Signatures and Dynamics of Compact Binary Coalescences and a Search in LIGO's S5 Data

Thesis by

Drew Garvin Keppel

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended April 20, 2009)

To my everloving wife, Lisa

Acknowledgements

Without the help of a multitude of people, all of the work presented here would not have been possible. I would like to thank my advisor, Alan Weinstein, for taking me under his wing as a student and sharing his passion for LIGO with me. He has given me guidance as a developing scientist, as well as the freedom to explore the paths I choose.

Additionally, I would like to thank my committee members: Yanbei Chen, Jay Marx, Tom Prince, and Kip Thorne, for all of their time and advice with regard to my thesis.

I would like to thank the review committee members for the Search for Low Mass CBCs in the First Year of LIGO's S5 Data: Marie-Anne Bizouard, Laura Cadonati, Eric Gustafson, Vicky Kalogera, Greg Mendell, Peter Saulson, and John Whelan, who have been an invaluable resource for validating this search.

I would like to thank the many postdoctoral scholars and research scientists who have walked the halls of the 6th floor of Millikan during my time at Caltech. Your examples have driven me to become a better scientist. I would especially like to thank Duncan Brown, Stephen Fairhurst, and Eirini Messaritaki for answering my endless questions about CBC searches and helping to guide me through the process; Kipp Cannon, Chad Hanna, and Antony Searle for sharing their visions of the future of CBC searches with me; and Parameswaran Ajith for his insight in IMR waveforms and help with parameter estimation.

I would also like to thank my fellow graduate students at Caltech: Diego Fazi, Lisa Goggin, Kari Hodge, Pinkesh Patel, and Becky Tucker; and throughout the LSC: Rahul Biswas, Collin Capano, Nick Fotopolus, Ian Harry, Gareth Jones, David McKechan, Craig Robinson, Andy Rodriguez, Lucia Santamaria, and Jake Slutsky; as well as other members of the CBC group: Sukanta Bose, Patrick

Brady, Jessica Clayton, Thomas Cokelaer, Jolien Creighton, Alex Dietz, Gaby Gonzalez, Romain Gouaty, B. S. Sathyaprakash, Anand Sengupta, and Ruslan Vaulin; and collaborators in the LSC-Virgo Collaboration (and anyone else I have failed to mention) for the friendships we have built and the many conversations we have had throughout the years while working, traveling, and playing together.

Pertaining to my work in relativity, I would like to thank Kip Thorne for leading us on a fascinating exploration that started from a homework problem; Yanbei Chen for his endless insights; and David Nichols, Jaff Kaplan, and the rest of the TAPIR group for the many discussions we have had.

I would also like to thank all of the administrative assistants: Irene Baldon, JoAnn Boyd, and Marylin Wright; and LIGO computer staff: Phil Ehrens, Dan Kozak, and Stuart Anderson, with whom I have interacted, for all of their aid and support, which allowed me to focus more strongly on my research.

This work would not be possible without the LIGO Laboratory and the LIGO Scientific Collaboration. I gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory and the Science and Technology Facilities Council of the United Kingdom, the Max-Planck-Society, and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector. As well, I gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the Council of Scientific and Industrial Research of India, the Istituto Nazionale di Fisica Nucleare of Italy, the Spanish Ministerio de Educación y Ciencia, the Conselleria d'Economia, Hisenda i Innovació of the Govern de les Illes Balears, the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, The National Aeronautics and Space Administration, the Carnegie Trust, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation. Finally, this research was supported in part by NSF grants PHY-0601459 and PHY-0653653, and by the David and Barbara Groce start-up fund at Caltech.

Last, but certainly not least, I would like to thank my parents, Lowell and B, and siblings, Alex and Cecily, for their love and support over the years and for encouraging me to pursue my dreams; my wife, Lisa, for her unending love and encouragement, years of companionship, patience with my work, and support on so many levels; and Maxwell's Equation for his perpetual energy and enthusiasm with which he tackles life. I also thank God for all of the opportunities and abilities He has given me to explore His creation.

Abstract

In this thesis, we probe several aspects of compact binary environments, focusing on results of orbits and collisions of compact objects. First, we describe a search for low-mass compact-binary-coalescence gravitational-wave signals in data from the LIGO detectors' most sensitive, longest-running science run to date (S5). We also go into detail on the interpretation of the results including its development. We then investigate the bounds on the mass of the graviton that could be achieved from the detection gravitational waves from a binary black hole merger. Last, we study the flow of momentum in compact binaries using the Landau-Lifshitz formalism.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

General relativity is a very rich theory that makes many exciting predictions about the universe's most extreme environment. One of those predictions is the emission of gravitational waves (GWs) from accelerating objects (e.g., orbiting compact binaries), which have yet to be detected directly. To that end, the Laser Interferomenter Graviational-wave Observatory (LIGO) detectors have been built and measure these waves as they pass through the earth. Recently, the LIGO detectors reached their design sensitivity and recorded data from their fifth science run, which accumulated a year's worth of coincident data where multiple detectors were in operation. The first part of this thesis will focus on a search for GW signals from compact binary coalescence (CBC) (LIGO's flagship search) in the first year of data from this science run. In those chapters we describe the signals we search for, the design of the GW detectors, how the search was performed, and its results. For this part of the thesis, my role was to tune and run the analysis mentioned above. This included contributing to the code base used for the analysis, implementing new features in many stages of the pipeline, and designing and implementing the postprocessing pipeline to calculate the detection statistic. Finally, this included managing the writing of and acting as the corresponding author on the resulting paper.

In the second part of this thesis, we focus on testing the effects of a massive graviton. That chapter describes a massive graviton's effects on the propagation of GWs and uses those effects to constrain the graviton's mass for a given CBC GW signal observation. The main result of my work is an application of previous methods to a new class of CBC waveforms that contain more information (i.e., complete CBC waveforms) and can more stringently constrain the graviton's mass.

The final part of this thesis focuses on the dynamics of compact objects as they interact and subsequently merge. Momentum flow between the bodies and the surrounding gravitational field is studied and found to explain interesting features seen in Numerical Relativity simulations of orbiting black holes that have nonzero spin angular momentum. For this part of the thesis, my role was to perform and verify the surface and volume integral calculations associated the momentum conservation results.

Chapter 2

Gravitational Waves and Their Sources, Including Compact Binary Coalescences

In this chapter we give a brief introduction to General Relativity, focusing on GW emission. We then focus our attention on GWs from CBCs. The following presentation uses the conventions found in [1].

2.1 Pillars of Relativity

Einstein developed the pillars of Relativity Theory in the early 20th Century, which can be summarized by two thought experiments, or “gedankenexperiments.” These gedankenexperiments reveal different implications of a single assumption, that the Laws of Physics are the same for everyone observing an experiment.

The first application of this assumption was to the speed of light. Einstein assumed that the speed of light was a constant, independent of an observer’s frame of reference. The conclusions Einstein drew from this led him to develop the Special Theory of Relativity, which is useful in Lorentz reference frames (i.e., those without gravity).

The second application of this assumption was the universality of the motion of objects in gravitational freefall. Einstein’ gedankenexperiment went as follows: suppose there are two observers observing the motion on a object in freefall in their local reference frame. Over small distances and times, they will observe the same motion of the object whether the motion is due to gravitational acceleration towards a gravitating body (e.g., for an observer whose frame is resting on the Earth), or due to a uniform acceleration of the reference frame (e.g., for an observer whose frame is in a rocket ship that is constantly accelerating in a specific direction). From this gedankenexperiment, Einstein concluded that the acceleration around a gravitating body is produced by a curvature of spacetime in the object's vicinity. Using this framework, Einstein developed the General Theory of Relativity.

2.2 Einstein's Equations

In the General Theory of Relativity, the equations that link the curvature of spacetime to the presence of matter and energy are called the Einstein Field Equation (EFE) given as

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ , \tag{2.1}$$

where $G_{\mu\nu}$ is the Einstein Tensor given by $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$, $R = g^{\mu\nu} R_{\mu\nu}$, $R_{\mu\nu} = R^{\gamma}{}_{\mu\nu\gamma}$, $R^{\gamma}{}_{\mu\nu\gamma}$ is the Riemann Curvature Tensor, $g_{\mu\nu}$ is the spacetime metric, and $T_{\mu\nu}$ is the stress-energy tensor. These equations tell matter how to move due to the curvature of spacetime and spacetime how to curve due to the presence of matter.

2.3 Gravitational Wave Solutions

The EFEs not only describe how spacetime curvature is influenced in the presence of matter, it also predicts how the curvature should propagate away from accelerating objects. Just as electromagnetic waves are emitted and travel away from accelerating electric charges, GWs are emitted and travel away from accelerating masses. If we are far away from the accelerating source, the EFEs simplify to

$$G_{\mu\nu} = 0 \ . \tag{2.2}$$

In this region, the spacetime is essentially Minkowski except for the perturbations do to GWs: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric. Assuming these perturbations are small so that all components of $h_{\mu\nu} \ll 1$, we can calculate how these perturbations propagate using a linearized version of the EFEs. In the following discussion, we follow closely the description found in reference [1].

2.3.1 Linearized Gravity

To first order in $h_{\mu\nu}$ the EFE become

$$-\bar{h}_{\mu\nu,\alpha}{}^{\alpha} - \eta_{\mu\nu}\bar{h}_{\alpha\beta,}{}^{\alpha\beta} + \bar{h}_{\mu\alpha,}{}^{\alpha}{}_{\nu} + \bar{h}_{\nu\alpha,}{}^{\alpha}{}_{\mu} = 0 \ , \tag{2.3}$$

where $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$ and $h \equiv \eta^{\alpha\beta}h_{\alpha\beta}$. We can eliminate all of the terms after the first by imposing the tensor analogue of the Lorentz gauge condition of electromagnetic theory, namely $\bar{h}_{\mu\alpha,}{}^{\alpha} = 0$. This leaves us with the wave equation

$$\bar{h}_{\mu\nu,\alpha}{}^{\alpha} = 0 \ . \tag{2.4}$$

This wave equation has plane wave solutions

$$\bar{h}_{\mu\nu} = \Re\left[A_{\mu\nu}e^{ik_\alpha x^\alpha}\right] \ , \tag{2.5}$$

where $\Re[\ldots]$ means the real part of $[\ldots]$. From the wave equation we find $k_\alpha k^\alpha = 0$, which means $\mathbf{k}$ is a null vector and the waves travel at the speed of light. From the gauge condition $\bar{h}_{\mu\alpha,}{}^{\alpha} = 0$ we find $A_{\mu\alpha}k^\alpha = 0$, which means the amplitude tensor $\mathbf{A}$ is orthogonal to $\mathbf{k}$.

At this point $\mathbf{A}$ has six degrees of freedom, however four of these can be pinned down with an additional gauge specification. Let us choose $A_{\mu\nu}u^\nu = 0$ where u^ν is a 4-velocity of a particular frame. This fixes an additional three degrees of freedom. For the final gauge choice, let us choose

$A^\mu{}_\mu = 0$. All of these constraints in a Lorentz frame with $u^0 = 1$ and $u^j = 0$ give

$$h_{\mu 0} = 0 \; , \tag{2.6}$$

$$h_{kj,}{}^{j} = 0 \; , \tag{2.7}$$

$$h_{kk} = 0 \; . \tag{2.8}$$

The first of these equations implies only the spatial components of $h_{\mu\nu}$ are nonzero. The second implies that the spatial components are divergence free. Finally, the third implies that $h_{\mu\nu}$ is trace free. This gauge is called the transverse-traceless (TT) gauge because in this gauge $h_{\mu\nu}$ is transverse to the time direction and the direction of propagation, and is trace free.

These gauge choices leave us with two degrees of freedom for the propagating waves. This can be interpreted as two polarizations of GWs, the plus (+) and cross (×) polarizations, named for their affect on test particles. In terms of the these polarizations and in the TT gauge, $A_{\mu\nu}$ takes the form

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.9}$$

2.3.2 Geodesic Deviation

Objects undergoing free motion will follow geodesics described by the geodesic equation

$$\frac{d^2x^\lambda}{dt^2} + \Lambda^\lambda{}_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt} = 0 \; . \tag{2.10}$$

Consider two particles, A and B, separated by n^j in a coordinate system that is a local Lorentz frame along A's world line. In this coordinate system the separation vector becomes $n^j = x_B{}^j - x_A{}^j = x_B{}^j$, since A is at the origin. In a TT coordinate system that moves with particle A and with its proper reference frame to first order in the metric perturbation h^{TT}_{jk}, the geodesic deviation equation is written as

$$\frac{d^2 n^j}{\partial t^2} = -R^{\mathrm{TT}}_{j0k0} n^k \ . \tag{2.11}$$

In the TT gauge the Riemann curvature tensor takes the simple form

$$R_{j0k0} = -\frac{1}{2} h^{\mathrm{TT}}_{jk,00} \ . \tag{2.12}$$

Using this along with the separation vector's definition, the geodesic deviation equation becomes

$$\frac{d^2 x_B{}^j}{\partial t^2} = \frac{1}{2} \frac{\partial^2 h^{\mathrm{TT}}_{jk}}{\partial t^2} x_B{}^k \ . \tag{2.13}$$

This equation of motion can be integrated, assuming particles initially at rest relative to each other, to obtain the location of B as a function of time in the proper reference frame of A

$$x_{Bk}(t) = x_B{}^j(0) \left(\delta_{jk} + \frac{1}{2} h^{\mathrm{TT}}_{jk} \right) \ , \tag{2.14}$$

where h^{TT}_{jk} should be evaluated at the location of A. The second in the parentheses on the right-hand side can be viewed as a strain and is proportional to the metric perturbation h_{jk}.

2.3.3 Gravitational-Wave Polarizations

As mentioned above, there are two degrees of freedom for a GW, which can be seen as distinct polarizations. Let us consider a plane, monochromatic wave propagating in the z direction. In the TT gauge of this wave, the gauge constraints (equation (2.6)) leave the only nonzero components of $h^{\mathrm{TT}}_{\mu\nu}$

$$h^{\mathrm{TT}}_{xx} = -h^{\mathrm{TT}}_{yy} = \Re \left[A_+ e^{-i\omega(t-z)} \right] \ , \tag{2.15}$$

$$h^{\mathrm{TT}}_{xy} = h^{\mathrm{TT}}_{yx} = \Re \left[A_\times e^{-i\omega(t-z)} \right] \ , \tag{2.16}$$

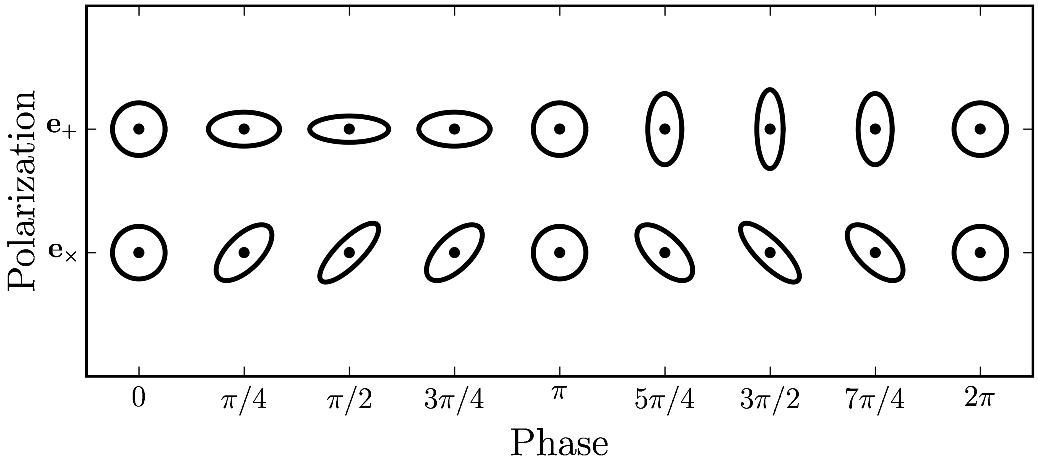

Figure 2.1: Effects of Passing Gravitational Waves
The effect of the + and × polarization GW passing perpendicular to a ring of test particles.

where A_+ and $A_\times$ are the amplitudes of the + and × polarizations, respectively, and ω is the frequency of the GW.

Defining the polarization tensors $\mathbf{e}_+ = (\mathbf{e}_x \oplus \mathbf{e}_x) - (\mathbf{e}_y \oplus \mathbf{e}_y)$ and $\mathbf{e}_\times = (\mathbf{e}_x \oplus \mathbf{e}_y) + (\mathbf{e}_y \oplus \mathbf{e}_x)$ we find a GW can be decomposed into its independent polarizations as

$$h_{jk}^{\mathrm{TT}} = h_+ e_{+jk} + h_\times e_{\times jk} \ . \tag{2.17}$$

The effects of each polarization can be seen independently by looking at a circular ring of test particles surrounding a central particle: let all of the particles lie in the plane perpendicular to the direction of propagation of the incident GW. For an incident +-polarized wave, the polarization tensor $\mathbf{e}_+$ shows that the test particles will oscillate about their circular equilibrium positions in such a way that as the distance from the origin along the x direction increases, the distance from the origin along the y direction decreases. This results in the circle being stretched and squeezed along the x and y axes at one moment, and vice versa half a wavelength later. For a ×-polarized wave, a similar thing happens, except rotated by 45°. This effect is visualized in figure 2.1.

2.4 Gravitational Wave Sources

In the case of electromagnetic radiation, a multipolar analysis reveals that electromagnetic waves are sourced by, in order of significance, accelerating electric-charge monopoles, electric-charge dipoles and electric-current dipoles, electric-charge quadrupoles and electric-current quadrupoles, etc. In the case of gravitational radiation, the conservation of energy prevents radiation due to the acceleration of mass monopoles, the conservation of linear momentum prevents radiation due to the acceleration of mass dipoles, and the conservation of angular momentum prevents radiation due to the acceleration of current dipoles. This results in the leading-order radiation coming from the acceleration of mass quadrupoles.

In reference [2], Thorne gives the relationships between GW radiation and its source. Radiation can be related to the source through the multipolar expansion. For Newtonian sources the relevant multipolar moments are given by

$$\mathcal{I}_{A_l} = \left[\int \rho X_{A_l} d^3x\right]^{\mathrm{STF}} , \tag{2.18a}$$

$$\mathcal{S}_{A_l} = \left[\int \left(\epsilon_{a_l pq} x_p \rho v_q\right) X_{A_{l-1}} d^3x\right]^{\mathrm{STF}} , \tag{2.18b}$$

where $X_{A_l} = x_{a_1} x_{a_2} \ldots x_{a_l}$, ρ is the Newtonian mass density, v is the velocity, ϵ_{ijk} is the Levi-Civita tensor, and STF stands for the symmetric trace-free part.

For these sources, the radiation h^{TT}_{jk} is given by

$$\begin{aligned} h^{\mathrm{TT}}_{jk} &= \Bigg[\sum_{l=2}^{\infty} \frac{4}{l!} r^{-1\,(l)}\mathcal{I}^{\mathrm{rad}}_{jkA_{l-2}}(t-r) N_{A_{l-2}} \\ &\quad + \sum_{l=2}^{\infty} \frac{8}{(l+1)!} r^{-1} \epsilon_{pq(j}{}^{(l)}\mathcal{S}^{\mathrm{rad}}_{k)pA_{l-2}}(t-r) n_q N_{A_{l-2}}\Bigg]^{\mathrm{TT}} , \end{aligned} \tag{2.19}$$

where ${}^{(l)}\mathcal{I}$ and ${}^{(l)}\mathcal{S}$ represent the lth time derivative of these moments, and $N_{A_l} = n_1 n_2 \ldots n_l$. The superscript "rad" denotes this as the moment in the radiation field, which is the same as the moment in the near zone if we ignore nonlinear corrections proportional to $\ln(\lambda/2\pi R)$ where R is the background curvature of the local universe.

The energy and momentum carried away from a radiating system by GWs is given simply in the form of the stress-energy tensor for GWs:

$$T^{\mathrm{GW}}_{\mu\nu} = \frac{1}{32\pi} \left\langle h^{\mathrm{TT}}_{jk,\mu} h^{\mathrm{TT}}_{jk,\nu} \right\rangle \ . \tag{2.20}$$

Expanding the energy term (i.e., $\mu = \nu = 0$) in terms of the multipole moments, we find the energy flux carried by the waves to be

$$\begin{aligned} \frac{dE}{dt} &= \sum_{l=2}^{\infty} \frac{(l+1)(l+2)}{(l-1)l} \frac{1}{l!(2l+1)!!} \left\langle {}^{l+1}\mathcal{I}_{A_l} {}^{l+1}\mathcal{I}_{A_l} \right\rangle \\ &\quad + \sum_{l=2}^{\infty} \frac{4l(l+2)}{(l-1)} \frac{1}{(l+1)!(2l+1)!!} \left\langle {}^{l+1}\mathcal{S}_{A_l} {}^{l+1}\mathcal{S}_{A_l} \right\rangle \ , \end{aligned} \tag{2.21}$$

where $l! = l(l-1)\cdots 2\cdot 1$ and $l!! = l(l-2)(l-4)\cdots(2\,\mathrm{or}\,1)$.

GWs can come any source with an accelerating mass or current multipole, however, in most situations the dominant term in the multipole expansion is the $l = 2$ mass quadrupole. Different types of sources can be characterized into the following categories: known-transient waveforms (e.g., compact binary signals), unknown-transient waveforms (i.e., burst signals), discernible and narrow-band continuous wave sources (e.g., pulsar and white-dwarf binary signals), and indiscernible or broad-band continuous wave sources (e.g., stochastic and confused white-dwarf binary sources). In section 2.4.1 we derive the waveforms associated with compact binary signals. In section 2.4.2 we briefly discuss additional examples and how each type of waveform is searched for.

2.4.1 Compact Binaries

One of the families of known transient waveforms is associated with compact binary systems (i.e., binary systems whose objects are either neutron stars or black holes). As these objects orbit each other, they act as a source for GWs.

Compact binaries are interesting signals, in part, because they are so well understood. These signals can act as standard candles for measuring astrophysical distances due to the fact that their amplitude is uniquely determined by their phase evolution and the luminosity distance to the source. In addition, GW signals from compact binaries are considered some of the most extreme probes of General Relativity as the signals originate from objects that massively curve space-time and travel at speeds approaching the speed of light.

2.4.1.1 Compact Binary Inspiral Waveforms

In this section we go over the leading-order effects in calculating a compact binary inspiral waveform. To aid in this calculation, we assume the objects are Newtonian point particles with masses m_A and m_B. In this case, the mass density term in equation (2.18a) become Dirac delta functions and the quadrupole moment $\mathcal{I}_{jk}$ is given as

$$\begin{aligned} \mathcal{I}_{jk} &= \left[\sum_A m_A x_{Aj} x_{Ak}\right]^{\mathrm{STF}} \\ &= \sum_A m_A \left(x_{Aj} x_{Ak} - \frac{1}{3}\delta_{jk} r_A^2 \right) \ . \end{aligned} \tag{2.22}$$

The leading-order term in the GW radiation (equation (2.19)) is due to the mass quadrupole and given as

$$h_{jk}^{\mathrm{TT}} = \frac{2}{D}\ddot{\mathcal{I}}_{jk}^{\mathrm{TT}}(t-D) \ , \tag{2.23}$$

where D is the distance from the source.

Let us assume these objects are in quasicircular orbits around each other described by Kepler's Laws where their orbital speed is related to their separation by

$$\Omega = \sqrt{\frac{M}{r^3}} \ , \tag{2.24}$$

where $M = m_A + m_B$ is the total mass of the system, and r is the separation.

As these objects orbit each other, in a coordinate system whose origin is at the center of mass with the plane of the orbit in the x-y plane, their locations are given by

$$x_A(t) = \frac{\mu}{m_A} r\cos(\Omega t) \; , \tag{2.25a}$$

$$y_A(t) = \frac{\mu}{m_A} r\sin(\Omega t) \; , \tag{2.25b}$$

$$x_B(t) = -\frac{\mu}{m_B} r\cos(\Omega t) \; , \tag{2.25c}$$

$$y_B(t) = -\frac{\mu}{m_B} r\sin(\Omega t) \; , \tag{2.25d}$$

where $\mu = m_A m_B/M$ is the reduced mass of the system.

Using these trajectories and equation (2.22) for $\mathcal{I}_{jk}$, we find quadrupole moments for this system as

$$\mathcal{I}_{xx} = r^2\mu\left(\cos^2(\Omega t) - \frac{1}{3}\right) \; , \tag{2.26a}$$

$$\mathcal{I}_{yy} = r^2\mu\left(\sin^2(\Omega t) - \frac{1}{3}\right) \; , \tag{2.26b}$$

$$\mathcal{I}_{jz} = r^2\mu\left(-\delta_{jz}\frac{1}{3}\right) \; , \tag{2.26c}$$

$$\mathcal{I}_{xy} = \mathcal{I}_{xy} = r^2\mu\sin(\Omega t)\cos(\Omega t) \; . \tag{2.26d}$$

As GWs leave the system, they carry away energy, which has a back-reaction effect on the orbiting binary. The energy carried away by quadrupole radiation is given from equation (2.21) as

$$\frac{dE}{dt}_{\text{mass quadrupole}} = \frac{1}{5}\left\langle \dddot{\mathcal{I}}_{jk}\dddot{\mathcal{I}}_{jk}\right\rangle \; , \tag{2.27}$$

which, substituting these moments in, we find to be

$$\begin{aligned}
\frac{dE}{dt}_{\text{mass quadrupole}} &= \frac{32r^4\mu^2\Omega^6}{5}\left\langle 4\cos^2(\Omega t)\sin^2(\Omega t) + \left(2\sin^2(\Omega t) - 1\right)^2\right\rangle \\
&= \frac{32r^4\mu^2\Omega^6}{5} \; .
\end{aligned} \tag{2.28}$$

Using the Keplerian relation equation (2.24), we find this to be

$$\frac{dE}{dt}_{\text{mass quadrupole}} = \frac{32\mu^2 M^3}{5r^5} \ . \tag{2.29}$$

The energy contained in the system as a function of their separation is given by the Newtonian formula

$$E = -\frac{\mu M}{2r} \ . \tag{2.30}$$

As energy leaves, the orbiting objects sink further into their respective gravitational potentials, causing their orbital separation to decrease. Assuming this energy loss occurs adiabatically, we can find the orbital separation as a function of time. First, we find the derivative of the separation with respect to time as

$$\begin{aligned} \frac{dr}{dt} &= \left(\frac{dE}{dt}\right) / \left(\frac{dE}{dr}\right) \\ &= -\frac{64\mu M^2}{5r^3} \ . \end{aligned} \tag{2.31}$$

Integrating this, we find the separation as a function of time to be

$$r(t) = \left(\frac{256\mu M^2}{5}\right)^{\frac{1}{4}} (t_c - t)^{\frac{1}{4}} \ , \tag{2.32}$$

where t_c denotes the time when $r = 0$.

We can use this result along with equation (2.24) to find the orbital frequency $\Omega(t)$ as a function of time

$$\Omega(t) = \left(\frac{256\mathcal{M}^{\frac{5}{3}}}{5}\right)^{\frac{-3}{8}} (t_c - t)^{\frac{-3}{8}} \ , \tag{2.33}$$

where $\mathcal{M} = \eta^{3/5} M$ is the chirp mass and $\eta = \mu/M$ is the symmetric mass ratio.

We can now use equation (2.23) to find the quadrupole radiation as

$$\begin{aligned} h_c(t) \equiv -h_{xx}^{\mathrm{TT}} = h_{yy}^{\mathrm{TT}} &= \frac{4r^2(t)\mu\Omega^2(t)}{D}\left(\cos^2(\Omega(t)t) - \sin^2(\Omega(t)t)\right) \\ &= \frac{4r^2(t)\mu\Omega^2(t)}{D}\cos(2\Omega(t)t) \; , \end{aligned} \tag{2.34a}$$

$$\begin{aligned} h_s(t) \equiv -h_{xy}^{\mathrm{TT}} = -h_{yx}^{\mathrm{TT}} &= \frac{8r^2(t)\mu\Omega^2(t)}{D}\sin(\Omega(t)t)\cos(\Omega(t)t) \\ &= \frac{4r^2(t)\mu\Omega^2(t)}{D}\sin(2\Omega(t)t) \; . \end{aligned} \tag{2.34b}$$

We can split the radiation into separate parts for the amplitude evolution $A(t)$ and the phase evolution $\Phi(t)$. The radiation then takes the form

$$h_c(t) = A(t)\cos(2\Phi(t)) \; , \tag{2.35a}$$

$$h_s(t) = A(t)\sin(2\Phi(t)) \; . \tag{2.35b}$$

Using the formulae for the separation and orbital frequency evolutions, we find $A(t)$ and $\Phi(t)$ to be

$$\begin{aligned} A(t) &= \frac{4r^2(t)\mu\Omega^2(t)}{D} \\ &= \frac{4\mu M}{Dr(t)} \\ &= \frac{5^{\frac{1}{4}}}{D}\mathcal{M}^{\frac{5}{4}}\left(t_c - t\right)^{\frac{-1}{4}} \; , \end{aligned} \tag{2.36a}$$

$$\begin{aligned} \Phi(t) &= \int \Omega(t)dt \\ &= \phi_0 - \left(\frac{t_c - t}{5\mathcal{M}}\right)^{\frac{5}{8}} \; , \end{aligned} \tag{2.36b}$$

where ϕ_0 is the orbital phase at coalescence.

The signal described by equations (2.36a and 2.36b) is characterized by a sinusoidal oscillation

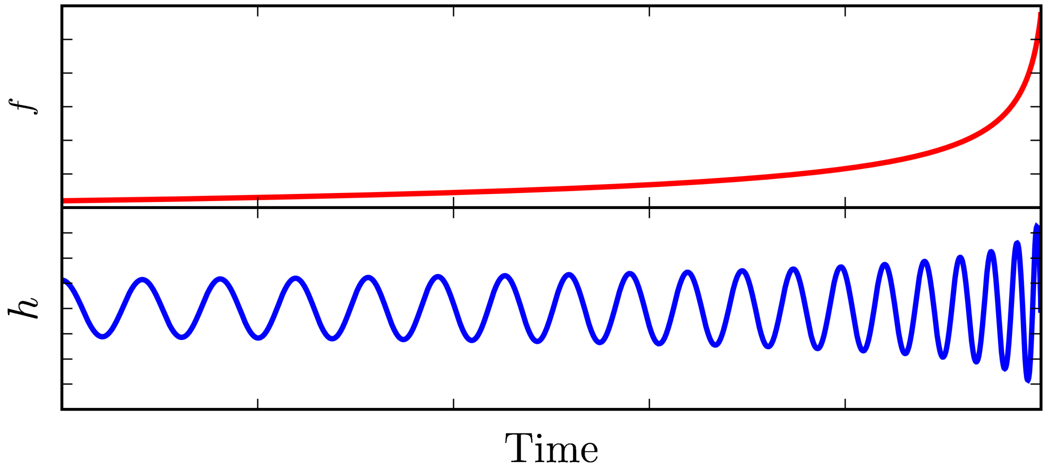

Figure 2.2: Inspiral Waveform
The frequency f and strain h evolution of an inspiral waveform given in arbitrary units.

that increase in both amplitude and frequency as a function of time (i.e., a chirp signal). This is illustrated in figure 2.2, which shows both the frequency evolution and the strain evolution over a few cycles near the end of an inspiral waveform.

This continues as long as there is a local minimum in the effective potential of the system. For the Schwarzschild potential, which is the unique potential outside a spherically symmetric mass distribution, there exists a local minimum for a test particle in a circular orbit around the object down to a separation of $6M$, called the innermost stable circular orbit (ISCO). At this point the above derivation breaks down and we terminate the calculation of the waveform. This occurs at a $(t_c - t)$ value of

$$(t_c - t) = \frac{405M}{16\eta} \ , \tag{2.37}$$

and a frequency of

$$f_{\mathrm{ISCO}} = \frac{1}{6\sqrt{6}\pi M} \ . \tag{2.38}$$

The bodies will then plunge together, merging, leaving a final perturbed black hole, which relaxes through quasi-normal mode oscillations. These final stages of the coalescence we call the “merger” and “ringdown.”

