ligo.org/`

[2] `https://lscsoft.docs.ligo.org/raven/`

For candidates originating from the unmodeled searches, cWB and oLIB, the public alert FAR threshold was 1 per 4 years. This includes a trials factor of $2 + 1$ accounting for the two searches and Raven [156]. This more strict FAR threshold was chosen to be more conservative to reflect our uncertainty about as yet undetected sources for transient GW bursts. There was 1 high significance alert associated with a candidate from the unmodeled search cWB in O4a, however as described below this candidate was subsequently determined unlikely to be astrophysical in origin and retracted.

For high significance alerts, an initial automated GCN notice was produced similar to that for the low significance alerts, however these alerts are additionally followed up with in depth source parameter inference, manual vetting by the Rapid Response Team (RRT), and detailed data quality checks [270]. The parameter inference is used to produce updated sky localizations which are released in subsequent GCN notices. Once the candidate is validated by human vetting, a GCN circular is produced to announce the detection.

If the candidate is determined to be instrumental or terrestrial in origin, the RRT can choose to retract it to reflect our doubt that it has an astrophysical origin. This is announced via GCN notice. Retractions are expected in low-latency gravitational wave detection for a few reasons. First, the LIGO detectors are highly dynamic instruments; as opposed to the stable operation of many EM telescopes, scientists at the LIGO sites are constantly tweaking the detector configuration, even throughout an observing run. This can lead to great successes as indicated by the continual improvement in sensitivity throughout the observation run, shown in Fig. 6.1. However, this also means that the detector noise characteristics are subject to change as often as on weekly timescales. This can lead to new and unexpected noise sources or glitch classes at any time, posing a problem for GW search algorithms in characterizing the background. Secondly, many retractions come from times when only a single detector was operating (or when only a single detector participated in the detection). Historically, GW searches did not allow for such single-detector candidates and coincidence among at least two detectors was a requirement. If GW searches can accurately identify signals in a single detector we gain a significant improvement in duty cycle and corresponding detection rate. Notably, the GstLAL configuration which allowed for ranking of single detector candidates in O2 was crucial in the low-latency detection of GW170817 [66]. The ability to identify single-detector candidates has clear science outcomes, albeit at the risk of greater uncertainties in significance and a higher possibility of false alarms.

Still, one of the performance goals in O4a was to reduce the number of retractions. In

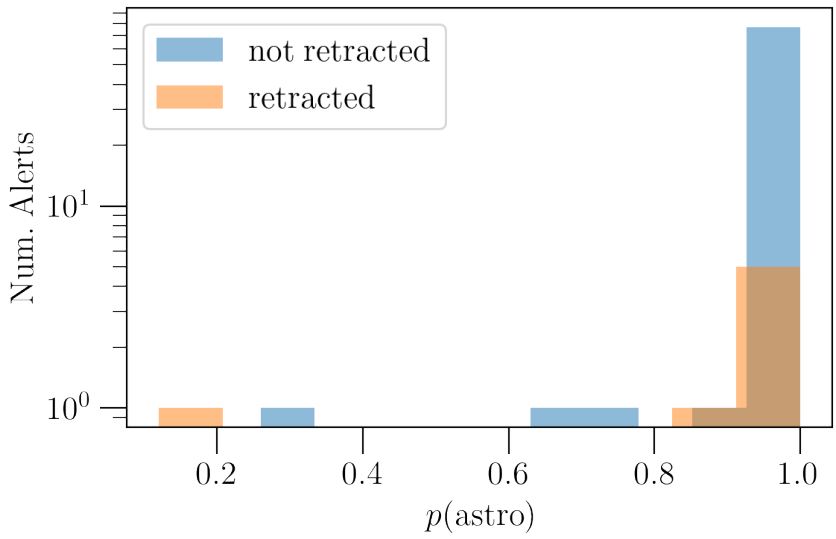

Figure 6.2. Histogram of p(astro) values for the 88 high significance alerts in O4a. The 7 retracted candidates are shown in blue; the remaining 81 are shown in orange.

O4a, there were 7[3], compared to 8 in O3a and 16 in O3b. Among these, S231112ag [304], S230712a [305], and S230708bi [306] were identified only by GstLAL. These will be discussed in more detail in Sec. 6.4. S230830b [307] and S230715bw [308] were identified below the FAR threshold only by SPIIR. Both candidates were identified as coincident in LIGO Hanford and Livingston data and retracted due to data quality issues. S230808i [309] was identified by the cWB pipeline in LIGO Hanford and LIGO Livingston data. This was the only high significance alert from an unmodeled search in O4a. Despite having a low FAR of 6.9×10^{-11} Hz (about one per 500 years), it was retracted due to being considered as having low significance by pipeline experts. S230622ba [310] was identified by MBTA with a modest significance in LIGO Hanford and Livingston data. The FAR was estimated as 5.18×10^{-8} Hz (about $1-2$ per year). The candidate was retracted after data quality checks identified the presence of a glitch near the signal.

6.3.1 Source classification

In addition to the FAR of public candidates, we compute and release the p(astro) as a measure of the candidate significance. The p(astro) is computed independently by each search pipeline, and as mentioned in Chapter 4 this calculation requires an estimate of

[3]Here, I am only considering the alerts from all sky searches. This excludes 4 retracted Early Warning candidates [300–303].

the background of noise triggers as well as the foreground signal distribution [205–210]. . Following [3], we can estimate the purity of a set of candidates by summing the $p(\text{astro})$ and dividing by the total number of candidates. This gives an estimate of the proportion of candidates in the set we can expect to be astrophysical, or conversely the fraction of contamination by terrestrial artifacts. The purity of O4a public alerts is estimated as:

$$\text{purity} = \frac{\sum_{N-R} p(\text{astro})}{N}, \tag{6.1}$$

where $N = 88$ is the total number of public alerts and $R = 7$ is the number of alerts which were retracted in low-latency. We do not include the retracted candidates in the $p(\text{astro})$ sum as we do not expect these candidates to originate from astrophysical sources. This can also be thought of as taking $p(\text{astro}) = 0$ for retracted candidates. We find a purity of 0.88 in O4a, or correspondingly a contamination fraction of about 0.12, indicating ~ 10 noise candidates. These include the 7 retracted candidates and ~ 2 candidates likely to originate from noise in the remaining 81.

Fig. 6.2 shows a histogram of the $p(\text{astro})$ values for all high significance alerts. The figure shows that most of the non-retracted candidates had $p(\text{astro}) \sim 1$, indicating a high likelihood of being astrophysical in origin. Although, there are several low $p(\text{astro})$ outliers which may be the most likely candidates to have originated from noise, as expected from the contamination fraction above. We find that many of the retracted candidates had high $p(\text{astro}) > 0.80$. This suggests that these candidates were caused by rare noise artifacts which our background and foreground models were not able to account for.

Information about the masses of GW candidates is not released to the public in low-latency. Still, it is important for astronomers to have an insight to the sources of the public GW candidates so that they can prioritize their observations according to which candidates are most likely to have EM counterparts. To this end, we provide the mulit-component $p(\text{astro})$ as a course-grained mass estimate [205, 206, 269, 311]. This formalism assumes that the occurrence of noise and astrophysical events are independent Poisson processes. In O4a, we take a conservative boundary of $3M_\odot$ to distinguish neutron star components and black holes components [156]. That is, $p(\text{BNS})$ represents the probability that a candidate has both masses $< 3M_\odot$, $p(\text{BBH})$ represents the probability that a candidate has both masses $> 3M_\odot$, and $p(\text{NSBH})$ represents the probability that a candidate has only its primary mass $> 3M_\odot$. By summing these probabilities over all the 81 not-retracted high significance candidates in O4a, we can find estimates of

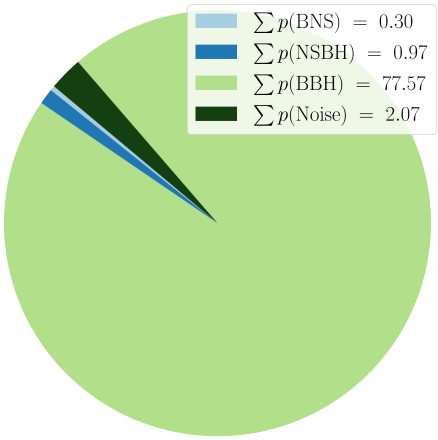

Figure 6.3. The $p(\mathrm{astro})$ source classification for the 81 non-retracted candidates in O4a. Each slice represents the $p(\mathrm{source})$ over all candidates, where source $\in \{\mathrm{BNS}, \mathrm{NSBH}, \mathrm{BBH}, \mathrm{Noise}\}$. The source categories BNS, NSBH, BBH are constructed based on a neutron star mass cut of $3M_{\odot}$.

the number candidates with sources of each category. Fig. 6.3 shows a pie chart of the multi-component $p(\mathrm{astro})$. We find that most of the candidates are consistent with BBH origins, ~ 1 may be consistent with an NSBH origin, ~ 2 are likely to originate from noise, and probably none originate from a BNS. The candidate which was most likely to originate from an NSBH was S230529ay [312] (now, GW230529 [63]). This candidate is discussed in more detail in Sec. 6.3.3 along with a highly significant NSBH candidate, S230518h [313], which was detected during the engineering run ahead of O4a.

6.3.2 Sky localization

The sky localization capabilities in O4a suffered from having only two detectors in the observing network. As mentioned previously, the Virgo detector was commissioning during O4a so only LIGO Hanford and LIGO Livingston could be used to to localize candidates. This led to overall worse sky localizations than what was observed in previous observing runs, namely O3. The CDF of 90% credible sky area for all 81 high significance not retracted candidates in O4a is shown in Fig. 6.4. The distributions highlight the difference in sky localization capabilities when a candidate is identified in a single detector versus in a pair of detectors.

The best localized candidate in O4a was S230627c [314] with a 90% credible sky area of 82 deg^2. The worst localized candidates in O4a were all identified in only a single detector, as indicated by Fig. 6.4. This is expected and comparable to the previous

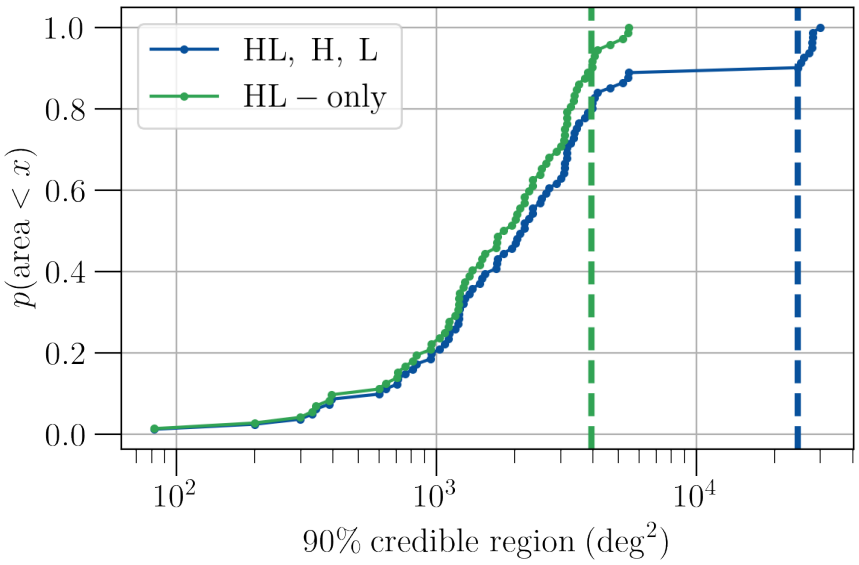

Figure 6.4. CDF of 90% credible area for the 81 not retracted high significance alerts in O4a. The blue curve includes candidates identified in all combinations of the detector network, including those candidates only identified in a single detector – LIGO Hanford or LIGO Livingston. The green curve excludes single detector candidates, showing only candidates identified in *both* LIGO Hanford and LIGO Livingston. The 90$^{\text{th}}$ percentiles for each distribution are indicated by the vertical dashed lines at $\sim 24,500$ deg^2 and ~ 3950 deg^2 for all candidates and only HL candidates, respectively.

observing run [3].

In addition to the 2D sky localization, the 3D localization is an important quantity for host galaxy identification. The 3D localization depends on the estimated luminosity distance. The distribution of all estimated luminosity distances for high significance not retracted candidates in O4a is shown in Fig. 6.5. The luminosity distances observed in O4a are consistent with those in previous observing runs. However, these are low-latency estimates and certainly subject to change following detailed parameter inference offline. To date, the closest candidate observed by the LVK Collaboration is still GW170817 at only ~ 40 Mpc. The closest candidates in O4a were probably GW230529 [63, 312], S230627c, and S230814ah [315] at 197 ± 62.0 Mpc , 291 ± 64.0Mpc , and 330 ± 105Mpc , respectively. These distances are comparable to that of GW200115 [115]. The furthest candidates observed in O4a were probably S230922q [316] and S231119u [317] with luminosity distances of 6.65 ± 2.35Gpc and 6.60 ± 2.56Gpc , respectively. These are most likely closer than GW190403, which is probably the farthest source in GWTC-3 to date [3, 138].

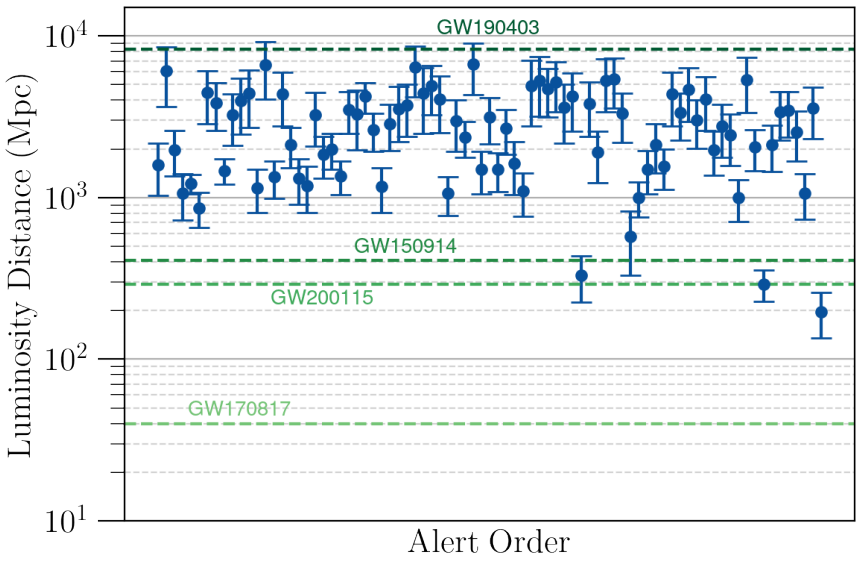

Figure 6.5. Estimated luminosity distances in Mpc for the 81 not retracted high significance alerts in O4a. The candidates are arranged on the horizontal axis by their order of detection. The blue markers indicate the mean luminosity distance estimate with errors bars indicating ± 1 standard deviation. For context, the estimated luminosity distances of notable events from previous observing runs are indicated: GW170817 – the first BNS detection and closest observed source at 40.0 Mpc [66], GW200115 – the first confident NSBH detection at 290 Mpc [115], GW150914 – the first detection of GWs from a BBH merger at 410 Mpc [12], and GW190403 – probably the furthest event detected to date [138].

6.3.3 NSBH candidates

There were three candidates which had significant probabilities of originating from an NSBH source. The first of these, S230518h [313], was observed during the engineering run prior to the start of O4a in both LIGO Hanford and LIGO Livingston data. Despite this, the data around the time of the event was confirmed to have good quality and there is no evidence of excessive noise which would cast doubt upon the astrophysical nature of the event. This candidate had p(NSBH) of 0.86 , which was the highest among all candidates in O4a. However, it was identified as having < 0.01 probability of having remnant mass outside the final compact object [299]. This is a useful metric to indicate the likelihood that the candidate has an EM counterpart. The negligible p(Has Remnant Mass) is consistent with the lack of coincident EM observations for this event. This is not unexpected for NSBH events as the presence of EM emission will be highly dependent on the source intrinsic properties [219–221, 318–322]. If the mass ratio is too high, the neutron star may plunge directly into the black hole without producing

any EM counterpart.

The candidate GW230529ay [63, 312] was observed in LIGO Livingston data with a FAR of 6×10^{-3} per year. It had a p(NSBH) of 0.62 , making it the most likely NSBH candidate in O4a (after the engineering run candidate, S230518h). This candidate was identified as having masses $3.6^{+0.8}_{-1.2}$ and $1.4^{+0.6}_{-0.2}$ $M_\odot$ [63]. This is the first observation of a compact object with mass confidently constrained to the lower mass gap, $3 - 5$ $M_\odot$. Historically it has not been clear whether compact objects can exist in this mass range, being more massive than typical neutron stars and lighter than observed x-ray black holes [166, 167, 235, 237, 323, 324]. This observation definitively fills the lower mass gap and parameter estimation finds that the primary mass is most likely a very light black hole. The existence of such a source will have implications for expected NSBH merger rates, the BH mass distribution, and the fraction of NSBH mergers expected to produce EM counterparts. Particularly, such a low mass black hole merging with a neutron star may have a higher probability of producing EM emission than a system with a more extreme mass ratio [219–221, 318–322]. However, in the case of GW230529 only LIGO Livingston was observing at the time and therefore the candidate's 90% credible sky area was very large at $24,534$ deg^2 . This large sky area would have made it difficult for many astronomers to perform detailed follow up efforts and indeed no EM counterpart was identified.

Finally, S230627c [314], which was the best localized candidate in O4a, had relatively high p(NSBH) of 0.49 , making it an event of particular interest for EM follow up. This candidate was identified by all six search pipelines in LIGO Hanford and LIGO Livingston data. It had a very low FAR of 0.01 per year. Despite the high p(NSBH), it was identified as having a < 0.01 probability of including at least one neutron star, p(Has NS). It is likely that S230627c's source is a BBH with a secondary mass in the lower mass gap, but it is not possible to make a conclusive statement.

6.4 GstLAL performance

As described in Chapter 5, a significant portion of dissertation work has involved the preparation of the GstLAL inspiral search pipeline towards O4a [113]. This work culminated in the incredible performance of the pipeline throughout O4a. The main goals of the GstLAL search pipeline ahead of O4a were as follows: (i) demonstrate state of the art sensitivity across the CBC parameter space, (ii) produce high purity results by reducing false alarms, and (iii) achieve extremely low latencies to enable multi-messenger

SID	Inst.	FAR (yr^{-1})	p(BNS)	p(NSBH)	p(BBH)	p(terrestrial)	GCN
S230630am	HL	2.4×10^{-8}	< 0.01	< 0.01	0.98	0.02	[325]
S230704f	HL	2.8×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[326]
S230706ah†	HL	4.3×10^{-8}	< 0.01	< 0.01	0.97	0.03	[327]
S230708t	HL	4.3×10^{-8}	< 0.01	< 0.01	0.97	0.03	[328]
S230708cf	HL	1.6×10^{-8}	< 0.01	< 0.01	0.99	0.01	[329]
S230726a	L	3.8×10^{-14}	< 0.01	< 0.01	> 0.99	< 0.01	[330]
S230729z	HL	3.4×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[331]
S230802aq	H	2.2×10^{-8}	< 0.01	0.06	0.90	0.03	[332]
S230807f	HL	7.1×10^{-8}	< 0.01	< 0.01	0.95	0.05	[333]
S230814ah	L	1.8×10^{-21}	< 0.01	< 0.01	> 0.99	< 0.01	[315]
S230822bm	HL	2.6×10^{-8}	< 0.01	< 0.01	0.98	0.02	[334]
S230825k	HL	2.4×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[335]
S230831e	HL	2.0×10^{-8}	< 0.01	< 0.01	0.98	0.02	[336]
S230911ae	H	1.9×10^{-12}	< 0.01	< 0.01	> 0.99	< 0.01	[337]
S230930al	HL	7.4×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[338]
S231005ah	HL	2.0×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[339]
S231008ap	HL	1.5×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[340]
S231014r	HL	1.0×10^{-8}	< 0.01	< 0.01	> 0.99	< 0.01	[341]
S231020bw	HL	3.5×10^{-10}	< 0.01	< 0.01	> 0.99	< 0.01	[342]
S231029y	L	2.2×10^{-10}	< 0.01	< 0.01	> 0.99	< 0.01	[343]
S231113bb	HL	5.6×10^{-8}	< 0.01	< 0.01	0.96	0.04	[344]
S231119u	HL	7.4×10^{-8}	< 0.01	< 0.01	0.95	0.05	[317]
S231129ac	HL	1.8×10^{-8}	< 0.01	< 0.01	0.99	0.01	[345]
S231231ag	H	8.4×10^{-15}	< 0.01	< 0.01	> 0.99	< 0.01	[346]
S240104bl†	H	3.6×10^{-17}	< 0.01	< 0.01	> 0.99	< 0.01	[347]
S240109a	H	7.3×10^{-9}	< 0.01	< 0.01	> 0.99	< 0.01	[348]

Table 6.2. High significance candidates in O4a found significantly only by the GstLAL search pipeline. Candidates in **bold** are those for which GstLAL was the only search pipeline to contribute a trigger; other candidates were found by additional pipelines but with FAR above the high significance threshold, i.e. 1 per 5 months. The two candidates indicated with a dagger (†) were identified by only one of the GstLAL checkerboarded analyses. The instruments observing at the time of the candidate are indicated using single-letter identifiers (H for LIGO Hanford and L for LIGO Livingston). These are not necessarily the same instruments that contributed triggers associated with the candidate. The FARs are given in inverse years. The last column gives the GCN circular.

followup.

The GstLAL search pipeline contributed a significant trigger to 80 out of the 81 high significance not retracted candidates. This demonstrates the sensitivity of the search throughout the parameter space. The only high significance not retracted candidate which was not found significantly by GstLAL was S231123cg [349]. This candidate was identified with high significance by the modeled search, PyCBC Live and unmodeled search, cWB. It was identified in coincidence in LIGO Hanford and LIGO Livingston data with a FAR of about 0.01 per year and a $p(\mathrm{BBH})$ of > 0.99 . The corresponding GstLAL trigger had marginal significance, with FAR between the thresholds of 2 per day and 1 per 5 months. If astrophysical, this candidate has parameters which are likely outside of the GstLAL template bank. This is currently thought to be the reason why the GstLAL candidate had lower significance. The significance as estimated by the GstLAL search pipeline may be expected to fluctuate as a result of further offline GstLAL analyses, particularly re-ranking candidates against the global O4a background. Therefore, it remains to be seen whether GstLAL will identify this candidate as part of the next catalog update, GWTC-4.

Among the 80 candidates identified with high significance by the GstLAL search pipeline, 26 were found significantly *only* by GstLAL. These are listed in Table 6.2 along with the instruments observing at the time of each candidate, their FARs and $p(\mathrm{astro})$ source classifications as estimated by GstLAL, and the corresponding GCN circulars. As indicated in bold in the table, these include 9 candidates where GstLAL was the only pipeline to contribute a trigger. These are all candidates identified in only a single detector (with the exception of S231020bw). Other modeled pipelines (e.g., PyCBC Liveand MBTAOnline) allowed single detector candidates for the first time in O4a. However, these pipelines place a constraint on the signal duration for single detector candidates, i.e. only signals longer than some minimum duration will be considered in single detector time. This is done to reduce the chance of falsely identifying a glitch as a signal, since glitches are typically short in duration and difficult to distinguish from high mass BBH signals, especially without the test of coincidence. On the other hand, the GstLAL search pipeline makes no minimum duration cut on single detector candidates. Therefore, as indicated by the $p(\mathrm{BBH})$ in Table 6.2 all of these are consistent with BBH mergers.