2.4.1.2 Inspiral Waveform Approximants

The above derivation was a rather simple one using only the leading-order energy-loss terms and the quadrupole approximation. All of these arguments have been expanded in Post-Newtonian (PN) theory where matching between the near-zone gravitational field and the wave-zone gravitational field is done. In various studies different assumptions and expansions are used to calculate these waveforms to varying orders in v/c, where an nPN order calculation is a controlled Taylor expansion up to $(v/c)^n$. The following are a few examples: reference [3] computes waveforms to Newtonian order in amplitude and 2PN order in phase, references [4, 5, 6, 7] use the Effective One Body (EOB) approximation for the Hamiltonian to compute waveforms to Newtonian order in amplitude and 3PN order in phase, references [8, 9] use the Padé resummation technique to compute waveforms to Newtonian order in amplitude and 3.5PN order in phase. All of the previous waveforms ignored spin effects associated with angular momentum of the bodies. The SpinTaylor approximant is computed to Newtonian order in amplitude and 3.5PN order in phase using formulae from references [10] and based on references [11, 12, 3, 13, 14, 15, 8, 16].

An example waveform constructed at Newtonian order in amplitude and 2PN order in phase is given similarly as above, however, in order to simply the expressions, we introduce a dimensionless time variable

$$\Theta = \left(\frac{\eta}{5M}\left(t_c - t\right)\right)^{1/8} . \tag{2.39}$$

In terms of this new variable, the orbital phase is given by

$$\begin{aligned}\Phi(t) \quad = \quad & \phi_0 - \frac{1}{\eta}\left[\Theta^5 + \left(\frac{3715}{8064} + \frac{55}{96}\eta\right)\Theta^3 - \frac{3\pi}{4}\Theta^2\right. \\ & \left. + \left(\frac{9\,275\,495}{14\,450\,688} + \frac{284\,875}{258\,048}\eta + \frac{1855}{2048}\eta^2\right)\Theta^1\right] .\end{aligned} \tag{2.40}$$

In this notation, equation (2.36b) is given by the first two terms on the right hand side of equation (2.40).

2.4.1.3 Inspiral Waveform Stationary Phase Approximation

When searching for these signals in GW detector data, it is convenient to convert these waveforms to the frequency domain using the Stationary Phase Approximation (SPA)

$$\begin{aligned}\tilde{h}(f) &= \frac{A(t)}{\sqrt{\dot{f}(t)}} e^{i\Psi(f)} && (2.41)\\ &= \mathcal{A}(f) e^{i\Psi(f)} \; . && (2.42)\end{aligned}$$

In this approximation, the Newtonian amplitude $\mathcal{A}(f)$ is given as

$$\mathcal{A}(f) = \sqrt{\frac{5}{24}} \left(\frac{\mathcal{M}^{5/6}}{\pi^{2/3} D f^{7/6}} \right) \; , \tag{2.43}$$

and the 2PN phase $\Psi(f)$ is given in terms of a dimensionless frequency variable,

$$\tau = (\pi M f)^{1/3} \; , \tag{2.44}$$

as

$$\begin{aligned}\Psi(f) &= 2\pi f t_c + \phi_0 - \frac{3}{128\eta} \left[\tau^{-5} + \left(\frac{3715}{756} + \frac{55}{9}\eta \right) \tau^{-3} - 16\pi\tau^{-2} \right. \\ &\quad \left. + \left(\frac{15\,293\,365}{508\,032} + \frac{27\,145}{504}\eta + \frac{3085}{72}\eta^2 \right) \tau^{-1} \right] \; . && (2.45)\end{aligned}$$

2.4.2 Other Types of Gravitational Waves

Additional examples of transient situations in which the gravitational waveform is well modelled include GWs from cusps traveling along cosmic strings, hyperbolic encounters between compact objects, and ringdowns from perturbed black holes. Since the waveform is well modeled, GW signals from these transients can be searched for using matched filtering, which we discuss in chapter 5, as was used in references [17, 18, 19, 20, 21, 22, 23].

We call transient GW events “bursts” when their waveform is not well modeled. Examples of situations generating these types of signals include GWs from supernova and neutron star glitches. In order to search for these types of signals, since the waveform is unknown, GW data is searched for transient excess power. In order to ensure that these transients are not just noise in the detectors, these types of searches require either coherent or coincident signals to be seen in multiple detectors. Examples of searches that have been performed from the LIGO Scientific Collaboration can be found in [24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 23, 34].

Stochastic waves come from sources that are continuously emitting in either an incoherent fashion or in such a large number that the signals at the detectors are confused and inseparable. Examples of these types of sources include relic GWs from inflation, GWs from Galactic binary white-dwarf systems, and GWs from slow-spinning Galactic pulsars. These types of signals are searched for by looking for correlations between multiple GW detectors. Examples of these searches can be found in [35, 36, 37, 38, 39, 40, 41, 42, 43].

Continuous wave sources are are defined as signals that are quasimonochromatic and are separable from other sources. Such signals are commonly emitted by asymmetric spinning neutron stars, such as pulsars. These signals are emitted from their source at a constant frequency, however when they detected, they have been modified by Doppler shifts from motion due to the Earth's rotation, the Earth's orbit around the sun, the orbit of the source (e.g., when the source is part of a binary system), etc. Searches for these types of signals then correct for these shifts, coherently or incoherently sum the data over larger periods of time, and also correlate data from multiple detectors in order to reduce noise. Examples can be found in [44, 45, 46, 47, 48].

The remainder of this thesis focuses on GWs from CBCs.

Chapter 3

Introduction to LIGO

In this chapter we describe the history of GW detectors and the design of the LIGO GW detectors, which have been built for the detection of GWs. This description is broken up into several parts: a brief history of GW detectors in given in section 3.1, the design of the detectors is described in section 3.2, the dominant noise sources for LIGO are described in section 3.3, the calibration of LIGO is described in section 3.4, the response of the detectors to GWs is given in section 3.5, a description of using multiple detectors in a global network is given in section 3.6, and a description of previous LIGO science runs are given in section 3.7.

3.1 History of Gravitational-Wave Detectors

There exists a rich, and somewhat controversial, history of searching for GWs [49]. In the late 1950s and early 1960s, Weber designed and built the first GW detector in the form of a large, suspended, metal bar [50]. This bar was fit with piezoelectric transducers around its center to measure the oscillations of this driven harmonic oscillator. This bar was isolated from the surrounding environment using a combination of seismic isolation suspensions and a vacuum chamber. As GWs passed through this bar, the dynamical Riemann tensor would induce oscillations of the bar's fundamental mode, giving it the name of a resonant mass detector. These oscillations would then be detected by the transducers and, looking at the fundamental mode oscillations using a Fourier Transform, recorded. Weber's group went on to build several additional detectors so that they could look for coincident triggers (i.e., times when the detectors exhibited excess noise at their resonance frequency).

Results from these searches can be found in Refs. [51, 52] where they find rates of coincident triggers well above the background.

After Weber built this detector and reported data from it, several groups around the world built their own detectors. These group (with detectors located in e.g., Argonne, Glasgow, Moscow, ISR, IBM, Reading, Rutherford, Tokyo, Munich, Frascati, Stanford, Bell Labs, and Rochester [53]) then went on to publish additional papers about the rate of coincident triggers not being consistent with Weber's observations.

The next set of advanced detectors were the ultralow temperature bars Allegro [54], AURIGA [55], Explorer [56], NAUTILUS [57], and Niobe [58]. These detectors further suppressed the noise in the detectors by reducing the temperature of the detectors to a few Kelvin using refrigeration, allowing detector noise temperatures of a few to submilliKelvin levels to be reached. Searches for GWs with these detectors also could not measure a rate above their background.

Another design for resonant mass detectors, a spherical mass rather than a cylindrical bar, was conceived by Forward in Ref. [59]. If we look at the antenna response of the different resonant mass detectors, we see that the bars are sensitive to strains parallel to their cylindrical axis with the response tensor

$$R^{ab} = u^a_X u^b_X \ , \tag{3.1}$$

where u^a_X is the unit vector along the cylindrical axis of the bar. This implies that these detectors will be relatively insensitive to GWs traveling along this axis as GWs are transverse(-traceless) (TT) waves (see chapter 2).

However, this is not the case for spherical resonant detectors. For spherical detectors, the response tensor would be dependent on the number and locations of transducers, with their locations not being limited to the locations conducive to measuring the fundamental mode. Since GWs are TTwaves, a basis of five tensors is all that is needed in order to determine all directions, polarizations, and magnitudes of incident GWs [59]. One independent set given by Forward is

$$A^{ab} = \frac{1}{\sqrt{2}}(u^a_X u^b_X - u^a_Y u^b_Y) \ , \tag{3.2a}$$

$$B^{ab} = \frac{1}{\sqrt{6}}(-u^a_X u^b_X - u^a_Y u^b_Y + u^a_Z u^b_Z) \ , \tag{3.2b}$$

$$C^{ab} = \frac{1}{\sqrt{2}}(u^a_X u^b_Y + u^a_Y u^b_X) \ , \tag{3.2c}$$

$$D^{ab} = \frac{1}{\sqrt{2}}(u^a_X u^b_Z + u^a_Z u^b_X) \ , \tag{3.2d}$$

$$E^{ab} = \frac{1}{\sqrt{2}}(u^a_Y u^b_Z + u^a_Z u^b_Y) \ . \tag{3.2e}$$

Several detectors of this sort are being pursued around the world. These include the MiniGRAIL project [60] in Leiden University, the Netherlands and the “Mario Schenberg” GW detector [61] in the University of São Paulo, Brazil.

A different detector technology that has also been designed and built is the interferometer detector. This type of detector, in its most basic form, consists of a Michelson Interferometer that measures differential length changes along its two arms. The advantage of this type of detector is that, as one is measuring actual length changes instead of the amplitude of oscillations on resonance of a harmonic oscillator, the bandwidth of the detector can be much larger than that of the resonant mass detectors. If the arms are orthogonal and oriented along the X and Y directions respectively, this type of detector has the response tensor

$$R^{ab} = u^a_X u^b_X - u^a_Y u^b_Y \ . \tag{3.3}$$

A meter size version of this detector was first used for this purpose by Forward in the 1970s [62], of comparable sensitivity to the bar detectors operating at that time. Around the same time, Weiss analyzed the noise sources of such a detector and developed the first ideas of kilometer-scale devices [63]. This inspiration is what has led to the modern day construction of TAMA300 [64], GEO600 [65], Virgo [66], and the three LIGO detectors [67], all kilometer-scale GW detectors.

Resonant mass detectors are typically sensitive to GWs in the frequency range of 10^3–10^4 Hz, and ground-based interferometer detectors are sensitive to GWs in the frequency range of 10^1–10^4 Hz. Other techniques are also currently being pursued to detect GWs in different frequency regions. These include space-based interferometers, such as LISA [68] and DECIGO [69], designed to be sensitive to GWs with frequencies between 10^{-3}–10^{-2} Hz and 10^{-1}–10^1 Hz respectively, pulsar timing experiments (first described in Refs. [70, 71]) aimed at detecting very low frequency GWs in the range of 10^{-9}–10^{-7} Hz, and Cosmic Microwave Background polarization measurements [72, 73, 74, 75, 76, 77, 78, 79, 80, 81] designed to detect GWs in the extremely-low frequency band of 10^{-18}–10^{-15} Hz.

3.2 Design of LIGO

The LIGO detectors were designed to detect differential strains from incident GWs, based on a simple Michelson Interferometer (see figure 3.1). This design starts with an input laser incident on a beam splitter located at the origin with the face at a 45° angle between the x- and y-axes. The beam splitter directs 50% of the light in two perpendicular directions, x and y. The light travels toward mirrors located at the end of the arms, $(L_x, 0)$ and $(0, L_y)$, which redirect the light back toward the beam splitter. The phase the light accumulates as it travels from the beam splitter to the end mirror is given by

$$\phi = \int_0^L \frac{2\pi}{\lambda} dx = \frac{2\pi L}{\lambda} \, , \tag{3.4}$$

where L is the distance traveled and λ is the wavelength of the light. The total phase of one round trip will then be the sum of the integral down and the integral back.

When the light returns to the beam splitter, the two beams interfere with each other in such a way that, depending on the differential arm length, the light is either directed back toward the laser (the symmetric port), or toward a photodetector (the antisymmetric port), or some combination of the two. The differential arm length $L_x - L_y$ can then be measured, modulo one wavelength, by looking at the power of the light coming from the antisymmetric port, P_{AS}. This is given by

$$\begin{aligned} P_{\mathrm{AS}} &= P_0 \sin^2(\phi_x - \phi_y) \\ &= P_0 \sin^2\left(\frac{4\pi}{\lambda}(L_x - L_y)\right) , \end{aligned} \tag{3.5}$$

where P_0 is the input laser power, ϕ_x is the phase accumulated for the beam in the x-arm, and ϕ_y is the phase accumulated for the beam in the y-arm. This shows that P_{AS} is dependent on the differential arm length.

Let us assume that the arms are of the same length L, and an incident GW with amplitude h is stretching the detector along the x-axis and squeezing it along the y-axis. We see from chapter 2 that this effect causes the distance the light needs to travel to be given by $L' \approx L(1 \pm h/2)$, giving ϕ_x and ϕ_y as

$$\begin{aligned} \phi_x &= \int_0^{L'} \frac{2\pi}{\lambda} dx - \int_{L'}^{0} \frac{2\pi}{\lambda} dx \\ &\approx \frac{4\pi L}{\lambda}\left(1 + \frac{h}{2}\right) , \end{aligned} \tag{3.6a}$$

$$\phi_y \approx \frac{4\pi L}{\lambda}\left(1 - \frac{h}{2}\right) . \tag{3.6b}$$

With these phase shifts, we find that P_{AS} is then given by

$$P_{\mathrm{AS}} = P_0 \sin\left(4\pi \frac{hL}{\lambda}\right)^2 , \tag{3.7}$$

such that for the same change in P_{AS}, the strain h is inversely proportional the length L of the arms. However, although there is a signal at the antisymmetric port for this interferometer, this is a poor way to detect GWs. For one, since the above power at the antisymmetric port is proportional to $\sin^2(\Delta\phi)$, which can be Taylor expanded about zero as $\Delta\phi^2$, the signal we are measuring is *quadratically* proportional to the small GW signal we are trying to sense. In addition, there is no information on which direction we need to move the mirrors to bring the signal back to the dark fringe.

In order to get around these problems, LIGO has adopted the use of the “heterodyne detection” technique, which combines a phase-modulated input laser with a Schnupp Asymmetry. In this setup, before the laser light enters the interferometer, the light is phase modulated with a frequency Ω such that instead of all the light entering the detector with one frequency, there are effectively three superposed laser beams entering with three different frequencies and electric field amplitudes: the original beam with frequency ω and amplitude E_0, also known as the *carrier*, and two *sidebands* with frequency $\omega \pm \Omega$ and amplitude E_1.

When the arm lengths are perfectly equal, all of these beams will come out at the symmetric port, which results in the same answer as above. However, there is a different response when a Schnupp Asymmetry is introduced between the arms. This occurs when the arms of the interferometer are *unequal* by an amount that is an integer number of wavelengths of the carrier light but *not* an integer of wavelengths for the sidebands. In this situation, in the absence of a GW signal, the carrier light will come out the symmetric port while the sidebands will be transmitted to the antisymmetric port. When there is a GW signal present, a phase shift will be introduced on all three beams such that a mixture of all three beams will exit the antisymmetric port. Processing the signal coming from the antisymmetric port with demodulation results in an error signal that is *linear* in the GW amplitude, just as is needed.

In order to increase the sensitivity of the detector, LIGO uses two modifications to the standard Michelson interferometer. The first is the use of Fabry-Perot cavities in the two arms in order to increase the storage time of light within the arms, effectively increasing the length of the arms, and thus the phase shift on the light, by the same factor. The second is to boost the light circulating within the interferometer for a given laser input power with the addition of a power-recycling mirror. A schematic for the design of LIGO, including all of the additions to the standard Michelson interferometer, can be seen in figure 3.1.

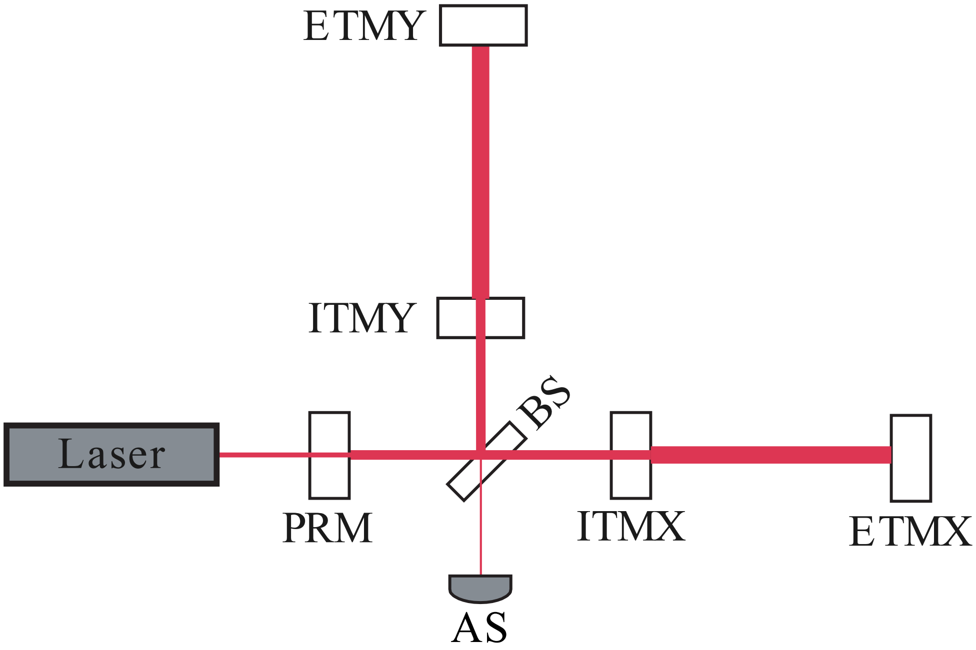

Figure 3.1: LIGO Schematic
A simple schematic of the LIGO detectors. The mirror labels are as follows: PRM (power recycling mirror), BS (beam splitter), ITMX/Y (x-/y- arm input test mass), ETMX/Y (x-/y- arm end test mass), AS (antisymmetric port). The width of the laser beam denotes power in different portions of the detector. A simple Michelson interferometer includes the mirrors BS and ETMX/Y. The addition of the ITMX/Y mirrors makes use of Fabry-Perot cavities for the arms. The use of the PRM boosts the power circulating in the whole interferometer.

3.3 Dominant Noise Sources

The design sensitivity of the LIGO detectors is limited by several fundamental noise sources for different frequency bands [82]: Seismic Noise for $f < 40\,\mathrm{Hz}$, suspension thermal noise for $40\,\mathrm{Hz} < f < 150\,\mathrm{Hz}$, and photon shot noise for $f > 150\,\mathrm{Hz}$.

Seismic Noise is the result of ground motion coupling to the motion of the mirrors. This type of noise is suppressed in multiple ways, both actively and passively. Passive isolation is achieved by suspending the mirrors as pendula, on seismic isolation stacks. This suspension acts as a low pass filter for motion coming from the ground, resulting in a pendulum response function for the mirror motion that dies as $1/f^2$ for noise above the resonant frequency. Active isolation is also used by sensing ground motion and feeding this back into the controls of the Hydraulic External Pre-Isolator for cancellation.

Suspension Thermal Noise results from the random thermal motions of the room-temperature suspension feeding into motions of the mirrors. This motion has been calculated by Saulson in Ref. [83].

Photon Shot Noise results from a finite number of photons arriving each second, informing us of the motion of the mirrors. Since we are limited in the number of photons that arrive at the photodetector per second by the power of the laser light, we end up not being able to detect strains better than [82]

$$\Delta h[\mathrm{Hz}]^{-1/2} = \frac{1}{L}\left(\frac{hc\lambda}{\pi^2 P}\right)^{1/2} , \tag{3.8}$$

where h is Planck's constant, c is the speed of light, λ is the wavelength of the laser light, and P is the power of the laser light. The arm cavity introduces an enhancement in P at the expense of reduced sensitivity at high frequencies due to the cavity pole. This results in a shot-noise equivalent strain that rises with f above the cavity pole frequency of $\sim$90 Hz.

These are not the only noise sources in the LIGO detectors. figure 3.2 shows all of the known noise sources during the S5 run [84]. There the noise in channel that carries the GW signal (DARM) is compared to the sum of all the known noise sources (Total). Above 100 Hz, the dominant noise

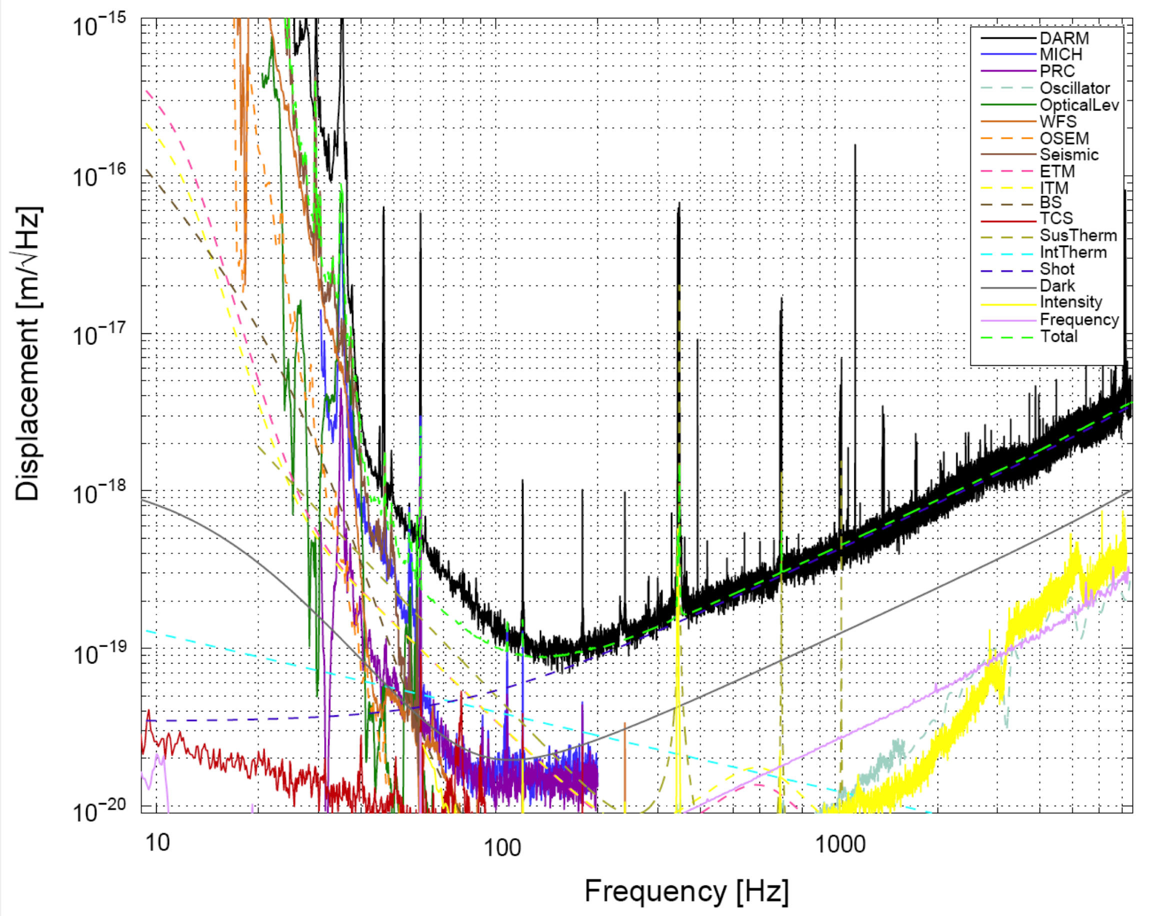

Figure 3.2: LIGO Noise Budget
The noise budget for the 4 km instrument at Hanford from Ref. [84]. The line labeled "Total" is the sum of all known noise sources. The black line (measured DARM noise) should be compared to this.

source is shot noise and DARM agrees very well with Total. Between 50 and 100 Hz there are noise sources that have not been accounted for in addition to the expected dominant source, which is suspension thermal noise. Below 50 Hz the noise is mainly due to man-made activities and is highly nonstationary, thus it is hard to estimate at any given point in time.

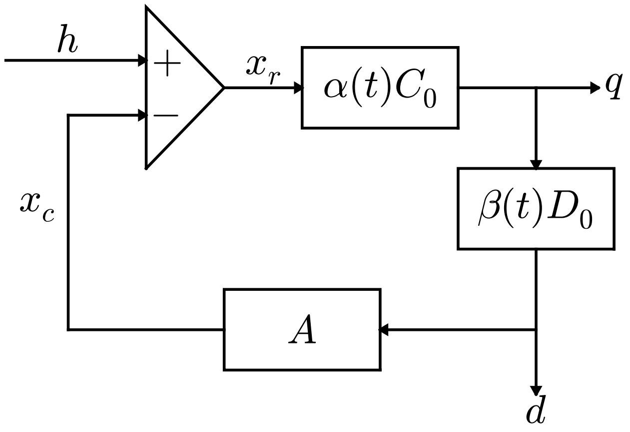

Figure 3.3: Model Control Loop
A simple model for the control loop of the LIGO length sensing and control system [85].

3.4 Calibration

In this setup, the light incident on the antisymmetric port's photodetector carries the information of the strain the detector is experiencing. However, in order to keep the detectors in a stable data-taking mode, this signal is fed back into detector such that the light arriving from the antisymmetric port is kept roughly constant. This means that the strain of the detector must be reconstructed from the signal arriving from the antisymmetric port, known as the error signal q, and the transformations made on q to control the detector, what we call the response function R. This reconstruction is easily understood in the frequency domain as [85]

$$h(f) = R(f)q(f)\ . \tag{3.9}$$

The feedback loop diagram for this control system is shown in figure 3.3. In the frequency domain, the relationships depicted in the diagram take the simple form

$$x_r(f) = h(f) + x_c(f) , \tag{3.10a}$$

$$q(f) = \alpha(t) C_0(f) x_r(f) , \tag{3.10b}$$

$$d(f) = \beta(t) D_0(f) q(f) , \tag{3.10c}$$

$$x_c(f) = A(f) d(f) , \tag{3.10d}$$

where $C_0(f)$ is a reference sensing function measured at some reference time whose frequency dependence is determined primarily by the Fabry-Perot arm cavities and is roughly $(1 - f/f_p)^{-1}$ with $f_p \approx 90$ Hz and $\alpha(t)$ is a real overall gain that depends on the light power stored in the arms and is measured every ~ 10 seconds, A is an actuation function that is largely determined by the pendulum suspension of the mirrors, and D_0 is the feedback filter and $\beta(t)$ is a real overall gain.

From these relations, $R(f)$ is found to be

$$R(f) = \frac{1 + \alpha(t)\beta(t)G_0(f)}{\alpha(t)C_0(f)} , \tag{3.11}$$

where $G_0(f)$ is the reference open loop gain given by

$$G_0(f) = A(f) C_0(f) D_0(f) , \tag{3.12}$$

which has a unity gain frequency of ~ 150 Hz. The time-dependent gain factors $\alpha(t)$ and $\beta(t)$ are measured by tracking the amplitude of calibration lines that are injected into the detectors at around 1 kHz, 150 Hz, and a few tens of Hz.

This procedure tells us how to calibrate the error signal from the detector in order to obtain the strain signal in the frequency domain. However, for LIGO's fifth science run (S5), an additional step was added to the calibration [85], giving the strain signal in the time domain $h(t)$. This is the main method used for analyses of the S5 data.

3.5 Interferometer Antenna Response

In chapter 2, we described the effect of incident GWs on test particles perpendicular to the direction of propagation of the waves. We noted that interferometric detectors have a response tensor, in the reference frame of the beam splitter, of the form

$$R^{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{3.13}$$

The resulting response of the detector is given by

$$\frac{\Delta l}{l} = \frac{1}{2} h_{\alpha\beta} R^{\alpha\beta} . \tag{3.14}$$

GWs from arbitrary directions will have different projections onto this response tensor. Here we describe those projections as the antenna response of the detector to GWs incident from directions given by θ and ϕ, and the polarization angle ψ, which bring equation (3.14) into the form

$$\frac{\Delta l}{l} = F_+ h_+ + F_\times h_\times , \tag{3.15}$$

where F_+ and $F_\times$ are the antenna factors we are evaluating. The above angles are identified in figure 3.4.

In this derivation we start with the assumption that the wavelength of the GW is much longer than the storage time of the light in the interferometer's arms times the speed of light $\lambda \gg cT$.

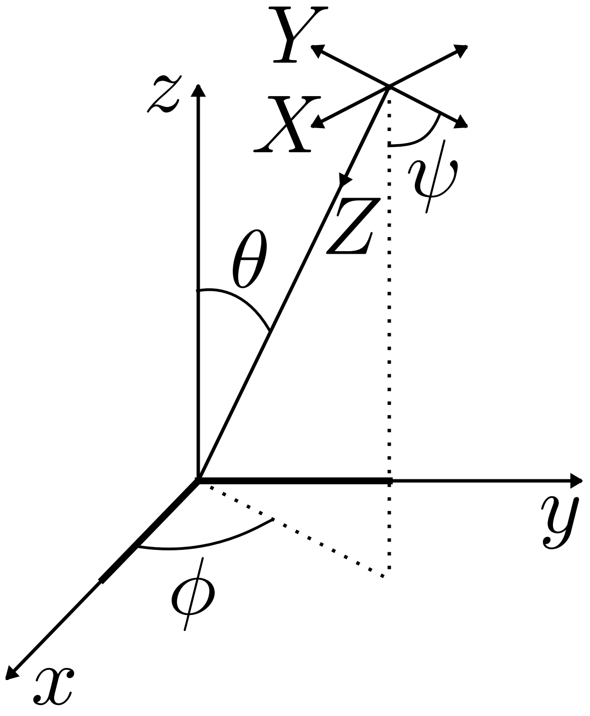

Figure 3.4: Euler Angles
The Euler angles $\{\phi, \theta, \psi\}$ used in converting from the GW propagation frame $\{X, Y, Z\}$ to the detector frame $\{x, y, z\}$.

In the radiation frame, denoted with "$'$"s above the indices, the GW $h_{\alpha'\beta'}$ is given by

$$h_{\alpha'\beta'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.16}$$

which can be split into

$$h_{\alpha'\beta'} = S^+_{\alpha'\beta'} h_+ + S^\times_{\alpha'\beta'} h_\times \ . \tag{3.17}$$

To transform this into frame of the detector (i.e., the frame of the beam splitter) we need to transform both indices of $h_{\alpha'\beta'}$. This transformation can be seen as the combination of different

Euler angle rotations with the combined rotation matrix given by

$$R_{\alpha}^{\alpha'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\cos\theta\cos\phi\cos\psi - \sin\phi\sin\psi & -\cos\theta\cos\phi\sin\psi + \sin\phi\cos\psi & \sin\theta\cos\phi \\ 0 & \cos\theta\sin\phi\cos\psi - \cos\phi\sin\psi & \cos\theta\sin\phi\sin\psi + \cos\phi\cos\psi & -\sin\theta\sin\phi \\ 0 & -\sin\theta\cos\psi & -\sin\theta\sin\psi & -\cos\theta \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A & B & C \\ 0 & D & E & F \\ 0 & G & H & I \end{pmatrix} . \tag{3.18}$$

Applying this to B_+ and $B_\times$, we find

$$S^{+}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A^2 - D^2 & AB - DE & AC - DF \\ 0 & AB - DE & B^2 - E^2 & BC - EF \\ 0 & AC - DF & BC - EF & C^2 - F^2 \end{pmatrix} , \tag{3.19a}$$

$$S^{\times}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2AD & AE + BD & AF + CD \\ 0 & BD + AE & 2EB & BF + CE \\ 0 & CD + AF & CE + BF & 2CF \end{pmatrix} . \tag{3.19b}$$

We can now calculate the antenna factors where $F_+ = (1/2)S^{+}_{\alpha\beta}R^{\alpha\beta}$ and $F_\times = (1/2)S^{\times}_{\alpha\beta}R^{\alpha\beta}$ resulting in

$$F_+ = \frac{1}{2}\left(1 + \cos^2\theta\right)\cos 2\phi\cos 2\psi + \cos\theta\sin 2\phi\sin 2\psi \ , \tag{3.20a}$$

$$F_\times = \frac{1}{2}\left(1 + \cos^2\theta\right)\cos 2\phi\sin 2\psi + \cos\theta\sin 2\phi\cos 2\psi \ . \tag{3.20b}$$

We can see the structure of these antenna factors when plotted with $\psi = 0$ in figure 3.5. From these we find that there are several regions of the sky where the detectors are sensitive to neither polarization, which we call nodes.

There is one additional angle that will affect the amplitude of the signal that a interferometer detector will receive from a quasi-circular inspiralling compact binary, the angle of inclination ι between the orbital plane and vector connecting the origin of the detector and the origin of the source. This will have an effect on resulting signal seen by the detector $h(t)$ such that

$$h(t) = h_+(t)F_+\left(1 + \cos^2\iota\right) + h_\times(t)F_\times\left(2\cos\iota\right) \ . \tag{3.21}$$

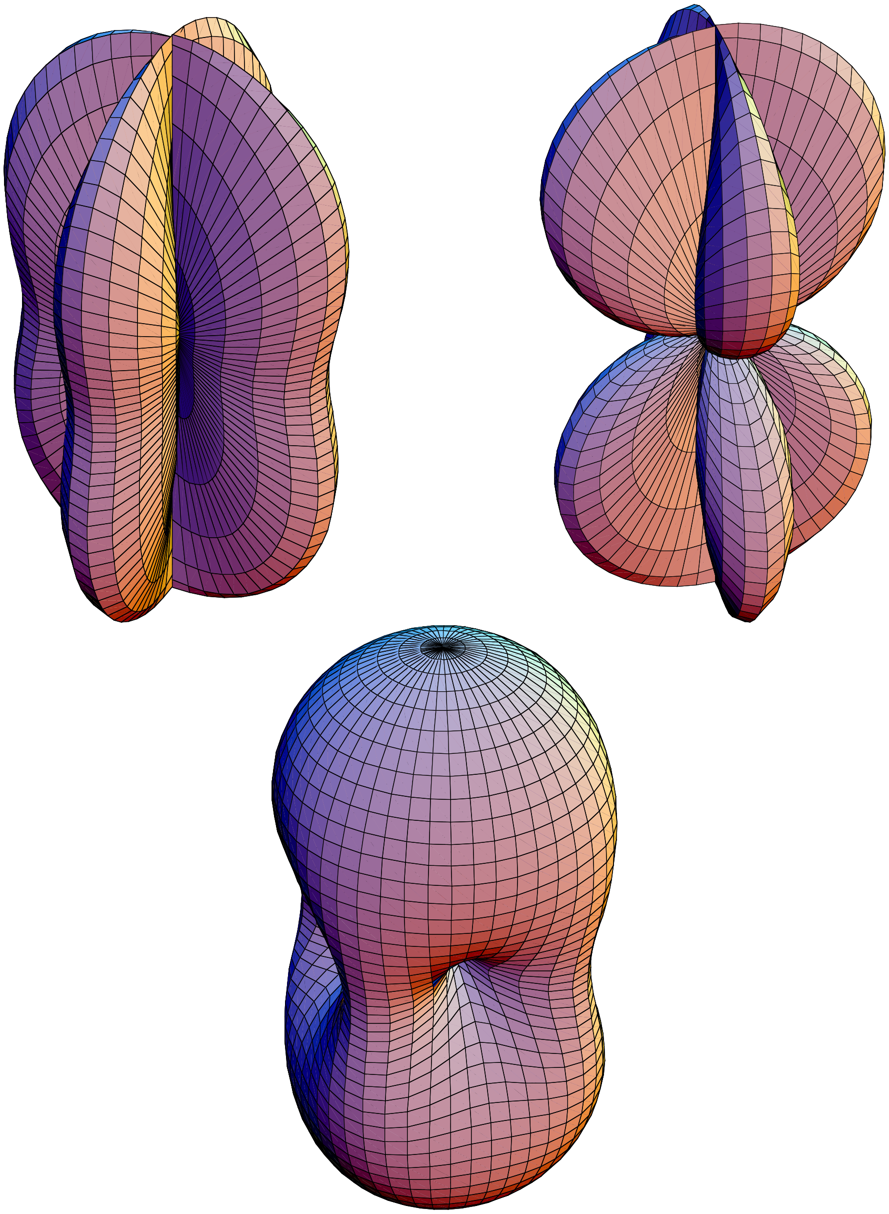

Figure 3.5: Detector Antenna Patterns
The antenna patterns for interferometer detectors for plus (top left), cross (top right), and combined (bottom) polarizations with polarization angle $\psi = 0$. For this figure, the detectors arms are oriented such that the first node in the bottom plot is found at 45° between the two arms of the detector and subsequent nodes occur every 90° as you rotate around the vertical axis.

Table 3.1: Combined Horizon Volumes

Detector Combination	Volume (Mpc^3)
H1H2	2.13×10^3
H2L1	2.11×10^3
H2V1	1.81×10^3
H2L1V1	8.84×10^3
H1L1 H1H2L1	1.23×10^4
H1V1 H1H2V1	8.68×10^3
H1L1V1 H1H2L1V1	1.54×10^4

The combined horizon volume to which each detector combination is sensitive. When calculating this, we compute the distance at which each detector would see a binary neutron star signal with inclination angle ι at an SNR of 8. For each point in the sky and each polarization ψ, we take the distance from the second most sensitive detector (i.e. only double coicident observations are required), volume-averaged over ψ and ι, and then integrate over the sky. These calculations are done using the design sensitivity of each detector.

3.6 Networks of Detectors

As was discussed above, coincidence can be used to verify a GW signal when it is seen by multiple detectors. However, this can only be done when at least two detectors see a signal. One common way to quantify the sensitivity of a detector is by looking at the distance at which a given detector can see a optimally located and oriented binary neutron star signal (i.e., an overhead, face on orbit; $\theta = \phi = \psi = \iota = 0$) with an SNR of 8. This is referred to as the horizon distance. Non-optimally located and oriented sources can be detected with SNR $= 8$ at shorter distances. Table 3.1 shows the volumes of space that are screened after averaging over all of these angles when different combinations of interferometric GW detectors are operating. The angle-averaged sky maps associated with each detector and combination of detectors can be seen in figure 3.6 and 3.7 respectively, which are Mollweide projections of the sky. As additional detectors are built around the world, we will be able to more completely cover the sky.

Just as combining the signals from different transducers on a spherical resonant mass detector allows us to determine the sky location and polarization of a given signal, coherently combining the output from multiple, misaligned interferometric detectors accomplishes the same thing [86]. Another advantage of coherent combinations of signals is the amplification of the signal relative to

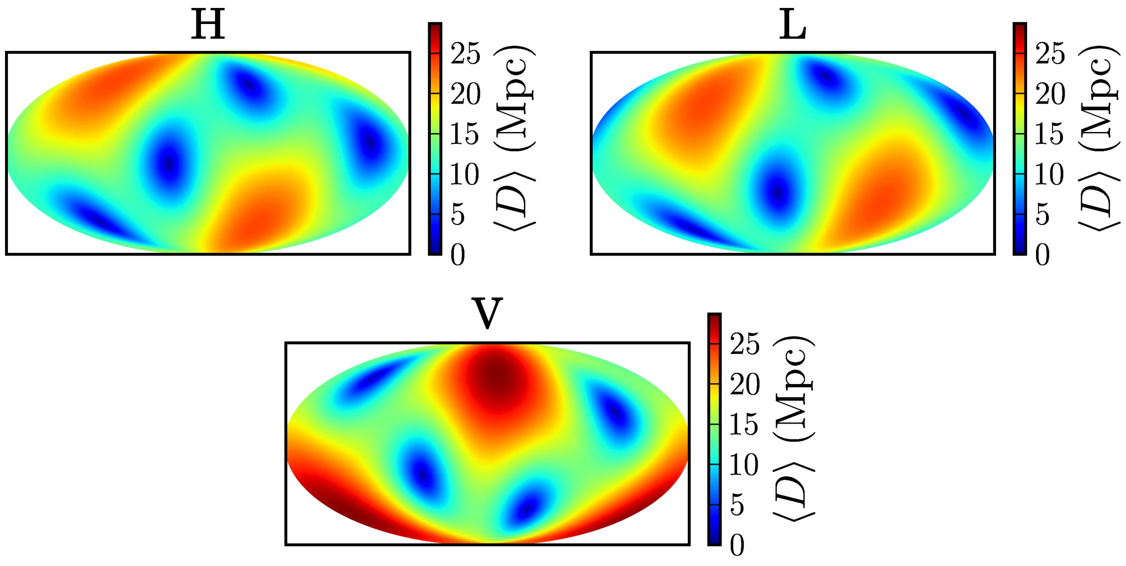

Figure 3.6: Individual Detector Angle-Average Sky Maps
The Hanford (top left), Livingston (top right), and Virgo (bottom) angle-averaged sky maps for their design sensitivities. The color scale shows the average distance at which each detector can see a binary neutron star signal with an SNR of 8, volume-averaged over the polarization and inclination angles.

the noise in the detectors. Because the noise in the different detectors is independent random noise, when the signals are combined, the random fluctuations are suppressed though incoherent addition while external GW signals add coherently.

A final advantage of coherent analysis is the overdetermination of the signal from multiple data streams. When at least three detectors are used, since there are only two independent polarizations of a GW signal, the data from the detectors can be combined in such a way that the signal should be cancelled out. This type of combination, called a null-combination, can be used to identify an accidental coincidence of noise triggers in the detector network since this will cause excess energy in all the detectors that is incoherent.

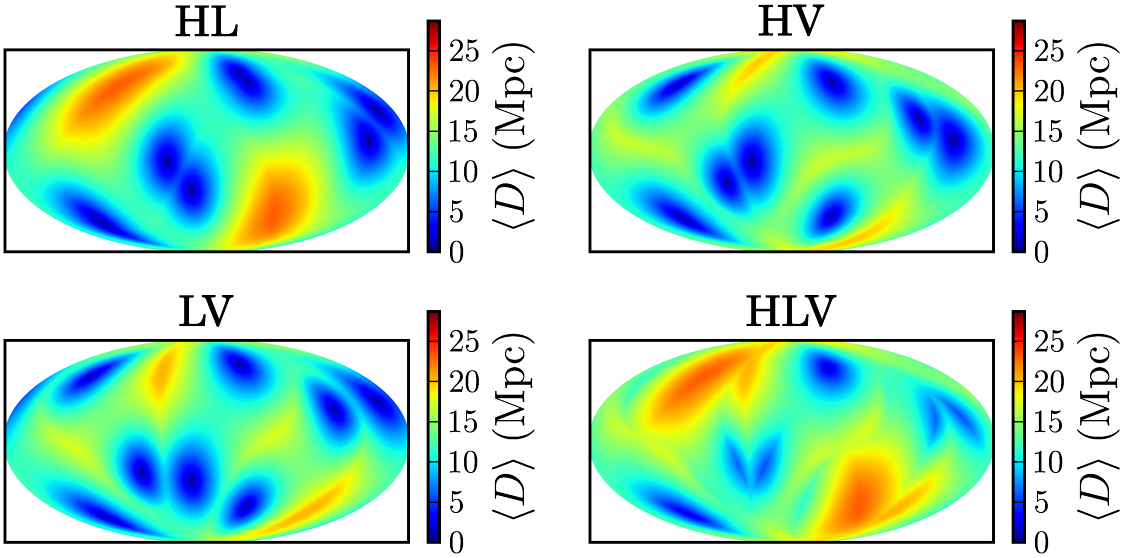

Figure 3.7: Combined Detector Angle-Average Sky Maps
The Hanford-Livingston (top left), Hanford-Virgo (top right), Livingston-Virgo (bottom left), and Hanford-Livingston-Virgo (bottom right) combined angle-averaged sky maps for their design sensitivities. The color scale shows the average distance at which each that detector combination can see a binary neutron star signal with an SNR of 8, volume-averaged over the polarization and inclination angles.

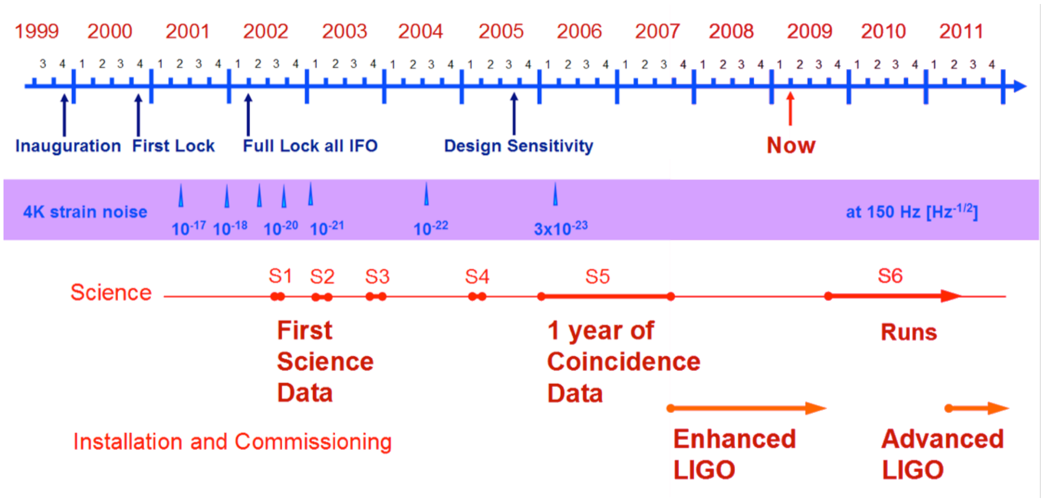

Figure 3.8: LIGO Science Runs Timeline
A timeline showing the progress made in the LIGO detectors strain sensitivity as well as the different periods of science data taking [84].

3.7 LIGO's Past and Future

After the LIGO detectors were built, there were several times when science data was taken in between commissioning periods on their way toward their design sensitivity (figure 3.8). These science runs (S1–S4) were analyzed for many different types of signals [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 23, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48]. One of the important aspects of these runs was finding times during which the detectors were not operating properly. During these times, many auxiliary channels of the detectors were recorded and analyzed, leading to the discovery of different ways the interferometers glitch and the suppression of such glitches during the next commissioning break.

In the upcoming months, the LIGO detectors will finish being upgraded from their initial design (Initial LIGO) to their enhanced design (called enhanced LIGO) and will start taking data for their sixth science run with an expected factor of two increase in sensitivity over S5. A few years later, the detectors will be taken offline again to undergo a more extensive reworking, returning with their advanced design (Advanced LIGO) in place, which should further improve the sensitivity by an order of magnitude with respect to Initial LIGO.

Chapter 4

Astrophysics of Compact Binaries

This chapter is broken up into two sections. The first (section 4.1) presents evidence for the existence of compact objects (i.e., neutron stars and black holes). The second (section 4.2) presents predicted rates for the coalescence of binary compact objects based on observations and theoretical considerations.

4.1 Evidence for Neutron Stars and Black Holes

To date, there have been many observations of compact objects, which confirm the existence of neutron stars and strongly suggest the existence black holes. The theoretical maximum mass for an electron-degenerate star is given by the Chandrasekhar mass limit of $\sim$1.4 $M_{\odot}$. This limit is implicated in multiple types of stellar explosions, including Type Ia/b/c and II supernovae, in which either the mass of an electron-degenerate, white dwarf approaches this limit (Type Ia), or the core of a massive a star approaches this limit (Type 1b/c II). For the latter situation, the observations of the remnant object has been linked to compact object observations, as in the case of Cassiopeia A [87].

Compact objects with a tightly orbiting companion are good testing grounds for differentiating between neutron stars and stellar mass black holes. If the companion object is close enough, it fills its Roche Lobe, the boundary at which the gas of the companion object is no longer gravitationally bound to its parent star, and leaks onto the compact object. This gas then forms an accretion disk around the compact object as it is transferred. By measuring properties of the binary system as well as of the accretion disk, we can learn many things about the compact object. Studying the orbital parameters of the binary will tell us the mass of both the compact object and companion object. If the mass of the compact object is larger than the Chandrasekhar Limit, then the compact object must be either a neutron star or a black hole. If thermonuclear explosions are seen as matter falls onto the compact object, it is determined to be a neutron star. In the case of black holes, studying the profile of Fe-lines in the accretion disk can allow the spin of the black hole to be determined [88].

In addition, observation of binary pulsars such the Hulse-Taylor pulsar [89] have also yielded evidence for compact objects. The two objects, roughly 1.44 and 1.39 $M_\odot$, in this system are orbiting each other once every 8 hours, corresponding to a separation of $\sim$1 $R_\odot$. Additionally, this system is seen to lose orbital energy at the rate predicted by GW emission in general relativity. At this separation, if the objects were ordinary stars, there should be some other much stronger force causing them to depart from this prediction. Electromagnetic observations have also found no counterpart, which at a distance of 8 kpc, excludes these objects from being white dwarfs.

4.2 Predicted Compact Binary Coalescence Rates

In this section we summarize a number of different ways the rate of CBC are predicted. These methods are broken up into three categories, namely extrapolations from merging binary neutron star (BNS) observations (section 4.2.1), population synthesis of field binaries (section 4.2.2), dynamical simulations of star clusters (section 4.2.3), and extrapolations from short Gamma Ray Burst (GRB) observations (section 4.2.4) [90].

4.2.1 Extrapolations from Merging Binary-Neutron-Star Observations

Observations of binary systems involving pulsars that will coalesce within a Hubble time can be extrapolated in order to obtain a BNS merger rate. There are currently four known systems of binary pulsars that will merge within a Hubble time, and one more possible system, all in our galaxy. These systems are PSR B1913+16 (the Hulse-Taylor pulsar), PSR B1534+12, J0737-3039A, J1756-2251, and possibly J1906+0746.

In [91, 92], the authors use these binary pulsar systems to calculate a combined probability on the rate of coalescences in our galaxy. To do this they combine the number $N_{\rm tot}$ of observed Galactic pulsars, the fraction of Galactic pulsars with pulse and orbital characteristics similar to those of a particular type, the lifetime $\tau_{\rm life}$ of each of the observed systems, and the upward correction factor f_b^{-1} due to pulsar beaming. For each type of pulsar population, with the ratio α of the mean number of observed pulsars $\langle N_{\rm obs}\rangle$ to the total number of pulsars $N_{\rm tot}$ for that population

$$\alpha = \frac{\langle N_{\rm obs}\rangle}{N_{\rm tot}} \,, \tag{4.1}$$

they obtain a probability density function $P(R)$ for the rate given by

$$P(R) = \left(\alpha\tau_{\rm life} f_b\right)^2 R e^{-(\alpha\tau_{\rm life} f_b)R} \,. \tag{4.2}$$

These probability distributions can be combined (Appendix A of [93]) to obtain the most probable rate of $R \approx 71$ MWEG^{-1} Myr^{-1} and a 90% confidence interval of ∼15–240 MWEG^{-1} Myr^{-1} [94] without the inclusion of J1906+0746 and a factor of 2 larger with its inclusion, where an MWEG is a Milky Way Equivalent Galaxy. This can be converted to 7.1×10^{-1} Mpc^{-3} Myr^{-1} using a galaxy number density of 10^{-2} MWEG Mpc^{-3}.

4.2.2 Population Synthesis of Field Binaries

Population synthesis simulations have been used to estimate the rate of binary coalescences for BNS, binary black-hole neutron-star (BHNS), and binary black hole (BBH) systems in the binary evolution scenario (i.e., stars formed as binaries during their stellar formation). These simulations start with stars distributed according to an initial mass function. These systems are then evolved keeping track of their evolutionary details. At the end of the simulation, observational constraints can be imposed on the results such that the results are consistent with the observed sample of merging Galactic BNS, wide Galactic BNS, white dwarf–neutron star binaries, the observed rate of Type Ib/c, Type II supernovae, etc.

With these constraints imposed, the final rates can be can be obtained by looking at the number of BNS, BHNS, and BBH with orbits tight enough to merge within a Hubble time. Results suggest that the range of merger rates are 1–10^3 MWEG^{-1} Myr^{-1} for BNS [91, 92], 5×10^{-2}–1×10^2 MWEG^{-1} Myr^{-1} for BHNS [95], and 4×10^{-2}–1×10^2 MWEG^{-1} Myr^{-1} for BBH [96] systems.

4.2.3 Dynamical Simulations of Star Clusters

In this section we summarize arguments of [97, 98, 99, 100] for predicting the CBC rate from different types of star clusters. Star clusters may provide the necessary breeding ground for BBH due to their increased star formation rate and their increased density. Studies have shown [101] that in such an environment, mass segregation occurs through dynamical interactions within the cluster, driving the higher mass objects toward the center, increasing the possibility of forming stellar mass BBH systems.

For a cluster of mass M_{cl}, since stars form according to a power-law mass distribution [102], the number of stars that are massive enough to form black holes after supernova (i.e., those with masses greater than 20 $M_\odot$) is roughly $3 \times 10^{-3}(M_{\mathrm{cl}}/M_\odot)$. These objects undergo rapid evolution forming their black holes quickly (on a timescale of about $t_{\mathrm{SN}} \sim (M)^{-2.5} \times 10^{10}$ yr). Then, since these are the most massive objects, they tend to sink toward the center with a timescale of

$$t_{\mathrm{relax}} \simeq t_{\mathrm{cross}} \times 0.1N/\ln N \ , \tag{4.3}$$

where t_{cross} is the typical crossing time of the cluster [103]. This happens through a mechanism known as equipartition, which is a statistical tendency for objects undergoing two-body interactions to equilibrate their kinetic energy. This tendency causes the lower mass object to leave a two-body interaction with a larger average speed than the higher mass object, thus causing the higher mass object to sink further in the potential well of the cluster.

Another effect that occurs in the mass segregation process is the accumulation of high mass binaries. When a binary interacts with a third object, the two most massive objects tend to leave the interaction as a binary [104]. In these interactions, the binding energy of the binary tends to increase. These interactions thus favor the formation of tight binaries from the highest mass objects in the cluster.

The combination of mass segregation and binary exchanges continue until the recoil speed from the three-body interactions is large enough such that even the BBH systems receive a recoil larger than the escape velocity of the cluster [105, 106]. At that point, the binary has hardened enough that even though it leaves the cluster, the time until merger due to gravitational radiation is less than a Hubble time.

These arguments have been applied to globular clusters in [98]. They find that the rate of mergers from globular clusters is given by $\sim g_{\rm cl} g_{\rm evap}$ Mpc^{-3} Myr^{-1}, where $g_{\rm cl}$ is the fraction of total star formation that occurs in clusters and $g_{\rm evap}$ is the fraction of cluster-forming mass that possesses the birth conditions necessary for this process to occur.

These arguments have also been applied to nuclear star clusters both in the presence of a supermassive black hole (SMBH) [99], and not [100]. In [99], they found a wide range of merger rates in the presence of a SMBH with rates varying between 1.5×10^{-6} and 2×10^{-4} Myr^{-1} per galaxy (not Milky Way equivalent galaxy) depending on the model chosen. In [100], they argue for merger rates in the absence of a SMBH of a few times 10^{-3} Mpc^{-3} Myr^{-1}.

4.2.4 Extrapolations from Short-Gamma-Ray-Burst Observations

Here we present arguments used in [107] for estimating the rate of mergers of dynamically formed BNS and BHNS systems in globular clusters from short GRB observations. Short GRBs are theorized to originate in the merger of two compact objects forming a black hole surrounded by an accretion disk [108, 109, 110, 111]. In order for an accretion disk to form, there must be matter present in the system from either one or both objects, which excludes the possibility of short GRB coming from BBH systems.

BNS and BHNS systems can form in one of two ways. The objects were originally binary stars that both underwent core collapse forming a binary of compact objects, or the objects separately underwent core collapse and then dynamically formed through companion capture and possible exchange interactions. The former we will refer to as "primordial" systems, which can be found in the field. Results from population synthesis have shown that primordial systems that will merge within a Hubble time do so soon after their formation, thus the redshift associated with such systems should closely follow the star formation rate of the universe [112, 113, 114]. On the other hand, dynamically formed systems will be delayed by the relaxation time t_{relax} of the cluster, which can be on the order of a Hubble time.

Using this time delay, the short GRB rate can be calculated from the star formation rate history using

$$R_{GRB}(z) \propto \int_0^{t(z)} R_{\mathrm{SFR}}(t-\tau)P(\tau)d\tau \ , \tag{4.4}$$

where $R_{\mathrm{SFR}}(t)$ is the star formation rate at time t, $P(\tau)$ is the distribution of time delays τ before merger, which goes as $P(\tau) \sim 1/\tau$ for primordial systems [115, 116] while it increases for increasing time delays for dynamically formed systems [117].

Using equation (4.4) and the distribution of observed luminosities for short GRB, [107] calculate the local rate of events to be 1.3×10^{-3} and 4.0×10^{-3} Mpc^{-3} Myr^{-1} assuming all observed events come from primordial or dynamically formed systems respectively. Given the observed distribution of redshifts associated with short GRB, the best fit of the data comes from a 60% contribution from dynamically formed systems. Reducing the contribution from dynamically formed systems such that the Kolmogorov-Smirnov probability that the observed distribution comes from the expected distribution reaches a value of 0.1, the contribution from dynamically formed systems is found to be at least 10%. These two combinations yield a local event rate of 2.9×10^{-3} and 1.6×10^{-3} Mpc^{-3} Myr^{-1} for 60% and 10% fraction of dynamically formed mergers respectively.

The above rate calculations have only taken into account the observed rate, which ignores the effects of beaming at the source. In [118] and [119], the authors have estimated the beaming factor to be on the order of $f_b^{-1} \sim 100$, where f_b is the fraction of the total solid angle within which the GRB is emitted. Taking this into account, the beaming-corrected rates are found to be $2.9(f_b^{-1}/100) \times 10^{-1}$

and $1.6(f_b^{-1}/100) \times 10^{-1}$ Mpc^{-3} Myr^{-1} for 60% and 10% fraction of dynamically formed mergers respectively.

Chapter 5

Gravitational Wave Signal Processing

In this chapter, we describe the matched filtering that is done in order to search for a signal of known form within data that also contains noise, as is the case for analyzing GW data.

This problem has been well studied within the field of signal-processing, and the optimal tool for extracting the signal from the noise has been found to be the Weiner Filter [120], also known as the matched or optimal filter.

5.1 The Optimal Filter

Let us assume we have a data stream $s(t)$ that is the sum of a stationary, Gaussian noise $n(t)$ and a signal $h(t)$

$$s(t) = n(t) + h(t) \ . \tag{5.1}$$

Since $n(t)$ is stationary, the mean of $n(t)$ can be taken to zero by defining

$$n(t) \equiv n_{\mathrm{raw}}(t) - \langle n_{\mathrm{raw}} \rangle \ , \tag{5.2}$$

where the brackets denote the expectation value. Since this is true for any stationary series, from here we will assume $n(t)$ has a mean of zero. The two-sided power spectral density (PSD) $S_n(f)$ of this Gaussian noise is defined by

$$\langle \tilde{n}(f)\tilde{n}(f')\rangle \equiv \delta(f-f')S_n(f)\ . \tag{5.3}$$

Let us define a real-valued filter $F(t)$ as

$$A = \int\limits_{-\infty}^{\infty} F(t)a(t)dt = \int\limits_{-\infty}^{\infty} \tilde{F}^*(f)\tilde{a}(f)df\ , \tag{5.4}$$

where A is the filtered value of $a(t)$. Above we have used the convolution theorem and transformed from the time domain to the frequency domain using the Fourier transform of a function, defined by

$$\tilde{a}(f) = \int\limits_{-\infty}^{\infty} a(t)e^{-2\pi i f t}dt\ . \tag{5.5}$$

Using the above properties of the noise, we find

$$\begin{aligned} \langle N^2 \rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{F}^*(f)\tilde{F}^*(f')\,\langle \tilde{n}(f)\tilde{n}(f')\rangle df' df \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{F}^*(f)\tilde{F}^*(f')\delta(f-f')S_n(f)df' df \\ &= \int_{-\infty}^{\infty} \left|\tilde{F}(f)\right|^2 S_n(f)df\ , \end{aligned} \tag{5.6}$$

where we have also used the property of real-valued functions that $\tilde{a}^*(f) = \tilde{a}(-f)$ in the last line.

We now wish to find the form of the filter $F(t)$, or in this case $\tilde{F}(f)$, which will optimally extract the signal from the noise. We measure this by maximizing the ratio of filtered values

$$\begin{aligned} \frac{H^2}{\langle N^2 \rangle} &= \frac{\left|\int_{-\infty}^{\infty} \tilde{F}^*(f)\tilde{h}(f)df\right|^2}{\int_{-\infty}^{\infty}\left|\tilde{F}(f)\right|^2 S_n(f)df} \\ &= \frac{\left|\int_{-\infty}^{\infty} \tilde{F}^*(f)\sqrt{S_n(f)}\ \tilde{h}(f)/\sqrt{S_n(f)}df\right|^2}{\int_{-\infty}^{\infty}\left|\tilde{F}(f)\right|^2 S_n(f)df}\ . \end{aligned} \tag{5.7}$$

We can use the Cauchy-Schwarz inequality,

$$\left| \int_{-\infty}^{\infty} A(f)B(f)df \right|^2 \leq \int_{-\infty}^{\infty} |A(f)|^2 df \int_{-\infty}^{\infty} |B(f)|^2 df \ , \tag{5.8}$$

to argue that in order for the ratio of equation (5.7) to be maximized, the functions corresponding to $A(f)$ and $B(f)$ must be equal up to a constant C, where $A(f) = \tilde{F}^*(f)\sqrt{S_n(f)}$ and $B(f) = \tilde{h}(f)/\sqrt{S_n(f)}$. Equation (5.7) then becomes

$$\begin{aligned} \frac{H^2}{\langle N^2 \rangle} &= \frac{C\left(\int_{-\infty}^{\infty} \left|\tilde{F}(f)\right|^2 S_n(f)df\right)\left(\int_{-\infty}^{\infty} \left|\tilde{h}(f)\right|^2 / S_n(f)df\right)}{\int_{-\infty}^{\infty} \left|\tilde{F}(f)\right|^2 S_n(f)df} \\ &= C\int_{-\infty}^{\infty} \frac{\tilde{h}^*(f)\tilde{h}(f)}{S_n(f)} df \ . \end{aligned} \tag{5.9}$$

From this exercise we find that

$$\tilde{F}^*(f) = C\frac{\tilde{h}(f)}{S_n(f)} \ , \tag{5.10}$$

which is the definition of the optimal filter for $h(t)$ embedded in stationary, Gaussian noise.

Let us use the optimal filter to define an inner product of the data s with the template h:

$$(h|s) = \int_{-\infty}^{\infty} \frac{\tilde{h}^*(f)\tilde{s}(f)}{S_n(f)} df \ . \tag{5.11}$$

An interesting interpretation of equation (5.11) is to split the PSD $S_n(f)$ into two amplitude spectral density (ASD) terms $\sqrt{S_n(f)}$ that can be associated with the data and the template separately,

$$(h|s) = \int_{-\infty}^{\infty} \frac{\tilde{h}^*(f)}{\sqrt{S_n(f)}} \frac{\tilde{s}(f)}{\sqrt{S_n(f)}} df \ . \tag{5.12}$$

This ends up weighting both the template and the data by the inverse of the ASD, which can be seen as “whitening” both the template and the data. The reason this is known as whitening is because for stationary, white, Gaussian noise the PSD is frequency independent resulting in an optimal filter of the form

$$(h|s) = \int_{-\infty}^{\infty} \tilde{h}^*(f)\tilde{s}(f)df \ . \tag{5.13}$$

If we define $\bar{h}(f) = \frac{\tilde{h}^*(f)}{\sqrt{S_n(f)}}$ and $\bar{s}(f) = \frac{\tilde{s}^*(f)}{\sqrt{S_n(f)}}$, equation (5.12) comes into the same form as equation (5.13)

$$(h|s) = \int_{-\infty}^{\infty} \bar{h}^*(f)\bar{s}(f)df \ . \tag{5.14}$$

5.2 The Waveform Overlap

Now that we have an optimal filter for extracting a signal for noisy data, let us use it to define the overlap M of two vectors a and b

$$\mathrm{M} = (a'|b') \ , \tag{5.15}$$

where the "′" denotes a normalization such that

$$a' = \frac{a}{\sigma_a} \ , \tag{5.16a}$$

and

$$\sigma_a^2 = (a|a) \ , \tag{5.16b}$$

and similarly for b.