The remaining 17 candidates in Table 6.2 had triggers from the various other pipelines, but none which met the high significance threshold of 1 per 5 months. These candidates were all identified as coincident signals in LIGO Hanford and LIGO Livingston data.

Again, all of these candidates are consistent with originating from BBH mergers. This excess of BBH signals detected by GstLAL will certainly inform our understanding of black hole mass and spin distributions as well as merger rates in subsequent publications.

Towards the second goal of reducing false alarms, I will discuss the three GstLAL retractions in O4a: S231112ag [304], S230712a [305], and S230708bi [306]. In O3a and O3b respectively, there were 6 and 8 retracted GstLAL candidates. Therefore, compared to the 3 in O4a, the retraction rate has reduced significantly from ~ 1 (~ 1.6) per month in O3a (O3b) to only ~ 0.3 per month in O4a.

S231112ag [304] and S230708bi [306] were both identified as single detector candidates in LIGO Livingston data even though LIGO Hanford was also observing at the time. Given its signal parameters and the lack of coincidence, S231112ag was difficult to distinguish from a glitch and thus retracted. The data near S230708bi was determined to be consistent with a known class of instrumental artifacts. It is likely that both of these candidates were consistent with rare instrumental artifacts which were not captured in the background histograms at the time of their detections. However, being single detector candidates during times when two detectors were observing, both of these triggers were included in the background histograms after their own detection. This means that subsequent similar candidates would be down-weighted in significance given their representation in the background. This demonstrates how the GstLAL search pipeline *learns* detector noise properties throughout its low-latency operation. Both of these candidates are expected to be heavily down-weighted in significance when ranked against the global O4a background offline.

S230712a [305] was identified in both LIGO Hanford and LIGO Livingston data. This candidate was identified during a time of instability in the GstLAL search pipeline due to transient computing failures. The trigger is thought to have been produced at a time when the PSD estimation had not yet settled (i.e. $\mathcal{O}$(sec) after job start-up, see Chapter 2 – Sec. 2.2 and Sec. 2.7) which falsely inflated its significance. Further, the trigger was not reproducible when re-analyzing the surrounding data in an offline mode. Therefore the candidate was determined to be caused by a rare computational instability and not consistent with a signal.

Towards the goal of producing extremely low-latency alerts in O4a Fig. 6.6 shows a histogram of reported latencies of GstLAL candidates compared to candidates from the other search pipelines. The latency is defined as, $t_{\mathrm{GDB}} - t_c$, i.e. the difference between the measured coalescence time and the GPS time at which the candidate is uploaded to GraceDB. The minimum upload latency from GstLAL was 12.1 sec compared to 14.2

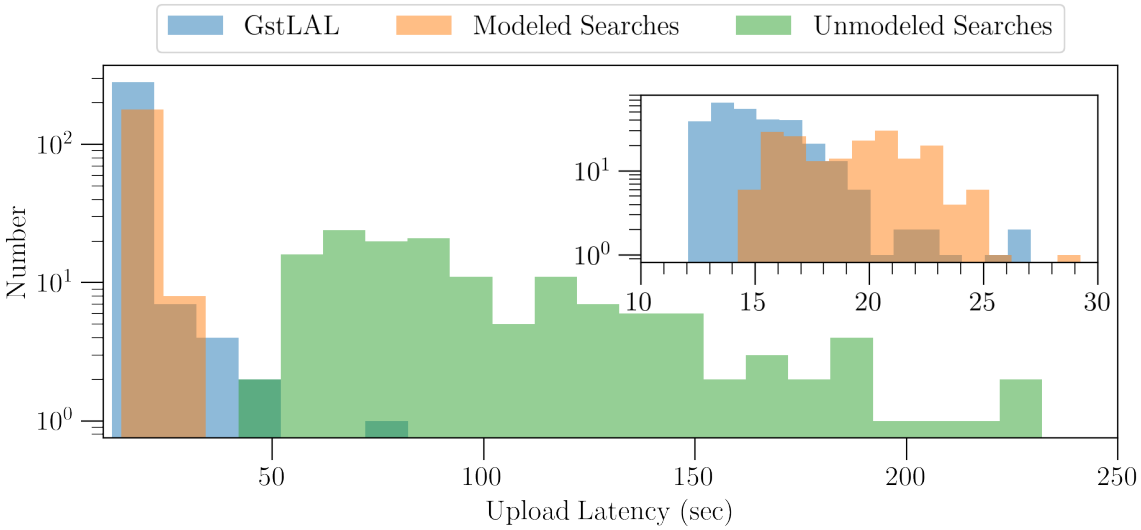

Figure 6.6. Upload latency (defined as the difference between the event coalescence time and the GPS time of upload to GraceDB. GstLAL latencies are shown in blue, all other modeled search latencies (i.e. PyCBC Live, MBTAOnline, and SPIIR) are shown in orange, and unmodeled search latencies (i.e. cWB and oLIB) are shown in green. There are outliers above 250 seconds in each distribution which are not shown. The inset axis compares GstLAL latencies to other modeled search latencies in the range of $10 - 30$ seconds.

sec for all other modeled search pipelines. This indicates that GstLAL was often the first search pipeline to report a candidate event. However, the maximum upload latencies of 87.4 and 28.8 sec for GstLAL and other modeled search pipelines respectively show that GstLAL has a larger spread in its latencies. This is a known feature of the pre-aggregation across background bins done by the pipeline (see Sec. 2.7 and Sec. 5.3.3.4) [113].

As described in Sec. 5.1 of Chapter 5 [113] and [107], the GstLAL search pipeline was operated as two analyses on separate computing clusters with a *checkerboarded* template bank. These were affectionately nicknamed as the "Edward" analysis which ran on the computing cluster at the California Institute of Technology and the "Jacob" analysis which ran on the computing cluster at the Pennsylvania State University's Institute for Computational and Data Sciences (ICDS). The construction of the bank using the binary tree geometric placement method of [58] makes it natural to split the full template bank into two redundant *checkerboards*, i.e. each checkerboard has distinct templates but both cover the full parameter space with similar effectualness.

We nominally expect both checkerboarded analyses to produce the same candidate list. However, differences can arise due to transient issues such as computing outages at each cluster and weekly staggered maintenance downtime for each analysis [113]. In O4a there

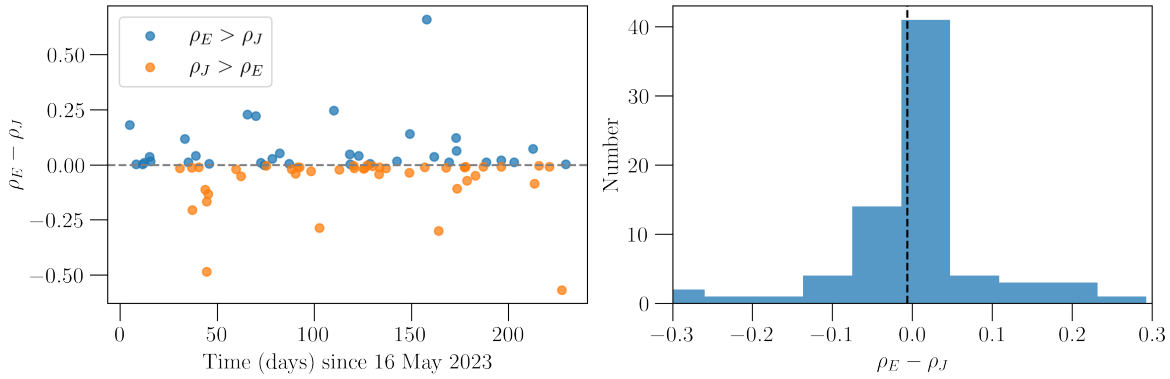

Figure 6.7. Fractional difference between the SNR of O4a GstLAL candidates as seen by the Edward (ρ_E) and Jacob (ρ_J) checkerboard analyses. The SNRs are taken from the minimum FAR candidate in each analysis. Negative (positive) values, shown in orange (blue) indicate that the SNR measured by the Jacob (Edward) analysis was greater.

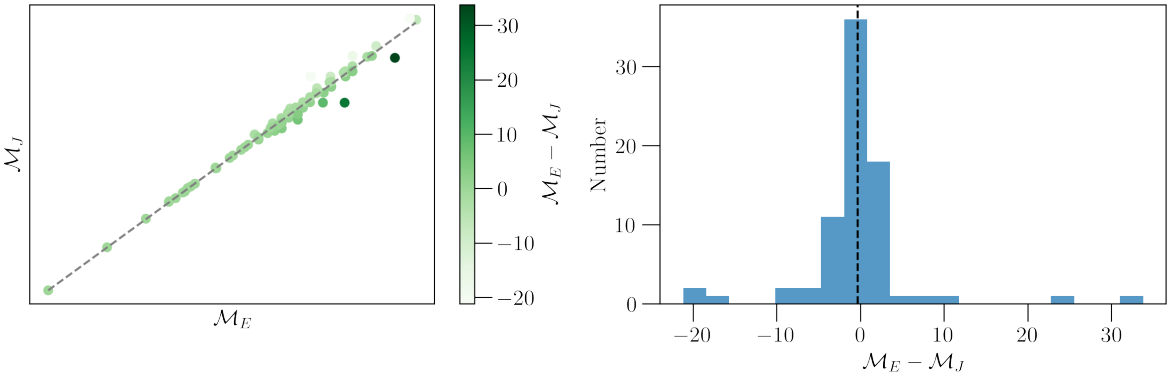

Figure 6.8. The chirp mass of the template producing the minimum FAR candidate in the Edward analysis ($\mathcal{M}_E$) vs in the Jacob analysis ($\mathcal{M}_J$). The color bar indicates the difference, $\mathcal{M}_E - \mathcal{M}_J$. The grey dashed line indicates the diagonal, i.e. $\mathcal{M}_E = \mathcal{M}_J$. Axis ticks are removed to obscure the exact values of the chirp masses recovered as at the time of writing, this is proprietary information.

were three candidates identified in GstLAL only by the Edward analysis (S240104bl [347], S230731an [350], and S230706ah [327]) and one candidate identified only by the Jacob analysis (S230814r [351]). S240104bl and S230706ah occurred during scheduled downtimes for the Jacob analysis. Further, these candidates were both identified significantly only by GstLAL; these are indicated with a dagger (†) in Table 6.2. S230731an (S230814r) occurred during a time of instability due to transient computing failures in the Jacob (Edward) analysis. These examples demonstrates the importance of running GstLAL as two redundant analyses on separate computing clusters. Without checkerboarding, these candidates may have been missed by GstLAL or altogether.

We expect that both checkerboarded analyses will produce similar results in terms of SNR recovery and estimated source parameters. Fig. 6.7 shows the fractional difference in SNR of the minimum FAR trigger in each analysis,

$$\frac{(\rho_E - \rho_J)}{\rho_E}. \tag{6.2}$$

The mean difference is about -0.11% with 43 events having a higher SNR as measured by the Jacob analysis compared to 34 with a higher SNR from the Edward analysis [4]. As expected, Fig. 6.8 shows that the source parameters as estimated by each checkerboard are largely consistent. The higher variation at larger masses is expected as it is well known that in this regime, where signals have a very short inspiral duration in the sensitive frequency band, the $\mathcal{M}_c$ is less well measured. Therefore, slight fluctuations in background statistics produced by the different templates may be expected to have a significant effect on which template produces the lowest FAR. Additionally, there is a much lower density of templates required to achieve the desired minimal match [107] at high masses, meaning that the difference in chirp mass of neighboring templates before the checkerboarding will have been much larger. Therefore, for a given signal at high mass the best matching template in one checkerboarded bank may have a significantly different chirp mass than in the other. Overall, these results confirm the expectation that both checkerboarded analyses will demonstrate similar sensitivity and performance across the parameter space.

6.5 Conclusion

This chapter presents results from the first part of LIGO, Virgo, and KAGRA's fourth observing run, O4a. In this time, there were 88 high significance alerts issued automatically via GCN notice. This number of detections is nearly as many as were detected in all three of the past observing runs combined [3], representing an incredible success owed to the detector upgrades bringing improved sensitivity as well as the developments in the search algorithms made throughout the commissioning break prior to the run. These observations represent a significant increase in the total population of GW signals observed to date and will certainly bring new insights to our understanding of the universe for years to come. Studying these populations will allow us to more carefully determine the mass and spin distributions of black holes and neutron stars, the channels through

[4] This indicates a very slight preference for *Team Jacob*.

which these objects form binaries, and their merger rates throughout the cosmic history.

Of the 88 alerts issued, 87 originated from one or more of the modeled searches for gravitational waves from CBC sources and 1 originated from an unmodeled search for transient GW "bursts". We estimate a purity among these alerts of $\sim 88\%$, corresponding to a contamination from terrestrial or instrumental artifacts of $\sim 12\%$. This indicates an expected ~ 10 noise events, including 8 which were retracted in low-latency. Further offline re-analysis of the data will allow us to more carefully validate these candidates and determine their origins and source properties. These will be published in the next update to the Gravitational Wave Transient Catalog, GWTC-4.

Following a short commissioning break, the second part of the fourth observing run (O4b) commenced in April 2024 and at the time of writing is expected to continue for approximately ten months until February 2025 [286]. One of the most exciting prospects for this continued run is the addition of Virgo, which is expected to observe with a BNS inspiral range of ~ 55 Mpc [286]. This addition will bring an improvement to the sky localization capabilities of the GW detector network, improving our prospects for identifying EM counterparts to potential BNS and NSBH observations [352].

Chapter 7
Archival searches for stellar-mass binary black holes in LISA

This chapter is a reprint of [353].

Stellar-mass binary black holes will sweep through the frequency band of the Laser Interferometer Space Antenna (*LISA*) for months to years before appearing in the audio-band of ground-based gravitational-wave detectors. One can expect several tens of these events up to a distance of 500 Mpc each year. The *LISA* signal-to-noise ratio for such sources even at these close distances will be too small for a blind search to confidently detect them. However, next generation ground-based gravitational-wave detectors, expected to be operational at the time of *LISA*, will observe them with signal-to-noise ratios of several thousands and measure their parameters very accurately. We show that such high fidelity observations of these sources by ground-based detectors help in archival searches to dig tens of signals out of *LISA* data each year.

7.1 Introduction and background

The discovery of GW150914 [12] by the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) [37] and the continued detections of stellar mass binary black holes [138, 164, 354] by Advanced LIGO and Virgo [39] set the stage for observing such systems with the Laser Interferometer Space Antenna (*LISA*) [355] when it comes online. GW150914 is the result of the merger of a pair of $36^{+5}_{-3}\,M_\odot$ and $31^{+3}_{-4}\,M_\odot$ black holes at a distance of 440^{+150}_{-170} Mpc [164]. The companion masses are larger than what was initially thought possible from stellar evolution [356] (see, however, [357]). It therefore earned the adjective *heavy* for black holes in the mass range $\sim [20\,M_\odot, 100\,M_\odot]$ now routinely observed by LIGO and Virgo [138, 164, 358–361]. Such heavy binary black holes (hBBHs)

within ~ 1 Gpc could also be visible in the *LISA* band [355] at an earlier stage in their evolution, albeit with a signal-to-noise ratio (SNR) of a few.

The search for stellar-mass binary black holes in *LISA* data could take formidable number of templates $\sim 10^{30-40}$ templates (faster chirping binaries requiring greater number of templates) and the associated computational resources [362]. The false alarm rate due to the large number of templates [363] would mean that only a handful of nearby sources with SNRs greater than ~ 15 might be detected for a p-value of 10^{-3}. Note, however, that third generation (3G) ground-based observatories, such as the Einstein Telescope [71, 364] and the Cosmic Explorer [365], operating at the same time as *LISA* would observe these sources some months to years after the signal passes the *LISA* band, with far greater SNRs compared to those in *LISA* and determine the source parameters to a good accuracy. Narrowing down the source parameters by ground-based detectors should then help in archival searches for such systems in *LISA* data by reducing the parameter space and, hence, the number of templates to $\sim 10^{12}$ while also decreasing the associated false alarm rates and computational costs. Current estimates still require an SNR threshold of ~ 8 [366] to ~ 14 [362].

Multiband observations of *h*BBH systems in *LISA* and ground-based detectors would greatly benefit the science return of these observatories [367–377]. This is because the parameter degeneracies that are present in the later part of the system's evolution in ground-based detectors could be resolved by observing the earlier part of the system's evolution in *LISA*. Several authors [368–372, 378] have demonstrated that this can principally yield tests of general relativity orders of magnitude better than what would be possible with either detector or detector-network by itself. What is critical to making that science possible is to unambiguously detect the signals in the *LISA* band.

In this paper we will show that 3G observatories will pin down the parameters of *h*BBH systems well enough to reduce the number of templates required for matched-filter searches to detect such systems in *LISA* data to a mere few $\times 10^4$ as opposed to previous estimates of $\sim 10^{12}$. This means that it will be possible to identify *h*BBH signals in *LISA* data with an SNR of 4 or more with a p-value of 10^{-2} or better. This will increase the number of sources that will be available for joint observation by both space-borne and ground-based observatories and hence enhance the science return of multiband observations.

The rest of the paper is organized as follows: In Sec. 7.2, we will compute for the joint *h*BBH population expected to be observed the visibility and measurement capabilities of 3G observatories. We will discuss, in particular, the uncertainties in the sky localizations,

masses, and spins of the companion black holes—parameters that would need to be searched for in *LISA* data. In Sec. 7.3, we show how the problem of assessing *LISA*'s performance in observing binary black holes can be mapped to the audio-frequency band. This is possible since there is no mass scale in general relativity: waveforms from binary black holes of different total mass will all look *exactly* the same as long as all other parameters remain the same except for a rescaling of time. This helps in using tools that have been developed for the analysis of ground-based detectors such as the LSC Algorithm Library [379]. In Sec. 7.4, we will estimate the number of templates required to search for *h*BBH systems in *LISA* data using the knowledge of parameter accuracies from 3G observatories. We will use two complementary methods to compute the number of templates. The first method works out the invariant volume of the signal manifold over the relevant range of parameters and then divides it by the fraction of volume covered by each template. This gives the minimal number of templates required for archival searches. In a realist data analysis pipeline, however, one needs to make a choice of templates based on a template placement algorithm [380]. We will use one such algorithm [381] to get a more realistic estimate of the number of templates. In Sec. 7.4.4, we characterize the efficiency of the template bank by computing the overlap of *h*BBH waveforms with random parameters maximized over the set of templates in the template bank. We will also discuss in Sec. 7.5 the distribution of the SNRs of the sub-population of sources that will be observed by both *LISA* and 3G observatories. We conclude in Sec. 7.6 with a summary of the results and future plans.

7.2 Stellar Mass Binary Black Holes in 3G Observatories

3G detectors like Cosmic Explorer (CE) and the Einstein Telescope (ET) will observe stellar-mass binary black holes all the way up to redshifts of $\sim$ 10–50 [74, 382, 383], depending on the intrinsic parameters of the source such as its masses and spins, as well as extrinsic parameters such as the position of the source on the sky and the orientation of the binary's orbit relative to the observer's line of sight. *LISA* could observe a small fraction of such systems if they are within $\sim 1\,\mathrm{Gpc}$ [355], but digging them out of the background noise in a blind search will take formidably large computational resources due to the large number of matched filters needed to cover the full parameter space of masses, spins, and sky position [362]. CE and ET will observe binary black holes that are this close with SNRs of several hundreds to several thousands and determine their parameters with extremely good precision. Such high-fidelity observations will narrow

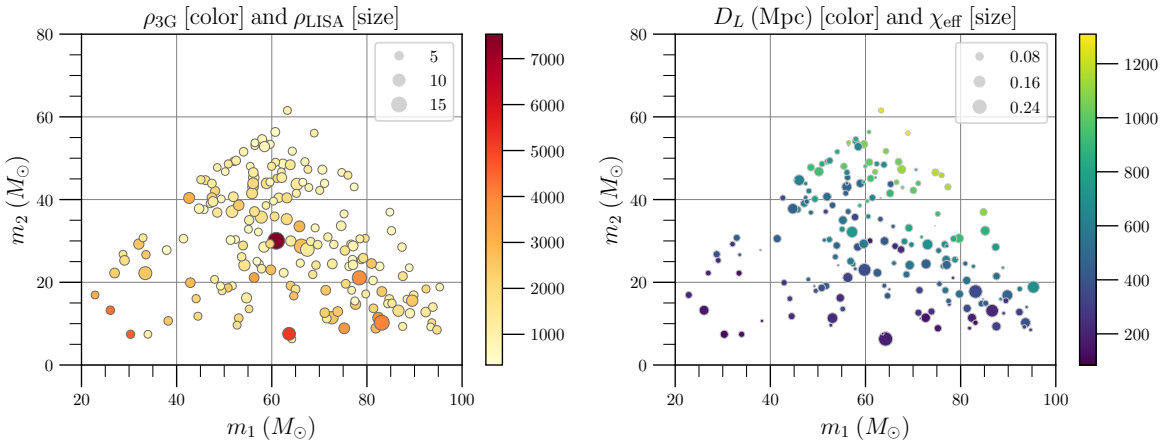

Figure 7.1. The plot shows the distribution of the 181 *h*BBHs with a *LISA* SNR $\rho_{LISA} \geq 4$, the two axes show the masses of the companions in both panels. In the left plot, color bar is the 3G SNR ρ_{3G} and the size of circles depict the *LISA* SNR ρ_{LISA}; in the right plot, color bar is the luminosity distance D_L and the size of the circles represent the effective spin parameter $\chi_{\rm eff}$.

down the search space in the *LISA* frequency band, which greatly reduces both the computational resources required but also the background noise from the large number of templates needed in a blind search.

In this Section, we will begin with the visibility of stellar-mass binary black holes in ground-based detectors and then go on to describe the precision with which the parameters can be measured. We shall show that all of the parameters but the chirp mass will be measured by a network of 3G observatories with an accuracy better than *LISA* which implies that it will only be necessary to construct templates in chirp mass for LISA.

7.2.1 Visibility

In order to assess what systems will be observable by both *LISA* and a network of 3G detectors, we simulated a population of 5×10^5 *h*BBHs which are uniformly distributed in co-moving volume up to a redshift of $z = 10$. The companion masses are chosen to follow a power law [280] for the larger companion $m_1 \in [5M_\odot 100M_\odot]$, $p(m_1) \propto m_1^{-\alpha}$ with exponent $\alpha = 1.6$, and a uniform distribution in $[5M_\odot, m_1]$ for the lighter companion m_2 [384]. The companion spins are assumed to be aligned or anti-aligned with the orbital angular momentum, and are drawn from a Gaussian distribution with 0 mean and a standard deviation of 0.1. Our choice of spin distribution is motivated by the source parameters of the 10 gravitational-wave detections of binary black holes reported in

GWTC-1 [164]. The root mean square χ_{eff} of these 10 events is 0.13, or excluding the high spin event GW170729, only 0.08.

We calculated the SNR, ρ_{LISA}, in *LISA* with the estimated power spectral density (PSD) provided in Ref. [385] and by marginalizing over the angular dependencies of the signal. We found 181 of the simulated signals to be visible in *LISA* with $\rho_{LISA} \geq 4$ and used these systems as candidates for our 3G-assisted archival search study for *LISA* data [371]. We will justify this choice of SNR in Sec. 7.5. Our choice for a 3G network consists of one ET and two CEs located at fiducial sites in Cascina (Italy), Idaho (USA), and New South Wales (Australia), respectively. The detector sensitivities are set to ET-D for the ET detector and CE1 (40 km, compact-binary optimized) for the two CEs [386].