This definition of the overlap will take its largest value of 1 when a' and b' are the same function, which can be seen by

$$\begin{aligned} (a'|a') &= \left(\frac{a}{\sigma_a}|\frac{a}{\sigma_a}\right) \\ &= \frac{1}{\sigma_a^2}(a|a) \\ &= 1 \ . \end{aligned}$$

Since we are interested in the strength of the signal that matches the template h in the data s, we compute the signal-to-noise ratio (SNR) by normalizing the overlap with respect to the template but not the data

$$\mathrm{SNR} = \frac{1}{\sigma_h} (s|h) \ . \tag{5.17}$$

The value of the SNR will then be proportional to the amplitude of the signal buried in the noise.

So far we have defined the overlap between s and h where we have assumed the template to be of the same length as our data. However, what we are actually interested in is the matched filter of s using h where the data $s(t)$ is an extended time series whose length is longer than the template time series $h(t)$. The matched filter output as a time series is given by

$$(s(t)|h) = \int_{-\infty}^{\infty} \tilde{s}^* \tilde{h} e^{-2\pi i f t} df \ . \tag{5.18}$$

For real-valued time series s and h, this filter outputs a real-valued time series that, in the absence of signal, is χ^2distributed with one degree of freedom.

5.2.1 Matched Filtering for Compact Binary Signals

As derived in chapter 2, the GWs we are searching for come in two polarizations: the plus (+) and cross (×) polarizations. The actual signal seen by a detector is a combination of the two polarizations that can be calculated using the antenna pattern of the detector and the parameters of the source. Specifically, the signal seen by the detector $h(t)$ is given by

$$h(t) = F_+ h_+(t) + F_\times h_\times(t) \ , \tag{5.19}$$

where h_+ and $h_\times$ are the plus and cross polarizations of the signal respectively, and F_+ and $F_\times$ are the antenna patterns of the detector giving the sensitivity to the plus and cross polarizations of GW signals respectively. Descriptions of these are given in chapter 3.

Due to the nature of GW signals from the inspiral phase of CBCs (see chapter 2 for a discussion), we find that the phase evolution of the cross polarization is 90° out of phase with the plus polarization, given by

$$h_+(t) = A(t)\cos\left(\phi(t)\right) \ , \tag{5.20}$$

$$h_\times(t) = A(t)\sin\left(\phi(t)\right) \ , \tag{5.21}$$

where $A(t)$ is the amplitude evolution of the signal and $\phi(t)$ is the phase evolution of the signal. From this, we find a simple relation between the Fourier transform of the two polarizations:

$$\tilde{h}_+ = i\tilde{h}_\times \ . \tag{5.22}$$

When we filter $s = (Xh_+/\sigma_h) + (Yh_\times/\sigma_h)$ with the template h_+, we obtain the matched-filter real-time series z_+

$$z_+ = (s|h_+) = \frac{X}{\sigma_h}\left(h_+|h_+\right) + \frac{Y}{\sigma_h}\left(h_\times|h_+\right) \ . \tag{5.23}$$

Using equation (5.22), we find

$$\begin{aligned}
(h_\times|h_+) &= \int_{-\infty}^{\infty} \frac{\tilde{h}^*_\times h_+ + \tilde{h}^*_+ h_\times}{S_n} df \\
&= \int_{-\infty}^{\infty} \frac{\left(i\tilde{h}_+\right)^* h_+ + \tilde{h}^*_+ i h_+}{S_n} df \\
&= \int_{-\infty}^{\infty} \frac{-i\tilde{h}^*_+ h_+ + \tilde{h}^*_+ i h_+}{S_n} df \\
&= 0 \ ,
\end{aligned}$$

which then implies

$$z_+ = X\sigma_h \ , \tag{5.24a}$$

where we have used equation (5.16b) for the definition of σ_h. A similar procedure shows that when we filter s with the template $h_\times$, we obtain the matched filter time series $z_\times$

$$z_\times = Y\sigma_h \ . \tag{5.24b}$$

From equation (5.24) we construct our SNR ρ using the combination

$$\rho = \frac{1}{\sigma_h}\sqrt{|z_+|^2 + |z_\times|^2}\ . \tag{5.25}$$

This is referred to as a *two-phase filter*, which has twice the degrees of freedom of a single-phase filter. The bonus of extracting information from both polarizations of the GW comes with the cost of increasing the expectation value of ρ when there is no signal present.

By combining the definition of the SNR (equation (5.25)) with the template normalization σ_h (equation (5.16b)) we can define an effective distance for a given trigger

$$D_{\text{eff}} = \frac{\sigma_h}{\rho}\ . \tag{5.26}$$

We choose a normalization distance of 1 Mpc to set the templates' amplitudes. This results in the units of the effective distance being Mpc.

5.2.2 Template Bank Construction

The manifold of waveforms is a continuous space in the component masses, of which we are only able to search discrete points. In order to make sure we do not miss a signal because its parameters are slightly different from what we are searching for, we construct a bank of templates in such a way as to minimize the loss of a signal's SNR. This is done by computing the overlap of waveforms with different parameters using equation (5.15).

As noted above, the overlap of a waveform with itself is unity. Any *mismatch*, $1 - \mathrm{M}$, between two waveforms will show up as a reduction of the recovered SNR when searching for one waveform with the other. When constructing a template bank, it is useful to view this mismatch as a measure of the distance between two templates. With this in mind, we can create a metric that will tell us the distance between two templates of different parameters. This is done by defining the parameter space in which the metric exists, and then computing the mismatch between templates. For two templates infinitesimally far apart in the parameter space θ^μ, the mismatch is [121, 122]

$$
\begin{aligned}
1-\mathrm{M} &\approx 1-\left(h\left(\theta^{\mu}+d\theta^{\mu}\right)|h\left(\theta^{\nu}\right)\right) \\
&\approx \frac{1}{2}\left[1+1-2\left(h\left(\theta^{\mu}+d\theta^{\mu}\right)|h\left(\theta^{\nu}\right)\right)\right] \\
&\approx \frac{1}{2}[\left(h\left(\theta^{\mu}\right)|h\left(\theta^{\nu}\right)\right)+\left(h\left(\theta^{\mu}+d\theta^{\mu}\right)|h\left(\theta^{\nu}+d\theta^{\nu}\right)\right) \qquad (5.27)\\
&\quad -\left(h\left(\theta^{\mu}+d\theta^{\mu}\right)|h\left(\theta^{\nu}\right)\right)-\left(h\left(\theta^{\mu}\right)|h\left(\theta^{\nu}+d\theta^{\nu}\right)\right)] \\
&\approx \frac{1}{2}\left[\left(h\left(\theta^{\mu}+d\theta^{\mu}\right)-h\left(\theta^{\mu}\right)|h\left(\theta^{\nu}+d\theta^{\nu}\right)-h\left(\theta^{\nu}\right)\right)\right] \\
&\approx \frac{1}{2}\left[\left(\frac{\partial h}{\partial\theta^{\mu}}d\theta^{\mu}|\frac{\partial h}{\partial\theta^{\nu}}d\theta^{\nu}\right)\right] \\
&\approx \frac{1}{2}\left[\left(h_{\mu}|h_{\nu}\right)\right]d\theta^{\mu}d\theta^{\nu} \\
&\approx \frac{1}{2}g_{\mu\nu}d\theta^{\mu}d\theta^{\nu}\ , \qquad (5.28)
\end{aligned}
$$

where $h_\mu \equiv \frac{\partial h}{\partial\theta^\mu}$, and we have used the symmetry $(a|b) = (b|a)$. The above equation defines the metric $g_{\mu\nu}$ to be of the quadratic form

$$g_{\mu\nu} \equiv \left(h_\mu|h_\nu\right)\ . \qquad (5.29)$$

The metric (equation (5.29)) can be used to predict to quadratic order how quickly the mismatch grows as we move in different directions in parameter space away from a particular point. Using this metric we place templates such that the furthest distance any point is from a template is less than a tolerance value ϵ. This will ensure that no more than ϵ of the SNR is lost due to our discretization of parameter space.

The question we must now ask is which parameter space to use. Answering this question specifies which directions we will use when taking partial derivatives of the templates. In the end, it is useful to choose those directions such that the metric is as flat as possible across the space. This helps in developing an algorithm that most efficiently covers the parameter space. Such a space has been found (equations (23) and (24) in Ref. [123]) whose directions correspond to the *chirp times* of a waveform. Using 2.0 PN order waveforms for nonspinning binary objects in quasi-circular orbital evolution, there are chirp times τ_i for $i = 0, 2, 3, 4$, corresponding to the chirp times proportional to $(v/c))^i$ in the PN expansion of the phase. These chirp times are given in terms of the mass parameters as [124, 125, 122]

$$\begin{aligned}
\tau_0 &= \frac{5}{256\,(\pi f_0)^{8/3}}\mathcal{M}^{-5/3}\,, && \text{(5.30a)}\\
\tau_2 &= \frac{5}{192\eta^{2/5}\,(\pi f_0)^2}\left(\frac{743}{336}+\frac{11}{4}\eta\right)\mathcal{M}^{-1}\,, && \text{(5.30b)}\\
\tau_3 &= \frac{\pi}{8\eta^{3/5}\,(\pi f_0)^{5/3}}\mathcal{M}^{-2/3}\,, && \text{(5.30c)}\\
\tau_4 &= \frac{5}{128\eta^{4/5}\,(\pi f_0)^{4/3}}\left(\frac{3\,058\,673}{1\,016\,064}+\frac{5429}{1008}\eta+\frac{617}{144}\eta^2\right)\mathcal{M}^{-1/3}\,, && \text{(5.30d)}
\end{aligned}$$

where M is the total mass of the binary, $\eta = m_1 m_2/M^2$ is the symmetric mass ratio, $\mathcal{M} = \eta^{3/5}M$ is the chirp mass, and f_0 is some fiducial frequency. If f_0 is chosen to be the starting frequency for the waveform generation, then the chirp times correspond to the 0PN length of the chirp for τ_0, and the $(i/2)$PN correction to the chirp time for $i = 2, 3, 4$. For the search described in later chapters, we have chosen to use $(\tau_0,\ \tau_3)$ as the parameter space in which we lay our templates, however we compute a 3D metric that additionally includes the difference in coalescence times of two waveforms for use in other portions of the search. Although the metric is not perfectly flat for PN order greater than 1.0PN, we assume it is slowly varying, and thus flat in the local vicinity of a point. When constructing the bank, we use the hexagonal placement algorithm [123], which is the most efficient placement algorithm for a flat, 2D space.

5.2.3 Signal-Based Vetoes

Signal-based vetoes are discriminators that can be used to separate triggers arising from either a random instantiation of the Gaussian noise or transient excess power in the data from triggers arising from actual signals. These vetoes work by comparing how a particular aspect of the data should behave in the presence of a signal and not.

In the following sections we discuss several signal-based vetoes that look at different aspects of the data. The first two signal-based vetoes described below, the $\chi^2$and r^2vetoes, use information from a single detector. Since we require triggers to be coincident between multiple detectors, which we will discuss in more detail in chapter 6, there are addition signal-based vetoes we can perform using information from more than one detector. The effective distance cut and the amplitude consistency check are examples of this, using information from two coaligned detectors.

5.2.3.1 The χ^2Veto Calculation

The χ^2veto tests the consistency of the data with what we expect from a signal by looking at how the SNR, ρ, is accumulated from different parts a template. This is done by breaking up the template into p continuous bins such that each is expected to provide ρ/p to the SNR calculation in the presence of a real signal. This calculation is described by

$$\chi^2 = \sum_{i=1}^{p} \left(\rho_i - \frac{\rho}{p} \right)^2 , \tag{5.31}$$

where p is the number of χ^2bins, ρ_i is the SNR from the ith bin of the template, and ρ is the total SNR.

With this formulation, in the presence of Gaussian noise, $\chi^2$is a quantity that is χ^2distributed with $2p - 2$ degrees of freedom:

$$\left\langle \chi^2 \right\rangle_{\mathrm{N}} = 2p - 2 \; . \tag{5.32}$$

Normally such a quantity would be χ^2distributed with $p - 1$ degrees of freedom, however there is an extra factor of 2 because ρ^2 is calculated from a two-phase filter, which has an expectation value of $\left\langle \rho^2 \right\rangle = 2$ in Gaussian noise.

The $\chi^2$is not actually the quantity we threshold on when applying the χ^2veto. This is because the expectation value of χ^2in the presence of signal is not the same as in Gaussian noise. Since we are using a discrete template bank to search for signals in a continuous parameter space, we will not recover a signal with a template of those exact parameters. If we calculate the expectation value of χ^2in the presence of signal without noise, evaluated at the time when the SNR is maximized, we find

$$\left\langle\chi^2\right\rangle_{\mathrm{S}} = \delta\rho^2 \ , \tag{5.33}$$

where δ is the mismatch between the template and the signal. Since a real signal can have a large value of $\chi^2$for high SNRs, we end up thresholding on a normalized χ^2, ξ^2, where we veto triggers with

$$\xi^2 \equiv \frac{\chi^2}{p + \delta\rho^2} > \xi_*^2 \ . \tag{5.34}$$

For the χ^2veto, p, δ, and ξ_* are tunable parameters.

It should be noted that in order to do this calculation, p additional inverse Fourier Transforms of the data need to be computed for the p different parts of the template.

5.2.3.2 The r^2Veto Calculation

Another signal-based veto that has been developed is the r^2veto. This veto is also a measure of how much the data looks like an actual signal. The difference between the r^2veto and the χ^2veto is that the r^2veto looks over a stretch of data rather than at a single point in time. Specifically, the r^2veto measures the amount of time the χ^2time series spends above a particular threshold for the T seconds prior to an trigger. The duration r^2_{duration} is given as

$$r^2_{\mathrm{duration}} = \frac{d}{f_{\mathrm{sampling}}} \ , \tag{5.35}$$

where d is the number of data points in the time window $(t_0 - T, T)$ that have $\chi^2/p > r_*^2$, and f_{sampling} is the sampling rate of the data. For the r^2veto, T and r_*^2 are tunable parameters.

5.2.3.3 The Effective Distance Cut

When looking for coincident triggers between detectors that are coaligned, we apply an effective distance cut. Coaligned detectors, such as the two collocated Hanford detectors H1 and H2, have the interesting property of sharing the same antenna patterns. The effective distance seen by a particular detector for a given CBC GW signal is

$$D_{\mathrm{eff}} = \frac{D}{\sqrt{F_+^2 \left(1+\cos^2\iota\right)^2/4 + F_\times^2 \cos^2\iota}}\,, \tag{5.36}$$

which is only dependent on the distance to the source D, the inclination angle ι between the orbital plane and the line of site between the source and the detector, and the detector antenna patterns F_+ and $F_\times$. In terms of measured quantities, the effective distance is defined in equation (5.26). Since coaligned detectors have the same antenna patterns, they should see the same D_{eff} for triggers coming from a real CBC GW signal.

Since noise fluctuations in the different detectors can separately change the effective distance seen by each detector, we do not require the effective distances to agree perfectly. Instead we allow for some variation and veto coincident triggers that have a fractional effective distance difference κ greater than a given threshold κ_* (i.e., $\kappa > \kappa_*$). We compute κ as

$$\kappa = \frac{2\left|D_{\mathrm{eff,A}} - D_{\mathrm{eff,B}}\right|}{D_{\mathrm{eff,A}} + D_{\mathrm{eff,B}}}\,, \tag{5.37}$$

where $D_{\mathrm{eff,A}}$ is the effective distance from detector A, $D_{\mathrm{eff,B}}$ is the effective distance from detector B.

This veto is only applied between coaligned detectors due to the fact that the same signal can be seen by two nonaligned detectors with very different amplitudes, and thus effective distances, since the blind spots of the detectors' antenna patterns do not coincide. The blind spots are a very small portion of the sky and thus it is not unreasonable to apply a similar veto between nonaligned detectors. In that case, the fraction of true signals being vetoed can be calculated by looking at the volume of the universe in which such a signal would originate, bounded by the distance at which each detector would see a signal with an SNR above a particular threshold.

5.2.3.4 The Amplitude Consistency Check

In this section we describe the calculation of an amplitude consistency check. This check is very similar to the effective distance cut, however instead of checking the consistency between two triggers that make up a coincident trigger, we are checking whether the lack of a trigger from one coaligned detector is consistent with the measured parameters of the trigger in the other detector.

As always, we use the PSD from each interferometer to calculate a template normalization σ (equation (5.16b)) for a given time. With this, we compute a horizon distance for a canonical mass BNS template as the distance to which we would see an optimally located and oriented signal with an SNR of 8. Similar to the effective distance (equation (5.26)), this is given as

$$D_{\mathrm{horizon}} = \frac{\sigma}{8} \; . \tag{5.38}$$

We can rearrange the effective distance cut (equation (5.37)) using the effective distance (equation (5.26)) for the trigger from detector A and the horizon distance (equation (5.38)) from detector B. Solving for ρ_{A}, we find the veto to be

$$\rho_{\mathrm{A}} > \left(\frac{2+\kappa_*}{2-\kappa_*}\right)\left(\frac{D_{\mathrm{horizon,A}}}{D_{\mathrm{horizon,B}}}\right)\rho_{\mathrm{threshold,B}} \; . \tag{5.39}$$

In order to use this equation to veto triggers from detector A when there is no coincident trigger in detector B, we have substituted the SNR threshold of detector B $\rho_{\mathrm{threshold,B}}$ in the place of ρ_{B}. This veto tells us to veto triggers from detector A when the ratio of the horizon distances in detectors A and B is small enough such that there should have also been a trigger above the SNR threshold of detector B. The factor $(2+\kappa_*)\,/\,(2-\kappa_*)$ allows for noise fluctuations to change ρ_{A}, $D_{\mathrm{horizon,A}}$, and $D_{\mathrm{horizon,B}}$ by some amount.

This veto is also only applied between coaligned detectors due to the fact that the there exist locations in the sky and polarizations of a signal such that for two different, nonaligned detectors, one of them should see the signal above the SNR threshold, while the other should not. As in the case for the effective distance cut, a similar veto can be constructed for nonaligned detectors if one is willing to veto a fraction of true signals.

5.3 Previous Detection Statistic: Effective SNR

For the previous searches, a detection statistic was developed to separate signals from background. This statistic is the combined effective SNR and is constructed as follows.

The single-detector SNR ρ is produced by matched filtering the data against our templates. From ρ we define the *effective SNR*, ρ_{eff}, as

$$\rho_{\mathrm{eff}}^2 = \frac{\rho^2}{\sqrt{\left(\frac{\chi^2}{2p-2}\right)\left(1+\frac{\rho^2}{250}\right)}} \,, \tag{5.40}$$

where p is the number of bins used by the χ^2veto, $2p-2$ is the number of degrees of freedom of the χ^2veto in Gaussian noise, and the 250 is a tunable parameter that helps to further separate signals from background. This definition of the effective SNR reduces the apparent significance of nonGaussian instrumental artifacts since it weights the SNR by the χ^2. This effectively reduces the significance of outliers to the expected SNR distribution due to Gaussian noise (i.e., nonGaussian instrument artifacts) while minimally affecting the apparent significance triggers of real signals.

We then combine the effective SNRs for the single-detector triggers that form a coincident trigger into the *combined effective SNR*, ρ_{c}, for that coincident trigger using

$$\rho_{\mathrm{c}}^2 = \sum_{i=1}^{N} \rho_{\mathrm{eff},i}^2 \,. \tag{5.41}$$

Chapter 6

Data Analysis Pipeline and Tuning

In this chapter we describe the standard pipeline used to analyze data from GW detectors, such as the LIGO detectors, for GW signals from CBC. In addition, we will also describe the tuning that was done for the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data, referred to as "this search."

The name of the executable that generates the standard pipeline is the Hierarchical Inspiral Pipeline Executable (HIPE); and we colloquially refer to this portion of the pipeline as the HIPE pipeline. The goal of the HIPE pipeline is to distill GW data in the form of a time series of strain data into a collection of candidate GW triggers. The HIPE pipeline is generic enough that it is used for many different matched filtering searches including searches for GW signals from low mass CBC, triggered CBC searches (i.e., GRBs), high mass CBC, CBC whose component objects are spinning, and black hole ringdowns.

The HIPE executable, LALAPPS_INSPIRAL_HIPE, generates a directed acyclic graph (DAG) of computational jobs that are run by the Condor High Throughput Computing environment, where a DAG, as described by the name, is a collection of jobs with interdependencies such that there are no loops. HIPE reads an input file containing all of the arguments to invoke at different stages of the pipeline as well as the locations of files containing lists of times associated with individual detectors' operating status. The time files for each detector included for this search were times the detector was in "science mode" (i.e., when the detectors were operating nominally), which we call *analyzed time files*, and four sets of time files indicating increasingly severe excursions from nominal operation, which we call *category veto files*.

Using these inputs, HIPE first computes all of the times an individual detector was in science mode for 2064 continuous seconds. Because of the way our filtering is done, HIPE throws out the beginning and end 72 seconds in order to calculate the “analyzable times”. It then takes intersections of these times across different detectors to obtain different kinds of “coincident times” during which two or more detectors were in operation. Finally, HIPE sets up job dependencies across the different stages of the pipeline and produces the DAG to be run by Condor.

The different stages of the HIPE pipeline employed in this search were:

- template bank generation (detailed in section 6.3);
- first-stage single-detector trigger generation (detailed in section 6.4);
- first-stage coincident trigger generation (detailed in section 6.5);
- template bank regeneration (detailed in section 6.6)
- second-stage single-detector trigger generation (detailed in section 6.7);
- second-stage coincident trigger generation (detailed in section 6.8).

The template bank generation stage generates a bank of templates in some specified portion of the parameter space. Those templates are then used for the first-stage single-detector trigger generation by performing matched filtering of the templates against the data, producing a list of single-detector triggers. The first-stage single-detector triggers from different detectors are then combined and used to search for coincident triggers in the first-stage coincident trigger generation by identifying triggers that are coincident across detectors in parameter space, including time. The first-stage coincident triggers are converted back into single-detector template banks for a second round of matched filtering in the template bank regeneration stage. This new template bank is used for the second-stage single-detector trigger generation by matched filtering the data against the templates again, this time also calculating and vetoing triggers based on signal-based vetoes. Finally, the second-stage single-detector triggers are used for second-stage coincident trigger generation, where triggers occurring during times in the category veto files are thrown out, coincident triggers are found, and more signal-based vetoes are calculated and used. The output of this stage of the pipeline is the final output of the HIPE pipeline. The pipeline flow is represented by figure 6.1.

Multiple stages of matched filtering and coincidence occur because the calculation of the signal-based vetoes in the second-stage single-detector trigger generation is very compute intensive. In order to speed the search up, we only compute these signal-based vetoes on triggers that have already passed coincidence.

The follow sections will go through each of these stages in detail, but first let us describe the blinding and tuning procedure we use and how we test our pipelines using software injections.

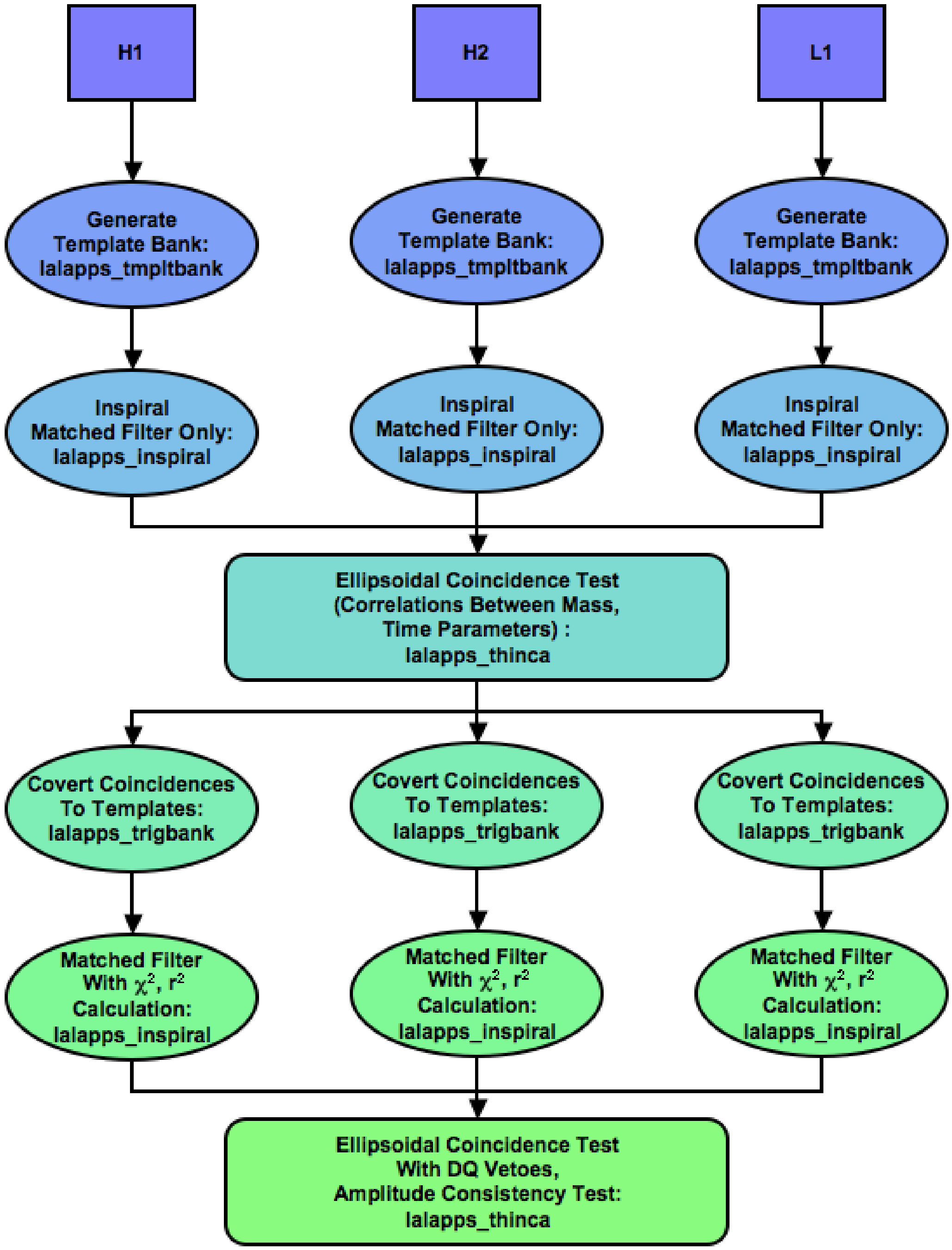

Figure 6.1: The HIPE Pipeline
A flowchart showing how GW detector data are analyzed in the HIPE pipeline.

6.1 Blinding the Analysis Pipeline

One of the goals of searching for GWs is to give a statistical probability that the triggers seen are real signals or due to background. This requires that the search pipeline's tuning not be informed by the *in-time* triggers that are present in the data. This can be achieved by *blinding* the data during the tuning process.

To do this, we only look at coincident triggers associated with 1) *time-shifted data* in which the detector data streams streams have been slid against each other such that no true GW signal could produce coincident triggers in that time shift (this data estimates the background of the pipeline), 2) *software injected data* in which GWs have been artificially introduced into the data in software (we use these to measure the efficiency of our pipeline at detecting real signals), and 3) *playground in-time data* that we have defined to be 600 of every 6370 seconds of data, which we use to check our pipeline to make sure nothing has gone awry in our analysis (we search for GW signals in this data but exclude it from statistical statements resulting from the search). Using these triggers, we tune the parameters of the pipeline and construct a detection statistic to rank the importance of the coincident triggers such that the time-shift coincident triggers are suppressed while the coincident triggers associated with software injections are emphasized.

Only at this point do we look at the in-time coincident triggers ("open the box"), unblinding the analysis to the part of the data in which putative GW detections can be found. In this way we avoid any bias in establishing the detection statistic that can come from tuning the search parameters after candidates are examined, which could lead to subjectively elevating or "killing" such candidates. However, we choose to apply an unavoidably subjective follow-up procedure (described in chapter 10) to develop sufficient confidence in a potential detection candidate to warrant the announcement of the first discovery of GWs.

6.2 Simulated Waveform Injections

In order to measure the efficiency of our pipeline to recovering GW signals from CBC, we inject several different PN families of waveforms into the data and check to see the fraction of signals that are recovered. The different waveform families used for injections in this search include GeneratePPN computed to Newtonian order in amplitude and 2PN order in phase using formulae from reference [3], EOB computed to Newtonian order in amplitude and 3PN order in phase using formulae from references [4, 5, 6, 7], PadéT1 computed to Newtonian order in amplitude and 3.5PN order in phase using formulae from references [8, 9], and SpinTaylor computed to Newtonian order in amplitude and 3.5PN order in phase using formulae from references [10] and based upon references [11, 12, 3, 13, 14, 15, 8, 16]; using code from reference [126]. Each of these families except for SpinTaylor ignores the effects of spin on the orbital evolution.

The parameters of the signals injected for each of the families are chosen from either an astrophysical distribution or a nonastrophysical distribution. The parameters that are astrophysically distributed include:

- the sky location (right ascension, declination) is distributed uniformly on the surface of a sphere, since we expect no preferred direction in the sky for extragalactic GW sources;
- the inclination angle (ι) where $\cos\iota$ is distributed uniformly between 0 and 1, since there is no preferred orientation of sources with respect to the observer's line of sight;
- the polarization azimuthal angle (ψ) distributed uniformly between 0 and 2π, since there is no preferred orientation of sources with respect to the observer's orientation;
- (for the SpinTaylor waveform family) the component objects' spin orientations relative to the initial orbital angular momentum are individually distributed uniformly on the surface of a sphere. No correlation is expected amongst the three orientations of the component objects' spin orientations and the orientation of the orbital angular momentum for captured binaries. However, there may be correlations for primordial binaries.

The parameters that are nonastrophysically distributed include:

- the total mass is distributed uniformly between 2 and 35 $M_{\odot}$, since there is no reliable theory for predicting the distribution in total mass, and we wish to measure the CBC rate for a range of masses;

- the mass ratio $q = m_1/m_2$ is distributed uniformly between 1 and the maximum value such that $m_1, m_2 > 1M_{\odot}$, since there is no reliable theory for predicting the distribution in mass ratio;

- the physical distance D is distributed uniformly in $\log_{10} D$, since we want to test our pipeline on a large range of signal amplitudes;

- (for the SpinTaylor waveform family) the component objects' unitless spin magnitudes $\hat{a} \equiv (cS) / \left(Gm^2\right)$ are individually distributed uniformly from 0 to 1, since there is no consensus on the physical distribution of the component objects' spin magnitudes nor the correlations between the component objects' spin magnitudes.

6.3 Template Bank Generation

The search we are performing for GW signals from CBC is a matched filter search where we compare the data against a model signal waveform to see how much the data looks like the model. In this case the model, which we call a "template," is a gravitational waveform calculated from PN theory for a CBC of a certain mass (as discussed in chapter 2). In section 5.2.2 we discussed theoretically the generation of a template bank with the use of a metric on the parameter space. In this section we will go over the specific choices that were made in the construction of template banks used in this search.