The left panel of Fig. 7.1 shows the distribution of the 3G network SNR ρ_{3G} and ρ_{LISA} against the companion masses of the binaries for the 181 systems. All signals will be detected with SNRs of order ~ 1000 in 3G, a few reaching values almost 10 times as large. The left panel also confirms the expectation that loud *h*BBH events in the *LISA* band produce loud signals in CE and ET detectors. The right panel presents in a similar fashion the distribution of the luminosity distance, D_L, and effective spin parameter, $\chi_{\rm eff} = (m_1\chi_1 + m_2\chi_2)/M$ [387, 388], indicating that most of the systems are found at luminosity distances $\lesssim 1\,$Gpc.

We draw attention, in particular, to the visibility of *h*BBH in *LISA*. The rate for *h*BBH systems is constrained by the rate of BBH mergers whose up-to-date value is $R = 23.8^{+14}_{-8.7}\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ [138, 389]. Thus, if we take into account that only a fraction f of the binaries contain at least one heavy black hole ($> 20M_\odot$), where

$$f \simeq 1.9 \int_{20\,M_\odot}^{100\,M_\odot} \mathrm{d}m_1\, m_1^{-1.6} \int_{5\,M_\odot}^{m1} \mathrm{d}m_2 \frac{1}{m_1 - 5\,M_\odot} \simeq 0.32, \tag{7.1}$$

and 1.9 is the normalization factor, we obtain a median rate $R_{h\mathrm{BBH}} = f\,R = 7.6\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$ for *h*BBH merger. Although heavier binaries can be seen to a greater distance their relatively lower prevalence means that it is more likely that we will observe lighter binaries in *LISA* more frequently. The detection of lighter binaries in *LISA* is also aided by the large SNRs and high fidelity measurements of the 3G network. We also note that the *h*BBH systems are likely to have larger mass ratios, which is due to the power-law distribution of the primary companion and flat-distribution of the secondary. The large mass ratio is also responsible for low effective spins of *h*BBH systems as seen in the right panel of Fig. 7.1.

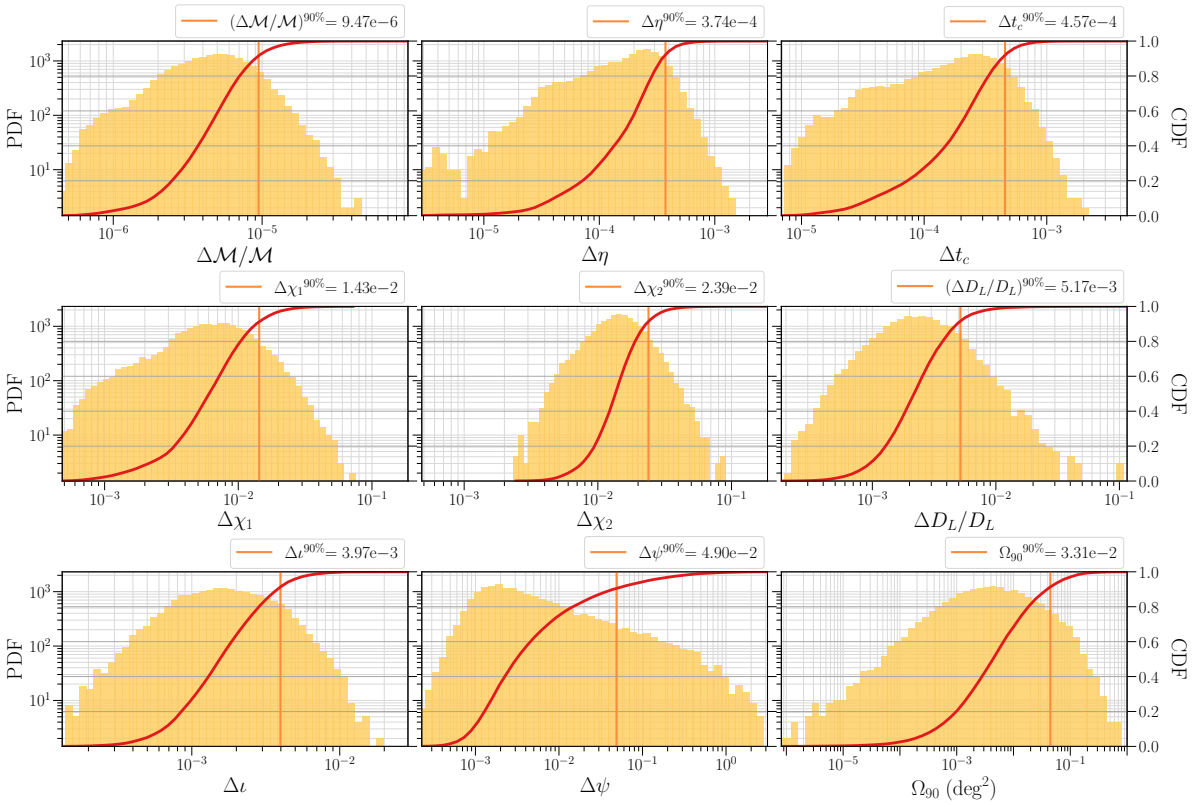

Figure 7.2. The plot shows the probability and cumulative density functions of the measurement error for the nine parameters of interest for the 181 binary black hole mergers that can be detected in the archival search of *LISA* data after they are identified and measured in the 3G data. We simulated 100 realizations for the sky location and binary orientation angles to show the range of parameter errors to be expected from 3G networks.

7.2.2 Measurability

Next, we want to assess the quality of the parameter estimation that a 3G network can achieve. This crucial information allows us to perform the archival search more efficiently by decreasing the dimension and volume of the parameter space for a template search.

We perform this assessment with `gwbench` [386], a Python package that implements the well-known Fisher information [233, 234, 390] formalism, and estimate the 1σ-error bounds for each of the 181 systems. The formalism provides an analytical approximation of the Gaussian noise likelihood around its maximum and thus allows us to estimate the measurement errors $\sigma_{\lambda_i} = \sqrt{\Sigma_{ii}}$ on a set of parameters $\boldsymbol{\lambda}$ from the covariance matrix Σ in the likelihood:

$$P(\boldsymbol{\lambda}) \sim \mathrm{e}^{-\frac{1}{2}\,\Sigma^{-1}_{ij}\,\Delta\boldsymbol{\lambda}_i\Delta\boldsymbol{\lambda}_j}. \tag{7.2}$$

Given a model for the detector response $\tilde{h}(f;\boldsymbol{\lambda})$, we can calculate the covariance matrix

as the inverse of the Fisher information matrix Γ

$$\Sigma^{-1}_{ij} \equiv \Gamma_{ij} = \langle \partial_{\lambda_i} \tilde{h}(f;\boldsymbol{\lambda}), \partial_{\lambda_j} \tilde{h}(f;\boldsymbol{\lambda}) \rangle. \tag{7.3}$$

The scalar product between two waveforms $\tilde{h}(f;\boldsymbol{\lambda}_1)$ and $\tilde{g}(f;\boldsymbol{\lambda}_2)$ is defined as

$$\langle \tilde{h}(\boldsymbol{\lambda}_1), \tilde{g}(\boldsymbol{\lambda}_2) \rangle = 2 \int_0^\infty \frac{\tilde{h}(f,\boldsymbol{\lambda}_1)\, \tilde{g}^*(f,\boldsymbol{\lambda}_2) + c.c.}{S_h(f)}\, \mathrm{d}f. \tag{7.4}$$

where $\tilde{g}^*$ denotes the complex conjugate of $\tilde{g}$. Note that although the limits in the integral range from 0 to ∞, in reality the detector noise power spectral density $S_h(f)$ outweighs the signal power outside a finite frequency range $[f_1, f_2]$ and often the waveform itself will have no support above a frequency $f_{\rm cut} = f(\boldsymbol{\lambda})$ determined by its intrinsic parameters. Thus, the integral gets most of its support over a finite range of frequency $[f_1, f_2]$.

The error bounds that a network of several detectors can achieve are readily computed via the network Fisher matrix

$$\Gamma_{\rm net} = \sum_d \Gamma_d, \tag{7.5}$$

which is the sum of the Fisher matrices Γ_d for all the detectors in the network. Hence, given a detector response model, we calculate all the detector Fisher matrices and invert their sum to obtain the network covariance matrix from which we extract the desired error bounds.

Lastly, we perform two sanity checks to avoid including faulty numerical data. We first disregard any Fisher matrix Γ whose condition number $c_\Gamma = e_M/e_m$ exceeds a threshold value of 10^{15} to avoid inverting matrices that are ill-conditioned for this numerical task. e_M and e_m are the maximum and minimum eigenvalues of Γ. Further, we scrutinize all inversions, if any calculated error bound is smaller than the inversion error $\epsilon = ||\Sigma \cdot \Gamma - I||_{\rm max}$. Here, I and $||\cdot||_{\rm max}$ represent the identity and maximum matrix norm, respectively.

The loudness of the *h*BBH signals in CE and ET detectors allows us to make use of waveform models that include higher-order spherical harmonic modes which capture more physical information and thus increase the accuracy of the parameter estimation. For this purpose we applied the Fisher formalism to the `lalsimulation` waveform `IMRPhenomHM` [391] for the full set of 11 parameters: chirp mass $\mathcal{M}$, symmetric mass ratio η, the aligned components of the companion spins $\chi_{1,z}$ and $\chi_{2,z}$, luminosity distance D_L, coalescence time t_c, phase of coalescence ϕ_c, inclination angle ι, right ascension α, declination δ, and polarization angle ψ. `IMRPhenomHM` is an aligned-spin waveform

model that does not include the spin components perpendicular to the orbital angular momentum of the binary, thus our Fisher analysis is four parameters short of the standard 15-parameter analyses.

Since we marginalized over the four angles for the calculation of the *LISA* SNRs, we randomly sampled 100 realizations of each angle for each of the 181 systems and performed the Fisher analysis on these 181×100 parameter sets. The resulting error bounds are shown in Fig. 7.2, where we show the errors on right ascension and declination combined in the 90%-credible sky area Ω_{90} and omitted the error for phase of coalescence.

The *LISA* parameter estimation has been explored with the Fisher formalism in [355] for signals with $\rho_{LISA} > 8$. The study reports the following bounds for the majority of its 1000 simulated events: $\Delta\mathcal{M}/\mathcal{M} \in [10^{-7}, 4 \times 10^{-6}]$ peaked at $\sim 10^{-6}$, $\Delta\eta/\eta \in [6\times10^{-4}, 3\times10^{-2}]$ peaked at $\sim 8\times10^{-3}$, $\Delta t_c \in [10^{-1}, 7\times10^{1}]$ peaked at $\sim 3\times10^{0}$, and $\Omega_{90} \in [2 \times 10^{-2}, 4 \times 10^{0}]$ peaked at $\sim \times10^{-1}$. In Ref. [392], the *LISA* parameter estimation has been explored for GW150914-like systems, with SNR $\rho_{LISA} \approx 10 - 13$. The study quotes fractional error in $\mathcal{M} < 10^{-4}$ to below 10^{-6}. For our population of *h*BBH we find that *LISA* will constrain the chirp mass even better than 3G observatories. We note that if this were not the case, we would only require a single template to dig out the *h*BBH from *LISA* data.

Comparing our error bounds to these estimates we can clearly see that a network of CE and ET observatories will outperform *LISA* for the estimation of most parameters: our 90% values are either well below (t_c) or of the order of (Ω_{90}) the lower bound of the reported ranges. The exceptions are the chirp mass and symmetric mass ratio which benefit from the long, many-cycle signals in the *LISA* band: the fractional $\mathcal{M}$ errors in the *LISA* band are better or the same compared to the 3G bounds and if we scale our absolute errors in η to relative errors—i.e. multiplication with factors between 4 to 17 in the case of our binaries with $\eta \in [0.06, 0.25]$—we obtain roughly the same ranges (the 3G network performs better on the lower end). Thus, LISA can only improve the chirp mass measurement, without adding significant information to the estimation of η.

There are two caveats to this comparison that favor the 3G results even more: The cited study performed the Fisher analysis only for six parameters which positively biases their results in comparison to our analysis over 11 parameters. Further, their reported errors come from a louder population with $\rho_{LISA} > 8$, whereas most of our signals have SNRs lower than that. The events considered in their study would result in even louder 3G events, as seen in Fig. 7.1, and thus better error estimates.

In conclusion, our findings show that a 3G network allows to estimate the parameters

of hBBHs with such high fidelity that we can assume most parameters to be known and focus the archival searches on $\mathcal{M}$ only.

7.3 Mapping the *LISA* data analysis problem to the audio band

The LSC Algorithm Library (LALSuite) [379] has many data analysis tools such as compact binary waveform models, template placement algorithms, filtering routines, etc., that are extremely useful, sometimes critical, in evaluating data analysis problems such as the ones explored in this paper. LALSuite was developed primarily for the analysis of data from LIGO and Virgo interferometers that operate in the audio frequency band from 1 Hz to 10 kHz. Unfortunately, some of the algorithms do not readily work at frequencies below 1 Hz and the effort required to rewrite the algorithms for the *LISA* band, $\sim 100\,\mu$Hz to 100 mHz, would be formidable. Luckily, it is possible to scale the *LISA* problem into the audio band owing to the fact that the fundamental quantity of interest, namely the strain measured by the gravitational-wave detectors which represents the change in proper length between 'free' test masses in response to a passing gravitational wave, has no physical dimension.

To illustrate the required scaling, let us consider gravitational waves emitted by an inspiraling binary system composed of a pair of black holes, but the argument would work no matter what type of source we consider. Furthermore, for the sake of clarity, we will consider the lowest order post-Newtonian (PN) waveform [203, 393] (often referred to as the 'Newtonian' waveform) from a binary system composed of non-spinning black holes. However, the arguments follow through irrespective of the PN order. At the Newtonian order, the strain response of an interferometer to gravitational waves from a binary system composed of black holes of masses (m_1, m_2) at a luminosity distance D_L is given by

$$h(t) = \frac{4G\eta M}{c^2 D_L}\left(\frac{\pi G M f(t)}{c^3}\right)^{2/3}\cos\left(\pi\int_{t_0}^{t} f(t)\,\mathrm{d}t\right) \tag{7.6}$$

where $M = (m_1 + m_2)$ is the total mass of the system, $\eta = m_1 m_2/M^2$ is the symmetric mass ratio, and $f(t)$ is the instantaneous gravitational-wave frequency:

$$f(t) = f_0\left(1 - \frac{t - t_0}{\tau}\right)^{-3/8}. \tag{7.7}$$

Here t_0 is a fiducial time when the frequency of the gravitational wave is f_0 and τ, called

the *chirp time*, is the duration of the signal from a time when its frequency was f_0 until the frequency (in this approximation) diverges[1] and the two black holes merge:

$$\tau = \frac{5}{256\,\eta}\frac{GM}{c^3}\left(\frac{c^3}{\pi\,GM f_0}\right)^{8/3}. \tag{7.8}$$

The chirp time of a binary of total mass $100\,M_\odot$, equal component masses (so that $\eta = 1/4$), and starting frequency of 12 mHz would be 5 years. Chirp time is a sharp function of the total mass as well as the starting frequency. A signal starting from a frequency that is a factor 2 (10) smaller would last a factor ~ 6.3 (respectively, 464) longer.

From Eq. (7.8), we can compute the duration Δt spent by the binary starting at frequency f_1 at time t_1 and reaching frequency f_2 at time t_2:

$$\Delta t \equiv t_2 - t_1 = \frac{5}{256\,\eta}\frac{GM}{c^3}\left(\frac{c^3}{\pi\,GM f_1}\right)^{8/3}\left[1-\left(\frac{f_1}{f_2}\right)^{8/3}\right]. \tag{7.9}$$

If $f_2 \gg f_1$, the second term in the equation above will be negligible, and Δt will essentially be the same as the chirp time starting at frequency f_1:

$$\Delta t \simeq \frac{5}{256\,\eta}\frac{GM}{c^3}\left(\frac{c^3}{\pi\,GM f_1}\right)^{8/3}. \tag{7.10}$$

We choose the starting frequency f_1 for the stellar mass binary black holes in the mass range observed by *LISA*, such that the signal lasts for a fixed duration in the *LISA* band. For a given Δt, the starting frequency f_1 depends on the total mass and mass ratio of the binary:

$$f_1(M,\eta) = \frac{c^3}{\pi\,GM}\left[\frac{5}{256\,\eta}\frac{GM}{c^3\Delta t}\right]^{3/8}. \tag{7.11}$$

The left half of Fig. 7.3 shows the starting frequency f_1 as a function of total mass and mass ratio $q = m_1/m_2$ for $m_1 > m_2$; f_1 is greatest for systems with small total mass but large mass ratio and smallest for systems with large total mass, but small mass ratio. Over the total mass range of [30, 130] $M_\odot$ the lower frequency cutoff is never smaller than 5 mHz and can be as large as 45 mHz at the lower end of the mass range and upper range of the mass ratio. The upper frequency cutoff for integration is chosen to be $f_2 = 1\,\mathrm{Hz}$.

[1] In reality, the merger occurs when the horizons of the two black holes merge which happens at a finite frequency but this is not relevant to our discussion.

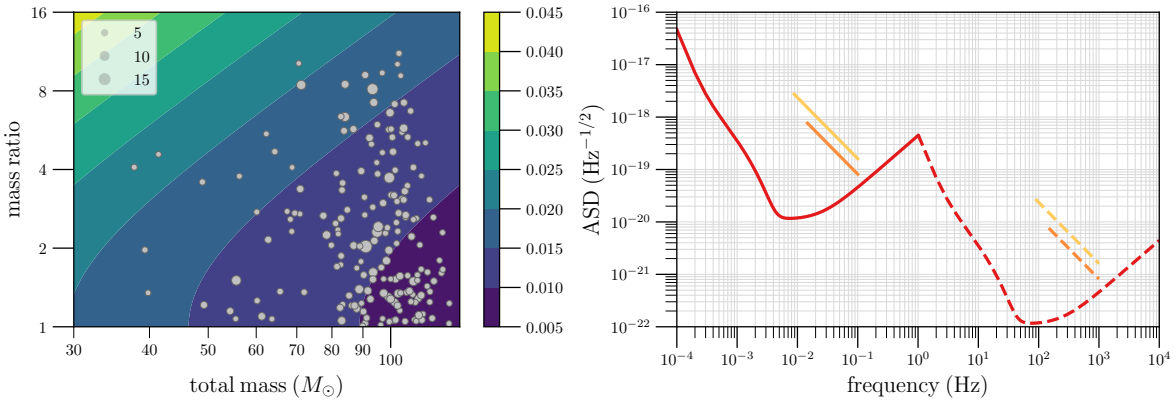

Figure 7.3. **Left:** Contour plot of the starting frequency for stellar mass binary black holes that last 5 years in the *LISA* band. The starting frequency scales with the mass ratio of each system and inversely with the total mass, as indicated by the colored contours, where the values of starting frequency are shown in mHz. The grey dots mark the positions of the 181 hBBHs for which we performed Fisher analysis in a 3G network to obtain parameter error estimates. The size of the dots represents the corresponding *LISA* SNR, ρ_{LISA}.
Right: *LISA*'s amplitude noise spectral density (ASD) $\sqrt{S_h(f)}$ is plotted before (red solid line) and after (red dashed line) applying frequency scaling with $\alpha = 10^4$. Also shown are the amplitude spectra $\sqrt{f}|\tilde{h}(f)|$ for a GW150914-like (in orange) and GW190521-like (yellow) systems at a distance of 500 Mpc with signal-to-noise ratio of 4.8 and 16, respectively. The integration time is assumed to be 5 years in the *LISA* band which translates to 1.54×10^5 s in the audio band.

Our goal is to scale up the frequency from *LISA*'s observing band to the audio band of ground-based detectors. Scaling up frequencies by a factor of $\alpha = 10^4$ would bring the lowest starting frequencies up to 50 Hz and the largest starting frequencies up to 450 Hz. This is the scaling we will use in this paper. An integration time of $\Delta t_{LISA} = 5\,\mathrm{yr}$ in the *LISA* band would correspond to an integration time of $\Delta t_{\mathrm{audio}} = \Delta t_{LISA}/\alpha \simeq 1.58 \times 10^4\,\mathrm{s}$.

We can see from Eq. (7.6) that gravitational-wave strain would remain unchanged if we simultaneously scale up the frequency by factor of α and scale down the chirp time, the total mass, and the luminosity distance by the same factor. Therefore, the signal would now last for a much shorter period of $\Delta t/\alpha$ with exactly the same amplitude as before but at a higher frequency. The SNR of the scaled up signal, but with a scaled up *LISA* noise spectral density $S_h^{LISA}(f)$, will also be the same as before. To see this, recall that the expectation value of the SNR of a signal is given by:

$$\rho^2(f_1, f_2) = 4\Re \int_{f_1}^{f_2} \frac{|\tilde{h}(f)|^2}{S_h^{LISA}(f)}\,\mathrm{d}f, \tag{7.12}$$

where $\tilde{h}(f) \equiv \int_{-\infty}^{\infty} h(t)\,\exp(2\pi i f t)\,\mathrm{d}t$ is the Fourier transform of the gravitational-wave

strain. Changing the variable $f \to \nu = \alpha f$ would scale down the Fourier mode strain by a factor of α, $\tilde{h}(f) \to \tilde{h}(\alpha f)/\alpha$, and similarly the *LISA* noise PSD $S_h^{LISA}(f) \to S_h^{LISA}(\alpha f)/\alpha \equiv S_h^{\rm audio}(\alpha f)/\alpha$. Thus the SNR remains unchanged:

$$\rho^2(\nu_1, \nu_2) = 4\Re \int_{\nu_1}^{\nu_2} \frac{|\tilde{h}(\alpha f)|^2/\alpha^2}{S_h^{\rm audio}(\alpha f)/\alpha}\, \alpha {\rm d}f. \tag{7.13}$$

The scaled version of the *LISA* noise PSD $S_h^{\rm audio}(\nu)/\alpha$ is shown in the right hand panel of Fig. 7.3.

In summary, the required scaling transformation are to:

1. Scale up the frequency: $f \to \alpha f$.
2. Scale down the time duration, total mass and distance: $\tau \to \tau/\alpha$, $M \to M/\alpha$, and $D_L \to D_L/\alpha$.
3. Transform the *LISA* power-spectral density into the audio band, i.e. $S_h^{\rm audio}(f) = \frac{1}{\alpha} S_h^{LISA}(f/\alpha)$.

7.4 Template banks for archival searches

In this section we present the number of templates required to detect hBBH in the *LISA* band by matched filtering. We use accuracies on the masses obtained from parameter estimation in 3G detectors. We use two independent methods to calculate template bank numbers. The first method assumes the metric in order to place templates, which would provide a minimum on the number of templates required. As a check on this method, we also calculate the number of templates using a stochastic placement algorithm, which overestimates the required number of templates.

7.4.1 Metric method

7.4.1.1 Metric on the signal manifold

The number of templates required for a search can be found using the geometric formulation of signal analysis [55, 380, 390]. The scalar product (7.4) can be used to define waveforms or signals of *unit norm*. A signal is said to be of unit norm if its scalar product with itself is unity and will be denoted by $\hat{a} : \langle \hat{a}, \hat{a} \rangle = 1$.

In the geometric formalism, the overlap $\mathcal{O}$ or *match* between two 'nearby' normalized signals $\hat{h}(\boldsymbol{\lambda})$ and $\hat{h}(\boldsymbol{\lambda}+\Delta\boldsymbol{\lambda})$ with slightly different parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}+\Delta\boldsymbol{\lambda}$ is given by:

$$\mathcal{O}(\boldsymbol{\lambda},\Delta\boldsymbol{\lambda}) \equiv \langle \hat{h}(\boldsymbol{\lambda}), \hat{h}(\boldsymbol{\lambda}+\Delta\boldsymbol{\lambda})\rangle \approx 1 - g_{\alpha\beta}\,\mathrm{d}\lambda^{\alpha}\,\mathrm{d}\lambda^{\beta}, \tag{7.14}$$

where $g_{\alpha\beta}$ is the metric on the signal manifold [55, 390]:

$$g_{\alpha\beta} \equiv -\frac{1}{2}\left.\frac{\partial^2 \mathcal{O}(\boldsymbol{\lambda},\Delta\boldsymbol{\lambda})}{\partial\lambda^{\alpha}\partial\lambda^{\beta}}\right|_{\Delta\boldsymbol{\lambda}=0} = \langle \hat{h}_{\alpha}, \hat{h}_{\beta}\rangle, \quad \hat{h}_{\alpha} \equiv \frac{\partial\hat{h}}{\partial\lambda^{\alpha}}. \tag{7.15}$$

When signals are nearby, in the sense that their overlap is close to unity, Eq. (7.14) is a good approximation for the overlap and the quantity $\mathrm{d}\ell^2 = g_{\alpha\beta}\,\mathrm{d}\lambda^{\alpha}\,\mathrm{d}\lambda^{\beta}$—the *proper distance* between them—obeys $\mathrm{d}\ell \ll 1$. Thus, two normalized signals at a proper distance of $d\ell$ from each other have an overlap of $1 - \mathrm{d}\ell^2$.

7.4.1.2 Minimal match and lower limit on the number of search templates

To filter signals out of data we must choose a bank of templates in the parameter space of interest such that any signal buried in the data within this parameter space has its overlap larger than a certain value called the *minimal match* M with at least one template in the bank. Of course, this requirement can be met by populating the parameter space with a dense set of templates but a higher density of templates would demand a greater computational cost. Thus, the density of templates must be chosen so that it is as sparse as possible while assuring minimal match with every signal in the parameter space of interest.

If $\boldsymbol{\lambda_k}$, $k = 1, \ldots N$, denotes the parameters of the templates in the bank then for a signal with arbitrary parameters $\boldsymbol{\lambda}$ the template bank must satisfy the following condition:

$$\max_{k} \langle \hat{h}(\boldsymbol{\lambda}), \hat{h}(\boldsymbol{\lambda}_k)\rangle \geq M, \tag{7.16}$$

the equality in the above equation giving the optimal number of templates. So we must choose the proper distance such that

$$\mathrm{d}\ell^2 = g_{\alpha\beta}\,\mathrm{d}\lambda^{\alpha}\,\mathrm{d}\lambda^{\beta} = (1 - M). \tag{7.17}$$

Given that the optimum matched filter signal-to-noise ratio that one can hope to achieve for a signal is $\rho^2_{\mathrm{opt}} = \langle h, h\rangle$, the above condition assures that the fractional drop in the signal-to-noise ratio between an arbitrary signal and the closest template in the bank is

no more than $\epsilon \equiv 1 - M$ called the *maximum mismatch,* i.e. $\rho \geq \epsilon \rho_{\rm opt}$.

Given the minimal match M how many templates are needed to cover the parameters with the smallest number of templates? This is the problem of template placement and to some extent the answer depends on what type of lattice is used to place the templates on the signal manifold. We will discuss a specific template placement algorithm used in this work in the next Section. We can get a lower limit on the number of templates needed by computing the total proper volume of the signal space divided by the fraction of volume covered by each template. Assuming that each template covers a proper volume $\mathrm{d}V = \mathrm{d}\ell^N = (1-M)^{N/2}$, where N is the dimension of the parameter space, the smallest number of templates N_T needed is:

$$N_T = \frac{V}{\mathrm{d}V} = \frac{1}{(1-M)^{N/2}} \int_{\boldsymbol{\lambda}_{\rm min}}^{\boldsymbol{\lambda}_{\rm max}} \sqrt{g}\, \mathrm{d}^N\lambda, \tag{7.18}$$

where $g = \det |g_{\alpha\beta}|$. The above estimate assumes that the templates are placed on a square lattice. The number of templates can be smaller with a more efficient lattice, e.g. a hexagonal lattice in two-dimensions, but this is unimportant for our considerations.

7.4.1.3 Newtonian approximation to number of templates

The orbital velocity of a stellar mass binary black hole of total mass M and gravitational-wave frequency $f_{\rm GW}$ is very small in the *LISA* frequency band compared to the speed of light (for clarity we include factors of G and c in the equation below):

$$\frac{v}{c} \equiv \left(\frac{\pi G M f}{c^3}\right)^{1/3} \simeq 0.025 \left(\frac{M}{10^2 M_\odot}\right)^{1/3} \left(\frac{f_{\rm GW}}{10\,\mathrm{mHz}}\right)^{1/3}. \tag{7.19}$$

For such a non-relativistic orbit the PN expansion parameter defined by $x \equiv v^2/c^2 \simeq 6.2 \times 10^{-4}$, meaning stellar-mass binary black holes in the *LISA* frequency band are in the adiabatic regime. The dynamics of such a system is essentially described by the Newtonian approximation (or 0 PN) given in Eqs. (7.7) and (7.8). In the stationary phase approximation the Fourier transform of the waveform in Eq. (7.6) is given by [394]:

$$h(f) = \frac{A}{D_L} f^{-7/6} \exp\left[2\pi i f t_C - i\phi_C + i\psi(f)\right], \tag{7.20}$$

where t_C is a fiducial time giving the time of arrival of the signal at the detector (the epoch at which the GW frequency reaches a pre-defined value), ϕ_C is a constant phase of the signal, and $\psi(f)$ is the PN approximation to the signal's phase evolution given at

the leading "Newtonian order" by:

$$\psi(f) = \frac{3}{128(\pi \mathcal{M} f)^{5/3}}. \tag{7.21}$$

At the leading order the phase depends only on the chirp mass $\mathcal{M}$ and not the mass ratio. It is useful to define a new parameter $\xi \equiv (3/128)(\pi\,\mathcal{M})^{-5/3}$, so that the phase is linear in this new parameter: $\psi(f) = \xi f^{-5/3}$. The parameter space of the signal consists of $\boldsymbol{\lambda} = (t_C, \phi_C, \xi)$. It turns out that the overlap $\mathcal{O}$ can be analytically maximized over the phase ϕ_C by using two quadrature filters, leaving just two parameters. Because of the analytic maximization over phase, the expression for the metric in the two-dimensional space of $\boldsymbol{\lambda} = (t_C, \xi)$ takes the form [55, 380]:

$$g_{\alpha\beta} = \frac{1}{2}\left(\mathcal{J}[\psi_\alpha \psi_\beta] - \mathcal{J}[\psi_\alpha]\mathcal{J}[\psi_\beta]\right), \tag{7.22}$$

where $\psi_\alpha \equiv \partial\psi(f)/\partial\lambda^\alpha$, and $\mathcal{J}$ is a functional of its argument defined for any function $a(f)$ by

$$\mathcal{J}[a(f)] \equiv \frac{1}{\rho^2}\langle f^{-7/3}, a(f)\rangle \tag{7.23}$$

As is well known, maximizing the overlap $\mathcal{O}$ over the parameter t_C is easily accomplished in the Fourier domain [395] and one needs a discrete lattice of templates only for the remaining one parameter ξ. The metric in the ξ-dimension is quite simply: $G_{22} = g_{22} - g_{21}^2/g_{11}$. Substituting for the various elements of the metric $g_{\alpha\beta}$ and simplifying one finds:

$$G_{22} = \frac{1}{2}\left[J_{17} - J_{12}^2 - \frac{(J_9 - J_4 J_{12})^2}{J_1 - J_4^2}\right] \tag{7.24}$$

where for any k, the moment J_k is given by $J_k \equiv \langle f^{-k/3}, 1\rangle$. In this notation the SNR is $\rho^2 = J_7$. G_{22} is demonstrably constant (in any case all one-dimensional spaces are flat) and ξ is already a Cartesian coordinate. The spacing between templates is constant in ξ and from Eq. (7.17) we have

$$\mathrm{d}\ell^2 = G_{22}\,\mathrm{d}\xi^2 = 1 - M \quad \Rightarrow \quad \mathrm{d}\xi = \sqrt{\frac{1 - MM}{G_{22}}} \tag{7.25}$$

Finally, the number of templates can be found using Eq. (7.18):

$$N_T = \sqrt{\frac{G_{22}}{(1-M)}}(\xi_{\max} - \xi_{\min}). \tag{7.26}$$

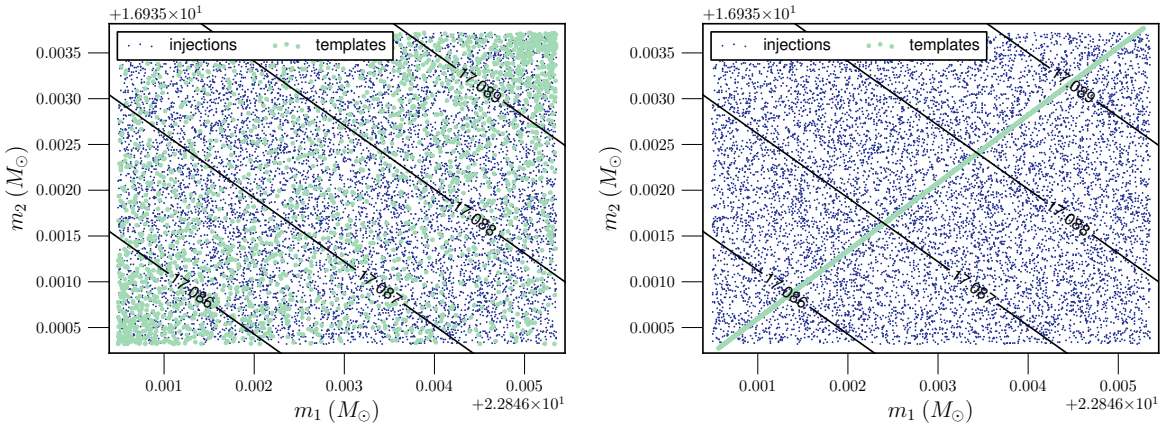

Figure 7.4. **Left:** For a *h*BBH source with component masses constrained by 3G to $m_1 = 22.84892$ and $m_2 = 16.93702\ M_\odot$ the template (green) and injection (blue) component masses are shown for a 3.5 PN template bank. Chirp mass contours are indicated by the black lines. **Right:** For the same source, the injection component masses are shown with the one-dimensional Newtonian template component masses.

In the next section we will compute the number of templates found using a template placement algorithm and compare it with the one found in this section.

7.4.2 Stochastic template placement algorithm

When the metric on the waveform manifold is not known exactly and cannot be approximated, we use a brute force approach [396, 397] to construct the template bank. In this method, the template bank is built by proposing templates in the desired parameter space until sufficient coverage is reached. For each proposed template, $\boldsymbol{\lambda_{prop}}$, the fitting factor (which is the maximum match of the proposed template with the templates in the bank) is calculated. If the fitting factor is greater than the minimal match required, we reject the proposed template and continue with a new proposal; if it is less than the required minimal match, we add the proposed template to our bank.

This method can get computationally expensive, but there are tricks we can employ to be able to use it in practice. One such trick is to define the “neighborhood" of the proposed templates. It can be defined in terms of a parameter chosen by the user, we used the chirp time for our purposes. When used, the fitting factor is calculated only for the templates in the neighborhood (set in form of units of the neighborhood parameter by the user), $\boldsymbol{\lambda_i}$, where $i : 1 - N$ for N templates near the proposal.

Another technique we use is to iteratively lower the frequency step in the calculation of the match [398]. The value of frequency spacing used in the match integral (Eq. (7.4))

is usually chosen to be $1/L$, where L is the closest power-of-2 greater than the length of the waveforms. This is required to measure the overlap between two waveforms in a time window of L s. But for the bank generation, we are interested in the maximum overlap between waveforms, which occurs near the time point corresponding to 0 displacement for two waveforms aligned at their peak amplitude. Therefore, we can increase the value of frequency spacing, *df*, which greatly reduces the cost of calculating the match. We can check if our chosen value for *df* is good enough by iteratively calculating matches by reducing *df* by half. If the last two overlaps agree to within 1%, we use that value or if $(1 - (\text{match}))$ is large, we continue to the next template. For our banks, we first calculate the matches with $df = 2.0$. If the mismatch is large, i.e.

$$(1 - \text{match}) > 0.05 + (1 - M), \tag{7.27}$$

where M = 0.98 is the minimum required match for template placement, then we move on to the next neighboring template $\boldsymbol{\lambda_{i+1}}$. Otherwise, we decrease the value of $df \to df/2$ and calculate the match again. We continue iteratively decreasing *df* until the last two matches converge,

$$\left|\text{match}_{df} - \text{match}_{df/2}\right| < 0.001 \times \text{match}_{df}. \tag{7.28}$$

Once convergence is reached, if the match $< M$ we add the proposal $\boldsymbol{\lambda_{prop}}$ to the template bank.

7.4.3 Number of templates

Using the stochastic placement algorithm, we calculate a template bank for each of the 181 *h*BBH sources, with component mass ranges constrained by 3G capabilities. These template banks use 3.5 PN TaylorF2 [394, 399–404] waveforms. The lower frequency cut-off for the signals ranges from 14.2–30.5 mHz with an upper frequency cut-off of 150 mHz. Scaling this to the audio band, we produce template banks with frequencies in the range of 142–1500 Hz. We use a minimal match of $M = 0.98$ for the template placement algorithm.

The left panel of Fig. 7.4 shows an example template bank in (m_1, m_2) space for a source of $m_1 = 22.85M_\odot$, $m_2 = 16.94M_\odot$, and $f_{low} = 305.44$ Hz, after scaling to the audio band. Fig. 7.5 shows the relationship between the source signal-to-noise ratio and the number of templates in the bank. The plot shows that the sources which require the largest number of templates are those for which both the *LISA* and CE signal-to-noise

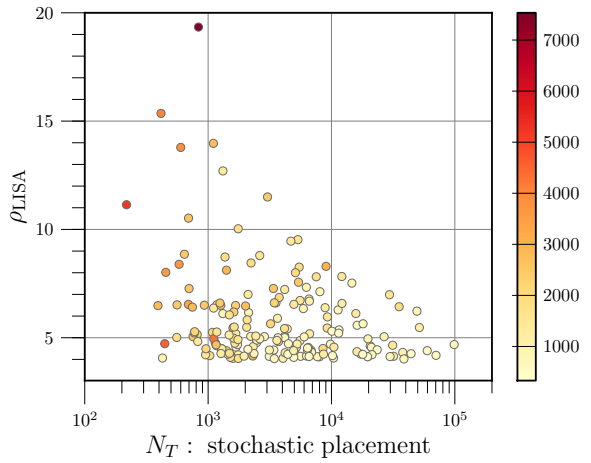

Figure 7.5. The number of templates from the stochastic placement algorithm is plotted against the *LISA* SNR, ρ_{LISA} for each source. The color bar indicates the CE SNR, ρ_{CE}.

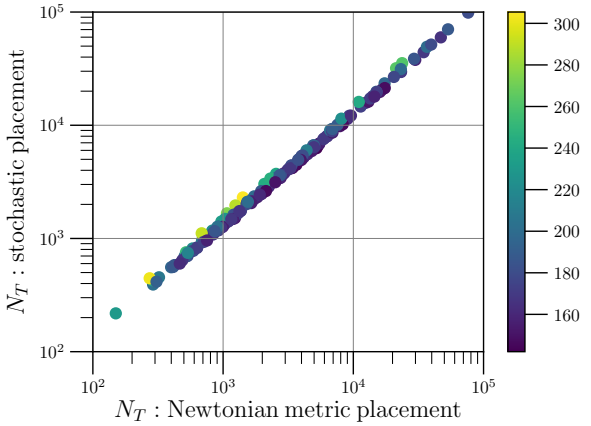

Figure 7.6. For 181 *h*BBH sources we plot the number of templates required for the minimal match condition, $M = 0.98$ calculated using the metric approach for one-dimensional Newtonian template banks (horizontal axis) and the number of 3.5 PN templates found using the stochastic placement algorithm (vertical axis). The color bar indicates the lower frequency cutoff, after scaling to the audio band. For a linear fit, $N_{T,\mathrm{stochastic}} = m \times N_{T,\mathrm{metric}}$, we find $m = 1.30$

ratio is small.

The cumulative distribution of stochastic placement template bank sizes for all 181 sources is shown in the top left panel of Fig. 7.7. The 50th percentile and 90th percentile bank sizes are 3.4×10^3 and 1.95×10^4 templates, respectively. The mean bank size, $\overline{N_T} = 8.02 \times 10^3$ with a standard deviation of 9.70×10^2. The smallest bank has only 218 templates while the largest bank has 9.88×10^4 templates. These bank sizes demonstrate a significant improvement from previous estimates of $\mathcal{O}(10^{12})$ templates.

Using the metric placement method, we again calculate template banks for each of the same 181 hBBH sources. The upper and lower frequency cut-offs are the same for these banks as for the template banks generated with the stochastic placement algorithm. We again require a minimal match, $M = 0.98$. The right panel of Fig. 7.4 shows the Newtonian template bank for the same source as shown on the left. This figure demonstrates that the Newtonian template banks are truly one-dimensional, where each template in (m_1, m_2) space is projected onto a line parameterized by the chirp mass, $\mathcal{M}$. The figure shows chirp mass contours indicated by the black lines. Using the one dimensional template bank it is clear that while component masses may not necessarily be tightly constrained by *LISA*, the chirp mass will be recovered very well. For this source, the two-dimensional bank had 2.3×10^3 templates and the one-dimensional bank had 1.4×10^3 templates. Therefore, the line in Fig. 7.4 is very densely packed.

The cumulative distribution of Newtonian template bank sizes is shown in the bottom left panel of Fig. 7.7. Here, the mean bank size is 6.2×10^3 templates. The 50th and 90th percentile bank sizes are 2.5×10^3 and 1.5×10^4 respectively.

A comparison of stochastic placement and metric method template bank sizes is given in Fig. 7.6. We find that $N_{T,2D} \approx 1.3 \times N_{T,1D}$. The two-dimensional PN template banks are about 1.3 times the size of the one-dimensional Newtonian template banks made using the metric method. Therefore, by using Newtonian template banks we can achieve an even further decrease in the computational cost required to dig signals out of the *LISA* data.

7.4.4 Efficiency of the Template Bank

To determine the bank efficiencies, we filter 1000 3.5 PN waveform injections against each bank to find the best matching templates. For each of the two-dimensional banks, we calculate the efficiency twice – first, using the 3.5 PN approximation in the match calculation and again using only the Newtonian approximation. A summary of all the bank efficiency calculations performed is given in Table 7.1. The top right panel of Fig. 7.7 shows the cumulative distribution of matches for each of the two-dimensional template banks, where green lines show matches calculated with the Newtonian template approximation and blue lines show matches calculated with the 3.5 PN template approximation. The mean 10^{th} percentile match for both sets are indicated by the vertical lines. These results demonstrate that using the Newtonian approximation in the match calculation is sufficient for our purposes.

This result is further proven by the one-dimensional Newtonian template bank

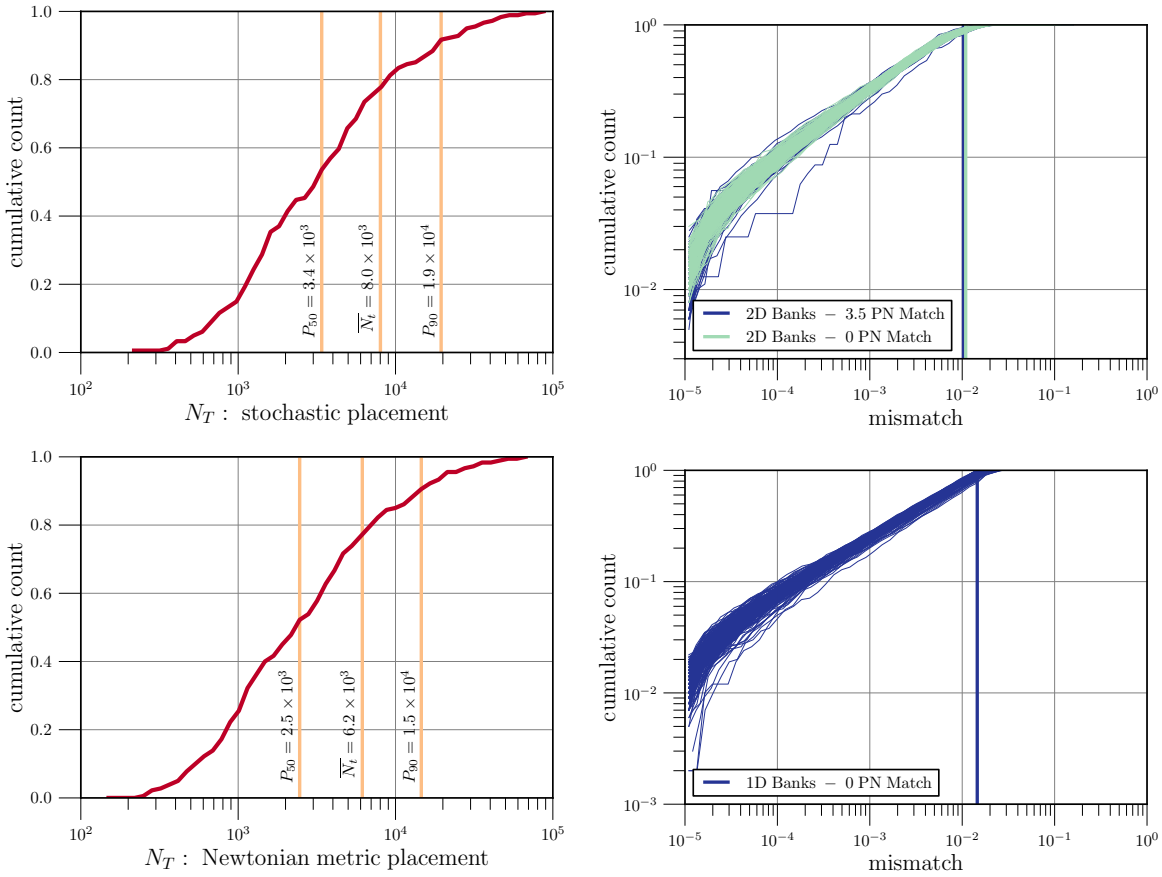

Figure 7.7. Top Left: The plot shows the cumulative distribution of the 3.5 PN template bank sizes. The vertical lines indicate, from left to right, the 50th percentile, mean, and 90th percentile bank sizes. The bins are evenly spaced in log. **Bottom Left:** The cumulative distribution of one-dimensional Newtonian template bank sizes. **Top Right:** The plot shows the cumulative distribution of mismatches of 1000 injections with each 3.5 PN template bank. In the match calculation, we use the Newtonian approximation (green) and the 3.5 PN approximation (blue). The vertical lines indicate the mean 10th percentile match for both sets of template banks. The position of the mean 10th percentile match value for 3.5 PN approximation has been artificially shifted to the left by 0.0005 so that the two lines are distinguishable. **Bottom Right:** The cumulative distribution of mismatches for one-dimensional Newtonian template banks.

efficiencies, shown in the bottom right panel of Fig. 7.7. As indicated by Table 7.1, for this set of bank efficiency calculations we use the Newtonian approximation both in the template waveforms and in the match calculation. Each of the three sets of bank efficiency calculations performed over the 181 *h*BBH systems had mean 10^{th}-percentile matches $M > 0.98$, so we conclude that these template banks will be effective in recovering *h*BBH signals from *LISA* data.