Based on the goals of the search, we know the masses of the templates should cover the range of total mass from 2 to 35 $M_{\odot}$with a minimum component mass of 1 $M_{\odot}$. A factor that governs the total number of template is that we are willing to lose at most 10% of the signals. A mismatch due to the discreteness of the template bank reduces the SNR we can recover, and thus the distance to which we can see, by a factor of the mismatch. Assuming an signal distribution uniform in volume, this translates the 10% tolerance in signal recovery to a match of $(90\%)^{1/3} = 97\%$ or a 3% mismatch.

The next choice we make is which waveform approximation is used to generate the templates. In interests of speed and simplicity of code, we choose to use 2.0 restricted PN templates, which are second PN order in the phase evolution and Newtonian order in the amplitude, generated in the frequency domain using the SPA. These choices are validated by studies that show these templates can recover time domain signals of different orders and different approximant families [127, 124, 128].

The final items we must choose in order to generate a template bank are the frequency range of our calculations and the calculation of the PSD of the data, S_n. The LIGO noise varies over many orders of magnitude for the frequency range we are interested in. The lowest mass signals we are searching for extend from arbitrarily low frequencies up to a few kHz. Because of the steep slope of the seismic noise in the detectors starting at 40Hz and below, we choose to suppress the data below 30Hz using two rounds of eighth-order Butterworth high-pass filters, and then analyze data above 40 Hz. This is shown in the spectra of the data after the different stages of filtering in figure 6.2. The gap between where our high-pass filters end and where we start analyzing data is to ensure we are not suppressing the signals in the band we analyze.

To choose our high frequency cutoff, we look at the frequencies our signals end on, a few kHz and below, and we choose 2048 Hz as our high frequency cutoff. In order to reduce the amount of data we work with, we down-sample the data from 16384 samples per second to 4096 samples per second, applying low pass filters above the Nyquist frequency of 2048 Hz in order to prevent aliasing of power from frequencies above this to low frequencies.

We calculate S_n using a 2048-second-long segment of data. This is broken up into 15 overlapping chunks of 256 seconds, which we Fourier transform to calculate 15 PSDs. S_n is then constructed by taking the median of these in each frequency bin. This PSD is used in the metric calculation of the chosen parameter space and, later on, in our matched filtering.

With the above procedures, we end up with $\sim$5000 templates covering the parameter space we are interested in. An example template bank can be seen in figure 6.3. The essentially flat metric in $(\tau_0,\ \tau_3)$ (discussed in chapter 5) produces an essentially uniform hexagonal spacing of templates

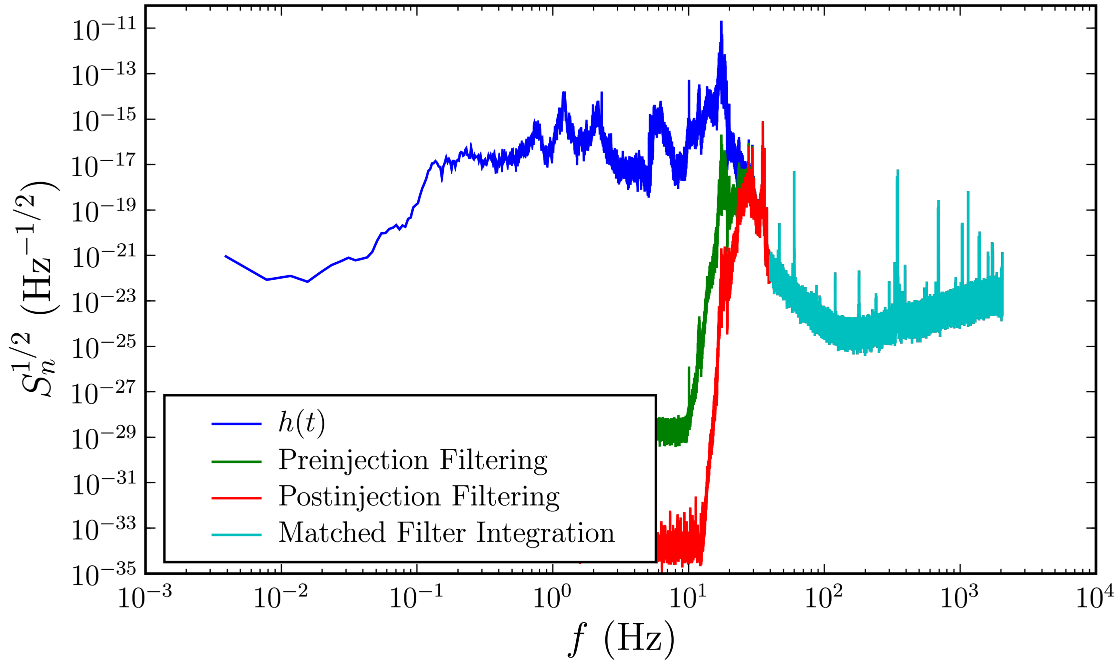

Figure 6.2: Data Preprocessing
Data spectra after the different levels of filtering we perform prior to matched filtering.

in that space, and a nonuniform spacing in $(m_1,\ m_2)$ space. Conversely, the straight bounds of the bank in $(m_1,\ m_2)$ space results in current boundaries in $(\tau_0,\ \tau_3)$ space.

6.4 First-Stage Single-Detector Trigger Generation

The first-stage single-detector trigger generation is the first round of matched filtering we perform on the data. In this stage we perform the same data conditioning as described in the template bank section with an addition step of injecting signals to the data when performing software injections. In that case, after the first round of high-pass filtering, the injection waveforms are added to the data, as noted in figure 6.2. The data conditioning continues after this as before with the second round of high-pass filtering and then resampling.

Once we get to this stage, we perform the matched filtering as described in chapter 5 in order to obtain an SNR time series for each template. Since the templates we are filtering with have a specific duration, there will be wrap-around effects near the boundaries of each chunk where the

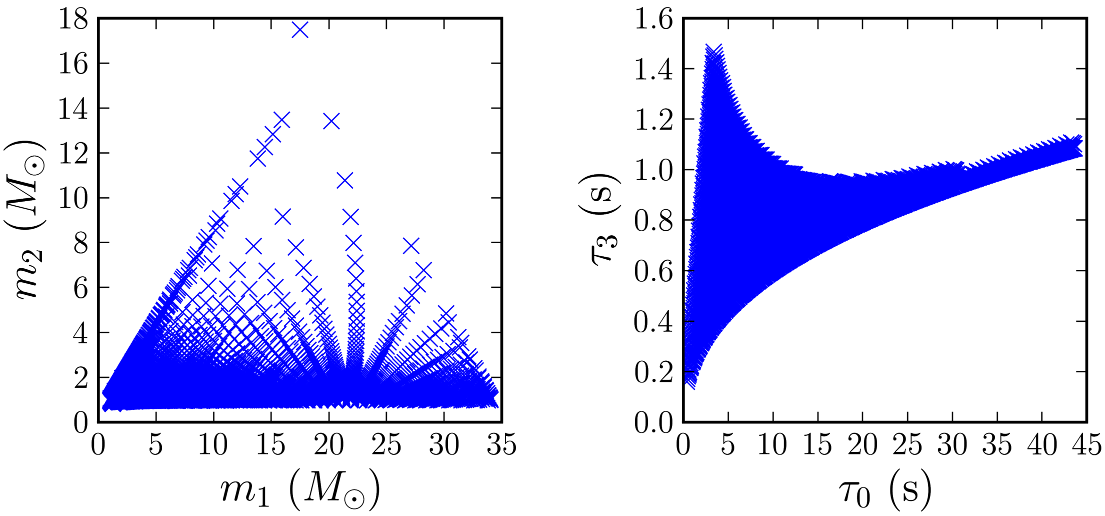

Figure 6.3: Sample Template Bank
A sample template bank is shown in both (m_1, m_2) space and (τ_0, τ_3) space.

SNR time series will be corrupted. We set the 256-second chunks to overlap by 128 seconds on each end, enabling us to throw away triggers within 64 seconds of each boundary containing the corrupted portion of the time series.

The SNR time series is then used to generate single-detector triggers by finding times when the SNR exceeds a threshold value of 5.5 for this search. This may happen for many adjacent points in time, so to reduce the amount of triggers we generate, we cluster the time series over the length of the template $t_{\rm tmplt}$ using a sliding window. This is implemented in the following way:

1. start at the beginning of the time series;
2. find the next time the SNR crosses the threshold;
3. generate a trigger, which records that time and SNR ρ^*;
4. search the time series over the next $t_{\rm tmplt}$ seconds for the first time when $\rho > \rho^*$;
5. if there exists such a time, update the trigger to contain the new time and SNR, then repeat step 4;
6. if there does not exist such a time, store the trigger in a trigger list and, if there is additional

time in the time series, repeat step 2.

After this procedure, we have a list of single-detector triggers from a single template. This process is repeated for each template in the bank. Once we are done filtering the data through all of the templates, we perform an additional round of clustering before we write these triggers to file and end this stage. In this final round of clustering, we cluster triggers coming from different templates when they satisfy some requirements. As we have noted before, adjacent templates have a match close to 1 in order for a signal falling between templates to have a mismatch with the nearest template of less than 3%. This results in templates near a particular template that produces triggers having a good chance of also crossing SNR threshold and producing triggers.

An algorithm called "trigscan" was devised that clusters these triggers across the template bank using the metric calculated during the template bank generation [129]. In generating the template bank, the portions of the metric that were used are those that ignored the time dimension of parameter space, as mentioned in chapter 5. In contrast, trigscan employs the full 3D metric to do the clustering. Trigscan works using the following algorithm:

1. sort the trigger list to be time ordered and start with the first trigger and use it to seed a cluster of one trigger;

2. create an error ellipse of size $\epsilon_{\text{trigscan}}$ around the trigger using the metric at that point in $(\tau_0,\ \tau_3,\ t)$ space;

3. for all triggers with a time $t \in (t_i - T, t_i + T)$, where t_i is the time of the current trigger and T is some time window, create error ellipses of size $\epsilon_{\text{trigscan}}$ at those points using their values for the metric;

4. for any triggers whose error ellipse overlaps the current trigger's, add them to the current cluster;

5. repeat steps 2–4 for any trigger added to the cluster;

6. when no additional triggers can be added to the cluster, save the trigger with the largest SNR

in the clustered trigger list;

7. if there is a next trigger in time, use it to create the next cluster of one trigger and repeating steps 2–5

We decided to use a trigscan metric scaling value of $\epsilon_{\text{trigscan}} = 0.06$. Ignoring the time dimension for a moment, let us assume we are in flat space with hexagonally spaced templates where the furthest distance any point can be from a template is 0.03. In this case, the templates have a separation of $2 \times 0.03 \cos(\pi/6) \approx 0.052$. If we are to create circles around each template and expand them to radius 0.06, templates slightly more than twice the templates separation distance will have circles that intersect. This can be seen pictorially in figure 6.4. Since the $(\tau_0,\ \tau_3)$ space in which we have chosen to place are templates is almost flat, our chosen value of $\epsilon_{\text{trigscan}} = 0.06$ allows triggers occurring at the same time (i.e., ignoring the time dimension of the 3D space) to be clustered when they are from next-to-nearest-neighbor templates in the template bank.

The trigscan-clustered trigger list is then saved to an xml file where each trigger contains information such as the GPS time it occurred, the SNR (ρ), the effective distance (D_{eff}), the phase at coalescence (ϕ_0), the mass parameters ($m_1,\ m_2,\ M,\ \eta,\ \mathcal{M}$), the chirp time parameters ($\tau_0,\ \tau_2,\ \tau_3,\ \tau_4,\ \tau_5$), the length of the template (t_{tmplt}), and the six metric components of the 3D metric at that point. Additional information associated with the time analyzed that is also calculated and saved to file is the horizon distance (equation (5.38)) for the detector using S_n for that 2048-second segment.

6.5 First-Stage Coincident Trigger Generation

A true GW signal from a CBC should affect all detectors around the world, and indeed, to verify a GW detection, it will have to be found by multiple means (i.e., associated electromagnetic, neutrino, or GW observation). For this search, we are only considering data from the LIGO detectors, so we rely on finding coincident GW triggers from different detectors.

Additional information useful in evaluating the performance of the pipeline is the background rate of triggers. The background is assumed to be due to accidental coincidences, which can be

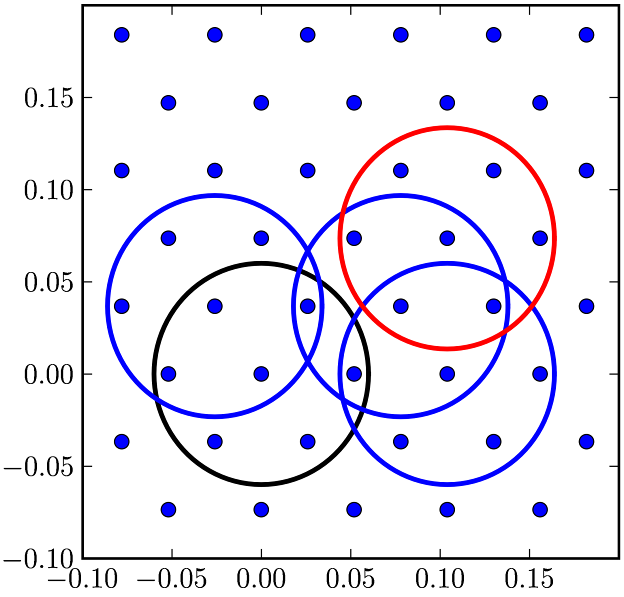

Figure 6.4: Trigscan Example
Templates are plotted as blue dots, which are spaced such that the furthest distance any point can be from a template is 0.03. We draw a black circle of radius 0.06 around a chosen template. Blue circles are drawn around three example templates that can be trigscan clustered with the chosen template (i.e., the blue circles intersect the black circle). A red circle is drawn around a single template that is slightly too far to be clustered with the chosen template.

estimated by time shifts (described in section 6.1). This estimation fails for collocated detectors, such as H1 and H2, due to time correlated background arising from the shared environment and crosstalk between the detectors (e.g., stray optical signals or electric control signals). Because of this, we can not reliably estimate the background in this way, thus H1H2 coincident triggers are excluded from this search. As these effects should only occur between collocated detectors, there is no in-time correlation between the Hanford and Livingston detectors.

Below, section 6.5.1 describes how the coincidence algorithm is executed, while section 6.5.2 describes how the time shifts are performed.

6.5.1 Coincidence Algorithm

The coincidence algorithm that has been developed for finding coincident triggers between multiple detectors is called The Hierarchical Inspiral Coicidence Algorithm (THINCA) and we commonly refer to this stage of the pipeline as the first THINCA stage. Here we take as input triggers from up to 3600 seconds of time when a unique set of detectors was generating triggers. Potential coincident triggers could lie across the boundary of two segments and thus not be recorded, however we find a negligible loss in efficiency due to this effect. The coincidence test used for this search in THINCA is similar to the trigscan clustering algorithm described above, called the “ellipsoidal” THINCA or “e-thinca” coincidence test. The main difference between the two is that trigscan builds up *clusters* of triggers from a *single detector*, while the e-thinca test checks whether triggers from *two* detectors are coincident *without agglomerating* coincident triggers into clusters.

First we will describe the e-thinca coincidence test before moving on to the full coincidence algorithm. The e-thinca test for two triggers from different detectors is as follows

1. create an error ellipse of size $\epsilon_{\mathrm{e-thinca}}$ around the triggers using their respective metrics;
2. if their error ellipses overlap, return yes, otherwise return no.

With the coincidence test defined, we now describe the full coincidence algorithm. It is as follows:

1. create a time sorted trigger list from all the triggers of the multiple detectors, and start with the first trigger;
2. for all triggers with a time $t \in (t_i, t_i + T)$, where t_i is the current trigger and T is some time window. If the trigger is from a different detector, perform the coincidence test between the two triggers;
3. if coincidence test return yes, create a coincident trigger in a coincident triggers list and link both single-detectors triggers to this coincident trigger;
4. repeat 2-3 for the next trigger in the list until end of list;
5. loop over coincident triggers, if all the single-detector triggers in the current coincidence are coincident with an additional trigger from a different detector, add that trigger to this coincidence;

6. loop over coincident triggers and remove coincident triggers that are subsets of higher-order coincident triggers.

After this procedure, the list of coincident triggers is saved to an xml file where the single-detector triggers of a coincidence all have the same event_id as well as all the information saved during the single-detector trigger generation.

6.5.2 Background Estimation: Time Shifts

As mentioned previously, time shifts are performed between different detectors in order to estimate the amount of background triggers we expect from the pipeline. These time shifts are chosen such that any coincidences coming from the time shifted data could not have come from a real GW signal. To ensure this, we have chosen to create a vector that we multiply by a shift number in order determine the amount of time to shift each detector. The vector contains a unit time shift for each detector, which must be unique.

For this search we have chosen the vector

$$v = \{0, 10, 5\} \ , \tag{6.1}$$

which we associate with the detectors H1, H2, and L1 respectively. This is used for 100 different time shifts ranging in shift number from -50 to 50, skipping 0. The maximum number of time shifts we could do with these shift vectors is 360, however we have found 100 to be adequate to estimate the false alarm rate for the loudest in-time triggers in the past. Additional time shifts could be performed with smaller shift vectors, however there is a time correlation for triggers coming from real signals that must be avoided.

As we said before, THINCA takes in as input up to 3600 seconds worth of triggers from a unique combination of multiple detectors. Using the above information, we find that when we shift triggers from different detectors against each other, we can end up with up to 2000 seconds where there are no H1 triggers to search for coincidence with H2 triggers. In order to avoid this but keep with the

3600-second limit, we have chosen to wrap the triggers on a 3600-second ring defined by the original start and end times of the coincident segment. In this approach, when a trigger's shift time becomes greater than the end time of the ring, we subtract the duration of the ring from the trigger's time, which moves it to the beginning of the ring. Conversely, when a trigger's shifted time becomes less than the start time of the ring, we add the duration of the ring to the trigger's time, which moves it to the end of the ring.

Once we have shifted the triggers in time, we can apply the above coincidence algorithm to produce coincident triggers for the individual time shifts. At the end of the algorithm, before we save the coincident triggers to file, we unshift the triggers' times so that they return to their original times.

There a several nice features of this approach. The first is that it keeps the amount of analyzed time in each time shift the same as the in time. In addition, this approach limits the maximum time offset between time-shifted coincident triggers to 3600 seconds. Since we expect the detectors' behaviors to be roughly stationary over this time period, this more accurately estimates the background for the in-time coincident triggers.

As discussed in section 6.1, there are noise correlations between H1 and H2 that bias the background estimation for H1H2 coincidences, which we find to also have an effect H1H2L1 coincidences. We have tested using a background estimation for H1H2L1 coincidences in which we shift H1 and H2 triggers by the same amount and found this to correct this bias. As this was done late in the analysis, it was not folded into the final results, however it has become a standard procedure for the other searches we are performing in the S5 data.

6.5.3 Coincidence Tuning

The threshold in the algorithm that controls the maximum e-thinca value ϵ we accept for coincidence we refer to as the e-thinca threshold ϵ_*. Larger values of ϵ_* correspond to accepting larger error ellipses for determining coincidence. We tuned the e-thinca threshold by looking at time-shifted coincident triggers and triggers associated with software injections. Figure 6.5 shows the fraction

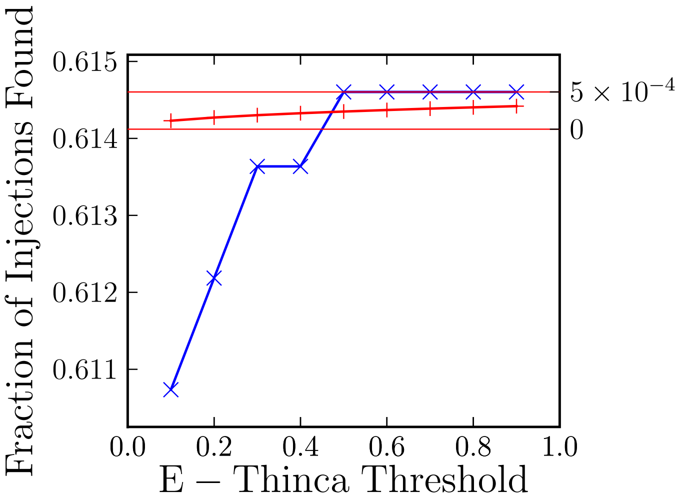

Figure 6.5: E-Thinca Tuning
The percentage of injections recovered (blue and left y-axis) and the expected percentage of injections recovered due to background coincidences (red and right y-axis) as a function of the e-thinca threshold. We have offset the blue and red curves for ease of comparison. The horizontal red lines show the efficiency gain per additional found injection. We see that over the whole range of e-thinca thresholds, there is less than one expected injection due to the background.

of injections recovered as a function of the e-thinca threshold. There we plot both the fraction of injections recovered and the expected fraction of injections whose time windows overlap time-shifted coincidences F_{false}, which is given as

$$F_{\mathrm{false}} = \frac{N_{\mathrm{time-shifts}} T_{\mathrm{window}}}{N_{\mathrm{inj}} T_{\mathrm{time-shifts}}} \,, \tag{6.2}$$

where N_{inj} is the number of injections made, $T_{\mathrm{inj}} = N_{\mathrm{inj}} t_{\mathrm{window}}$ is the total amount of time we were searching for injections, t_{window} is the window used to determine whether or not an injection is recovered, $N_{\mathrm{time-shifts}}$ is the number of time-shift coincidences for a given e-thinca threshold, and $T_{\mathrm{time-shifts}}$ is the total amount of time analyzed for time-shifted coincidences.

Figure 6.5 shows that for injections into H1H2L1 time, we find that the last additional injection is found at an $\epsilon_* = 0.5$ and that there is less than one expected injection due to the background at that value. Because of this, we decided to use $\epsilon_* = 0.5$ as the threshold for this search.

6.6 Trigger to Template

In this stage of the pipeline, we create a new template bank from triggers found in coincidence between detectors. This is done in the following manner:

1. use the coincident segments that a particular 2048-second single-detector segment participated in;
2. keep triggers from only that detector;
3. search through triggers keeping only one trigger for each combination of mass parameters;

Once this is done, we save the remaining triggers to an xml file for use as the template bank in the second-stage single-detector trigger generation.

One thing to note about this stage is that triggers from different single-detector segments can wind up as templates for nearby single-detector segments. This is due to the fact that coincidence segments can take in more than one single-detector segment from a given detector. But are we not trying to reduce the number of templates for the second, computationally expensive stage of matched filtering whereas this effect is causing an *increase* in the template bank? Even though this is the case, the increase in the template bank due to this effect is much smaller than the reduction we are trying to accomplish.

6.7 Second-Stage Single-Detector Trigger Generation

Here we describe the second-stage single-detector trigger generation. As noted above, for this stage we use the triggers found in coincidence with other detectors as the template bank for the second round of matched filtering. This requirement significantly reduces the number of templates with which we perform this stage of matched filtering. This reduction is necessary because in addition to performing the standard matched filtering we have described above, we also calculate two signal-based vetoes that test the consistency of the data to what we expect from a signal, one of which is computationally intensive. This veto requires an additional p inverse Fourier transforms, which causes this stage to be the time-limiting stage of the pipeline.

However, there are also drawbacks to this two-stage approach. Adding in the second stage of the pipeline adds significant complexity to the pipeline. In addition, coincident triggers found after the first stage of coincidence may not be the same as after the second stage of coincidence. This adds ambiguity when trying to track specific triggers through the pipeline.

Before the veto calculations begin, and after they are finished, the same algorithms are used as in the first-stage single-detector trigger generation described in section 6.4. Below, in section 6.7.1, we describe how we tuned the $\chi^2$and r^2signal-based vetoes (described in sections 5.2.3.1 and 5.2.3.2 respectively).

6.7.1 Signal-Based Veto Tuning

As in the e-thinca threshold, the veto parameters for both the χ^2veto and r^2veto are tuned by observing how they respond to software injection triggers and how they respond to time-shifted coincident triggers. The goal of the vetoes is to find an area of parameter space where there is a separation of time-shifted coincident triggers from software injection triggers that can be removed, thus reducing the chance of a background trigger occurring while not affecting the chance of recovering a true signal.

The χ^2statistic, as described in section 5.2.3.1, has an expected behavior for software injection triggers that grows proportional to ρ^2, where ρ is the SNR. This can be seen in figure 6.6. We are trying to tune this search such that triggers associated with signals whose waveforms differ slightly from the templates are still recovered. The most extreme divergence from the template waveforms we hope to recover come from CBC waveforms whose component objects have significant spin angular momentum, as discussed in section 6.2. By injecting these waveforms in as software injections, we find that by loosening the χ^2veto parameters as compared to other injection waveforms, we are still able recover these spinning signals. This is shown in figure 6.7. The parameters for the χ^2veto that were chosen for this search were a mismatch parameter of $\delta = 0.2$ and a threshold of $\xi^2 = 10.0$.

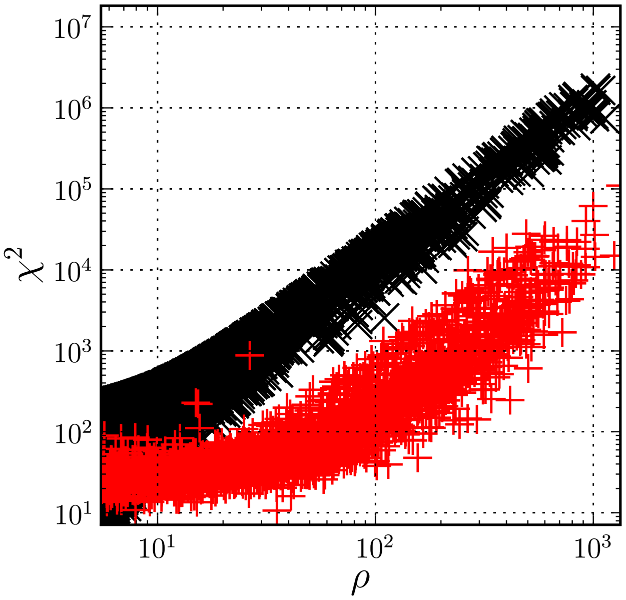

Figure 6.6: χ^2Tuning
χ^2values versus SNR for triggers associated with nonspinning software injections (red) and time-shifted coincident triggers (black). We can clearly see the ρ^2 dependence of the χ^2value for both categories of triggers.

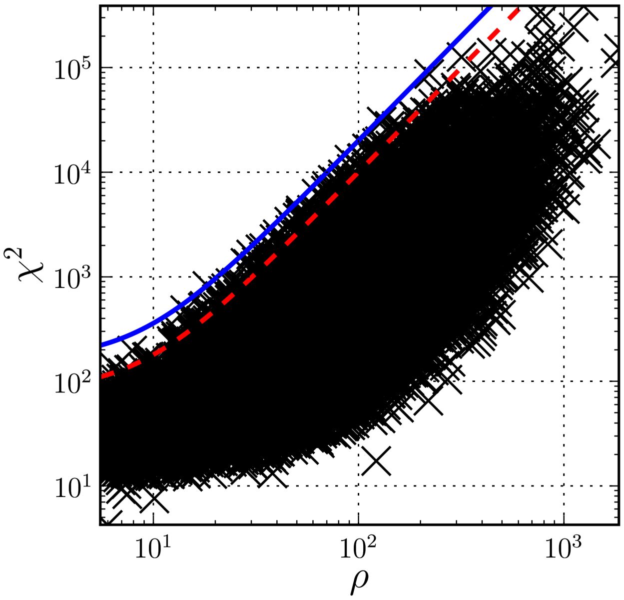

Figure 6.7: χ^2Threshold
χ^2versus SNR for software injections that include spin (black). The old χ^2threshold is plotted in red (tuned with nonspinning injections) along with the chosen χ^2threshold in blue (tuned with spinning injections).

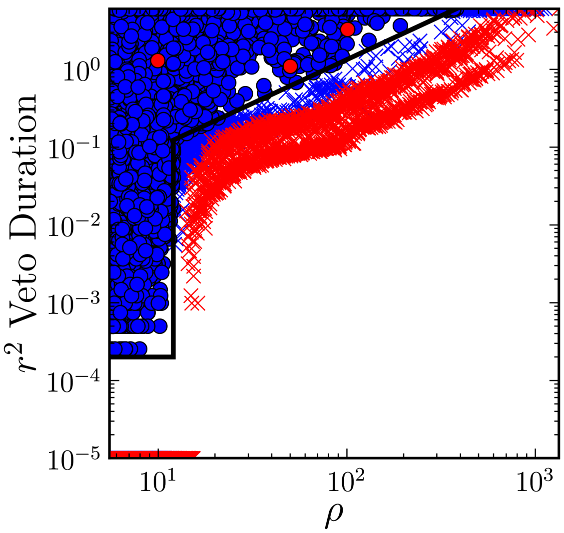

Figure 6.8: r^2Tuning

r^2veto duration versus SNR for software injection triggers (red) and time-shifted coincident triggers (blue). The black line denotes the r^2veto that was chosen for this search.

The r^2veto (described in section 5.2.3.2) also is expected to have a dependence on ρ because it derived from the χ^2statistic. Since there is a threshold below which the χ^2can fluctuate without contributing to the r^2_{duration} calculation, we expect the r^2_{duration} to be zero below a certain value of $\rho = \rho_*$. Above that value of ρ, since we have not modeled the r^2_{duration} dependence on ρ, we choose to veto triggers with an $r^2_{\text{duration}} > A \times \rho^B$, where A and B are tuned in this search. For this search we have chosen $T = 6$ sec. and $r^2_* = 15.0$. With these values fixed, we have tuned ρ_*, A, and B using figures such as figure 6.8. Using this information, we have chosen $\rho_* = 12$, below which we veto triggers with $r^2_{\text{duration}} > 0$, and above which we veto triggers with $r^2_{\text{duration}} > A \times \rho^B$. We have also chosen $A = 7.5 \times 10^{-3}$ and $B = 1.12764$. All of these choices are loose enough such that the leave some space for signals to move due to noise fluctuations or waveform uncertainties without being vetoed.

In figure 6.8 there are several interesting features that require explanation. The main feature that jumps out is the bimodal distribution of software injection triggers. We have found a clean

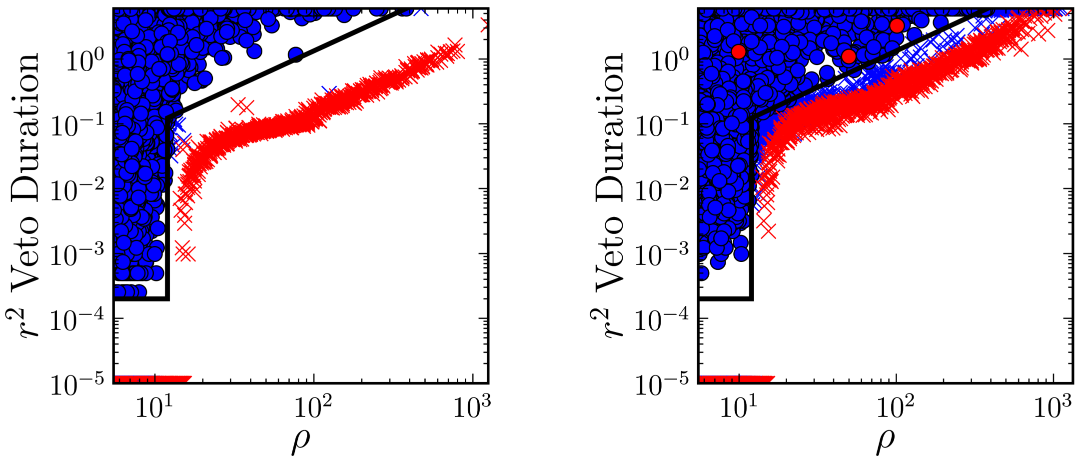

Figure 6.9: Different r^2Veto Populations
r^2veto duration versus SNR for software injections with total mass less than 20 $M_\odot$ (left) and greater than 20 $M_\odot$ (right).

break between to two distributions where the distribution with lower r^2_{duration} is from software injections with $M < 20M_\odot$ while the distribution with higher r^2_{duration} is from software injections with $M > 20M_\odot$, where M is the total mass of the system, as can be seen in figure 6.9. The total mass is used to define the frequency at which to stop the waveform calculation since the PN approximation breaks down at the frequency of the ISCO. From equation (2.38), the ending frequency f_{ISCO} corresponding to $M = 20M_\odot$ is $\sim 220Hz$, which is roughly when the f_{ISCO} is in the "sweet spot" of the detection band (i.e., where the noise is lowest).