$\boldsymbol{\lambda}_T$	$\boldsymbol{\lambda}_{inj}$	Match approx.	$M(P_{10})$
3.5 PN (2D)	3.5 PN	3.5 PN	$0.989 \pm 7.6 \times 10^{-4}$
3.5 PN (2D)	3.5 PN	0 PN	$0.989 \pm 8.7 \times 10^{-4}$
0 PN (1D)	3.5 PN	0 PN	$0.985 \pm 1.5 \times 10^{-3}$

Table 7.1. A summary of the bank efficiency calculations done. The first column indicates the PN order of the template waveforms, and in parenthesis the dimensionality of the template banks. The second column similarly shows the PN order of the injected waveforms, highlighting that we have always used 3.5 PN injection waveforms. The third column shows the PN order used in the calculation of the match between template and injected waveforms. The last column shows the mean 10^{th} percentile match across all 181 *h*BBH sources. The one standard deviation error is quoted.

7.5 Visibility of stellar-mass BBH in *LISA*

We now calculate, for template banks of $\mathcal{O}(10^2) - \mathcal{O}(10^5)$ templates, the minimal signal-to-noise ratio required to claim a detection in a matched filter search. We first assume that we have a segment of data consisting only of noise, $d(t) = n(t)$. If we filter this data with one template, the SNR would be a random value which, for a Gaussian background, follows the Rayleigh distribution [362]:

$$p(\rho) = \rho \exp\left(\frac{-\rho^2}{2}\right). \tag{7.29}$$

The probability of obtaining a value of ρ greater than some threshold ρ_{thr} can be computed by integrating the distribution above the threshold,

$$p(\rho > \rho_{\text{thr}}) = \int_{\rho_{\text{thr}}}^{\infty} \rho \exp\left(\frac{-\rho^2}{2}\right) \mathrm{d}\rho = \exp\left(\frac{-\rho_{\text{thr}}^2}{2}\right). \tag{7.30}$$

We would like to find the value of ρ_{thr} above which we can confidently claim a detection. The false alarm probability (FAP) is then the chance of finding $\rho > \rho_{\text{thr}}$ in the noise hypothesis. We can choose an acceptable value of FAP, say 10^{-2} and use Eq. (7.30) to solve for ρ_{thr}:

$$\rho_{\text{thr}} = \sqrt{-2 \ln \text{FAP}}. \tag{7.31}$$

Therefore, with FAP $= 10^{-2}$ the threshold is $\rho_{\text{thr}} \approx 3.0$. However, the above discussion assumes that we filter the data with only one template. With N_T templates, the FAP is the probability that at least one template has $\rho > \rho_{\text{thr}}$. First, consider the probability that none of the N_T templates have $\rho > \rho_{\text{thr}}$, that is $(1 - p(\rho > \rho_{\text{thr}}))^{N_T}$. The FAP is

the compliment of this, or

$$\mathrm{FAP} = 1 - (1-p)^{N_T}. \tag{7.32}$$

For very small p, we can approximate this as $\mathrm{FAP} \approx 1 - (1 - p \cdot N_T) = p(\rho > \rho_{\rm thr}) \cdot N_T$. We see that the number of templates becomes a trials factor on the false alarm probability. We must then scale down our required FAP by the number of templates in Eq. (7.31),

$$\rho_{\rm thr} = \sqrt{-2 \ln \frac{\mathrm{FAP}}{N_T}}. \tag{7.33}$$

Finally, using $\mathrm{FAP} = 10^{-2}$ and $N_T = 10^2 - 10^5$, we find a threshold SNR of $\rho_{\rm thr} = 4.3$–5.7. This is a significant improvement over the previously quoted $\rho_{\rm thr} = 14$ [362] or $\rho_{\rm thr} = 8$ quoted in Ref. [366]. The lower threshold further improves the feasibility of archival matched-filter searches in *LISA*, increasing the number of detectable stellar mass binary black holes in *LISA* by a factor of ~ 6.

7.6 Conclusions and Future Directions

The simultaneous operation of third generation ground-based gravitational-wave observatories, such as the Einstein Telescope and Cosmic Explorer, and a space-based detector, *LISA*, provides a unique opportunity for multiband observations of stellar-mass and intermediate-mass binary black hole inspirals and mergers. Such multiband observations greatly improve the bounds one can place on general relativity and alternative theories of gravity [368, 370–372]. However, the computational cost required to dig the signals out of *LISA* data in a blind search would be formidable, requiring signal-to-noise ratios greater than 14 to make confident detections.

We have shown that 3G observatories would constrain most of the signal parameters—the component masses, spins, sky position, and time of coalescence—so tightly that it would be possible to carry out the search for these signals over a vastly reduced parameter space in *LISA* data. Indeed, high fidelity measurements enabled by the 3G observatories considered in this paper imply that archival searches in *LISA* data require only a one-dimensional template bank over the binary's chirp mass—the only parameter that is measured better by *LISA* than 3G observatories. Thus, the volume of parameter space necessary to search over is greatly reduced. With template banks containing only 10^2–10^5 templates, archival searches in *LISA* are feasible for signals of SNR as low as $\sim$ 4–6. This would allow for the possibility of $\mathcal{O}(100)$ multiband gravitational-wave

detections per year, a number that can vastly improve the quality of tests of general relativity that can be performed by 3G observatories or *LISA* by themselves [371, 372].

We have made several implicit approximations in this study that would need to be reexamined to ascertain that the principal conclusions of this paper remain valid. Firstly, we have assumed that the observed binaries would have negligible residual eccentricity when their signals enter the *LISA* sensitivity band. This is true for most binaries that spend millions of years to slowly spiral-in and merge during which radiation reaction tends to circularize the orbits. However, black hole binaries that form in rich clusters could have non-negligible eccentricities when *LISA* observes them. 3G observatories should be able to constrain residual eccentricities to a pretty good accuracy which could then be used to reduce the search space. Even if the eccentricity in the audio band of 3G observatories is vanishingly small it will be possible to evolve the orbits back and limit the search parameter space in the *LISA* band. Eccentricity will likely be a parameter that would not be measured very well by 3G observatories (as any residual eccentricity would have largely decayed) and *LISA* will likely constrain it better.

Secondly, we have assumed that companion black hole spins are aligned with the orbital angular momentum of the system. Spin-orbit and spin-spin couplings will significantly alter the orbital evolution only when spins are large and misaligned with the orbit. This will be a relatively small part of the parameter space. Moreover, spin effects occur at higher PN orders and are measured better by 3G observatories than *LISA*. It is, however, important to confirm that precessional effects are negligible in archival searches especially since they are cumulative effects and the signals spend many more cycles in the *LISA* band than they do in the audio band.

Finally, if the binaries live in a gas-rich environment then the inertial drag could accelerate the rate at which the companions spiral in. Such drags might not be relevant when the binaries enter the audio band of 3G detectors and hence it would not be possible to correct for the presence of environment. In the same spirit, we have not included any ultra-light boson clouds or other fields that might surround the companion black holes in the low-frequency phase of the evolution. This could alter the orbital evolution but would not be relevant at later stages [405, 406].

While not all of these effects are equally important, it is necessary to investigate the cost of including them in an archival search as they could potentially reveal important mechanisms in play in the formation, evolution, and environments of stellar-mass and intermediate-mass black hole binaries.

Acknowledgements

We thank Anuradha Gupta for providing access to her simulation of binary black holes that can be observed both in *LISA* and the 3G network. BE is supported by the Eberly Graduate Fellowship of Penn State. SS is supported by the Eberly Postdoctoral Fellowship of Penn State. BSS and SB are supported in part by NSF Grant No. PHY-1836779. This paper has the LIGO document number LIGO-P20xxxxx.

Chapter 8
Conclusion

8.1 Summary

Black holes and neutron stars are some of the most extreme objects in the known universe, containing the entire mass of our sun in the radius of a metropolitan area. When these objects orbit one another, their extreme densities cause them to dynamically bend the space and time around them sending ripples across space known as gravitational waves [15]. These gravitational waves siphon energy from the binary causing the two objects to inspiral until they eventually merge. The effect of these waves as they travel throughout the Universe is to cause distances between fixed points to oscillate, although the strain induced is typically on the order of 10^{-21} or smaller. Still, we can detect these faint signals using a network of laser interferometers across the world, which act as extremely precise rulers. The interferometric gravitational wave detectors known as LIGO, Virgo and KAGRA transformed our understanding of the universe by directly detecting gravitational waves from merging black holes for the first time in 2015 [12]. Gravitational waves from compact binary mergers give us a new window through which to observe the universe, allowing us to push the boundaries of what we know about fundamental physics, astrophysics, and cosmology. For example, gravitational waves have allowed us to put Einstein's theory of general relativity to the test more than one hundred years later. Through gravitational waves, we can learn about how, when, and where in the universe black holes and neutron stars are formed from stellar collapse and how matter behaves under the most extreme conditions. Even more lucrative is the opportunity to observe the same astrophysical event with both gravitational and electromagnetic radiation, which is possible when a gamma ray burst accompanies the merger. This is the object of the burgeoning field of multi-messenger astronomy, which was kickstarted with the now famous merging pair of neutron stars, GW170817, along

with its coincident gamma ray burst, GRB170817A [66, 67, 407].

Much of my dissertation work has focused on the development and operation of the low-latency GstLAL inspiral search pipeline [4, 111–113]. In Chapter 2, I detail the methods of this search pipeline in its most up to date form, at the time of writing. Further, I give an overview of the technical software implementation to describe how the pipeline was operated in the fourth observing run.

In Chapter 3, I discuss a novel strategy for the joint detection of sub-threshold gravitational waves and gamma ray bursts. One of the most enlightening detections we have made is the joint observation of gravitational waves from GW170817 and electromagnetic radiation from GRB170817A [67]. By observing the same system with both gravitational waves and electromagnetic radiation we can learn much more than would be possible with either messenger alone. By searching for gravitational waves on a day-to-day basis and making detections as rapidly as possible we can increase our chances of making a multi-messenger observation overall. However, to enable deep follow up, we must be able to precisely identify where on the sky a gravitational wave signal originated from. Our best chance of achieving this is when a GRB is associated with the merger, since the localization abilities of GRB telescopes such as *Swift* are much better than what we can achieve with gravitational waves alone. As we have seen over several years of observation, the chances of finding a loud gravitational wave event associated with strong GRB emissions are very slim. For this reason, I have led an effort to leverage "sub-threshold" candidates. These are possible gravitational wave events with relatively low significance. By using these sub-threshold events to trigger targeted sub-threshold GRB searches with the *Swift* telescope, we can search for coincident events too quiet to be detected in either instrument alone [154]. In this way, we can benefit from Swift's localization abilities, even for quiet gravitational wave signals or weak GRB emission, and enable follow up observations all across the electromagnetic spectrum. This kind of collaborative approach may be the key to detecting the next multi-messenger event.

While multi-messenger astronomy focuses on the science potential of multiple observations of a single event, many other areas of study rely upon analysis of large populations of diverse signals. The Gravitational Wave Transient Catalog (GWTC-3) includes all detections made throughout each observing run to date [**?**, 138, 164, 165]. These catalogs are produced by re-analyzing the data after the observing run has ended. In Chapter 4, I overview my contributions to the production of this catalog and the approximately 70 black hole and neutron star mergers added to it during LIGO and Virgo's third observing run (O3) [**?**, 138]. Building up this catalog of diverse gravitational wave signals is critical

for answering long standing questions of astrophysics. For example, how often do pairs of black holes and neutron stars merge? How and where are these binaries formed? What are the bulk properties of the populations of these objects in the universe? As we transition from analysis on an event-by-event basis to analysis on populations of events, we can start to constrain the answers to these questions, but a full understanding will only come with more detections.

Much of my work has focused on improving the GstLAL low-latency search pipeline towards LIGO and Virgo's fourth observing run (O4) which took place from May 2023 – January 2024 [113]. This work is given in Chapter 5. I developed a software suite to quantify and monitor the real-time performance of any compact binary gravitational wave search [283]. With this software, rich performance metrics are available for visualization via live web-based dashboards and mining via machine-readable databases. Using this software, we can understand the performance of the analysis and predict its behavior and sensitivity in the presence of astrophysical signals. This work led to the incredible success of the search in O4a, which I present in Chapter 6. We have seen detection rates more than double, with 81 new highly significant gravitational wave candidates. In addition, I analyze the performance of the GstLAL search pipeline individually, demonstrating its significant contribution to the overall success of the run.

As we look forward to the coming years of observation, it's important to prepare methods for gravitational wave detection with next-generation detectors. The Laser Interferometer Space-based Antenna (LISA) is a detector planned to launch into orbit around the sun in the next decade [78]. With isolation from the environmental noise of Earth, LISA will be extremely sensitive to the very early portions of compact binary inspirals, where LIGO and Virgo are completely blind. This means that binary black hole mergers will sweep through LISA's sensitive frequency band long before we detect them on the ground. In Chapter 7, I present results of a study where we find that it will be computationally infeasible to detect these signals in LISA's data with a blind search, however we can perform targeted archival searches for them after their detection in ground-based detectors [353,362]. My work has demonstrated that this method could lead to tens of black hole merger detections in LISA each year. By observing these signals in both the low and high frequency bands of space-based and ground-based interferometers, we can glean more information about their sources than would be possible in either instrument alone.

8.2 Future directions

Although the rate of growth in the field of gravitational wave astronomy has been unprecedented, there is still much to learn and many open questions remain. The $\mathcal{O}(100)$ gravitational wave detections so far have come along with surprises and unexpected new information about the universe and this trend will surely continue as we build up larger catalogs. Already, the second part of the fourth observing run (O4b) is ongoing. This run is expected to last for approximately ten months bringing the full duration of O4 to 18 months. This is by far the longest observing run by the LIGO, Virgo, and KAGRA Collaboration to date. Paired with the advances in detector sensitivity, O4 will cover a much larger sensitive time-volume than any previous run and this will come along with a corresponding increase in detections. In O4b, we look forward to the possibility of detecting the next multi-messenger BNS, especially with the possibility of an early-warning detection. There are still regions of the CBC parameter space which have yet to be fully probed: the upper mass gap where pair-instability supernova predict a dearth of black holes, the heavy IMBH mass region, or systems with high or un-aligned spin to name a few. Detections in any of these areas would bring a wealth of new possibilities for study and surely uncover new information about the universe.

Having nearly doubled the size of the existing catalog, the new gravitational wave observations from O4a presented in this dissertation will bring advances in our understanding of physics in many areas. By adding to the population, we will be able to reduce uncertainty in estimated merger rates and mass and spin populations [62]. With improved sensitivities, we expect observed signals in O4a and beyond to become louder in the LIGO and Virgo data than previous observing runs, leading to better tests of waveform accuracy and general relativity [65,254,255]. The extended sub-threshold list of candidates presents opportunities for deep targeted follow-up for coincident GRBs [408]. These are just a few examples of the studies that can be done using our gravitational wave detections. We look forward to many years of continued detection with LIGO, Virgo, and KAGRA and all the understanding and surprises that they bring.

Bibliography

[1] Collaboration, L. S., "LIGO's Interferometer," .
URL https://www.ligo.caltech.edu/page/ligos-ifo

[2] Adhikari, R. X. (2014) "Gravitational Radiation Detection with Laser Interferometry," *Rev. Mod. Phys.*, **86**, p. 121, 1305.5188.

[3] Abbott, R. et al. (2023) "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run," *Phys. Rev. X*, **13**(4), p. 041039, 2111.03606.

[4] Messick, C. et al. (2017) "Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data," *Phys. Rev. D*, **95**(4), p. 042001, 1604.04324.

[5] Joshi, P., L. Tsukada, and C. Hanna (2023) "Method for removing signal contamination during significance estimation of a GstLAL analysis," *Phys. Rev. D*, **108**(8), p. 084032, 2305.18233.

[6] Zhang, Z.-B. and C.-S. Choi (2008) "An analysis of the durations of Swift Gamma-Ray Bursts," *Astron. Astrophys.*, **484**, p. 293, 0708.4049.

[7] Abbott, B. P. et al. (2017) "Multi-messenger Observations of a Binary Neutron Star Merger," *Astrophys. J. Lett.*, **848**(2), p. L12, 1710.05833.

[8] Einstein, A. (1916) "Näherungsweise Integration der Feldgleichungen der Gravitation," *Sitzungsberichte der Köuml;niglich Preussischen Akademie der Wissenschaften*, pp. 688–696.

[9] ——— (1915) "Die Feldgleichungen der Gravitation," *Sitzungsberichte der Köuml;niglich Preussischen Akademie der Wissenschaften*, pp. 844–847.

[10] Denson Hill, C. and P. Nurowski (2016) "How the green light was given for gravitational wave search," 1608.08673.

[11] Cervantes-Cota, J. L., S. Galindo-Uribarri, and G.-F. Smoot (2016) "A Brief History of Gravitational Waves," *Universe*, **2**(3), p. 22, 1609.09400.

[12] Abbott, B. P. et al. (2016) "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.*, **116**(6), p. 061102, 1602.03837.

[13] Schutz, B. (2009) *A First Course in General Relativity*, 2 ed., Cambridge University Press.

[14] Creighton, J. D. E. and W. G. Anderson (2011) *Gravitational-wave physics and astronomy: An introduction to theory, experiment and data analysis*.

[15] Sathyaprakash, B. S. (2013) "Gravitational waves and astrophysical sources," *Comptes Rendus Physique*, **14**, pp. 272–287.

[16] Hulse, R. A. and J. H. Taylor (1975) "Discovery of a pulsar in a binary system," *Astrophys. J. Lett.*, **195**, pp. L51–L53.

[17] Weisberg, J. M. and J. H. Taylor (2005) "Relativistic binary pulsar B1913+16: Thirty years of observations and analysis," *ASP Conf. Ser.*, **328**, p. 25, astro-ph/0407149.

[18] Dimmelmeier, H., J. A. Font, and E. Muller (2002) "Relativistic simulations of rotational core collapse. 2. Collapse dynamics and gravitational radiation," *Astron. Astrophys.*, **393**, pp. 523–542, astro-ph/0204289.

[19] Ott, C. D., A. Burrows, L. Dessart, and E. Livne (2006) "A New Mechanism for Gravitational Wave Emission in Core-Collapse Supernovae," *Phys. Rev. Lett.*, **96**, p. 201102, astro-ph/0605493.

[20] Dimmelmeier, H., C. D. Ott, H.-T. Janka, A. Marek, and E. Mueller (2007) "Generic Gravitational Wave Signals from the Collapse of Rotating Stellar Cores," *Phys. Rev. Lett.*, **98**, p. 251101, astro-ph/0702305.

[21] Abbott, B. P. et al. (2017) "All-sky search for short gravitational-wave bursts in the first Advanced LIGO run," *Phys. Rev. D*, **95**(4), p. 042003, 1611.02972.

[22] ——— (2019) "All-Sky Search for Short Gravitational-Wave Bursts in the Second Advanced LIGO and Advanced Virgo Run," *Phys. Rev. D*, **100**(2), p. 024017, 1905.03457.

[23] Abbott, R. et al. (2021) "All-sky search for short gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run," *Phys. Rev. D*, **104**(12), p. 122004, 2107.03701.

[24] Bildsten, L. (1998) "Gravitational radiation and rotation of accreting neutron stars," *Astrophys. J. Lett.*, **501**, p. L89, astro-ph/9804325.

[25] Nayyar, M. and B. J. Owen (2006) "R-modes of accreting hyperon stars as persistent sources of gravitational waves," *Phys. Rev. D*, **73**, p. 084001, astro-ph/0512041.

[26] Hutchins, T. J. and D. I. Jones (2023) "Gravitational radiation from thermal mountains on accreting neutron stars: sources of temperature non-axisymmetry," *Mon. Not. Roy. Astron. Soc.*, **522**(1), pp. 226–251, 2212.07452.

[27] Abadie, J. et al. (2011) "Beating the spin-down limit on gravitational wave emission from the Vela pulsar," *Astrophys. J.*, **737**, p. 93, 1104.2712.

[28] Abbott, R. et al. (2021) "Diving below the Spin-down Limit: Constraints on Gravitational Waves from the Energetic Young Pulsar PSR J0537-6910," *Astrophys. J.*, **913**, p. L27, 2012.12926.

[29] Allen, B. (1996) "The Stochastic gravity wave background: Sources and detection," in *Les Houches School of Physics: Astrophysical Sources of Gravitational Radiation*, pp. 373–417, gr-qc/9604033.

[30] Christensen, N. (2019) "Stochastic Gravitational Wave Backgrounds," *Rept. Prog. Phys.*, **82**(1), p. 016903, 1811.08797.

[31] Romano, J. D. and N. J. Cornish (2017) "Detection methods for stochastic gravitational-wave backgrounds: a unified treatment," *Living Rev. Rel.*, **20**(1), p. 2, 1608.06889.

[32] Agazie, G. et al. (2023) "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background," *Astrophys. J. Lett.*, **951**(1), p. L8, 2306.16213.

[33] Caprini, C. and D. G. Figueroa (2018) "Cosmological Backgrounds of Gravitational Waves," *Class. Quant. Grav.*, **35**(16), p. 163001, 1801.04268.

[34] Sathyaprakash, B. S. and B. F. Schutz (2009) "Physics, Astrophysics and Cosmology with Gravitational Waves," *Living Rev. Rel.*, **12**, p. 2, 0903.0338.

[35] Abbott, B. P. et al. (2017) "The basic physics of the binary black hole merger GW150914," *Annalen Phys.*, **529**(1-2), p. 1600209, 1608.01940.

[36] Weber, J. (1960) "Detection and Generation of Gravitational Waves," *Phys. Rev.*, **117**, pp. 306–313.

[37] Aasi, J. et al. (2015) "Advanced LIGO," *Class. Quant. Grav.*, **32**, p. 074001, 1411.4547.

[38] Buikema, A. et al. (2020) "Sensitivity and performance of the Advanced LIGO detectors in the third observing run," *Phys. Rev. D*, **102**(6), p. 062003, 2008.01301.

[39] Acernese, F. et al. (2015) "Advanced Virgo: a second-generation interferometric gravitational wave detector," *Class. Quant. Grav.*, **32**(2), p. 024001, 1408.3978.

[40] Abe, H. et al. (2022) "The Current Status and Future Prospects of KAGRA, the Large-Scale Cryogenic Gravitational Wave Telescope Built in the Kamioka Underground," *Galaxies*, **10**(3), p. 63.

[41] Abbott, B. P. et al. (2009) "LIGO: The Laser interferometer gravitational-wave observatory," *Rept. Prog. Phys.*, **72**, p. 076901, 0711.3041.