Another feature in figure 6.8 is a handful of injections above the line denoting the veto. Closer inspection of these triggers showed there was a glitch several seconds before the injections. An example is seen in figure 6.10 showing the r^2time series for the same data stretch with and without the injection present. This figure also shows how there is nonzero r^2_{duration} expected for injections above a certain ρ.

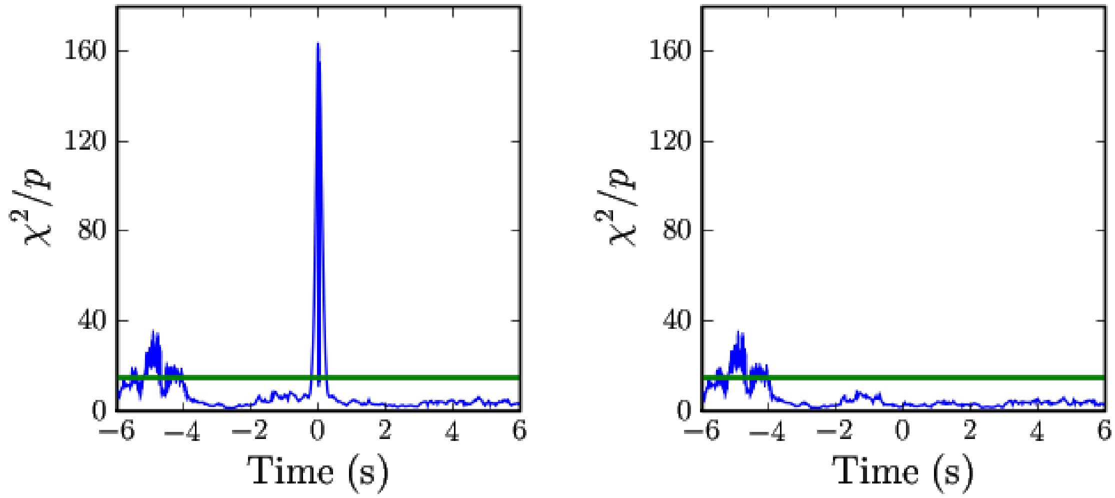

Figure 6.10: r^2Sample Time Series
The r^2time series for the same example stretch of data with (left) and without (right) a software injection present. The green line denotes the threshold value above which we count time samples in calculating the r^2veto duration. Comparison of the two plots reveal features that are due to the injection and present without the injection.

6.8 Second-Stage Coincident Trigger Generation

This is the final stage of the HIPE pipeline. We perform the second-stage coincident trigger generation using the same algorithm as in the first stage (described in section 6.5) with the addition of several vetoes. These different vetoes include vetoing times the instruments were not behaving nominally, what we call Data Quality (DQ) vetoes, as well as some signal-based vetoes using information from coincident triggers.

In the following sections, we discuss the use of DQ vetoes, and the tuning of two signal-based vetoes based on the collocation of the Hanford detectors, the H1H2 effective distance cut and the H1H2 consistency check, described in section 5.2.3.3.

6.8.1 Application of Data Quality Vetoes

While operating the LIGO detectors, we monitor many auxiliary channels that measure the state of the detectors or their local environment, in addition to the GW channel. We use these channels to determine whether or not we want to analyze data from the detectors at those times. This is based on how severe the excursions are from nominal observation mode, what we call "science mode," due to known artifacts introduced into the GW channel data by instrumental or environmental effects coupling to the GW channel. DQ flags are then defined as time segments during which these effects are present in the instrumental or environmental channels. Based on several considerations, we define several different categories of DQ flags for use as vetoes in this analysis.

Category 1 DQ flags are times when we know of a severe problem with the detector, bringing in to question whether the detector was actually in science mode. Some example cases are when calibration for the data is not present, or when loud vibrations were caused in the detector environment in order to test the response of the seismic isolation systems. We choose not to analyze these times because the noise from those times is too high for the data to be useful and we do not want it to affect our median PSD for adjacent times.

We then analyze the data, producing single-detector triggers for use in the rest of the categorization of the DQ flags. In the categorization we take into account several considerations. The first is the "efficiency ratio." For this, we calculate the fraction of triggers from an individual detector that occur during a particular proposed DQ flag and divide this by the "dead time," which is the fraction of the time flagged by the DQ flag. When the efficiency ratio is much greater than one, this DQ flag is highly correlated with false trigger production and is thus a good candidate for eliminating excess triggers. Another criteria we use in the categorization procedure is the "use percentage" of a DQ flag. The use percentage is defined as the fraction time segments from a particular DQ flag that contain single-detector triggers. This value is a measure of how strongly a particular channel couples to the GW channel. Ideally, the use percentage for a good DQ flag veto would be one, meaning that background triggers always occurred whenever that particular DQ flag was on. Finally, we measure the "safety" of using these DQ flags as vetoes by comparing them with the times of hardware injections, which measure the response of the detector to a simulated GW signal. This helps to confirm that the DQ flags will not be correlated with real signals. We require all of the DQ flags used as vetoes to pass this safety check.

Category 2 DQ flags are defined as times when there is a reasonably well established coupling between the GW channel and an auxiliary channel, the flag has a high efficiency ratio, particularly with high SNR triggers, and when there is a use percentage of 50% or greater. An example is when any of the data channels in the length-sensing and control servo reach their digital limit. These are times when we do not trust the data from the detector, however the data were not bad enough to mess up the PSD calculation. We allow these times to be included in our segment generation, but remove triggers that come from these times while calculating bounds on the rate of CBCs.

Category 3 DQ flags are defined as times when the coupling between the auxiliary channel and the GW channel is less well established or when the use percentage is low, but we still find high efficiency ratio for the particular DQ flag. An example is when the winds near the detectors are over 30 Mph. We allow these times to be included in our segment generation, but, as with the Category 2 vetoes, remove triggers that come from these times while calculating bounds on the rate of CBCs.

Category 4 DQ flags are defined as times when the coupling between the auxiliary channel and the GW channel is not well established, when the use percentage is low, when the overall dead time is several percent or greater, or the efficiency ratio is not much greater than one. An example is when nearby aircraft pass overhead. We allow these times to be included in our segment generation and also retain triggers that come from these times. This category is used in following up the loudest events from the search.

Summarizing the uses stated above, we do not analyze data vetoed by category 1 DQ flags, we veto but still follow up triggers in times defined by category 2 and 3 DQ flags, and we use the category 4 DQ flags as information when following up the loudest events from the search. These veto categories significantly reduce the SNR of outlying triggers, as can be seen in figure 6.11.

One could argue that the procedure used to define these vetoes biases the search since it uses in-time triggers. However, the triggers that are used are single-detector triggers *before* coincidence. There are many more triggers at this stage than after coincidence so that triggers associated with real GWs make up a small fraction and do not affect the tuning procedure to a significant degree. This procedure is in the process of being automated (as much as possible) to bias and to minimize the chance that GWs will be vetoed.

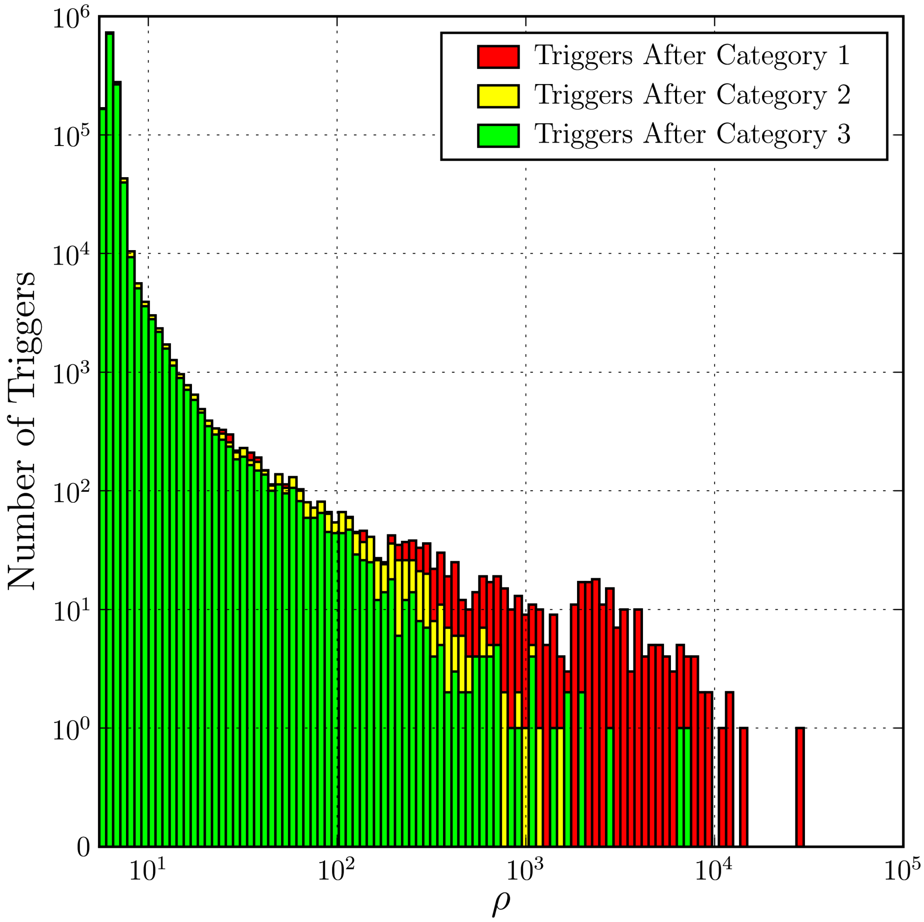

Figure 6.11: Effects of DQ Vetoes
Histogram of triggers for the H2 detector at the single-detector level before any coincidence requirement, clustered by the trigger with maximum SNR within 10 seconds, plotted after veto categories 1, 2, and 3. That is, category 2 vetoes remove the triggers in red, category 3 vetoes remove the triggers in yellow, and triggers in the green histogram remain.

6.8.2 Multidetector Signal-Based Veto Tuning

As described in section 5.2.3.3, there are additional signal-based vetoes we can perform for the collocated Hanford detectors. These vetoes, the effective distance cut and the amplitude consistency check, determine whether coincident triggers from two coaligned detectors satisfy certain amplitude conditions given the detectors aligned antenna patterns.

The veto parameter for both the effective distance cut and the amplitude consistency check is the same κ_* (cf., equation (5.37)), which is tuned by observing the fractional effective distance difference κ distribution for software injection triggers and time-shifted coincident triggers.

As described in section 5.2.3.3, κ should be near zero for triggers associated with signals (e.g., from software injections), while this is not necessarily the case for coincident noise triggers from time-shifted data. This can be seen in figure 6.12. From this figure, we see that the distribution of the time-shifted coincident H1H2 triggers peaks around $\kappa \sim 0.7$. We choose our threshold value to be $\kappa_* = 0.6$. At this value we veto the majority of the time-shifted coincident triggers, while vetoing a very small fraction of the software injections.

In figure 6.12 there is an interesting feature in κ plotted versus H1 SNR for the time-shifted coincident triggers. Here we see strip of time-shifted coincident triggers that extends up to high SNRs. We have found that this strip is consistent with what you would expect if you had found an H1 trigger coincident with an H2 trigger with an SNR of 5.5, while H2 had a horizon distance of 16 Mpc and H1 had a horizon distance of 32 Mpc. This situation is shown with the thin black line.

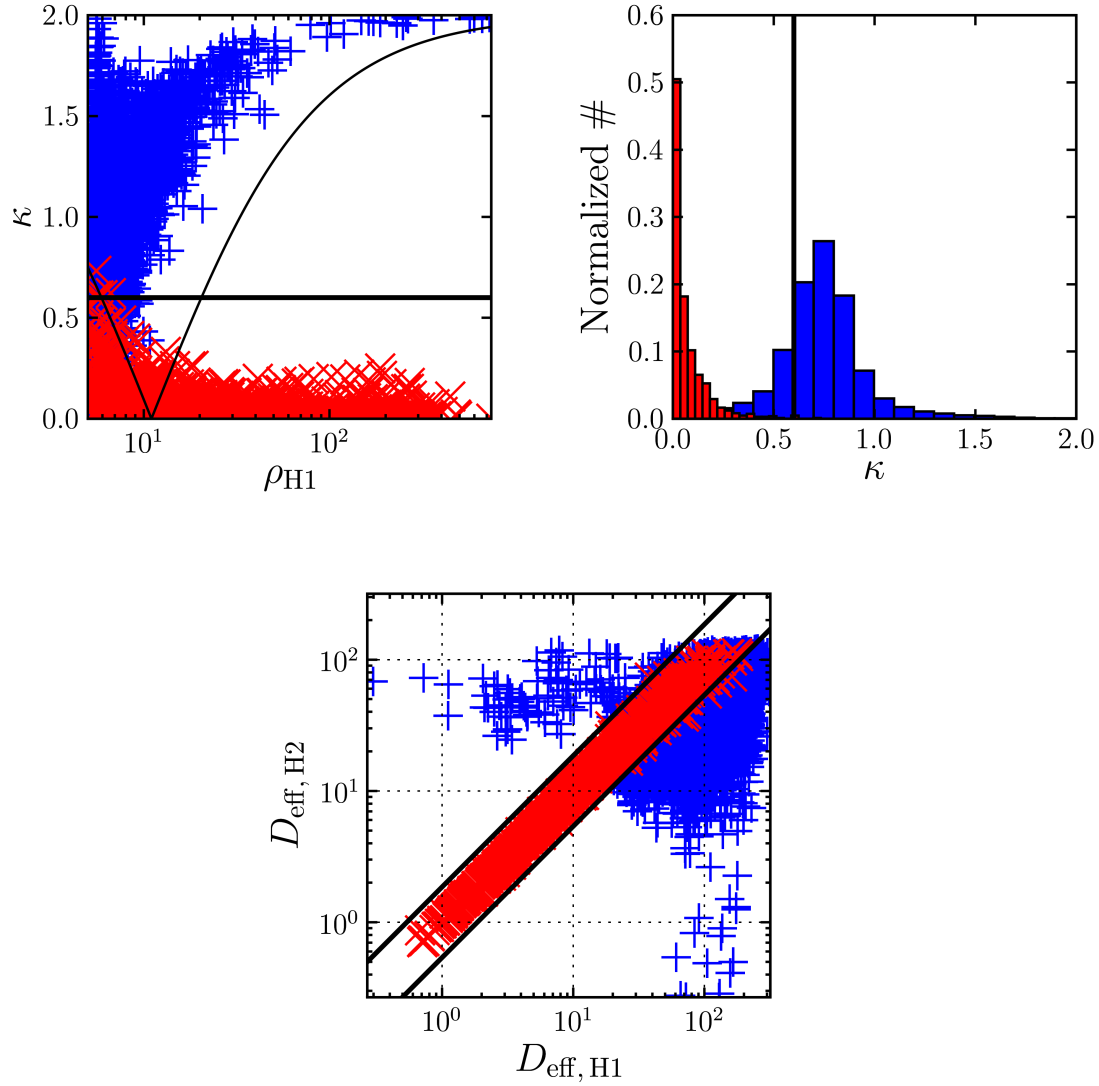

Figure 6.12: H1H2 Effective Distance Cut
The fractional effective distance difference κ versus the H1 SNR (top left), the histogram of κ (top right), and the H2 effective distance versus the H1 effective distance (bottom) for time-shifted coincident triggers (blue) and software injection triggers (red) in H1H2L1 time. The thick black lines show the veto value of $\kappa_* = 0.6$ and the thin black line shows a fit discussed in the text.

Chapter 7

False Alarm Rate Detection Statistic

In this chapter, we describe the development of the false alarm rate (FAR) as the detection statistic for the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data, referred to as "this search." This new detection statistic allows us to search over a large region of parameter space without being limited by a high background FAR from a smaller subregion. In reference [21], coincident triggers were ranked by combined effective SNR. However, for this search, we use a statistic derived from the background FAR, thus enabling a much improved efficiency for a given FAR.

The FAR is a statistic that can be defined for triggers when there is a measure of the background trigger rate. The time-shifted triggers provide an estimate of the FAR for each in-time coincident trigger. By counting the number of time-shifted triggers with a combined effective SNR greater than or equal to the in-time coincident triggers' combined effective SNR, and dividing by the total amount of time we searched for time-shifted triggers, we calculate the FAR for each in-time coincident trigger. This procedure is done separately for different trigger categories, which have different background rates. The result is a varying mapping of combined effective SNR to FAR for the different trigger categories. The FARs from the different categories can then be recombined into a combined FAR (FARc) for the final combined set of triggers.

Section 7.1 discusses how different trigger categories were chosen for this search, section 7.2 explains the FAR calculation, section 7.3 details how trigger categories are recombined, section

7.4 shows the comparison of different detection statistics, section 7.5 presents comparisons of in-time triggers to the many different background trials, and section 9.3.2 describes implications for measuring the background probability for the loudest triggers.

7.1 Definition of Trigger Categories

In this search, there are several aspects that influence the background rate. These include the type of coincidence a trigger was found in (triple coincident triggers versus double coincident triggers), the vetoes that were applied to the triggers, and the mass of the templates that correspond to the triggers. All of these aspects combine in order to define the trigger categories.

7.1.1 Coincidence Types

The first aspect we will focus on is the coincidence type. In this search we are searching for coincident triggers in time and two mass parameters between three GW detectors, the Hanford 4 km interferometer (H1), the Hanford 2 km interferometer (H2), and the Livingston 4 km interferometer (L1). A trigger is deemed a double coincidence if it passes the coincidence requirements between two detectors. A trigger is promoted to a triple coincidence if there are three double coincident triggers that are all found in coincidence with each other. Since triple coincident triggers have an extra coincidence requirement, there is a much lower background rate for triple coincident triggers than for double coincident triggers.

Say each detector has a probability of producing a background trigger for a particular mass template at a particular time of ϵ_1. Adding the double coincident requirement then reduces that probability by an additional factor of ϵ_1, so that $\epsilon_2 = \epsilon_1\epsilon_1 = \epsilon_1^2$. Adding an additional coincident requirement added another factor of ϵ_1, so the probability of getting a triple coincidence trigger at a particular time is then $\epsilon_3 = \epsilon_1\epsilon_2 = \epsilon_1^3$.

This background rate reduction can be illustrated by looking at background H1H2L1 triggers and H1L1 triggers produced during triple coincident time (figure 7.1). For this search the different trigger types we decided to use were H1H2L1, H1L1, and H2L1 triggers. We exclude H1H2 triggers

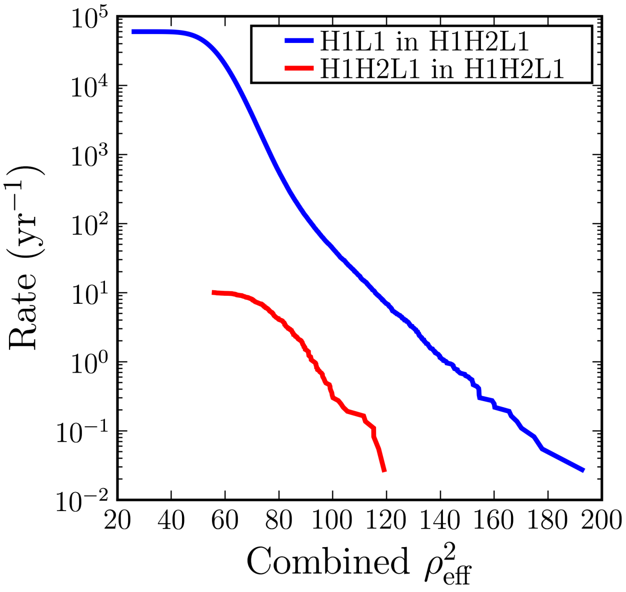

Figure 7.1: Effect of Coincidence on Backround Trigger Rate
The cumulative rate versus combined effective SNR for background H1L1 double coincident triggers during triple coincident time (blue) and H1H2L1 triple coincident triggers during triple coincident time (red). The additional coincidence requirement for H1H2L1 triggers is seen to drastically reduce the background trigger rate.

from this analysis as their background cannot be effectively estimated (section 6.5).

7.1.2 Veto Differences

Another aspect that affects the background rate is the differences in vetoes that are applied to different triggers. One set of vetoes that are trigger dependent are the amplitude vetoes applied between H1 and H2 triggers. These vetoes compare the effective distances found between coincident H1H2 triggers in H1H2L1 coincidences or, for H1L1 or H2L1 triggers, determine whether the other Hanford detector should have also produced a trigger at that time. These vetoes can only be applied when both Hanford detectors were in operation resulting in background rate differences between H1L1 triggers from triple coincident time versus H1L1 triggers from double coincident time, and similarly for H2L1 triggers. This effect is shown for H2L1 trigger in figure 7.2, which clearly shows broader background distributions for H2L1 time where the veto in not applied even though there is

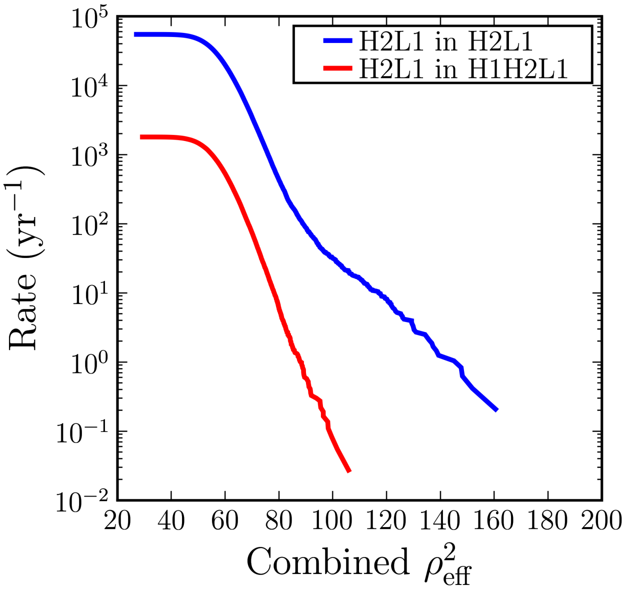

Figure 7.2: Effect of Amplitude Veto on Backround Trigger Rate
The cumulative rate versus combined effective SNR for background H2L1 double coincident triggers during double coincident time (blue) and H2L1 double coincident triggers during triple coincident time (red). The Hanford amplitude veto drastically reduces the brackground rate during triple coincident times, even though there is an order of magnitude more triple coincident time than double coincident time.

an order of magnitude more triple coincident time. This effect is also stronger for H2L1 triggers in triple coincident time than for H1L1 triggers in triple coincident time (figure 7.3). This is due to the fact that the H2 detector is roughly half as sensitive as the H1 detector, which means that if there is a trigger in H2, most of the time H1 should also have produced a trigger at that time.

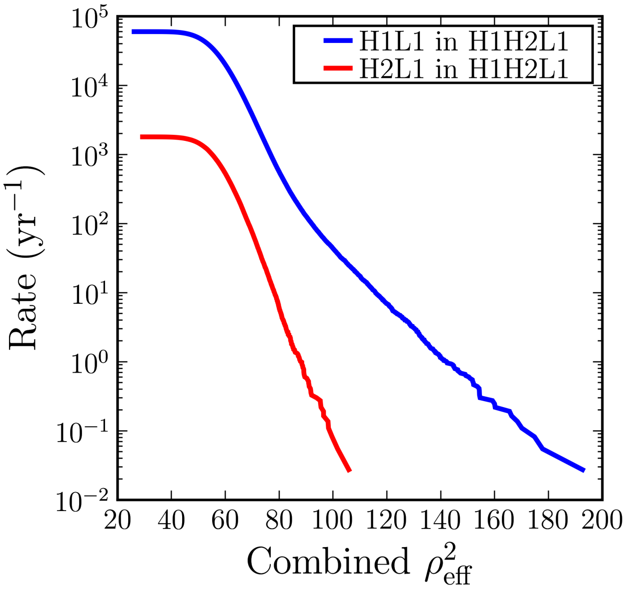

Figure 7.3: Effects of Different Amplitude Vetoes for H1 and H2
The cumulative rate versus combined effective SNR for background H1L1 double coincident triggers during triple coincident time (blue) and H2L1 double coincident triggers during triple coincident time (red). The Hanford amplitude veto is seen to be more stringent for H2L1 triggers than for H1L1 triggers due to the sentivitiy asymmetry between H1 and H2.

7.1.3 Mass Differences

This search is over a much larger parameter space than a single search has covered before. Specifically, the mass space we cover with templates goes from a minimum total mass $M = m_1 + m_2$ of 2 $M_{\odot}$to a maximum total mass of 35 $M_{\odot}$with a minimum component mass of 1 $M_{\odot}$. In order to cover this large space, we use around 6000 templates whose time durations vary from over 44 seconds to under 0.34 seconds.

Due to the large variation in the length of the templates, we find that there is a larger variation in the templates' responses to noise glitches in the detectors where high mass templates tend to pick up the noise glitches as triggers with higher effective SNR. This can be seen in figure 7.4 where we plot the cumulative distributions of background H1L1 double coincident triggers during triple coincident time versus combined effective SNR separated by small bands in the mean chirp mass $\mathcal{M} = M\eta^{3/5}$, where η is the symmetric mass ratio $\eta = m_1 m_2/M^2$. Three populations are found with different slopes for the distributions signifying different background rates, thus, we divide the mass space into three regions, for simplicity. The chirp mass divisions for these regions occur at $\mathcal{M} = 3.48$ $M_{\odot}$and $\mathcal{M} = 7.40$ $M_{\odot}$.

After dividing the triggers of a given type into these three different populations, we find that, indeed, the higher chirp mass templates have a broader combined effective SNR distribution for background triggers, signifying different background rates for the different chirp mass regions (figure 7.5).

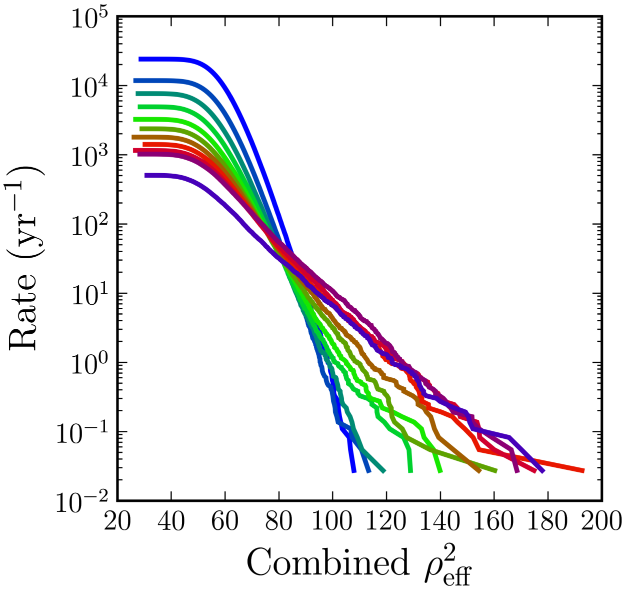

Figure 7.4: Combined ρ^2_{eff} Distributions for Background Triggers for Different Chirp Mass Bands
The cumulative rate versus combined effective SNR for background H1L1 double coincident triggers during triple coincident for eleven separate bands of "mean chirp mass" of equal widths. Lines colored toward the blue end of the spectrum are lower mass bands while line colored toward the purple end of the spectrum are high mass bands. Three populations are found with different slopes for the distributions signifying different background rates. Similar results are obtained by looking at plots for different coincident trigger types.

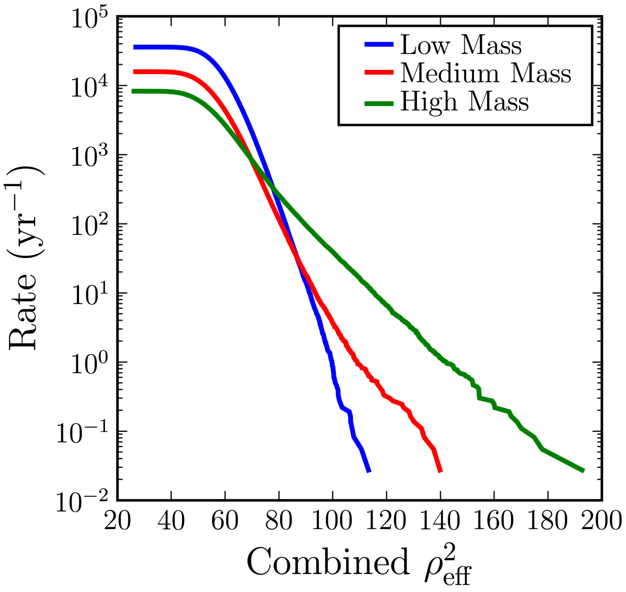

Figure 7.5: Combined ρ^2_{eff} Distributions for Background Triggers for Different Chirp Mass Bins
The cumulative rate versus combined effective SNR for background H1L1 double coincident triggers during triple coincident time for the low chirp mass bin (blue), the medium chirp mass bin (red), and the high chirp mass bin (green). Higher chirp mass templates have a shorter time duration, which leads to a larger response to noise glitches, and thus a broader combined effective SNR distribution for background triggers.

7.1.4 Trigger Categories

Due to all of the above effects, we end up dividing the triggers from this search into 15 different categories during our three different observation times. During triple coincident times we have 3 coincidence types subdivided into 3 chirp mass bins, resulting in 9 trigger categories. During each of the double coincident times we have 1 coincidence type subdivided into 3 chirp mass bins, resulting in 3 trigger categories for each double coincident time.

7.2 FAR: False Alarm Rate

The FAR can be calculated using any intermediate statistic ranking the "loudness" of triggers (such as the combined effective SNR for this search), where a higher intermediate statistic corresponds to a larger excursion for background noise behaviour. The FAR for any trigger is calculated using

$$\mathrm{FAR} = \frac{N}{\sum\limits_i T_i} \ , \tag{7.1}$$

where N is the total number of background triggers with intermediate statistic greater than or equal to the one in question, and T_i is the amount of analyzed time in background trial i.

For individual trigger categories, the maximum FAR is found by

$$\mathrm{FAR}_{\mathrm{max},j} = \frac{N_j}{\sum\limits_i T_i} \ , \tag{7.2}$$

where N_j is the total number of background triggers in the jth trigger category. This can be converted into a minimum inverse false alarm rate (IFAR) trivially by taking the inverse.

The expected number of triggers below a particular FAR due to background is FAR × T_0 where T_0 is the foreground time analyzed. Another way to say this is the expected number of triggers above a particular IFAR due to background is T_0/IFAR where T_0 is the foreground time analyzed, as seen in figure 7.6.