[42] Accadia, T. et al. (2012) "Virgo: a laser interferometer to detect gravitational waves," *JINST*, **7**, p. P03012.

[43] Meers, B. J. (1988) "Recycling in Laser Interferometric Gravitational Wave Detectors," *Phys. Rev. D*, **38**, pp. 2317–2326.

[44] Arain, M. A. and G. Mueller (2008) "Design of the Advanced LIGO recycling cavities," *Opt. Express*, **16**(14), pp. 10018–10032.
URL https://opg.optica.org/oe/abstract.cfm?URI=oe-16-14-10018

[45] Mueller, C. L. et al. (2016) "The Advanced LIGO Input Optics," *Rev. Sci. Instrum.*, **87**(1), p. 014502, 1601.05442.

[46] Nguyen, P. et al. (2021) "Environmental noise in advanced LIGO detectors," *Class. Quant. Grav.*, **38**(14), p. 145001, 2101.09935.

[47] Acernese, F. et al. (2022) "The Virgo O3 run and the impact of the environment," *Class. Quant. Grav.*, **39**(23), p. 235009, 2203.04014.

[48] Aiello, L., E. Cesarini, V. Fafone, M. Lorenzini, Y. Minenkov, I. Nardecchia, A. Rocchi, and V. Sequino (2019) "Thermal compensation system in advanced and third generation gravitational wave interferometric detectors," *J. Phys. Conf. Ser.*, **1226**(1), p. 012019.

[49] Brooks, A. F. et al. (2021) "Point absorbers in Advanced LIGO," *Appl. Opt.*, **60**(13), pp. 4047–4063, 2101.05828.

[50] Caves, C. M. (1981) "Quantum Mechanical Noise in an Interferometer," *Phys. Rev. D*, **23**, pp. 1693–1708.

[51] Caves, C. (1980) "Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer," *Phys. Rev. Lett.*, **45**(2), pp. 75–79.

[52] Yap, M. J. et al. (2020) "Broadband reduction of quantum radiation pressure noise via squeezed light injection," *Nature Photon.*, **14**(1), pp. 19–23, 1812.09804.

[53] Ganapathy, D. et al. (2023) "Broadband Quantum Enhancement of the LIGO Detectors with Frequency-Dependent Squeezing," *Phys. Rev. X*, **13**(4), p. 041021.

[54] Blanchet, L. (2014) "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries," *Living Rev. Rel.*, **17**, p. 2, 1310.1528.

[55] Owen, B. J. (1996) "Search templates for gravitational waves from inspiraling binaries: Choice of template spacing," *Phys. Rev. D*, **53**, pp. 6749–6761, gr-qc/9511032.

[56] Manca, G. M. and M. Vallisneri (2010) "Cover art: Issues in the metric-guided and metric-less placement of random and stochastic template banks," *Phys. Rev. D*, **81**, p. 024004, 0909.0563.

[57] Roy, S., A. S. Sengupta, and P. Ajith (2019) "Effectual template banks for upcoming compact binary searches in Advanced-LIGO and Virgo data," *Phys. Rev. D*, **99**(2), p. 024048, 1711.08743.

[58] Hanna, C. et al. (2023) "Binary tree approach to template placement for searches for gravitational waves from compact binary mergers," *Phys. Rev. D*, **108**(4), p. 042003, 2209.11298.

[59] Abadie, J. et al. (2010) "Sensitivity to Gravitational Waves from Compact Binary Coalescences Achieved during LIGO's Fifth and Virgo's First Science Run," 1003.2481.

[60] Finn, L. S. and D. F. Chernoff (1993) "Observing binary inspiral in gravitational radiation: One interferometer," *Phys. Rev. D*, **47**, pp. 2198–2219, gr-qc/9301003.

[61] Abbott, R. et al. (2020) "Properties and Astrophysical Implications of the 150 $M_\odot$ Binary Black Hole Merger GW190521," *Astrophys. J. Lett.*, **900**(1), p. L13, 2009.01190.

[62] ——— (2023) "Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3," *Phys. Rev. X*, **13**(1), p. 011048, 2111.03634.

[63] (2024) "Observation of Gravitational Waves from the Coalescence of a $2.5-4.5\ M_\odot$ Compact Object and a Neutron Star," 2404.04248.

[64] Abbott, B. P. et al. (2017) "A gravitational-wave standard siren measurement of the Hubble constant," *Nature*, **551**(7678), pp. 85–88, 1710.05835.

[65] Abbott, R. et al. (2021) "Tests of General Relativity with GWTC-3," 2112.06861.

[66] Abbott, B. P. et al. (2017) "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.*, **119**(16), p. 161101, 1710.05832.

[67] ——— (2017) "Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J. Lett.*, **848**(2), p. L13, 1710.05834.

[68] Drout, M. R. et al. (2017) "Light Curves of the Neutron Star Merger GW170817/SSS17a: Implications for R-Process Nucleosynthesis," *Science*, **358**, pp. 1570–1574, 1710.05443.

[69] Tanvir, N. R. et al. (2017) "The Emergence of a Lanthanide-Rich Kilonova Following the Merger of Two Neutron Stars," *Astrophys. J. Lett.*, **848**(2), p. L27, 1710.05455.

[70] Ellis, J., B. D. Fields, and R. Surman (2024) "Do we Owe our Existence to Gravitational Waves?" 2402.03593.

[71] Punturo, M. et al. (2010) "The Einstein Telescope: A third-generation gravitational wave observatory," *Class. Quant. Grav.*, **27**, p. 194002.

[72] Reitze, D. et al. (2019) "The US Program in Ground-Based Gravitational Wave Science: Contribution from the LIGO Laboratory," *Bull. Am. Astron. Soc.*, **51**, p. 141, 1903.04615.

[73] ——— (2019) "Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO," *Bull. Am. Astron. Soc.*, **51**(7), p. 035, 1907.04833.

[74] Sathyaprakash, B. et al. (2012) "Scientific Objectives of Einstein Telescope," *Class. Quant. Grav.*, **29**, p. 124013, [Erratum: Class.Quant.Grav. 30, 079501 (2013)], 1206.0331.

[75] Maggiore, M. et al. (2020) "Science Case for the Einstein Telescope," *JCAP*, **03**, p. 050, 1912.02622.

[76] Evans, M. et al. (2021) "A Horizon Study for Cosmic Explorer: Science, Observatories, and Community," 2109.09882.

[77] Amaro-Seoane, P. et al. (2012) "Low-frequency gravitational-wave science with eLISA/NGO," *Class. Quant. Grav.*, **29**, p. 124016, 1202.0839.

[78] Baker, J. et al. (2019) "The Laser Interferometer Space Antenna: Unveiling the Millihertz Gravitational Wave Sky," 1907.06482.

[79] Littenberg, T. B. and N. Yunes (2019) "Binary White Dwarfs as Laboratories for Extreme Gravity with LISA," *Class. Quant. Grav.*, **36**(9), p. 095017, 1811.01093.

[80] Gair, J. R., M. Vallisneri, S. L. Larson, and J. G. Baker (2013) "Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors," *Living Rev. Rel.*, **16**, p. 7, 1212.5575.

[81] Barausse, E. et al. (2020) "Prospects for Fundamental Physics with LISA," *Gen. Rel. Grav.*, **52**(8), p. 81, 2001.09793.

[82] Auclair, P. et al. (2023) "Cosmology with the Laser Interferometer Space Antenna," *Living Rev. Rel.*, **26**(1), p. 5, 2204.05434.

[83] Oppenheim, A. V., R. W. Schafer, and C. K. Yuen (1978) "Digital Signal Processing," *IEEE Transactions on Systems, Man, and Cybernetics*, **8**(2), pp. 146–146.

[84] Proakis, J. G. and D. G. Manolakis (1996) *Digital Signal Processing: Principles, Algorithms, and Applications*, Prentice Hall.

[85] Cannon, K., A. Chapman, M. Crispin-Ortuzar, C. Hanna, D. Keppel, A. Searle, and A. Weinstein (2009), "LLOID!? Low-Latency Online Inspiral Data Analysis," .
URL `https://dcc.ligo.org/DocDB/0005/G0900765/001/LIGO-G0900765.pdf`

[86] Abbott, B. P. et al. (2016) "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914," *Class. Quant. Grav.*, **33**(13), p. 134001, `1602.03844`.

[87] ——— (2020) "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals," *Class. Quant. Grav.*, **37**(5), p. 055002, `1908.11170`.

[88] Essick, R., P. Godwin, C. Hanna, L. Blackburn, and E. Katsavounidis (2020) "iDQ: Statistical Inference of Non-Gaussian Noise with Auxiliary Degrees of Freedom in Gravitational-Wave Detectors," *Mach. Learn.: Sci. Technol.*, **2**(1), p. 015004, `2005.12761`.

[89] Zevin, M. et al. (2017) "Gravity Spy: Integrating Advanced LIGO Detector Characterization, Machine Learning, and Citizen Science," *Class. Quant. Grav.*, **34**(6), p. 064003, `1611.04596`.

[90] Soni, S. et al. (2021) "Discovering features in gravitational-wave data through detector characterization, citizen science and machine learning," *Class. Quant. Grav.*, **38**(19), p. 195016, `2103.12104`.

[91] Cabero, M., A. Mahabal, and J. McIver (2020) "GWSkyNet: a real-time classifier for public gravitational-wave candidates," *Astrophys. J. Lett.*, **904**(1), p. L9, `2010.11829`.

[92] Cannon, K. et al. (2012) "Toward Early-Warning Detection of Gravitational Waves from Compact Binary Coalescence," *Astrophys. J.*, **748**, p. 136, `1107.2665`.

[93] Metzger, B. D. and A. L. Piro (2014) "Optical and X-ray emission from stable millisecond magnetars formed from the merger of binary neutron stars," *Mon. Not. Roy. Astron. Soc.*, **439**, pp. 3916–3930, `1311.1519`.

[94] Ciolfi, R. and D. M. Siegel (2015) "Short gamma-ray bursts in the "time-reversal" scenario," *Astrophys. J. Lett.*, **798**(2), p. L36, `1411.2015`.

[95] Nicholl, M. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. III. Optical and UV Spectra of a Blue Kilonova From Fast Polar Ejecta," *Astrophys. J. Lett.*, **848**(2), p. L18, 1710.05456.

[96] Margalit, B. and B. D. Metzger (2017) "Constraining the Maximum Mass of Neutron Stars From Multi-Messenger Observations of GW170817," *Astrophys. J. Lett.*, **850**(2), p. L19, 1710.05938.

[97] Metzger, B. D. (2017) "Welcome to the Multi-Messenger Era! Lessons from a Neutron Star Merger and the Landscape Ahead," 1710.05931.

[98] ——— (2020) "Kilonovae," *Living Rev. Rel.*, **23**(1), p. 1, 1910.01617.

[99] Most, E. R. and A. A. Philippov (2020) "Electromagnetic precursors to gravitational wave events: Numerical simulations of flaring in pre-merger binary neutron star magnetospheres," *Astrophys. J. Lett.*, **893**(1), p. L6, 2001.06037.

[100] Shannon, C. E. (1949) "Communication in the Presence of Noise," .

[101] Cannon, K., A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein (2010) "Singular value decomposition applied to compact binary coalescence gravitational-wave signals," *Phys. Rev. D*, **82**, p. 044025, 1005.0012.

[102] Cannon, K., C. Hanna, D. Keppel, and A. C. Searle (2011) "Composite gravitational-wave detection of compact binary coalescence," *Phys. Rev. D*, **83**, p. 084053, 1101.0584.

[103] Golub, G. H. and C. F. Van Loan (2013) *Matrix Computations*, 4th ed., Johns Hopkins University Press.

[104] 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," 1901.08580.

[105] Mukherjee, D. et al. (2021) "Template bank for spinning compact binary mergers in the second observation run of Advanced LIGO and the first observation run of Advanced Virgo," *Phys. Rev. D*, **103**(8), p. 084047, 1812.05121.

[106] Morisaki, S. and V. Raymond (2020) "Rapid Parameter Estimation of Gravitational Waves from Binary Neutron Star Coalescence using Focused Reduced Order Quadrature," *Phys. Rev. D*, **102**(10), p. 104020, 2007.09108.

[107] Sakon, S. et al. (2024) "Template bank for compact binary mergers in the fourth observing run of Advanced LIGO, Advanced Virgo, and KAGRA," *Phys. Rev. D*, **109**(4), p. 044066, 2211.16674.

[108] Neyman, J. and E. S. Pearson (1933) "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Phil. Trans. Roy. Soc. Lond. A*, **231**(694-706), pp. 289–337.

[109] Biswas, R. et al. (2012) "Likelihood-ratio ranking of gravitational-wave candidates in a non-Gaussian background," *Phys. Rev. D*, **85**, p. 122008, 1201.2959.

[110] Cannon, K., C. Hanna, and J. Peoples (2015), "Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation," 1504.04632.

[111] Tsukada, L. et al. (2023) "Improved ranking statistics of the GstLAL inspiral search for compact binary coalescences," *Phys. Rev. D*, **108**(4), p. 043004, 2305.06286.

[112] Cannon, K. et al. (2020), "GstLAL: A software framework for gravitational wave discovery," 2010.05082.

[113] Ewing, B. et al. (2024) "Performance of the low-latency GstLAL inspiral search towards LIGO, Virgo, and KAGRA's fourth observing run," *Phys. Rev. D*, **109**(4), p. 042008, 2305.05625.

[114] Hanna, C. et al. (2022) "Metric assisted stochastic sampling search for gravitational waves from binary black hole mergers," *Phys. Rev. D*, **106**(8), p. 084033, 2110.15463.

[115] Abbott, R. et al. (2021) "Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences," *Astrophys. J. Lett.*, **915**(1), p. L5, 2106.15163.

[116] Forbush, S. E. (1946) "Three Unusual Cosmic-Ray Increases Possibly Due to Charged Particles from the Sun," *Phys. Rev.*, **70**, pp. 771–772.

[117] Kunkel, W. and Madore, B. and Shelton, Ian and others (1987), "IAUC 4316: 1987A; N Cen 1986," .
URL http://www.cbat.eps.harvard.edu/iauc/04300/04316.html

[118] Herald, D. and McNaught, R. H. and Madore, B. and others (1987), "IAUC 4317: 1987A," .
URL http://www.cbat.eps.harvard.edu/iauc/04300/04317.html

[119] Burrows, A. and J. M. Lattimer (1987) "Neutrinos from SN 1987A," *Astrophys. J. Lett.*, **318**, pp. L63–L68.

[120] Keivani, A. et al. (2018) "A Multimessenger Picture of the Flaring Blazar TXS 0506+056: implications for High-Energy Neutrino Emission and Cosmic Ray Acceleration," *Astrophys. J.*, **864**(1), p. 84, 1807.04537.

[121] Mészáros, P., D. B. Fox, C. Hanna, and K. Murase (2019) "Multi-Messenger Astrophysics," *Nature Rev. Phys.*, **1**, pp. 585–599, 1906.10212.

[122] Klebesadel, R. W., I. B. Strong, and R. A. Olson (1973) "Observations of Gamma-Ray Bursts of Cosmic Origin," *Astrophys. J. Lett.*, **182**, pp. L85–L88.

[123] Kouveliotou, C., C. A. Meegan, G. J. Fishman, N. P. Bhyat, M. S. Briggs, T. M. Koshut, W. S. Paciesas, and G. N. Pendleton (1993) "Identification of two classes of gamma-ray bursts," *Astrophys. J. Lett.*, **413**, pp. L101–104.

[124] Galama, T. J. et al. (1998) "Discovery of the peculiar supernova 1998bw in the error box of GRB 980425," *Nature*, **395**, p. 670, astro-ph/9806175.

[125] Woosley, S. E. and J. S. Bloom (2006) "The Supernova Gamma-Ray Burst Connection," *Ann. Rev. Astron. Astrophys.*, **44**, pp. 507–556, astro-ph/0609142.

[126] Savchenko, V. et al. (2017) "INTEGRAL Detection of the First Prompt Gamma-Ray Signal Coincident with the Gravitational-wave Event GW170817," *Astrophys. J. Lett.*, **848**(2), p. L15, 1710.05449.

[127] Coulter, D. A. and Kilpatrick, C. D., and Siebert, M. R., and Foley, R. J. and others (2017) *GCN*, **21529**.
URL https://gcn.nasa.gov/circulars/21529

[128] Coulter, D. A. et al. (2017) "Swope Supernova Survey 2017a (SSS17a), the Optical Counterpart to a Gravitational Wave Source," *Science*, **358**, p. 1556, 1710.05452.

[129] Lien, A. et al. (2016) "The Third Swift Burst Alert Telescope Gamma-Ray Burst Catalog," *Astrophys. J.*, **829**(1), p. 7, 1606.01956.

[130] von Kienlin, A. et al. (2020) "The Fourth Fermi-GBM Gamma-Ray Burst Catalog: A Decade of Data," *Astrophys. J.*, **893**, p. 46, 2002.11460.

[131] Zhang, B. (2011) "Open Questions in GRB Physics," *Comptes Rendus Physique*, **12**, pp. 206–225, 1104.0932.

[132] ——— (2014) "Gamma-Ray Burst Prompt Emission," *Int. J. Mod. Phys. D*, **23**, p. 1430002, 1402.7022.

[133] Ciolfi, R. (2018) "Short gamma-ray burst central engines," *Int. J. Mod. Phys. D*, **27**(13), p. 1842004, 1804.03684.

[134] ——— (2020) "Binary neutron star mergers after GW170817," *Front. Astron. Space Sci.*, **7**, p. 27, 2005.02964.

[135] Gottlieb, O., E. Nakar, and O. Bromberg (2020) "The structure of hydrodynamic γ-ray burst jets," *Mon. Not. Roy. Astron. Soc.*, **500**(3), pp. 3511–3526, 2006.02466.

[136] Breschi, M., A. Perego, S. Bernuzzi, W. Del Pozzo, V. Nedora, D. Radice, and D. Vescovi (2021) "AT2017gfo: Bayesian inference and model selection of multicomponent kilonovae and constraints on the neutron star equation of state," *Mon. Not. Roy. Astron. Soc.*, **505**(2), pp. 1661–1677, 2101.01201.

[137] Pillas, M. (2023) *Exploring the physics of neutron star mergers with gravitational waves and gamma-ray bursts*, Ph.D. thesis, U. Paris-Saclay.

[138] Abbott, R. et al. (2021) "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," *Phys. Rev. X*, **11**, p. 021053, 2010.14527.

[139] Fong, W.-f., E. Berger, R. Margutti, and B. A. Zauderer (2015) "A Decade of Short-duration Gamma-ray Burst Broadband Afterglows: Energetics, Circumburst Densities, and jet Opening Angles," *Astrophys. J.*, **815**(2), p. 102, 1509.02922.

[140] LIGO Scientific Collaboration and Virgo Collaboration (2017) *GCN*, **21505**.
URL https://gcn.gsfc.nasa.gov/other/G298048.gcn3

[141] Fermi-GBM (2017) *GCN*, **524666471**.
URL https://gcn.gsfc.nasa.gov/other/524666471.fermi

[142] Soares-Santos, M. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. I. Discovery of the Optical Counterpart Using the Dark Energy Camera," *Astrophys. J. Lett.*, **848**(2), p. L16, 1710.05459.

[143] Cowperthwaite, P. S. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. II. UV, Optical, and Near-infrared Light Curves and Comparison to Kilonova Models," *Astrophys. J. Lett.*, **848**(2), p. L17, 1710.05840.

[144] Chornock, R. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. IV. Detection of Near-infrared Signatures of r-process Nucleosynthesis with Gemini-South," *Astrophys. J. Lett.*, **848**(2), p. L19, 1710.05454.

[145] Margutti, R. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. V. Rising X-ray Emission from an Off-Axis Jet," *Astrophys. J. Lett.*, **848**(2), p. L20, 1710.05431.

[146] Alexander, K. D. et al. (2017) "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/VIRGO GW170817. VI. Radio Constraints on a Relativistic Jet and Predictions for Late-Time Emission from the Kilonova Ejecta," *Astrophys. J. Lett.*, **848**(2), p. L21, 1710.05457.

[147] Gehrels, N. et al. (2004) "The Swift Gamma-Ray Burst Mission," *Astrophys. J.*, **611**, pp. 1005–1020, [Erratum: Astrophys.J. 621, 558 (2005)], astro-ph/0405233.

[148] Barthelmy, S. D. et al. (2005) "The Burst Alert Telescope (BAT) on the Swift MIDEX mission," *Space Sci. Rev.*, **120**, p. 143, astro-ph/0507410.

[149] Burrows, D. N. et al. (2005) "The Swift X-ray Telescope," *Space Sci. Rev.*, **120**, p. 165, astro-ph/0508071.

[150] Roming, P. W. A. et al. (2005) "The Swift Ultra-Violet/Optical Telescope," *Space Sci. Rev.*, **120**, pp. 95–142, astro-ph/0507413.

[151] Sakamoto, T. et al. (2008) "The First Swift BAT Gamma-Ray Burst Catalog," *Astrophys. J. Suppl.*, **175**, pp. 179–190, 0707.4626.

[152] ——— (2011) "The Second Swift BAT Gamma-Ray Burst Catalog," *Astrophys. J. Suppl.*, **195**, p. 2, 1104.4689.

[153] Fenimore, E., D. Palmer, M. Galassi, T. Tavenner, S. Barthelmy, N. Gehrels, A. Parsons, and J. Tueller (2003) "The Trigger algorithm for the Burst Alert Telescope on Swift," *AIP Conf. Proc.*, **662**(1), p. 491, astro-ph/0408514.

[154] Tohuvavohu, A., J. A. Kennea, J. DeLaunay, D. M. Palmer, S. B. Cenko, and S. Barthelmy (2020) "Gamma-Ray Urgent Archiver for Novel Opportunities (GUANO): Swift/BAT Event Data Dumps on Demand to Enable Sensitive Subthreshold GRB Searches," *Astrophys. J.*, **900**(1), p. 35, 2005.01751.

[155] DeLaunay, J. and A. Tohuvavohu (2022) "Harvesting BAT-GUANO with NITRATES (Non-Imaging Transient Reconstruction and Temporal Search): Detecting and Localizing the Faintest Gamma-Ray Bursts with a Likelihood Framework," *Astrophys. J.*, **941**(2), p. 169, 2111.01769.

[156] LIGO Scientific Collaboration and Virgo Collaboration (2023), "EMFollow User Guide," .
URL https://emfollow.docs.ligo.org/userguide/

[157] Singer, L. P. and L. R. Price (2016), "Rapid Bayesian position reconstruction for gravitational-wave transients," 1508.03634.

[158] Raman, G., S. Ronchini, J. DeLauney, A. Tohuvavohu, J. Kennea, and T. Parsotan (2024) "Swift-BAT GUANO follow-up of gravitational-wave triggers in the third LIGO–Virgo–KAGRA observing run," .
URL https://dcc.ligo.org/DocDB/0190/P2300364/005/BAT_GUANO_O3.pdf

[159] LIGO Scientific Collaboration (2024), "Raven Documentation: Joint FAR," .
URL https://lscsoft.docs.ligo.org/raven/joint_far.html

[160] Ashton, G., E. Burns, T. D. Canton, T. Dent, H. B. Eggenstein, A. B. Nielsen, R. Prix, M. Was, and S. J. Zhu (2018) "Coincident detection significance in multimessenger astronomy," *Astrophys. J.*, **860**(1), p. 6, 1712.05392.

[161] Zevin, M., S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow (2021) "One Channel to Rule Them All? Constraining the Origins of Binary Black Holes Using Multiple Formation Pathways," *Astrophys. J.*, **910**(2), p. 152, 2011.10057.

[162] Fishbach, M. and L. van Son (2023) "LIGO–Virgo–KAGRA's Oldest Black Holes: Probing Star Formation at Cosmic Noon With GWTC-3," *Astrophys. J. Lett.*, **957**(2), p. L31, 2307.15824.

[163] Vitale, S., W. M. Farr, K. Ng, and C. L. Rodriguez (2019) "Measuring the star formation rate with gravitational waves from binary black holes," *Astrophys. J. Lett.*, **886**(1), p. L1, 1808.00901.

[164] Abbott, B. P. et al. (2019) "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X*, **9**(3), p. 031040, 1811.12907.

[165] Abbott, R. et al. (2024) "GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run," *Phys. Rev. D*, **109**(2), p. 022001, 2108.01045.

[166] Bailyn, C. D., R. K. Jain, P. Coppi, and J. A. Orosz (1998) "The Mass distribution of stellar black holes," *Astrophys. J.*, **499**, p. 367, astro-ph/9708032.

[167] Kreidberg, L., C. D. Bailyn, W. M. Farr, and V. Kalogera (2012) "Mass Measurements of Black Holes in X-Ray Transients: Is There a Mass Gap?" *Astrophys. J.*, **757**, p. 36, 1205.1805.

[168] Woosley, S. E. and A. Heger (2021) "The Pair-Instability Mass Gap for Black Holes," *Astrophys. J. Lett.*, **912**(2), p. L31, 2103.07933.

[169] Fowler, W. A. and F. Hoyle (1964) "Neutrino Processes and Pair Formation in Massive Stars and Supernovae," *Astrophys. J. Suppl.*, **9**, pp. 201–319.

[170] LIGO Scientific Collaboration, Virgo Collaboration and KAGRA Collaboration (2021), "GWTC-3 Data Release," www.gw-openscience.org/GWTC-3/.

[171] Chen, H.-Y., D. E. Holz, J. Miller, M. Evans, S. Vitale, and J. Creighton (2021) "Distance measures in gravitational-wave astrophysics and cosmology," *Class. Quant. Grav.*, **38**(5), p. 055010, 1709.08079.

[172] Allen, B., W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton (2012) "FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries," *Phys. Rev. D*, **85**, p. 122006, gr-qc/0509116.

[173] Abbott, B. P. et al. (2018) "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA," *Living Rev. Rel.*, **21**(1), p. 3, 1304.0670.

[174] Barsotti, L., J. Harms, and R. Schnabel (2019) "Squeezed vacuum states of light for gravitational wave detectors," *Rept. Prog. Phys.*, **82**(1), p. 016905.

[175] Tse, M. et al. (2019) "Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy," *Phys. Rev. Lett.*, **123**(23), p. 231107.

[176] McCuller, L. et al. (2021) "LIGO's quantum response to squeezed states," *Phys. Rev. D*, **104**(6), p. 062006, 2105.12052.

[177] Ottaway, D. J., P. Fritschel, and S. J. Waldman (2012) "Impact of upconverted scattered light on advanced interferometric gravitational wave detectors," *Opt. Express*, **20**(8), pp. 8329–8336.

[178] Soni, S. et al. (2020) "Reducing scattered light in LIGO's third observing run," *Class. Quant. Grav.*, **38**(2), p. 025016, 2007.14876.

[179] Brooks, J., M. Mantovani, A. Allocca, J. C. Diaz, V. Dattilo, A. Masserot, and P. Ruggi (2020) "Temperature Control for an Intra-Mirror Etalon in Interferometric Gravitational Wave Detector Fabry–Perot Cavities," *Galaxies*, **8**(4), p. 80.

[180] Accadia, T. et al. (2010) "Noise from scattered light in Virgo's second science run data," *Class. Quant. Grav.*, **27**, p. 194011.

[181] Acernese, F. et al. (2019) "Increasing the Astrophysical Reach of the Advanced Virgo Detector via the Application of Squeezed Vacuum States of Light," *Phys. Rev. Lett.*, **123**(23), p. 231108.

[182] ——— (2020) "Quantum Backaction on kg-Scale Mirrors: Observation of Radiation Pressure Noise in the Advanced Virgo Detector," *Phys. Rev. Lett.*, **125**(13), p. 131101.

[183] Viets, A. et al. (2018) "Reconstructing the calibrated strain signal in the Advanced LIGO detectors," *Class. Quant. Grav.*, **35**(9), p. 095015, 1710.09973.

[184] Sun, L. et al. (2021) "Characterization of systematic error in Advanced LIGO calibration in the second half of O3," 2107.00129.

[185] Acernese, F. et al. (2018) "Calibration of Advanced Virgo and Reconstruction of the Gravitational Wave Signal $h(t)$ during the Observing Run O2," *Class. Quant. Grav.*, **35**(20), p. 205004, 1807.03275.

[186] ——— (2022) "Calibration of advanced Virgo and reconstruction of the detector strain h(t) during the observing run O3," *Class. Quant. Grav.*, **39**(4), p. 045006, 2107.03294.

[187] Abbott, B. P. et al. (2018) "Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run," *Class. Quant. Grav.*, **35**(6), p. 065010, 1710.02185.

[188] Davis, D. et al. (2021) "LIGO detector characterization in the second and third observing runs," *Class. Quant. Grav.*, **38**(13), p. 135014, 2101.11673.

[189] Acernese, F. et al. (2022) "Virgo Detector Characterization and Data Quality during the O3 run," 2205.01555.

[190] Davis, D. et al. (2021) *Data Quality Vetoes Applied to the Analysis of LIGO Data from the Third Observing Run*, *Tech. Rep. DCC-T2100045*, LIGO.
URL https://dcc.ligo.org/LIGO-T2100045/public

[191] Arnaud, N. et al. (2021) *Category 1 (CAT1) data quality vetoes applied to the analysis of the O3 run Virgo data*, *Tech. Rep. VIR-0560A-21*, Virgo.
URL https://tds.virgo-gw.eu/?content=3&r=18844

[192] Davis, D., T. B. Littenberg, I. M. Romero-Shaw, M. Millhouse, J. McIver, F. Di Renzo, and G. Ashton (2022) "Subtracting glitches from gravitational-wave detector data during the third LIGO-Virgo observing run," *Class. Quant. Grav.*, **39**(24), p. 245013, 2207.03429.

[193] Klimenko, S. and G. Mitselmakher (2004) "A wavelet method for detection of gravitational wave bursts," *Class. Quant. Grav.*, **21**, pp. S1819–S1830.

[194] Klimenko, S. et al. (2016) "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors," *Phys. Rev. D*, **93**(4), p. 042004, 1511.05999.

[195] Adams, T., D. Buskulic, V. Germain, G. M. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, and G. Wang (2016) "Low-latency analysis pipeline for compact binary coalescences in the advanced gravitational wave detector era," *Class. Quant. Grav.*, **33**(17), p. 175012, 1512.02864.

[196] Aubin, F. et al. (2021) "The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run," *Class. Quant. Grav.*, **38**(9), p. 095004, 2012.11512.

[197] Usman, S. A. et al. (2016) "The PyCBC search for gravitational waves from compact binary coalescence," *Class. Quant. Grav.*, **33**(21), p. 215004, 1508.02357.

[198] Nitz, A. H., T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown (2017) "Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search," *Astrophys. J.*, **849**(2), p. 118, 1705.01513.

[199] Luan, J., S. Hooper, L. Wen, and Y. Chen (2012) "Towards low-latency real-time detection of gravitational waves from compact binary coalescences in the era of advanced detectors," *Phys. Rev. D*, **85**, p. 102002, 1108.3174.

[200] Chu, Q. et al. (2022) "SPIIR online coherent pipeline to search for gravitational waves from compact binary coalescences," *Phys. Rev. D*, **105**(2), p. 024023, 2011.06787.

[201] Bohé, A. et al. (2017) "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**(4), p. 044028, 1611.03703.

[202] Pürrer, M. (2016) "Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins," *Phys. Rev. D*, **93**(6), p. 064041, 1512.02248.

[203] Buonanno, A., B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash (2009) "Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors," *Phys. Rev. D*, **80**, p. 084043, 0907.0700.

[204] Was, M., M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis (2010) "On the background estimation by time slides in a network of gravitational wave detectors," *Class. Quant. Grav.*, **27**, p. 015005, 0906.2120.

[205] Farr, W. M., J. R. Gair, I. Mandel, and C. Cutler (2015) "Counting And Confusion: Bayesian Rate Estimation With Multiple Populations," *Phys. Rev. D*, **91**(2), p. 023005, 1302.5341.

[206] Kapadia, S. J. et al. (2020) "A self-consistent method to estimate the rate of compact binary coalescences with a Poisson mixture model," *Class. Quant. Grav.*, **37**(4), p. 045007, 1903.06881.

[207] Tiwari, V. (2018) "Estimation of the Sensitive Volume for Gravitational-wave Source Populations Using Weighted Monte Carlo Integration," *Class. Quant. Grav.*, **35**(14), p. 145009, 1712.00482.

[208] Fong, H. K. Y. (2018) *From simulations to signals: Analyzing gravitational waves from compact binary coalescences*, Ph.D. thesis, Toronto U.
URL https://tspace.library.utoronto.ca/handle/1807/91831

[209] Andres, N. et al. (2022) "Assessing the compact-binary merger candidates reported by the MBTA pipeline in the LIGO–Virgo O3 run: probability of astrophysical origin, classification, and associated uncertainties," *Class. Quant. Grav.*, **39**(5), p. 055002, 2110.10997.

[210] Dent, T. (2021) *Extending the PyCBC pastro calculation to a global network*, *Tech. Rep. DCC-T2100060*, LIGO.
URL https://dcc.ligo.org/LIGO-T2100060/public

[211] LIGO Scientific Collaboration and Virgo Collaboration (2019) *GCN*, **26350**.
URL https://gcn.gsfc.nasa.gov/other/S191205ah.gcn3

[212] ——— (2020) *GCN*, **27042**.
URL https://gcn.gsfc.nasa.gov/other/S200213t.gcn3

[213] McBrien O. and others (2019) *GCN*, **26485**.
URL https://gcn.gsfc.nasa.gov/gcn/gcn3/26485.gcn3

[214] LIGO Scientific Collaboration and Virgo Collaboration (2020) *GCN*, **26640**.
URL https://gcn.gsfc.nasa.gov/other/S200105ae.gcn3

[215] ——— (2020) *GCN*, **26734**.
URL https://gcn.gsfc.nasa.gov/other/S200114f.gcn3

[216] Abbott, R. et al. (2022) "Search for intermediate-mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo," *Astron. Astrophys.*, **659**, p. A84, 2105.15120.

[217] Miller, M. C. and E. J. M. Colbert (2004) "Intermediate - mass black holes," *Int. J. Mod. Phys. D*, **13**, pp. 1–64, astro-ph/0308402.

[218] Greene, J. E., J. Strader, and L. C. Ho (2020) "Intermediate-Mass Black Holes," *Ann. Rev. Astron. Astrophys.*, **58**, pp. 257–312, 1911.09678.

[219] Foucart, F. (2012) "Black Hole-Neutron Star Mergers: Disk Mass Predictions," *Phys. Rev. D*, **86**, p. 124007, 1207.6304.

[220] Foucart, F., T. Hinderer, and S. Nissanke (2018) "Remnant baryon mass in neutron star-black hole mergers: Predictions for binary neutron star mimickers and rapidly spinning black holes," *Phys. Rev. D*, **98**(8), p. 081501, 1807.00011.

[221] Barbieri, C., O. S. Salafia, A. Perego, M. Colpi, and G. Ghirlanda (2020) "Electromagnetic counterparts of black hole-neutron star mergers: dependence on the neutron star properties," *Eur. Phys. J. A*, **56**(1), p. 8, 1908.08822.

[222] ——— (2019) "Light-curve models of black hole – neutron star mergers: steps towards a multi-messenger parameter estimation," *Astron. Astrophys.*, **625**, p. A152, 1903.04543.

[223] Ossokine, S. et al. (2020) "Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construction and Validation," *Phys. Rev. D*, **102**(4), p. 044055, 2004.09442.

[224] Aasi, J. et al. (2013) "Parameter estimation for compact binary coalescence signals with the first generation gravitational-wave detector network," *Phys. Rev. D*, **88**, p. 062001, 1304.1775.

[225] Berry, C. P. L. et al. (2015) "Parameter estimation for binary neutron-star coalescences with realistic noise during the Advanced LIGO era," *Astrophys. J.*, **804**(2), p. 114, 1411.6934.

[226] Smith, R. J. E., G. Ashton, A. Vajpeyi, and C. Talbot (2020) "Massively parallel Bayesian inference for transient gravitational-wave astronomy," *Mon. Not. Roy. Astron. Soc.*, **498**(3), pp. 4492–4502, 1909.11873.

[227] Christensen, N. and R. Meyer (2022) "Parameter estimation with gravitational waves," *Rev. Mod. Phys.*, **94**(2), p. 025001, 2204.04449.

[228] Feigelson, E. D. and G. J. Babu (2012) *Modern Statistical Methods for Astronomy: With R Applications*, Cambridge University Press, Cambridge.

[229] Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013) *Bayesian Data Analysis*, CRC Press, Boca Raton, FL.

[230] Pratten, G. et al. (2021) "Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes," *Phys. Rev. D*, **103**(10), p. 104056, 2004.06503.

[231] Thompson, J. E., E. Fauchon-Jones, S. Khan, E. Nitoglia, F. Pannarale, T. Dietrich, and M. Hannam (2020) "Modeling the gravitational wave signature of neutron star black hole coalescences," *Phys. Rev. D*, **101**(12), p. 124059, 2002.08383.

[232] Matas, A. et al. (2020) "Aligned-spin neutron-star–black-hole waveform model based on the effective-one-body approach and numerical-relativity simulations," *Phys. Rev. D*, **102**(4), p. 043023, 2004.10001.

[233] Cutler, C. and E. E. Flanagan (1994) "Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?" *Phys. Rev. D*, **49**, pp. 2658–2697, gr-qc/9402014.

[234] Poisson, E. and C. M. Will (1995) "Gravitational waves from inspiraling compact binaries: Parameter estimation using second postNewtonian wave forms," *Phys. Rev. D*, **52**, pp. 848–855, gr-qc/9502040.

[235] Shao, D.-S., S.-P. Tang, J.-L. Jiang, and Y.-Z. Fan (2020) "Maximum mass cutoff in the neutron star mass distribution and the prospect of forming supramassive objects in the double neutron star mergers," *Phys. Rev. D*, **102**(6), p. 063006, 2009.04275.

[236] Lim, Y., A. Bhattacharya, J. W. Holt, and D. Pati (2021) "Radius and equation of state constraints from massive neutron stars and GW190814," *Phys. Rev. C*, **104**(3), p. L032802, 2007.06526.

[237] Kalogera, V. and G. Baym (1996) "The maximum mass of a neutron star," *Astrophys. J. Lett.*, **470**, pp. L61–L64, astro-ph/9608059.

[238] Gerosa, D. and M. Fishbach (2021) "Hierarchical mergers of stellar-mass black holes and their gravitational-wave signatures," *Nature Astron.*, **5**(8), pp. 749–760, 2105.03439.

[239] Ziegler, J. and K. Freese (2021) "Filling the black hole mass gap: Avoiding pair instability in massive stars through addition of nonnuclear energy," *Phys. Rev. D*, **104**(4), p. 043015, 2010.00254.

[240] Costa, G., A. Ballone, M. Mapelli, and A. Bressan (2022) "Formation of black holes in the pair-instability mass gap: Evolution of a post-collision star," *Mon. Not. Roy. Astron. Soc.*, **516**(1), pp. 1072–1080, 2204.03492.

[241] Abbott, R. et al. (2020) "GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses," *Phys. Rev. D*, **102**(4), p. 043015, 2004.08342.

[242] Farrow, N., X.-J. Zhu, and E. Thrane (2019) "The mass distribution of Galactic double neutron stars," *Astrophys. J.*, **876**(1), p. 18, 1902.03300.

[243] Özel, F. and P. Freire (2016) "Masses, Radii, and the Equation of State of Neutron Stars," *Ann. Rev. Astron. Astrophys.*, **54**, pp. 401–440, 1603.02698.

[244] Hannam, M., P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer (2014) "Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms," *Phys. Rev. Lett.*, **113**(15), p. 151101, 1308.3271.

[245] Schmidt, P., F. Ohme, and M. Hannam (2015) "Towards models of gravitational waveforms from generic binaries II: Modelling precession effects with a single effective precession parameter," *Phys. Rev. D*, **91**(2), p. 024043, 1408.1810.

[246] Apostolatos, T. A., C. Cutler, G. J. Sussman, and K. S. Thorne (1994) "Spin induced orbital precession and its modulation of the gravitational wave forms from merging binaries," *Phys. Rev. D*, **49**, pp. 6274–6297.

[247] Fishbach, M., D. E. Holz, and B. Farr (2017) "Are LIGO's Black Holes Made From Smaller Black Holes?" *Astrophys. J. Lett.*, **840**(2), p. L24, 1703.06869.

[248] Talbot, C. and E. Thrane (2017) "Determining the population properties of spinning black holes," *Phys. Rev. D*, **96**(2), p. 023012, 1704.08370.

[249] Rodriguez, C. L., M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio (2016) "Illuminating Black Hole Binary Formation Channels with Spins in Advanced LIGO," *Astrophys. J. Lett.*, **832**(1), p. L2, 1609.05916.

[250] Callister, T. A., W. M. Farr, and M. Renzo (2021) "State of the Field: Binary Black Hole Natal Kicks and Prospects for Isolated Field Formation after GWTC-2," *Astrophys. J.*, **920**(2), p. 157, 2011.09570.

[251] Mckernan, B. et al. (2018) "Constraining Stellar-mass Black Hole Mergers in AGN Disks Detectable with LIGO," *Astrophys. J.*, **866**(1), p. 66, 1702.07818.

[252] Yang, Y. et al. (2019) "Hierarchical Black Hole Mergers in Active Galactic Nuclei," *Phys. Rev. Lett.*, **123**(18), p. 181101, 1906.09281.

[253] Tagawa, H., Z. Haiman, I. Bartos, and B. Kocsis (2020) "Spin Evolution of Stellar-mass Black Hole Binaries in Active Galactic Nuclei," *Astrophys. J.*, **899**(1), p. 26, 2004.11914.

[254] Pürrer, M. and C.-J. Haster (2020) "Gravitational waveform accuracy requirements for future ground-based detectors," *Phys. Rev. Res.*, **2**(2), p. 023151, 1912.10055.

[255] Abbott, B. P. et al. (2017) "Effects of waveform model systematics on the interpretation of GW150914," *Class. Quant. Grav.*, **34**(10), p. 104002, 1611.07531.

[256] Dálya, G., G. Galgóczi, L. Dobos, Z. Frei, I. S. Heng, R. Macas, C. Messenger, P. Raffai, and R. S. de Souza (2018) "GLADE: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era," *Mon. Not. Roy. Astron. Soc.*, **479**(2), pp. 2374–2381, 1804.05709.

[257] Dálya, G. et al. (2022) "GLADE+ : an extended galaxy catalogue for multi-messenger searches with advanced gravitational-wave detectors," *Mon. Not. Roy. Astron. Soc.*, **514**(1), pp. 1403–1411, 2110.06184.

[258] Abbott, B. P. et al. (2019) "Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1," *Phys. Rev. D*, **100**(10), p. 104036, 1903.04467.

[259] Yagi, K. and L. C. Stein (2016) "Black Hole Based Tests of General Relativity," *Class. Quant. Grav.*, **33**(5), p. 054001, 1602.02413.

[260] Barack, L. et al. (2019) "Black holes, gravitational waves and fundamental physics: a roadmap," *Class. Quant. Grav.*, **36**(14), p. 143001, 1806.05195.

[261] Abbott, R. et al. (2022) "Search for Subsolar-Mass Binaries in the First Half of Advanced LIGO's and Advanced Virgo's Third Observing Run," *Phys. Rev. Lett.*, **129**(6), p. 061104, 2109.12197.

[262] ——— (2023) "Search for subsolar-mass black hole binaries in the second part of Advanced LIGO's and Advanced Virgo's third observing run," *Mon. Not. Roy. Astron. Soc.*, **524**(4), pp. 5984–5992, [Erratum: Mon.Not.Roy.Astron.Soc. 526, 6234 (2023)], 2212.01477.

[263] ——— (2023) "Search for gravitational-lensing signatures in the full third observing run of the LIGO-Virgo network," 2304.08393.

[264] ——— (2021) "All-sky search for long-duration gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run," *Phys. Rev. D*, **104**(10), p. 102001, 2107.13796.

[265] ——— (2020) "GW190521: A Binary Black Hole Merger with a Total Mass of $150M_{\odot}$," *Phys. Rev. Lett.*, **125**(10), p. 101102, 2009.01075.

[266] Lo, K.-L., S. Sachdev, K. Blackburn, and A. Weinstein (2016), "Searching for Gravitational Waves from the Coalescence of High Mass Black Hole Binaries," . URL https://dcc.ligo.org/LIGO-T1600261

[267] Cannon, K., C. Hanna, and D. Keppel (2013) "Method to estimate the significance of coincident gravitational-wave observations from compact binary coalescence," *Phys. Rev. D*, **88**(2), p. 024025, 1209.0718.

[268] Moe, B., P. Brady, B. Stephens, E. Katsavounidis, R. Williams, and F. Zhang (2014), "GraceDB: A Gravitational Wave Candidate Event Database," . URL https://dcc.ligo.org/LIGO-T1400365

[269] Ray, A. et al. (2023), "When to Point Your Telescopes: Gravitational Wave Trigger Classification for Real-Time Multi-Messenger Followup Observations," 2306.07190.

[270] Chaudhary, S. S. et al. (2023) "Low-latency gravitational wave alert products and their performance in anticipation of the fourth LIGO-Virgo-KAGRA observing run," 2308.04545.

[271] LIGO Scientific Collaboration and Virgo Collaboration (2020) *GCN*, **26641**.
URL https://gcn.gsfc.nasa.gov/other/S200106au.gcn3

[272] ——— (2020) *GCN*, **26665**.
URL https://gcn.gsfc.nasa.gov/other/S200108v.gcn3

[273] ——— (2020) *GCN*, **26785**.
URL https://gcn.gsfc.nasa.gov/other/S200116ah.gcn3

[274] Dietrich, T., A. Samajdar, S. Khan, N. K. Johnson-McDaniel, R. Dudi, and W. Tichy (2019) "Improving the NRTidal model for binary neutron star systems," *Phys. Rev. D*, **100**(4), p. 044003, 1905.06011.

[275] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration (2023), "LIGO/Virgo/KAGRA Public Alerts User Guide," https://emfollow.docs.ligo.org/userguide/.

[276] Fairhurst, S., R. Green, C. Hoy, M. Hannam, and A. Muir (2020) "Two-harmonic approximation for gravitational waveforms from precessing binaries," *Phys. Rev. D*, **102**(2), p. 024055, 1908.05707.

[277] Ashton, G. et al. (2019) "BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy," *Astrophys. J. Suppl.*, **241**(2), p. 27, 1811.02042.

[278] Hogg, D. W. (1999), "Distance measures in cosmology," astro-ph/9905116.