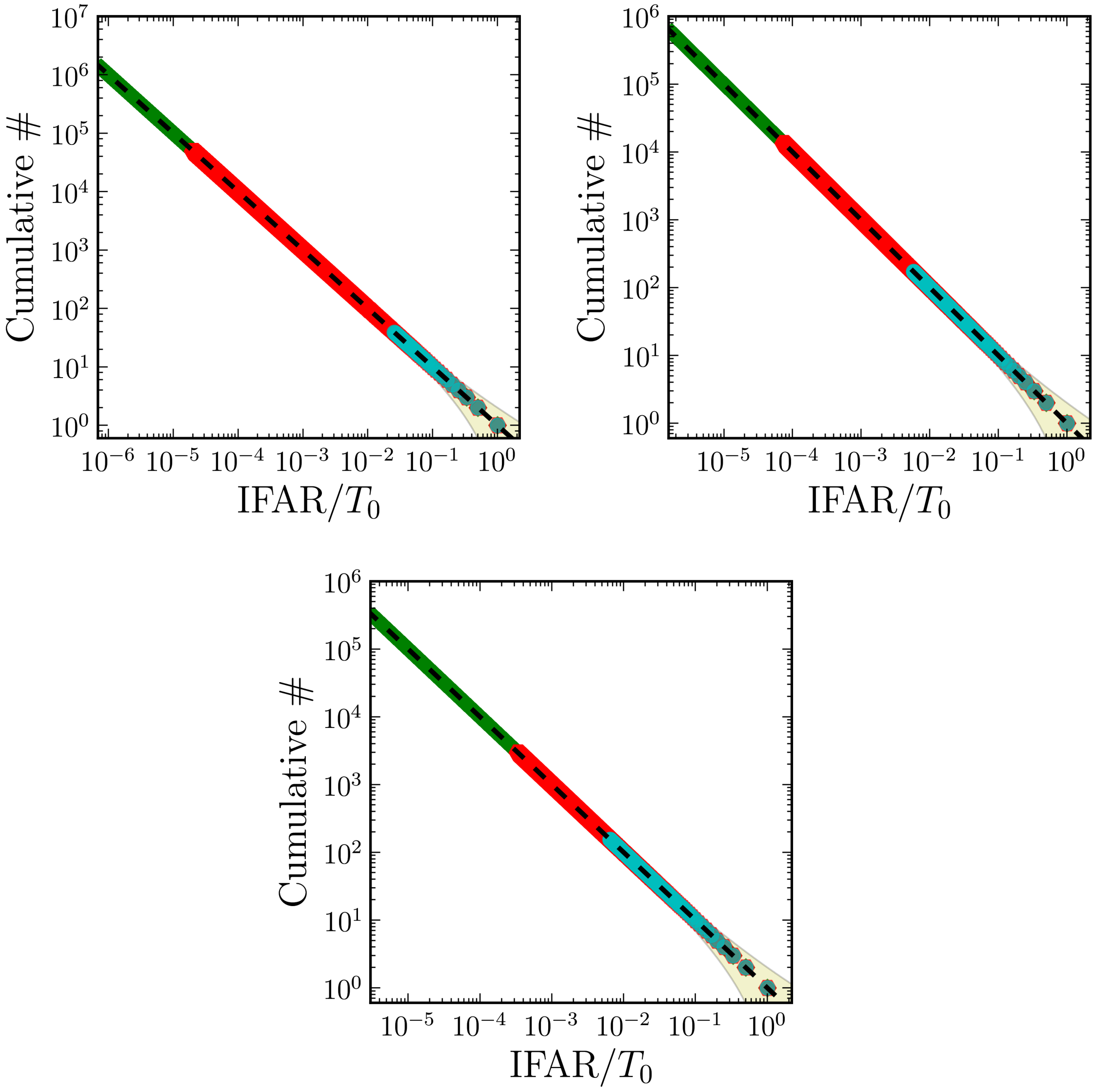

Figure 7.6: Background IFAR Plots

The cumulative number of triggers versus IFAR divided by the analysis time for the nine categories of background triggers in triple coincident time. The top left, top right, and bottom plots show the three different coincident trigger types for the low, medium, and high mass bins respectively. Because they are background triggers, they should, and do, fall exactly on the background line. The different category triggers in each plot all start at the lower right corner of the plot, but since they overlap, only the top color is shown.

7.3 FARc: Combined False Alarm Rate

When combining categories, we think of the different categories (coincidence type, trigger mass bin) as different trials. If there are m trials, then we expect that of those trials there will be m triggers with an IFAR $\geq T_0$. A simple way to normalize the FAR to bring the expected number of triggers with an IFAR of T_0 back to one is

$$\mathrm{FAR}' = m \times \mathrm{FAR}\ . \tag{7.3}$$

This can be seen in figure 7.7.

The simple way of combining shows that, moving from right to the left, there is a kink in the background triggers whenever you reach a minimum IFAR for a trigger category. The reason these kinks occur is because once we reach a minimum IFAR for a trigger category, there is then one less "trial" for us to combine. A better thing to do at that point is to normalize the FAR by the remaining number of categories rather than the total. In analytic form, the FARc becomes

$$\mathrm{FARc} = \left[\sum_{j=1}^{p} \Theta\left(\mathrm{FAR}_{\max,j} - \mathrm{FAR}\right)\right] \times \mathrm{FAR} + \sum_{j=1}^{p}\left[\Theta\left(\mathrm{FAR} - \mathrm{FAR}_{\max,j}\right)\mathrm{FAR}_{\max,j}\right]\ , \tag{7.4}$$

where p is the number of categories of triggers and $\Theta\left(x\right)$ is the Heaviside function. Calculating the FARc in this way can be seen in figure 7.8.

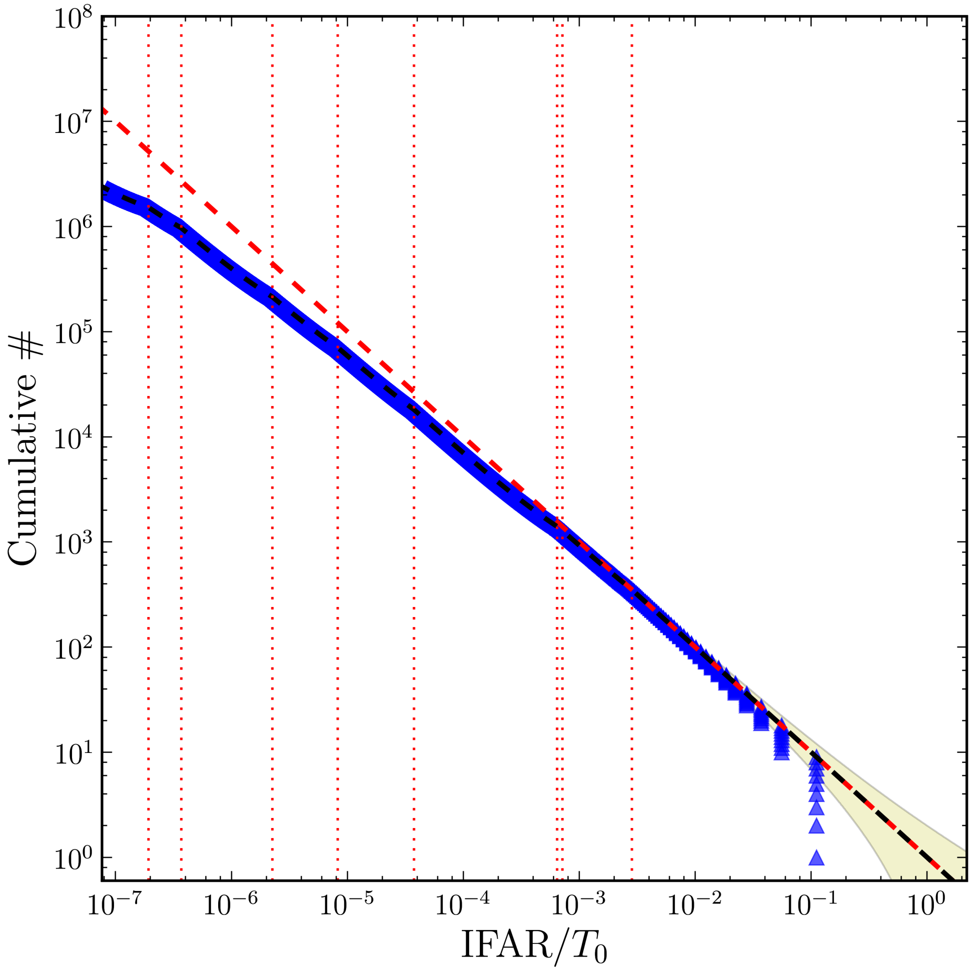

Figure 7.7: Combined Background IFAR Plot
The cumulative number of background triggers versus IFAR divided by the analysis time for nine categories of background triggers combined without correcting for the minimum IFAR of each category. As in the previous plot, because they are background triggers, they should, and do, fall exactly on the background line. The vertical, red, dotted lines denote the minimum IFAR for eight different trigger categories, which are the locations of the kinks in the background line. The ninth is not shown as that is where the background line terminates.

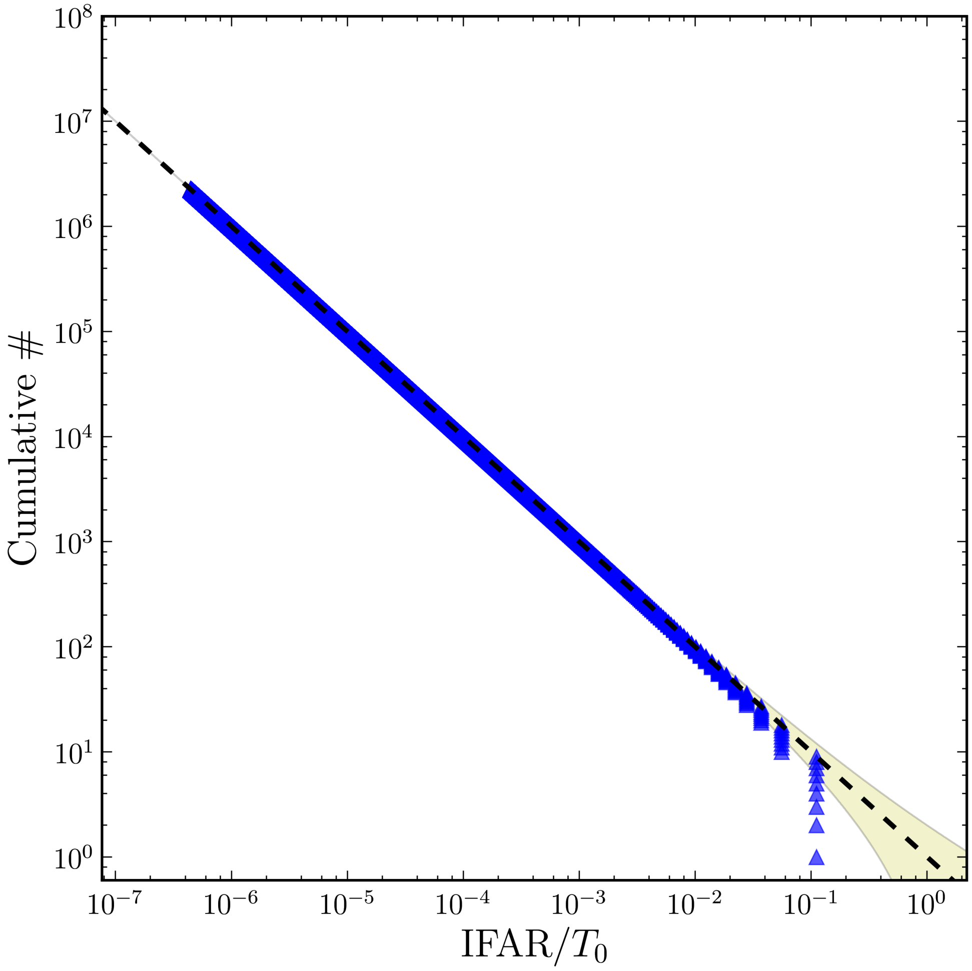

Figure 7.8: Background IFARc Plot
The cumulative number versus the combined IFAR (IFARc) divided by the analysis time for nine categories of background triggers. Again, because they are background triggers, they should, and do, fall exactly on the background line.

7.4 Comparison of Detection Statistics

Now that we have a method for calculating the FARc detection statistic, let us compare this statistic against the combined SNR and combined effective SNR in recovering software injections signals. This is done using a common comparison known as the Receiver Operating Characteristic (ROC) curve, shown in figure 7.9.

The FARc statistic out performs both the combined SNR and combined effective SNR for all mass regions. It is interesting to note that the largest improvement is seen for the low mass bin, and the least improvement in the high mass bin. This lends support to our assumption that the high mass templates have a higher false alarm rate.

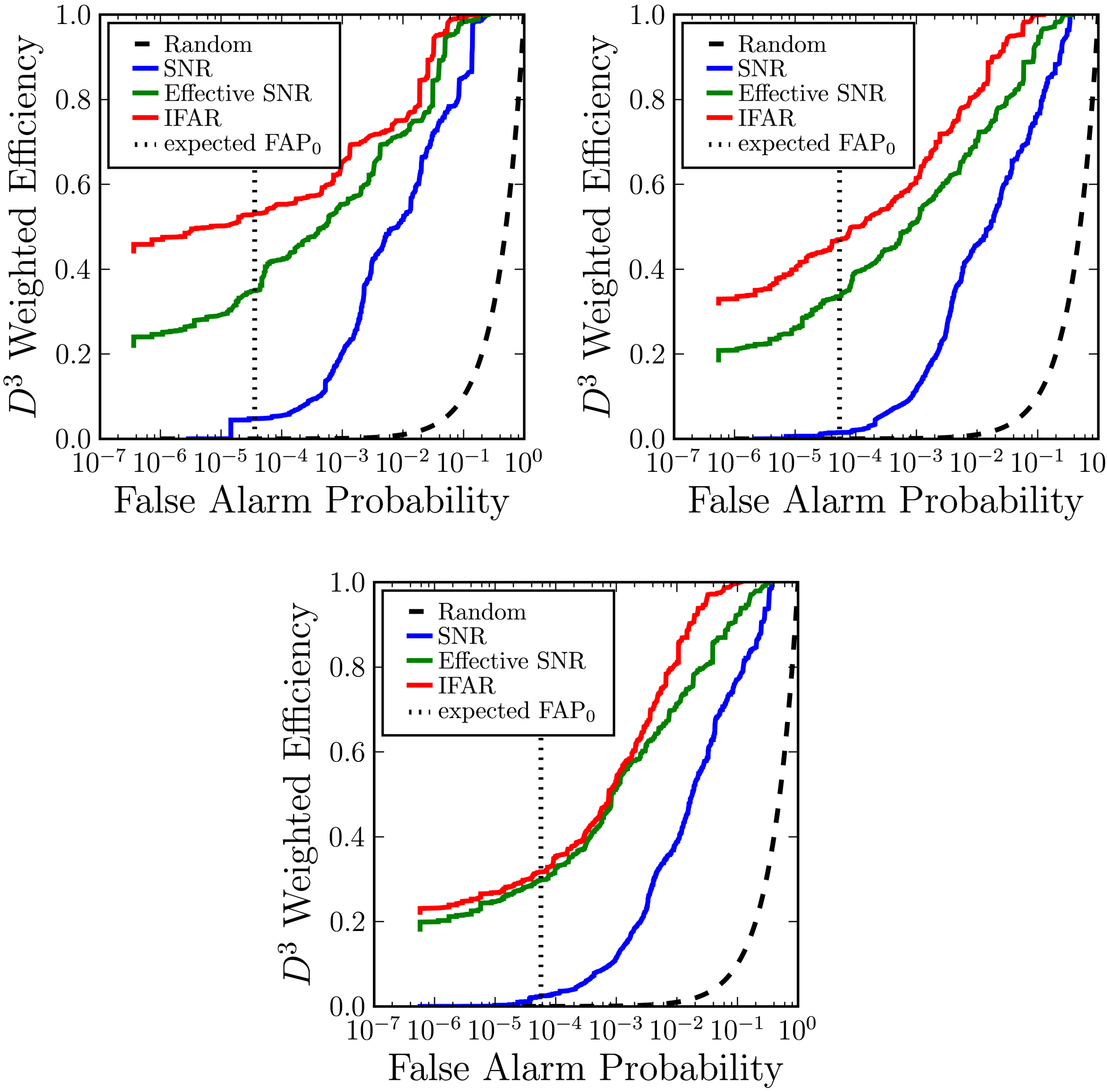

Figure 7.9: Comparison of Detection Statistics

We compare the efficiency of recovering injected signals versus false alarm probability (FAP) for three different detection statistics in the low (top left), medium (top right), and high (bottom) mass regions. The vertical line in each plot shows where the loudest in-time trigger is expected to occur when it is consistent with background. The injected signals were spaced uniformily in $\log(D)$. The plot has been made by reweighting each injection such that they represent a population uniform in the volume (D^3).

7.5 Plotting Background Trials: Lightning Bolt Plots

In addition to plotting the IFARc of foreground triggers, we can also plot the IFARc of each of the individual background trials (i.e., time shifts). In this case, the IFARc has been normalized by

$$\mathrm{IFARc}_i = \mathrm{IFARc} \times \frac{T_0}{T_i} \, , \tag{7.5}$$

where i corresponds the different background runs. An example plot of this is shown in figure 7.10. The plots give 100 examples of the expected distribution of background-only triggers to better judge the significance (inconsistency with background) of the in-time coincident triggers.

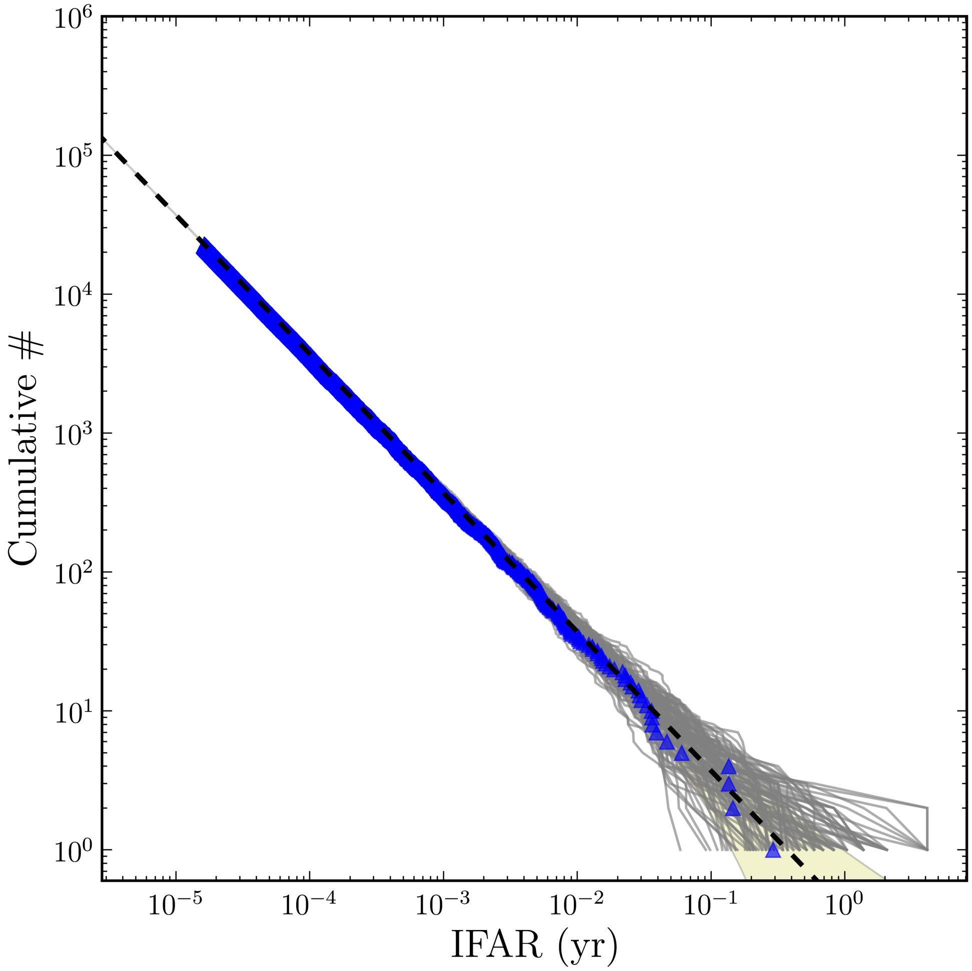

Figure 7.10: In-Time IFARc Plot
The cumulative of coincident triggers with IFARc greater than the value of IFARc on the x-axis for nine categories of foreground triggers from H1H2L1 observation time in this search. The dashed line shows the background distribution estimate averaged over all the time slides. The grey line give the distribution for each of the 100 background trial taken separately. The yellow band shows the $N^{1/2}$ one-standard deviation errors on the average background estimate.

7.6 False Alarm Probability

Using the FARc as the detection statistic allows a very easy calculation of the false alarm probability (FAP) and the background probability P_0 [130, 131]. Assuming a Poisson distribution, the FAP is the probability of getting any (zero) background triggers louder than the loudest zero-lag trigger (i.e., with a FARc lower than the FARc of the loudest zero-lag trigger since low FARcs are more significant) and the background probability is the probability of getting zero background triggers louder than the loudest zero-lag trigger, given by

$$\mathrm{FAP} = 1 - e^{x} \tag{7.6a}$$

$$P_0 = e^{x} \ , \tag{7.6b}$$

where $x = -\mathrm{FAR} \times T$, FAR is the FARc of the loudest zero-lag trigger, and T is the total analyzed time searching for zero-lag triggers. In-time coincident triggers with low FAP are GW detection candidates, and are subject to follow-up, as described in section 10.1.

Chapter 8

Ranking of Detection Candidates

In this chapter we describe the postprocessing pipeline that has been developed as an extension to the standard pipeline (discussed in chapter 6) to analyze data from GW detectors, such as the LIGO detectors, for GW signals from CBCs. In addition, we will also describe the tuning that was done for the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data, described in this thesis, and referred to as "this search."

At the end of the standard pipeline, we have distilled the data down from one data stream sampled at 16,384 Hz for each GW detector into files containing coincident triggers that have passed our thresholds and vetoes. These files are organized by coincident time segments defined by times when different combinations of detectors' data were being analyzed.

Recalling section 6.8, during these coincident time segments triggers are discarded if they occur during times flagged by the category 1, 2, or 3 DQ flag veto files. As mentioned in section 7.1, these vetoes affect the application of signal-based vetoes, which, in turn, significantly affect the background distribution of triggers. Because of this, we separate the coincident times (and triggers within those times) according to the detectors that were operating without DQ vetoes active. We then recalculate the amount of analyzed time in each type of coincidence time for the in-time data as well as for each time shift, accounting for the dead time incurred from the application of DQ flag vetoes. Next we cluster coincident triggers in order to reduce the final number of triggers to a manageable level, and separate them by categories (mass regions and coincident trigger types). Finally, we calculate the FAR for each of the in-time, software injection, and time-shift coincident

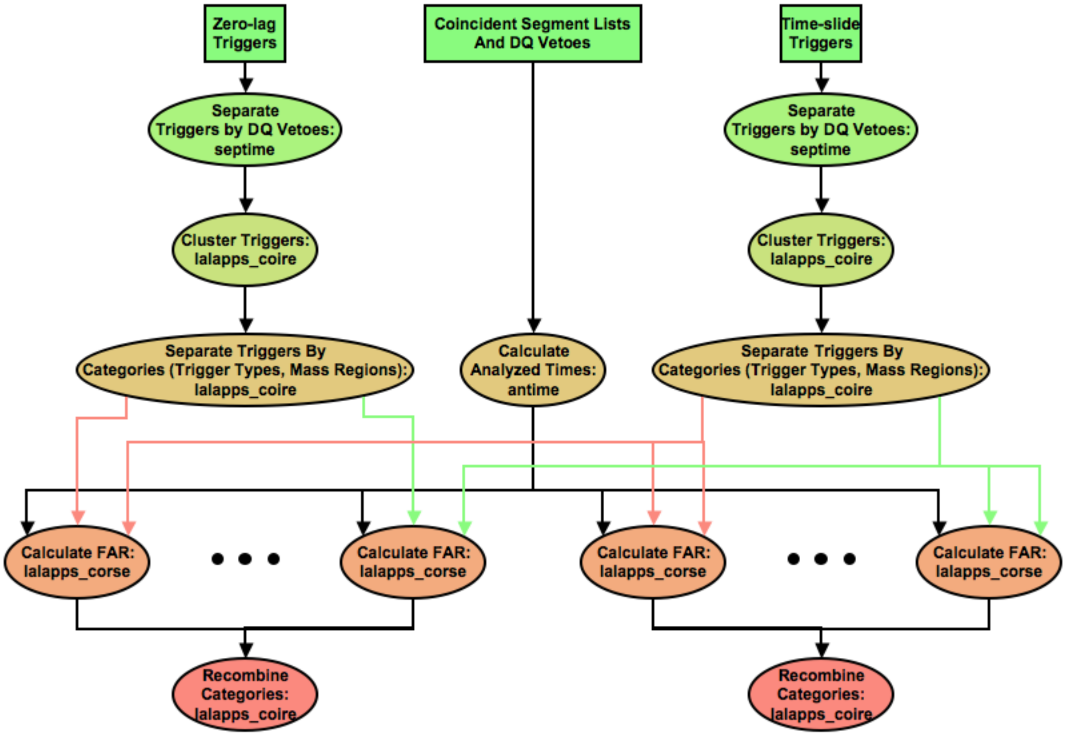

Figure 8.1: Postprocessing Pipeline
The stages of the postprocessing producing the final candidate triggers.

triggers, utilizing the in-time and time shifts' amounts of analyzed time. The postprocessing pipeline is represented in figure 8.1.

The different stages of the postprocessing pipeline used in this search are discussed below:

- separate triggers by veto times (detailed in section 8.1);
- calculate nonvetoed analyzed times (detailed in section 8.2);
- cluster triggers and separate them by categories (detailed in section 8.3);
- calculate the FAR for the triggers (detailed in section 8.4).

8.1 Separating Triggers by Data Quality Flag Vetoes

The application of DQ flag vetoes causes a loss of detector live time that was not accounted for in previous searches (it was instead counted as a loss of efficiency). Accounting for it as a loss of live time requires us to reclassify the coincident category of coincident triggers in the in-time and time-shifted data.

The first thing that is done in the postprocessing pipeline is to separate triggers according to which detectors were producing triggers that did not have DQ flag vetoes active. As mentioned above, because the DQ flag vetoes affect which signal-based vetoes we apply, we want to group double coincident triggers occurring during triple coincident time where the third detector had an active DQ flag veto with double coincident triggers coming from true double coincident time.

We work with single coincident science segments in which we separate triggers in a manner appropriate for both time-shift and in-time triggers. The output is a number of SEPTIME files for each type of time when two detectors were not being vetoed (for H1H2L1 times, the output files can be H1H2 times, H1L1 times, H2L1 times and H1H2L1 times; for H1L1 times, the output is only H1L1 time). In the case of in-time data, it is also dependent on whether or not there was a nonzero amount of time of that type.

Figure 8.2 shows a visual example of extracting the correct time-shift triggers from a single triple coincident segment. The vector used for the time shifts here is $v = \{0, 2, 1\}$ for H1, H2, and L1 respectively. The top black bar at $y = 10$ shows the segment we are working with is $[0 - 10)$, the next bars show the times when each of H1 (red), H2 (blue), and L1 (green) are not vetoed (H1 not vetoed $[1 - 10)$, H2 not vetoed $[1 - 10)$, L1 not vetoed $[1 - 6)$). Just below those lines show the in-time triggers that are shifted when computing coincidence in time-shifted data.

The middle lines and triggers from $y = 5$ to $y = -5$ show the time shifts numbered 5,4, . . . ,1,-1, . . . -5. The colored bars again denote the times the detectors were not vetoed, shifted by the appropriate amount for that time shift and detector. The triggers under the time-shift bars are the in-time triggers that were coincident in that time shift, shifted to their appropriate location. If you pick a coincident trigger in a specific time shift, you can look just above it to tell which detectors were producing nonvetoed triggers.

The final four lines at the bottom denote the four different times triggers could be generated (i.e., H1H2L1, H1H2, H1L1, and H2L1 times from top to bottom). The unshifted trigger times, which can be compared to the input trigger times at the top of the plot, are shown on each of those four lines. The septime program outputs those four lines are four separate trigger files.

At the end of this stage, for each coincident segment, we are left with files whose contents are coincident in-time or time-shifted triggers that occurred during coincident times defined as times when those detectors were analyzed and not vetoed by DQ flag vetoes.

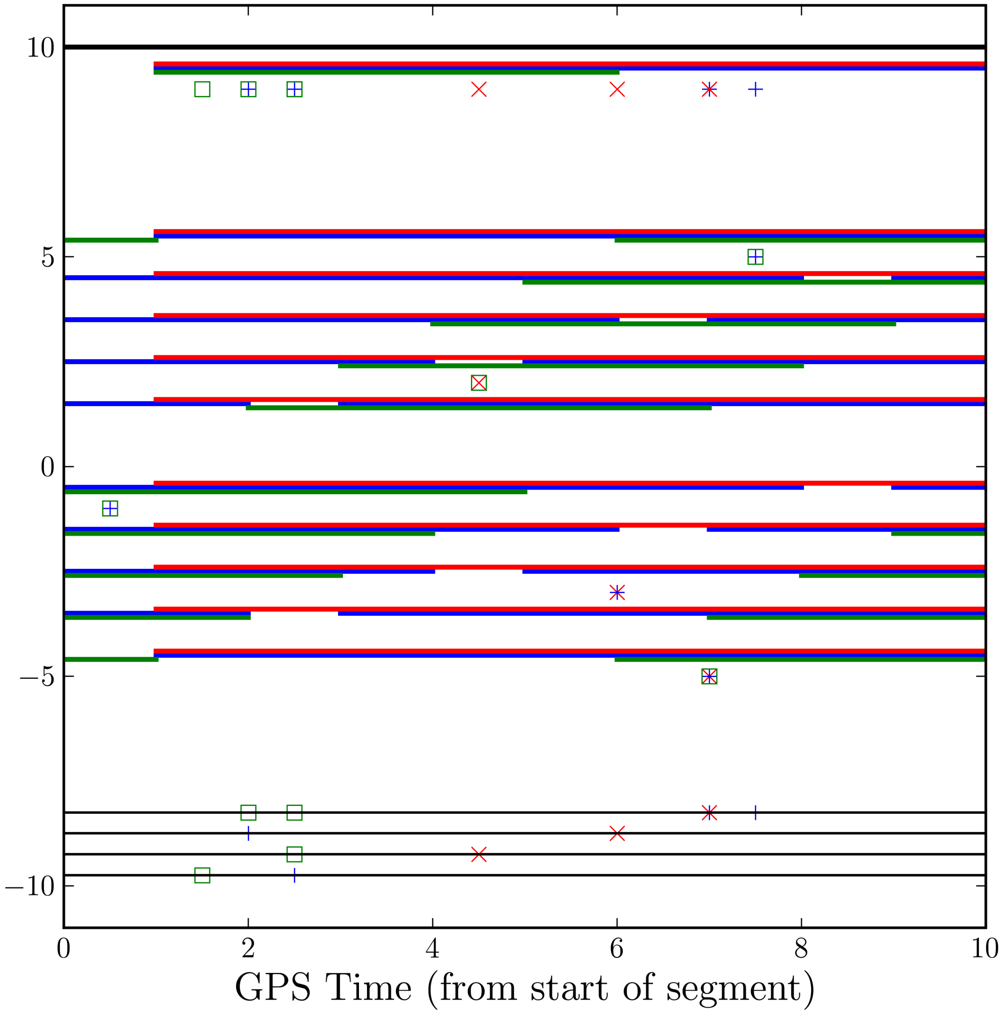

Figure 8.2: SEPTIME Example
An example of septime extracting time-shift triggers from the appropriate coincidence times once DQ flag vetoes are taken into account.

8.2 Calculate Analyzed Times

Now that we have sorted the triggers according to times when detectors were analyzed and not vetoed by DQ vetoes, what we will now refer to as "analyzed times," we also want to calculate the amount of analyzed times per in-time and time-shifted data. We do this using a program called antime that uses the science segment files and appropriate veto files. When detectors are vetoed, that time is removed from that type of time and added to the type of time defined by the detectors not being vetoed (e.g., H1H2L1 time where H2 is vetoed is added to H1L1 time).

Summarizing the in-time analyzed times, before the application of category 2 ans 3 DQ flag vetoes, the in-time analyzed times were 13,237,711 s for H1H2L1 times, 1,179,439 s for H1L1 times, 1,487,038 s for H2L1 times, and 7,337,970 s for H1H2 times. Once the category 2 and 3 vetoes are taken into account we have the following times from triple coincident times: 11,696,445 s for H1H2L1 times (88.4% of the original), 76,446 s for H1L1 times (0.6%), 470,764 s for H2L1 times (3.6%), and 790,885 s for H1H2 times (6.0%). In addition, from the double coincident times we have: 615,994 s for H1L1 times (52.3%), 949,899 s for H2L1 times (63.9%), and 6,951,276 s for H1H2 times (94.7%).

In total, we have 11,696,445 s of H1H2L1 time, 692,440 s of H1L1 time, 1,420,663 s of H2L1 time, and 7,742,161 s of H1H2 time.