[279] Singer, L. P. et al. (2016) "Going the Distance: Mapping Host Galaxies of LIGO and Virgo Sources in Three Dimensions Using Local Cosmography and Targeted Follow-up," *Astrophys. J. Lett.*, **829**(1), p. L15, 1603.07333.

[280] Salpeter, E. E. (1955) "The Luminosity function and stellar evolution," *Astrophys. J.*, **121**, pp. 161–167.

[281] (2023), "GstLAL," https://git.ligo.org/lscsoft/gstlal.

[282] (2023), "gwcelery," https://git.ligo.org/emfollow/gwcelery.

[283] Ewing, B. (2023), "gw-lts," https://git.ligo.org/rebecca.ewing/gw-lts.

[284] Pace, A. (2023), "igwn-alert," https://igwn-alert.readthedocs.io/.

[285] Godwin, P. (2023), "ligo-scald," https://git.ligo.org/gstlal-visualisation/ligo-scald.

[286] Collaboration, L. S. (2024), "LIGO, Virgo, and KAGRA Observing Run Plans," .
URL https://observing.docs.ligo.org/plan/

[287] Sachdev, S. et al. (2020) "An Early-warning System for Electromagnetic Follow-up of Gravitational-wave Events," *Astrophys. J. Lett.*, **905**(2), p. L25, 2008.04288.

[288] Magee, R. et al. (2021) "First demonstration of early warning gravitational wave alerts," *Astrophys. J. Lett.*, **910**(2), p. L21, 2102.04555.

[289] Wei, W. and E. A. Huerta (2021) "Deep learning for gravitational wave forecasting of neutron star mergers," *Phys. Lett. B*, **816**, p. 136185, 2010.09751.

[290] Magee, R. and S. Borhanian (2022) "Observing Scenarios for the Next Decade of Early Warning Detection of Binary Neutron Stars," *Astrophys. J.*, **935**(2), p. 139, 2201.11841.

[291] Nitz, A. H., M. Schäfer, and T. Dal Canton (2020) "Gravitational-wave Merger Forecasting: Scenarios for the Early Detection and Localization of Compact-binary Mergers with Ground-based Observatories," *Astrophys. J. Lett.*, **902**, p. L29, 2009.04439.

[292] Ewing, B. (Accessed 2024), "GW Interferometer Status," https://online.igwn.org.
URL https://git.ligo.org/lscsoft/gwistat/-/wikis/home

[293] Zweizig, J. (Accessed 2024), "SenseMonitor - LIGO DMT Monitors," .
URL https://labcit.ligo.caltech.edu/~jzweizig/dmt/Monitors/SenseMonitor/index.html

[294] Capote, E. (2024), "The Performance of the Advanced LIGO Detectors during the Fourth Observing Run," .
URL https://dcc.ligo.org/DocDB/0191/G2302344/004/Capote_APS_2024.pdf

[295] Yu, H. et al. (2020) "Quantum correlations between light and the kilogram-mass mirrors of LIGO," *Nature*, **583**(7814), pp. 43–47, 2002.01519.

[296] Dal Canton, T., A. H. Nitz, B. Gadre, G. S. Cabourn Davies, V. Villa-Ortega, T. Dent, I. Harry, and L. Xiao (2021) "Real-time Search for Compact Binary Mergers in Advanced LIGO and Virgo's Third Observing Run Using PyCBC Live," *Astrophys. J.*, **923**(2), p. 254, 2008.07494.

[297] Klimenko, S., S. Mohanty, M. Rakhmanov, and G. Mitselmakher (2005) "Constraint likelihood analysis for a network of gravitational wave detectors," *Phys. Rev. D*, **72**, p. 122002, gr-qc/0508068.

[298] Lynch, R., S. Vitale, R. Essick, E. Katsavounidis, and F. Robinet (2017) "Information-theoretic approach to the gravitational-wave burst detection problem," *Phys. Rev. D*, **95**(10), p. 104046, 1511.05955.

[299] Chatterjee, D., S. Ghosh, P. R. Brady, S. J. Kapadia, A. L. Miller, S. Nissanke, and F. Pannarale (2020) "A Machine Learning Based Source Property Inference for Compact Binary Mergers," *Astrophys. J.*, **896**(1), p. 54, 1911.00116.

[300] LIGO Scientific Collaboration and Virgo Collaboration (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S231030av.lvc

[301] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230918aq.lvc

[302] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230810af.lvc

[303] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230524x.lvc

[304] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S231112ag.lvc

[305] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230712a.lvc

[306] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230708bi.lvc

[307] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230830b.lvc

[308] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230715bw.lvc

[309] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230808i.lvc

[310] ——— (2023) *GCN*.
URL https://gcn.gsfc.nasa.gov/notices_l/S230622ba.lvc

[311] Gaebel, S. M., J. Veitch, T. Dent, and W. M. Farr (2019) "Digging the population of compact binary mergers out of the noise," *Mon. Not. Roy. Astron. Soc.*, **484**(3), pp. 4008–4023, 1809.03815.

[312] LIGO Scientific Collaboration and Virgo Collaboration (2023) *GCN*, **33889**.
URL https://gcn.nasa.gov/circulars/33889.txt

[313] ——— (2023) *GCN*, **33813**.
URL https://gcn.nasa.gov/circulars/33813.txt

[314] ——— (2023) *GCN*, **34086**.
URL https://gcn.nasa.gov/circulars/34086.txt

[315] ——— (2023) *GCN*, **34429**.
URL https://gcn.nasa.gov/circulars/34429.txt

[316] ——— (2023) *GCN*, **34756**.
URL https://gcn.nasa.gov/circulars/34756.txt

[317] ——— (2023) *GCN*, **35120**.
URL https://gcn.nasa.gov/circulars/35120.txt

[318] Shibata, M. and K. Taniguchi (2011) "Coalescence of Black Hole-Neutron Star Binaries," *Living Rev. Rel.*, **14**, p. 6.

[319] Pannarale, F., A. Tonita, and L. Rezzolla (2011) "Black hole-neutron star mergers and short GRBs: a relativistic toy model to estimate the mass of the torus," *Astrophys. J.*, **727**, p. 95, 1007.4160.

[320] Krüger, C. J. and F. Foucart (2020) "Estimates for Disk and Ejecta Masses Produced in Compact Binary Mergers," *Phys. Rev. D*, **101**(10), p. 103002, 2002.07728.

[321] Cipolletta, F., C. Cherubini, S. Filippi, J. A. Rueda, and R. Ruffini (2015) "Fast Rotating Neutron Stars with Realistic Nuclear Matter Equation of State," *Phys. Rev. D*, **92**(2), p. 023007, 1506.05926.

[322] Breu, C. and L. Rezzolla (2016) "Maximum mass, moment of inertia and compactness of relativistic stars," *Mon. Not. Roy. Astron. Soc.*, **459**(1), pp. 646–656, 1601.06083.

[323] Shapiro, S. L. and S. A. Teukolsky (1983) *Black holes, white dwarfs, and neutron stars: The physics of compact objects*.

[324] Bethe, H., G. Brown, and C.-H. Lee (2003) "Formation And Evolution Of Black Holes In The Galaxy," .

[325] LIGO Scientific Collaboration and Virgo Collaboration (2023) *GCN*, **34124**.
URL https://gcn.nasa.gov/circulars/34124.txt

[326] ——— (2023) *GCN*, **34140**.
URL https://gcn.nasa.gov/circulars/34140.txt

[327] ——— (2023) *GCN*, **34147**.
URL https://gcn.nasa.gov/circulars/34147.txt

[328] ——— (2023) *GCN*, **34194**.
URL https://gcn.nasa.gov/circulars/34194.txt

[329] ——— (2023) *GCN*, **34173**.
URL https://gcn.nasa.gov/circulars/34173.txt

[330] ——— (2023) *GCN*, **34264**.
URL https://gcn.nasa.gov/circulars/34264.txt

[331] ——— (2023) *GCN*, **34293**.
URL https://gcn.nasa.gov/circulars/34293.txt

[332] ——— (2023) *GCN*, **34314**.
URL https://gcn.nasa.gov/circulars/34314.txt

[333] ——— (2023) *GCN*, **34360**.
URL https://gcn.nasa.gov/circulars/34360.txt

[334] ——— (2023) *GCN*, **34520**.
URL https://gcn.nasa.gov/circulars/34520.txt

[335] ——— (2023) *GCN*, **34546**.
URL https://gcn.nasa.gov/circulars/34546.txt

[336] ——— (2023) *GCN*, **34605**.
URL https://gcn.nasa.gov/circulars/34605.txt

[337] ——— (2023) *GCN*, **34666**.
URL https://gcn.nasa.gov/circulars/34666.txt

[338] ——— (2023) *GCN*, **34783**.
URL https://gcn.nasa.gov/circulars/34783.txt

[339] ——— (2023) *GCN*, **34801**.
URL https://gcn.nasa.gov/circulars/34801.txt

[340] ——— (2023) *GCN*, **34807**.
URL https://gcn.nasa.gov/circulars/34807.txt

[341] ——— (2023) *GCN*, **34816**.
URL https://gcn.nasa.gov/circulars/34816.txt

[342] ——— (2023) *GCN*, **34857**.
URL https://gcn.nasa.gov/circulars/34857.txt

[343] ——— (2023) *GCN*, **34904**.
URL https://gcn.nasa.gov/circulars/34904.txt

[344] ——— (2023) *GCN*, **35010**.
URL https://gcn.nasa.gov/circulars/35010.txt

[345] ——— (2023) *GCN*, **35211**.
URL https://gcn.nasa.gov/circulars/35211.txt

[346] ——— (2023) *GCN*, **35445**.
URL https://gcn.nasa.gov/circulars/35445.txt

[347] ——— (2024) *GCN*, **35480**.
URL https://gcn.nasa.gov/circulars/35480.txt

[348] ——— (2024) *GCN*, **35502**.
URL https://gcn.nasa.gov/circulars/35502.txt

[349] ——— (2023) *GCN*, **35168**.
URL https://gcn.nasa.gov/circulars/35168.txt

[350] ——— (2023) *GCN*, **34303**.
URL https://gcn.nasa.gov/circulars/34303.txt

[351] ——— (2023) *GCN*, **34411**.
URL https://gcn.nasa.gov/circulars/34411.txt

[352] Messick, C. et al. (2020) "Automating the Inclusion of Subthreshold Signal-to-Noise Ratios for Rapid Gravitational-Wave Localization," 2011.02457.

[353] Ewing, B., S. Sachdev, S. Borhanian, and B. S. Sathyaprakash (2021) "Archival searches for stellar-mass binary black holes in LISA data," *Phys. Rev. D*, **103**(2), p. 023025, 2011.03036.

[354] Abbott, B. P. et al. (2016) "Binary Black Hole Mergers in the first Advanced LIGO Observing Run," *Phys. Rev. X*, **6**(4), p. 041015, [Erratum: Phys.Rev.X 8, 039903 (2018)], 1606.04856.

[355] Sesana, A. (2016) "Prospects for Multiband Gravitational-Wave Astronomy after GW150914," *Phys. Rev. Lett.*, **116**(23), p. 231102, 1602.06951.

[356] Fryer, C. L. and K. C. B. New (2011) "Gravitational waves from gravitational collapse," *Living Rev. Rel.*, **14**, p. 1.

[357] Belczynski, K., M. Dominik, T. Bulik, R. O'Shaughnessy, C. Fryer, and D. E. Holz (2010) "The effect of metallicity on the detection prospects for gravitational waves," *Astrophys. J. Lett.*, **715**, p. L138, 1004.0386.

[358] Abbott, B. P. et al. (2017) "GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2," *Phys. Rev. Lett.*, **118**(22), p. 221101, [Erratum: Phys.Rev.Lett. 121, 129901 (2018)], 1706.01812.

[359] Chatziioannou, K. et al. (2019) "On the properties of the massive binary black hole merger GW170729," *Phys. Rev. D*, **100**(10), p. 104015, 1903.06742.

[360] Nitz, A. H., T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai (2020) "2-OGC: Open Gravitational-wave Catalog of binary mergers from analysis of public Advanced LIGO and Virgo data," *Astrophys. J.*, **891**, p. 123, 1910.05331.

[361] Venumadhav, T., B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga (2019) "New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO," *Phys. Rev. D*, **100**(2), p. 023011, 1902.10341.

[362] Moore, C. J., D. Gerosa, and A. Klein (2019) "Are stellar-mass black-hole binaries too quiet for LISA?" *Mon. Not. Roy. Astron. Soc.*, **488**(1), pp. L94–L98, 1905.11998.

[363] Dhurandhar, S. V. and B. S. Sathyaprakash (1994) "Choice of filters for the detection of gravitational waves from coalescing binaries. 2. Detection in colored noise," *Phys. Rev. D*, **49**, pp. 1707–1722.

[364] Hild, S. et al. (2011) "Sensitivity Studies for Third-Generation Gravitational Wave Observatories," *Class. Quant. Grav.*, **28**, p. 094013, 1012.0908.

[365] Abbott, B. P. et al. (2017) "Exploring the Sensitivity of Next Generation Gravitational Wave Detectors," *Class. Quant. Grav.*, **34**(4), p. 044001, 1607.08697.

[366] Wong, K. W. K., E. D. Kovetz, C. Cutler, and E. Berti (2018) "Expanding the LISA Horizon from the Ground," *Phys. Rev. Lett.*, **121**(25), p. 251102, 1808.08247.

[367] Vitale, S. (2016) "Multiband Gravitational-Wave Astronomy: Parameter Estimation and Tests of General Relativity with Space- and Ground-Based Detectors," *Phys. Rev. Lett.*, **117**(5), p. 051102, 1605.01037.

[368] Barausse, E., N. Yunes, and K. Chamberlain (2016) "Theory-Agnostic Constraints on Black-Hole Dipole Radiation with Multiband Gravitational-Wave Astrophysics," *Phys. Rev. Lett.*, **116**(24), p. 241104, 1603.04075.

[369] Cutler, C. et al. (2019) "What we can learn from multi-band observations of black hole binaries," 1903.04069.

[370] Carson, Z. and K. Yagi (2020) "Parametrized and inspiral-merger-ringdown consistency tests of gravity with multiband gravitational wave observations," *Phys. Rev. D*, **101**(4), p. 044047, 1911.05258.

[371] Gupta, A., S. Datta, S. Kastha, S. Borhanian, K. Arun, and B. Sathyaprakash (2020) "Multiparameter tests of general relativity using multiband gravitational-wave observations," 2005.09607.

[372] Datta, S., A. Gupta, S. Kastha, K. Arun, and B. Sathyaprakash (2020) "Tests of general relativity using multiband observations of intermediate mass binary black hole mergers," 2006.12137.

[373] Gerosa, D., S. Ma, K. W. K. Wong, E. Berti, R. O'Shaughnessy, Y. Chen, and K. Belczynski (2019) "Multiband gravitational-wave event rates and stellar physics," *Phys. Rev. D*, **99**(10), p. 103004, 1902.00021.

[374] Jani, K., D. Shoemaker, and C. Cutler (2019) "Detectability of Intermediate-Mass Black Holes in Multiband Gravitational Wave Astronomy," *Nature Astron.*, **4**(3), pp. 260–265, 1908.04985.

[375] Grimm, S. and J. Harms (2020) "Multiband gravitational-wave parameter estimation: A study of future detectors," *Phys. Rev. D*, **102**(2), p. 022007, 2004.01434.

[376] Ng, K. K., M. Isi, C.-J. Haster, and S. Vitale (2020) "Multiband gravitational-wave searches for ultralight bosons," *Phys. Rev. D*, **102**(8), p. 083020, 2007.12793.

[377] Liu, C., L. Shao, J. Zhao, and Y. Gao (2020) "Multiband Observation of LIGO/Virgo Binary Black Hole Mergers in the Gravitational-wave Transient Catalog GWTC-1," *Mon. Not. Roy. Astron. Soc.*, **496**(1), pp. 182–196, 2004.12096.

[378] Toubiana, A., S. Marsat, E. Barausse, S. Babak, and J. Baker (2020) "Tests of general relativity with stellar-mass black hole binaries observed by LISA," *Phys. Rev. D*, **101**(10), p. 104038, 2004.03626.

[379] LIGO Scientific Collaboration (2018), "LIGO Algorithm Library - LALSuite," free software (GPL).

[380] Owen, B. J. and B. S. Sathyaprakash (1999) "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement," *Phys. Rev. D*, **60**, p. 022002, gr-qc/9808076.

[381] Babak, S., R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash (2006) "A Template bank to search for gravitational waves from inspiralling compact binaries. I. Physical models," *Class. Quant. Grav.*, **23**, pp. 5477–5504, gr-qc/0604037.

[382] Hall, E. D. and M. Evans (2019) "Metrics for next-generation gravitational-wave detectors," *Class. Quant. Grav.*, **36**(22), p. 225002, 1902.09485.

[383] Vitale, S. and M. Evans (2017) "Parameter estimation for binary black holes with networks of third generation gravitational-wave detectors," *Phys. Rev. D*, **95**(6), p. 064052, 1610.06917.

[384] Abbott, B. P. et al. (2019) "Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo," *Astrophys. J. Lett.*, **882**(2), p. L24, 1811.12940.

[385] Robson, T., N. J. Cornish, and C. Liu (2019) "The construction and use of LISA sensitivity curves," *Class. Quant. Grav.*, **36**(10), p. 105011, 1803.01944.

[386] Borhanian, S. (2021) "GWBENCH: a novel Fisher information package for gravitational-wave benchmarking," *Class. Quant. Grav.*, **38**(17), p. 175014, 2010.15202.

[387] Racine, E. (2008) "Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction," *Phys. Rev. D*, **78**, p. 044021, 0803.1820.

[388] Ajith, P. et al. (2011) "Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins," *Phys. Rev. Lett.*, **106**, p. 241101, 0909.2867.

[389] Abbott, R. et al. (2021) "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog," *Astrophys. J. Lett.*, **913**(1), p. L7, 2010.14533.

[390] Balasubramanian, R., B. S. Sathyaprakash, and S. V. Dhurandhar (1996) "Gravitational waves from coalescing binaries: Detection strategies and Monte Carlo estimation of parameters," *Phys. Rev. D*, **53**, pp. 3033–3055, [Erratum: Phys.Rev.D 54, 1860 (1996)], gr-qc/9508011.

[391] London, L., S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale (2018) "First higher-multipole model of gravitational waves from spinning and coalescing black-hole binaries," *Phys. Rev. Lett.*, **120**(16), p. 161102, 1708.00404.

[392] Toubiana, A., S. Marsat, S. Babak, J. Baker, and T. Dal Canton (2020) "Parameter estimation of stellar-mass black hole binaries with LISA," *Phys. Rev. D*, **102**, p. 124037, 2007.08544.

[393] Blanchet, L., T. Damour, B. R. Iyer, C. M. Will, and A. G. Wiseman (1995) "Gravitational radiation damping of compact binary systems to second postNewtonian order," *Phys. Rev. Lett.*, **74**, pp. 3515–3518, gr-qc/9501027.

[394] Sathyaprakash, B. S. and S. V. Dhurandhar (1991) "Choice of filters for the detection of gravitational waves from coalescing binaries," *Phys. Rev. D*, **44**, pp. 3819–3834.

[395] Schutz, B. F. (ed.) (1989) *Gravitational wave data analysis. Proceedings: NATO Advanced Research Workshop, Cardiff, UK, Jul 6-9, 1987*, vol. 253.

[396] Harry, I. W., B. Allen, and B. S. Sathyaprakash (2009) "A Stochastic template placement algorithm for gravitational wave data analysis," *Phys. Rev. D*, **80**, p. 104014, 0908.2090.

[397] Ajith, P., N. Fotopoulos, S. Privitera, A. Neunzert, and A. J. Weinstein (2014) "Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins," *Phys. Rev. D*, **89**(8), p. 084041, 1210.6666.

[398] Capano, C., I. Harry, S. Privitera, and A. Buonanno (2016) "Implementing a search for gravitational waves from binary black holes with nonprecessing spin," *Phys. Rev. D*, **93**(12), p. 124007, 1602.03509.

[399] Poisson, E. (1998) "Gravitational waves from inspiraling compact binaries: The Quadrupole moment term," *Phys. Rev. D*, **57**, pp. 5287–5290, gr-qc/9709032.

[400] Mikoczi, B., M. Vasuth, and L. A. Gergely (2005) "Self-interaction spin effects in inspiralling compact binaries," *Phys. Rev. D*, **71**, p. 124043, astro-ph/0504538.

[401] Bohé, A., S. Marsat, and L. Blanchet (2013) "Next-to-next-to-leading order spin–orbit effects in the gravitational wave flux and orbital phasing of compact binaries," *Class. Quant. Grav.*, **30**, p. 135009, 1303.7412.

[402] Bohé, A., G. Faye, S. Marsat, and E. K. Porter (2015) "Quadratic-in-spin effects in the orbital dynamics and gravitational-wave energy flux of compact binaries at the 3PN order," *Class. Quant. Grav.*, **32**(19), p. 195010, 1501.01529.

[403] Arun, K. G., A. Buonanno, G. Faye, and E. Ochsner (2009) "Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiraling compact binaries: Ready-to-use gravitational waveforms," *Phys. Rev. D*, **79**, p. 104023, [Erratum: Phys.Rev.D 84, 049901 (2011)], 0810.5336.

[404] Mishra, C. K., A. Kela, K. G. Arun, and G. Faye (2016) "Ready-to-use post-Newtonian gravitational waveforms for binary black holes with nonprecessing spins: An update," *Phys. Rev. D*, **93**(8), p. 084054, 1601.05588.

[405] Toubiana, A. et al. (2021) "Detectable environmental effects in GW190521-like black-hole binaries with LISA," *Phys. Rev. Lett.*, **126**(10), p. 101105, 2010.06056.

[406] Caputo, A., L. Sberna, A. Toubiana, S. Babak, E. Barausse, S. Marsat, and P. Pani (2020) "Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart," *Astrophys. J.*, **892**(2), p. 90, 2001.03620.

[407] Goldstein, A. et al. (2017) "An Ordinary Short Gamma-Ray Burst with Extraordinary Implications: Fermi-GBM Detection of GRB 170817A," *Astrophys. J. Lett.*, **848**(2), p. L14, 1710.05446.

[408] Abbott, R. et al. (2022) "Search for Gravitational Waves Associated with Gamma-Ray Bursts Detected by Fermi and Swift during the LIGO–Virgo Run O3b," *Astrophys. J.*, **928**(2), p. 186, 2111.03608.

Vita

Becca Ewing

Education

The Pennsylvania State University *University Park, PA*
Ph.D. in Physics 2024

Wright State University *Dayton, OH*
B.S. in Physics and Mathematics 2018

Selected Publications

Ewing, Becca, et al. (2024). "Performance of the low-latency GstLAL inspiral search towards LIGO, Virgo, and KAGRA's fourth observing run".

Abbott, R., et al. (2023). "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run."

Ewing, B., Sachdev, S., Borhanian, S., & Sathyaprakash, B. S. (2021). "Archival searches for stellar- mass binary black holes in LISA data."

Abbott, R., et al. (2021). "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run."

Abbott, R., et al. (2024). "GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run."

Selected Awards and Fellowships

Peter Eklund Award for Scientific Communication 2024

David C. Duncan Graduate Fellowship in Physics 2024

Science Achievement Graduate Fellowship 2018 - 2023

The Nellie H. and Oscar L. Roberts Scholarship Fund 2018 - 2021