Figure 8.3 shows the in-time and time-shifted analyzed times after category 2 and 3 DQ flag vetoes have been removed. The strong dependence on time-shift number seen in H1L1 and H2L1 time is due to correlations between H1 and H2 DQ vetoes. Since these vetoes are positively correlated due to their shared environment, as H1 and H2 times are shifted against each other during H1H2L1 times, time that was initially being vetoed by both H1 and H2 DQ vetoes in the in-time turns into time where only one of the Hanford detectors is being vetoed for the time shifts. Since the H1H2L1 time initially has an order of magnitude more time than H1L1 or H2L1 time, the small effect of sliding correlated H1 and H2 vetoes against each other becomes a more significant amount of time transfered to H1L1 and H2L1 analyzed times in comparison to their initial amount of analyzed times. This effect is more pronounced as one moves away from the nonshifted in-time data because the relative shift between H1 and H2 (10 s) is comparable to the typical duration of DQ flags.

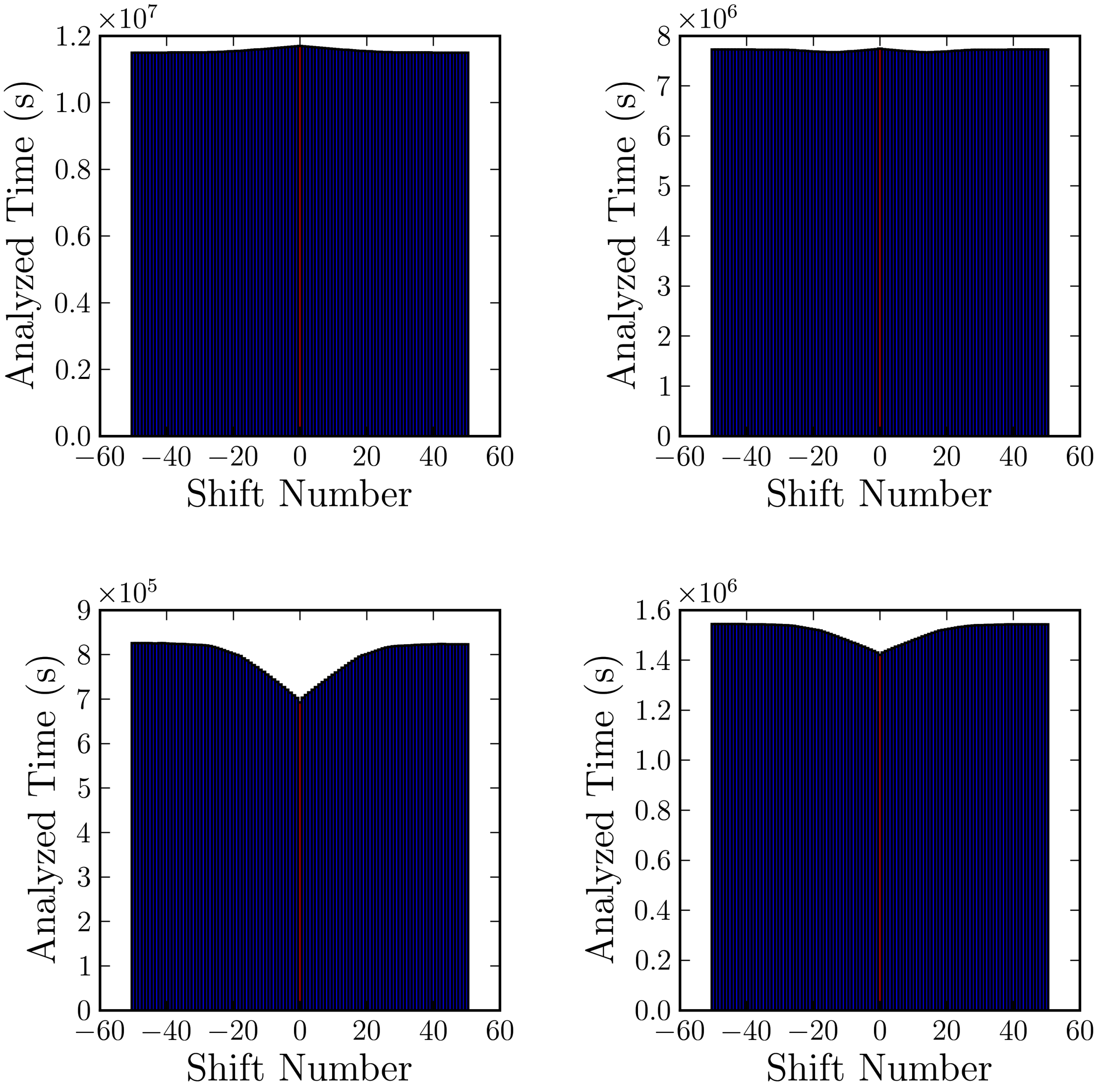

Figure 8.3: Background Analyzed Time
The analyzed times as a function of time-shift number can be seen here after the category 2, and 3 DQ flag vetoes are taken into account for H1H2L1 (top left), H1H2 (top right), H1L1 (bottom left), and H2L1 (bottom right) times. The red bar in each plot denotes the in-time analyzed time.

8.3 Trigger Clustering and Separation by Category

In the next stage of the postprocessing pipeline, we prepare the triggers for the FAR calculation. In order to reduce the total number of triggers from the search to a reasonable size, we cluster the SEPTIME files using a sliding cluster window of 10 s, using the combined effective SNR as the intermediate detection statistic. The output of this thep is one COIRE file per SEPTIME file. This incurs an insignificant loss of signal recovery efficiency since, generally, when there is a glitch in the data, the whole template bank produces triggers, masking possible signals during that time. The clustering we do focuses the glitch into the loudest trigger during that time, which generally has the largest combined effective SNR in the high mass region of the template bank. This effect could be the cause of the louder background combined effective SNR distribution for the high mass region, described in chapter 7.1.

Next we separate triggers into different categories due to the differences between the background trigger combined effective SNR distributions (chapter 7.1). In that chapter we found that the background combined effective SNR distributions can be affected by multiple factors including the coincidence level (i.e., double versus triple coincidence), the signal-based vetoes applied, and the mass of the templates producing the triggers. The first two of these three factors leads us to separate triggers by the different observation times and the different coincidence types (i.e., H1H2L1, H1H2, H1L1, and H2L1 triggers in triple coincident time and H1H2, H1L1, and H2L1 triggers in double coincident times). In addition, we separate these triggers at the chirp mass divisions of $\mathcal{M} = 3.48$ $M_{\odot}$and $\mathcal{M} = 7.40$ $M_{\odot}$.

These trigger categories are implemented by taking the above clustered COIRE files, grouping them together by observation time, separating the triggers into the three mass regions, and subdividing the triggers in triple times further into coincidence types (H1H2L1 triggers, H1H2 triggers, H1L1 triggers, and H2L1 triggers).

8.4 FAR Calculation

The final stage of the postprocessing pipeline is the FAR calculation. The FAR is calculated using the time shifts and analyzed times as described in section 7.2. This is done separately for each trigger category using LALAPPS_CORSE (the COincident Rate Statistic Estimator). Once the FAR is calculated, the categories for a particular observation time are recombined as separate trials and their FARc is calculated (described in section 7.3). In the end, the IFARc is used as our detection statistic where triggers with larger IFARc values are louder.

Chapter 9

Rate (Upper Limit) Calculation

Whether the search results in a small number of detection candidates or more, we can bound the rate of CBCs in the (nearby) universe. Depending on the significance of the loudest event, the confidence band we establish may or may not exclude zero. In either case, we can establish bounds on the rate for different source mass ranges.

In this discussion, we follow the loudest event formalism described in [132, 133, 130, 131]. One advantage of calculating bounds on a rate using the loudest event formalism is that for searches that have steep background distributions, the expected upper limit you could get is always more stringent than in fixed threshold formalisms. Another advantage is that one does not need to fix a threshold before looking at the in-time triggers. It should be noted that in the end, the formalism chosen does not affect the significance of the in-time triggers and is only necessary for calculating rate bounds.

In previous CBC searches [17, 19, 18], we calculated rates of CBCs in terms of Milky Way Equivalent Galaxies (MWEGs). However, as the horizon distances of our detectors expanded to include more galaxies, we found it was prudent [134] to report rates in terms of L_{10}s, where $1\,L_{10} \equiv 10^{10}\,L_{\odot,B}$, $L_{\odot,B}$ is the solar blue-light luminosity, and 1 MWEG $\approx 1.7\,L_{10}$. We have used these units for our rate calculations in [21, 22], and we shall use it here as well. The reason for this change is because L_{10} track the rate of star formation in a given galaxy and we think CBCs are proportional to this [135]. One drawback of this is that old elliptical galaxies are ignored since they do not have significant blue-light luminosity [136].

In order to calculate an upper limit on a rate of gravitational wave signals due to a particular search, there are several things we need to know. These are related to how sensitive the search is to signals from the universe, the background of the search, and the loudest event from the search. Below we detail the calculations performed in arriving at a rate (figure 9.1).

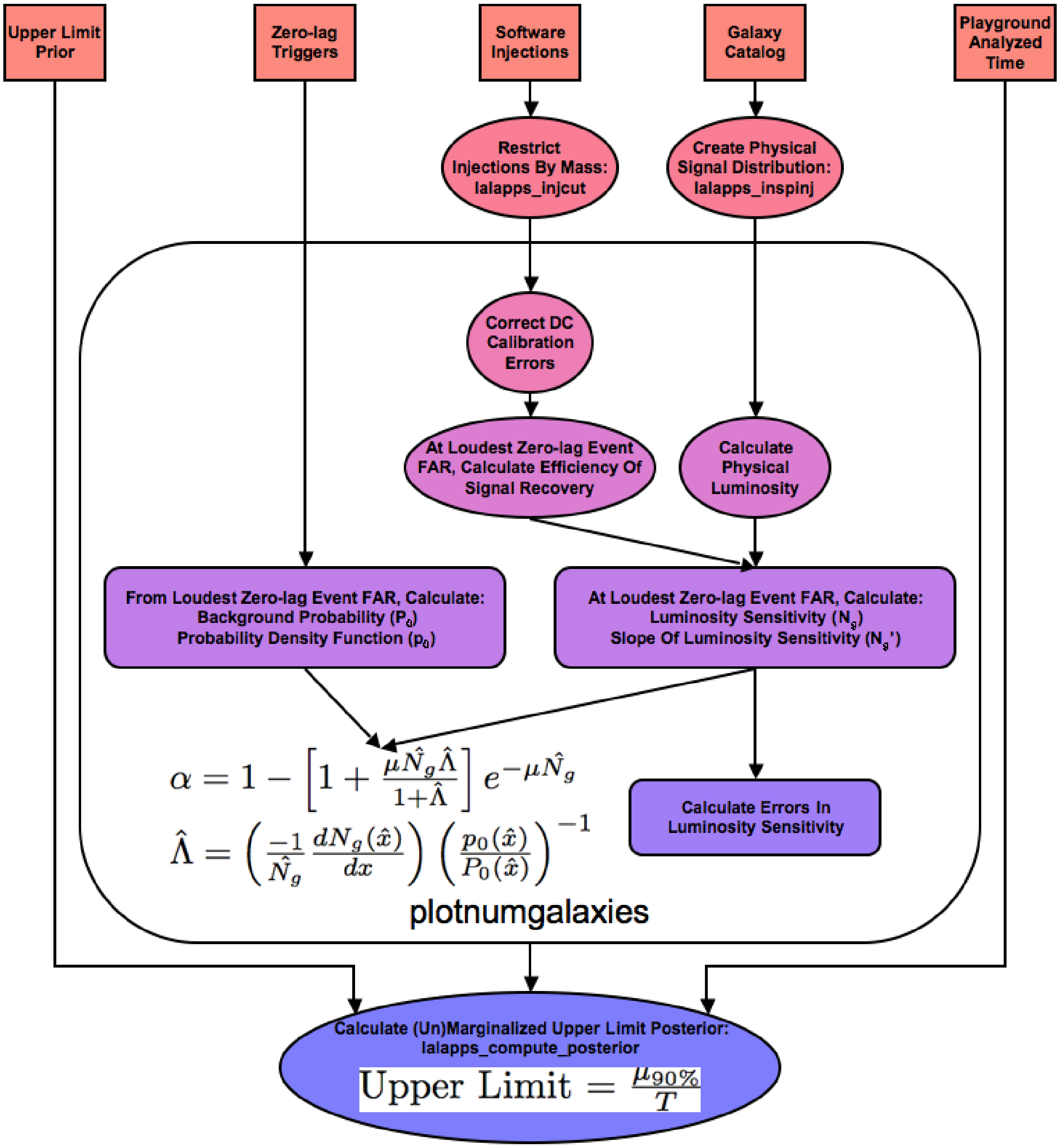

Figure 9.1: Upper Limit Calculation Pipeline
The stages of the upper limit calculation, in which we information from the top of the figure to establish bounds on the CBC rate in the nearby universe in terms of events yr^{-1} L_{10}^{-1}.

9.1 Posterior and Upper Limit Calculation

Calculating an upper limit on a rate of coalescences in the loudest-event formalism requires knowledge of the cumulative luminosity to which the search is sensitive and a measure of the likelihood that the loudest event was due to the observed background, described in section 9.3. We combine these with the time analyzed to calculate the posterior on the rate for the search. Assuming a prior on the rate $p_0(\mu)$, the posterior is given by [130]

$$p\left(\mu|\mathcal{C}_L, T, \Lambda\right) = p_0(\mu)\frac{\mathcal{C}_L T}{1+\Lambda}\left(1+\mu\mathcal{C}_L T\Lambda\right)e^{-\mu\mathcal{C}_L T} \ , \tag{9.1}$$

where μ is the rate in events $\text{yr}^{-1}\ L_{10}^{-1}$, $\mathcal{C}_L$ is the cumulative luminosity in L_{10}, T is the analyzed time in years, and Λ is a measure of the likelihood of detecting a single event with loudness parameter x (which in this analysis is FARc, cf., sections 7.3 and 8.4) versus such an event occurring due to the experimental background, given by [130]

$$\Lambda\left(x\right) = \left(\frac{-1}{\mathcal{C}_L}\frac{d\mathcal{C}_L}{dx}\right)\left(\frac{1}{P_0}\frac{dP_0}{dx}\right)^{-1} \ , \tag{9.2}$$

and P_0 the background probability given by equation (7.6b). In this search, $P_0 = dP_0/dx = e^x$, thus the term involving these quantities disappears, which we will discuss in section 9.3.2.

In the limit that $\Lambda \to 0$ (i.e., when the loudest event is absolutely due to background), the 90% confidence limit we set on the rate $\mu_{90\%}$ is given by

$$90\% = \int_0^{\mu_{90\%}} p(\mu)d\mu \Rightarrow \mu_{90\%} = \frac{2.303}{C_L T} \ , \tag{9.3}$$

and in the limit that $\Lambda \to \infty$ (i.e., when the loudest event is absolutely due to signal), the 90% confidence limit we set on the rate $\mu_{90\%}$ is given by

$$\mu_{90\%} = \frac{3.9}{C_L T} \ . \tag{9.4}$$

9.2 Incorporating Systematic Errors

The posterior (9.1) assumes a known value of $\mathcal{C}_L$ associated with the search. In reality, $\mathcal{C}_L$ has associated with it systematic uncertainties, which we model as unknown multiplicative factors ζ (fractional errors), each log-normally distributed about 1 with errors described in section 9.4. In order to marginalize over the effects of the systematic errors on the cumulative luminosity, we first calculate a probability distribution for the cumulative luminosity for the search, $p_d\left(\mathcal{C}_L\right)$. This is calculated using the log-normal error distributions of the cumulative luminosity multiplier $p_{d,\mathrm{error}}\left(\zeta\right)$ given by

$$p_{d,\mathrm{error}}\left(\zeta\right) = \frac{1}{\zeta\sigma_{\mathrm{error}}\sqrt{2\pi}} e^{\frac{-\ln(\zeta)}{2\sigma^2_{\mathrm{error}}}} \ , \tag{9.5}$$

where σ_{error} is the fractional error of the cumulative luminosity calculated in section 9.4. The $p_{d,\mathrm{error}}\left(\zeta\right)$ distributions are then convolved with each other to obtain a combined distribution $p_{d,\mathrm{comb}}\left(\zeta\right)$. The distribution of the cumulative luminosity is then given as

$$p_d\left(\mathcal{C}_L\right) = \int \delta\left(\mathcal{C}_L - \zeta\mathcal{C}_{L,\mathrm{m}}\right) p_{d,\mathrm{comb}}\left(\zeta\right) d\zeta \ , \tag{9.6}$$

where $\mathcal{C}_{L,\mathrm{m}}$ is the measured cumulative luminosity. To obtain a marginalized posterior, we integrate the above posterior (equation (9.1)) times the distribution of cumulative luminosities (equation (9.6)) over the cumulative luminosity [130]

$$p\left(\mu|T,\zeta\right) = \int p_d\left(\mathcal{C}_L\right) p\left(\mu|\mathcal{C}_L,T,\zeta\right) d\mathcal{C}_L \ . \tag{9.7}$$

The results of several experiments (e.g., different types of S5 observing time and previous runs such as S3 and S4) can be combined by taking the product of their likelihood functions; in the case of uniform priors, this is equivalent to taking the product of their posteriors, allowing us to define the rate upper limit μ at a confidence level α by numerically solving

$$\alpha = \int_0^{\mu} \prod_i p_i\left(\mu'\right) d\mu' \ , \tag{9.8}$$

where the $p_i(\mu')$ are the marginalized posteriors from different experiments calculated using a uniform prior on the rate. This is equivalent to using the posterior from previous observations as the prior for subsequent ones, sequentially.

9.3 Cumulative Luminosity and Background Probability

The standard quantities we need to have before we can calculate an upper limit on the rate of coalescences are the sensitivity of the search (i.e., the cumulative luminosity), and the probability that no background triggers are louder than the loudest event.

9.3.1 Cumulative Luminosity and its Derivative

We measure the *cumulative luminosity* that the search is sensitive to in units of L_{10}. The cumulative luminosity quantifies the potential sources of observable CBC, as measured by blue-light luminosity of the galaxies containing CBCs, which can be detected by our search. This can be calculated by integrating the efficiency $\epsilon(D)$ and physical luminosity $L(D)$ both as a functions of distance over distance

$$\mathcal{C}_L \equiv \int \epsilon(D) L(D) dD \, . \tag{9.9}$$

In practice, we do this as a discrete sum

$$\mathcal{C}_L = \sum_i \epsilon_i L_i \, , \tag{9.10}$$

where ϵ_i is the efficiency in distance bin i, and L_i is the luminosity in distance bin i.

We calculate the *efficiency* ϵ_i in distance bin i as

$$\epsilon_i = \frac{N_{\mathrm{found},i}}{N_{\mathrm{injected},i}} \, , \tag{9.11}$$

where $N_{\mathrm{injected},i}$ is the number of injections injected with distance $D \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$, $D_{i,\mathrm{min}}$ is the minimum distance in distance bin i, $D_{i,\mathrm{max}}$ is the maximum distance in distance bin i, and $N_{\mathrm{found},i}$

is the number of found injections in that distance bin. Distance bins in which there are no injections are taken to have $\epsilon_i = 0$. An injection is counted as found if there is a trigger associated with the injection with a detection statistic louder than the detection statistic of the loudest in-time trigger.

The *physical luminosity* is calculated by creating an "injection" file using a galaxy catalog. The "injections" are distributed according to the locations and luminosities of galaxies in the catalog. More luminous galaxies end up have more "injections" associated with them. The luminosity weight W each "injection" carries is determined by the most luminous galaxy in the catalog, galaxy j. W is given by

$$W = N_j/L_j \ , \tag{9.12}$$

where N_j is the number of "injections" in galaxy j, and L_j is the luminosity of galaxy j. The physical luminosity L_i in distance bin i is then given as

$$L_i = n_i W \ , \tag{9.13}$$

where n_i is the number of "injections" from all galaxies with distance $D \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$.

The *derivative of the cumulative luminosity* is computed numerically by recalculating $\mathcal{C}_L$ using several values of the detection statistic x around the largest value from the in-time triggers.

9.3.2 Background Probability

The *background probability* takes a simple form with the choice of the FARc for the detection statistic. Assuming a Poisson distribution, the probability of getting zero background triggers louder than the loudest in-time trigger (i.e., with a FAR lower than the FAR of the loudest in-time trigger since low FAR are more significant) is given by equation (7.6b) where, as before, $x = -\mathrm{FAR} \times T$, FAR is the FAR of the loudest in-time trigger, and T is the total analyzed time searching for in-time triggers.

Since more significant triggers are given by lower FAR, triggers with higher values of x are louder, so the derivative with respect to the detection statistic in the direction of louder triggers is in the positive x direction (d/dx). The *derivative of the background probability* with respect to the detection statistic in the direction of louder triggers dP_b/dx is

$$\frac{dP_b}{dx} = e^x \; . \tag{9.14}$$

Combining equations (7.6b) and (9.14), we find that $(1/P_0)(dP_0/dx) = 1$, and this term in equation (9.2) goes away.

We therefore can compute all of the quantities in equation (9.1) and 9.2 needed to determine the posterior on the rate μ, before marginalizing over systematic errors.

9.4 Systematic Error Calculation

Systematic errors associated with CBC searches for GW signals include errors associated with detector calibrations, simulation waveforms, Monte Carlo statistics, and galaxy catalog distances and magnitudes. Calculating these errors in terms of the cumulative luminosity is described below [131]. We can marginalize over these uncertainties using equation (9.7), to obtain the final posterior on the rate, and use equation (9.8) to obtain the upper limits. This analysis can be repeated as function of source mass.

9.4.1 Monte Carlo Errors

We refer to statistical errors associated with the efficiency calculation as *Monte Carlo errors.* Since we calculate the efficiency as a function of distance, we calculate the error for a particular distance bin i using the binomial formula, which gives an error of zero when the efficiency is zero or one, or when there are no injections in that bin:

$$\epsilon_{\mathrm{Monte\,Carlo},i} = \sqrt{\frac{N_{\mathrm{found},i}\left(N_{\mathrm{injected},i} - N_{\mathrm{found},i}\right)}{N^3_{\mathrm{injected},i} + 10^{-5}}} \; , \tag{9.15}$$

where $\epsilon_{\mathrm{Monte\,Carlo},i}$ is the error in the efficiency, $N_{\mathrm{injected},i}$ is the number of injections injected into distance bin i, and $N_{\mathrm{found},i}$ is the number of injections found in distance bin i. Finally, the associated fractional luminosity error $\sigma_{\mathrm{Monte\,Carlo}}$ is given in terms of $\epsilon_{\mathrm{Monte\,Carlo},i}$, the physical luminosity in each distance bin L_i, and the measured cumulative luminosity $\mathcal{C}_{L,\mathrm{m}}$ as

$$\sigma_{\mathrm{Monte\,Carlo}} = \frac{1}{\mathcal{C}_{L,\mathrm{m}}} \sum_i \epsilon_{\mathrm{Monte\,Carlo},i} L_i \ . \tag{9.16}$$

9.4.2 Calibration Errors

Calibration errors in the detectors are errors due to uncertainties in the absolute calibration of the detector response to differential displacements caused by GWs [137]. These errors affect the amplitude, and in turn the distance, at which we made injections to calculate the efficiency of our search, since the injections we made assuming a specific value of the noise floor. During S5, the one-sigma uncertainty in the amplitude (and thus the distance) associated with the calibration was 8.1% for H1, 7.2% for H2, and 6.0% for L1 [137]. To calculate the luminosity error due to a calibration error $\Delta_{\mathrm{calibration}}$, all of the injections are moved further in that detector's distance by $\Delta_{\mathrm{calibration}}$

$$D' = D\left(1 + \Delta_{\mathrm{calibration}}\right) \ , \tag{9.17}$$

and the efficiency is recalculated using equation (9.11) where $N_{\mathrm{injected},i}$ is the number of injections injected with distance $D' \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$, $D_{i,\mathrm{min}}$ is the minimum distance in distance bin i, $D_{i,\mathrm{max}}$ is the maximum distance in distance bin i, and $N_{\mathrm{found},i}$ is the number of found injections with distance $D' \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$. This gives us $\epsilon_{\mathrm{calibration},i}$ where the i denotes a particular bin in distance. The standard efficiency ϵ_i is then subtracted from $\epsilon_{\mathrm{calibration},i}$ on a bin-by-bin basis

$$d\epsilon_{\mathrm{calibration},i} = \epsilon_{\mathrm{calibration},i} - \epsilon_i \ . \tag{9.18}$$

Finally, the associated fractional luminosity error $\sigma_{\mathrm{calibration}}$ is given in terms of $d\epsilon_{\mathrm{calibration},i}$, L_i, and $\mathcal{C}_{L,\mathrm{m}}$ as

$$\sigma_{\mathrm{calibration}} = \frac{1}{\mathcal{C}_{L,\mathrm{m}}} \left(\sum_i d\epsilon_{\mathrm{calibration},i} L_i \right) \ . \tag{9.19}$$

9.4.3 Waveform Errors

Waveform errors are associated with how different the true signals are from the signals we use to measure the efficiency of our pipeline (i.e., the mismatch between the true signals and our injections). This error effectively reduces the distances in our efficiency calculation since we do not recover all of the power available in the signal due to the mismatch between the signal and our injections. To calculate this, all of the injections are moved further by the error $\Delta_{\mathrm{waveform}}$

$$D' = D\left(1 + \Delta_{\mathrm{waveform}}\right) \ , \tag{9.20}$$

and the waveform fractional error is calculated as above (section 9.4.2). We calculate the waveform error in units of luminosity assuming a waveform mismatch of 10%. This value is obtained by a detailed comparison of a variety of PN waveform approximants [8, 138], and it represents a conservative estimate of this theoretical error on the true waveform.

9.4.4 Galaxy Errors

Galaxy errors are errors associated with our galaxy catalog [134] used to construct the physical luminosity. Galaxy errors come in two types: *distance errors* and *magnitude errors.*

Future searches will be normalized per unit volume (Mpc^3) to obtain a rate density, rather than per unit L_{10}, in order to avoid these errors (which are not intrinsic to this search method). In the large volume limit (i.e., above a physical distance of 30 Mpc), there is a $0.02\, L_{10}\, \mathrm{Mpc}^{-3}$ conversion that can be used to compare the two rate units.

9.4.4.1 Distance Errors

To calculate the error on the luminosity due to distance errors, the physical luminosity calculation is changed such that the distance to each of the “injections” in a galaxy is increased

$$D' = D\left(1 + \Delta_D\right) \ , \tag{9.21}$$

where Δ_D is the fractional distance error for a given galaxy. Also, the "injections" in a particular galaxy's weighting is changed to be

$$W' = W \left(1 + \Delta_D\right)^2 \ , \tag{9.22}$$

where W is the original weighting for the galaxy. This form of weighting is chosen because the galaxy's luminosity is only known in terms of its magnitude and distance. If there is an error in the distance, the luminosity must change by $\left(1 + \Delta_D\right)^2$ in order to maintain the same magnitude. The luminosity $L_{\mathrm{distance},i}$ is calculated with weighting W' and with "injections" at distances D' as

$$L_{\mathrm{distance},i} = \sum_{i=1}^{n'} W' \ , \tag{9.23}$$

where n' is the number of "injections" from all galaxies with distance $D' \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$. $L_{\mathrm{distance},i}$ is then multiplied by the standard efficiency ϵ_i as a function of distance and integrated over distance to obtain the cumulative luminosity $\mathcal{C}_{L,\mathrm{distance}}$ at 1σ error in distance

$$\mathcal{C}_{L,\mathrm{distance}} = \left(\sum_i \epsilon_i L_{\mathrm{distance},i} \right) \ . \tag{9.24}$$

The associated fractional luminosity error $\sigma_{\mathrm{distance}}$ in terms of $\mathcal{C}_{L,\mathrm{m}}$ as

$$\sigma_{\mathrm{distance}} = \frac{1}{\mathcal{C}_{L,\mathrm{m}}} \left(\mathcal{C}_{L,\mathrm{distance}} - 1 \right) \ . \tag{9.25}$$

9.4.4.2 Magnitude Errors

To calculate the error on the luminosity due to magnitude errors, the physical luminosity calculation is changed such that the "injections" in a particular galaxy are weighted by

$$W' = W 10^{\Delta_M / 2.5} \ , \tag{9.26}$$

where Δ_M is the (absolute) magnitude error for a given galaxy. This follows from the formula used to calculate the luminosity of a galaxy

$$L_{B,i} = L_{B,\odot} \left(\frac{D_i}{10\mathrm{pc}} \right)^2 10^{(M_{B,\odot} - m_{B,i})/2.5} \,, \tag{9.27}$$

where $L_{B,\odot}$ is the blue solar luminosity, D_i is the distance to the galaxy, $M_{B,\odot}$ is the (absolute) blue solar magnitude, and $m_{B,i}$ is the (measured) apparent blue magnitude of the galaxy. The luminosity $L_{\mathrm{magnitude},i}$ is calculated with weighting W' as

$$L_{\mathrm{magnitude},i} = \sum_{i=1}^{n} W' \,, \tag{9.28}$$

where n is the number of "injections" from all galaxies with distance $D \in (D_{i,\mathrm{min}}, D_{i,\mathrm{max}})$. The magnitude fractional error is calculated as above (section 9.4.4.1).

Chapter 10

Results of the Search for Low Mass Compact Binary Coalescences in the First Year of Data from LIGO's Fifth Science Run

In this chapter we describe the results from the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data. In section 10.1 we discuss the follow-up procedures we use to build confidence in a possible GW event. We discuss the loudest events from the search in section 10.2. Finally, we present upper limits on the rate of coalescences in section 10.3.

During the analysis, and prior to unblinding the nonplayground data, we discovered an error in the computation of the template metric. This metric is used in the placement of the bank and the coincidence test. The error caused the metric distance between templates to be overestimated for the higher mass signals. This has the effect of causing the template placement algorithm to over-cover the higher mass region (i.e., to produce a bank with less than the requested 3% loss in SNR). This increased the computational cost of the search, but did not significantly reduce the sensitivity. However, this error also affected the coincidence algorithm by overestimating the distance between triggers for high mass signals. Since the coincidence window was empirically tuned on software injections and time-shifted coincidences the impact on the sensitivity of the search was not significant. Consequently, the decision was taken to unblind the data using the original, suboptimal analysis in order to begin studying any possible detection candidates and to use this result to compute the upper limit (in the absence of a detection). The decision was also taken to perform a complete reanalysis of the data with the corrected metric to verify the (non)detection statement from the original search by checking for any change in the list of the loudest in-time coincident triggers (the detection candidates).

10.1 Follow-up Procedure for Coincident Triggers

The distribution of IFARc in-time coincident triggers is examined to identify the events with the largest IFARc (the "loudest" events). Events with a FARc much smaller than 1 per observation time are deemed significant. However, this takes into account only the GW channel data filtered by the search pipeline. Information from other channels are only used coarsely in the definition of the category 2 and 3 DQ flags (section 6.8.1). Much more information is available in both the GW channel and the auxiliary channels, but it is not yet automatically folded into the detection statistic (this is a work in progress, which will be discussed more in section 11.1). We therefore have developed a rather detailed follow-up procedure to examine this additional information in detail, albeit, in a necessarily subjective manner. Automating this procedure for time-shifts and simulated GW signals (i.e., injections) to provide an improved, unbiased detection statistic is in progress.

In the meantime, we check our loudest coincident triggers with a checklist of tests designed to see if a statistically significant trigger is believable as a detection candidate. The methods employed in this checklist are tested against injection and time-shift coincidence triggers. Reference [139] describes the tests that we perform on the coincident triggers and the data surrounding them. At present, our standard tests include the following: We check the integrity of the data for corruption. We also check the status of the detectors and the presence of any data quality flags in the surrounding data. We assess whether there could have been environmental or instrumental causes found in auxiliary channels at the time of the trigger. We check the appearance of the data at the time of the trigger in the form of the strain, SNR, and χ^2 time-series, and time-frequency spectrograms (an example is shown in figure 10.1).

In addition, for any statistically significant candidate that survives the tests listed above, we plan

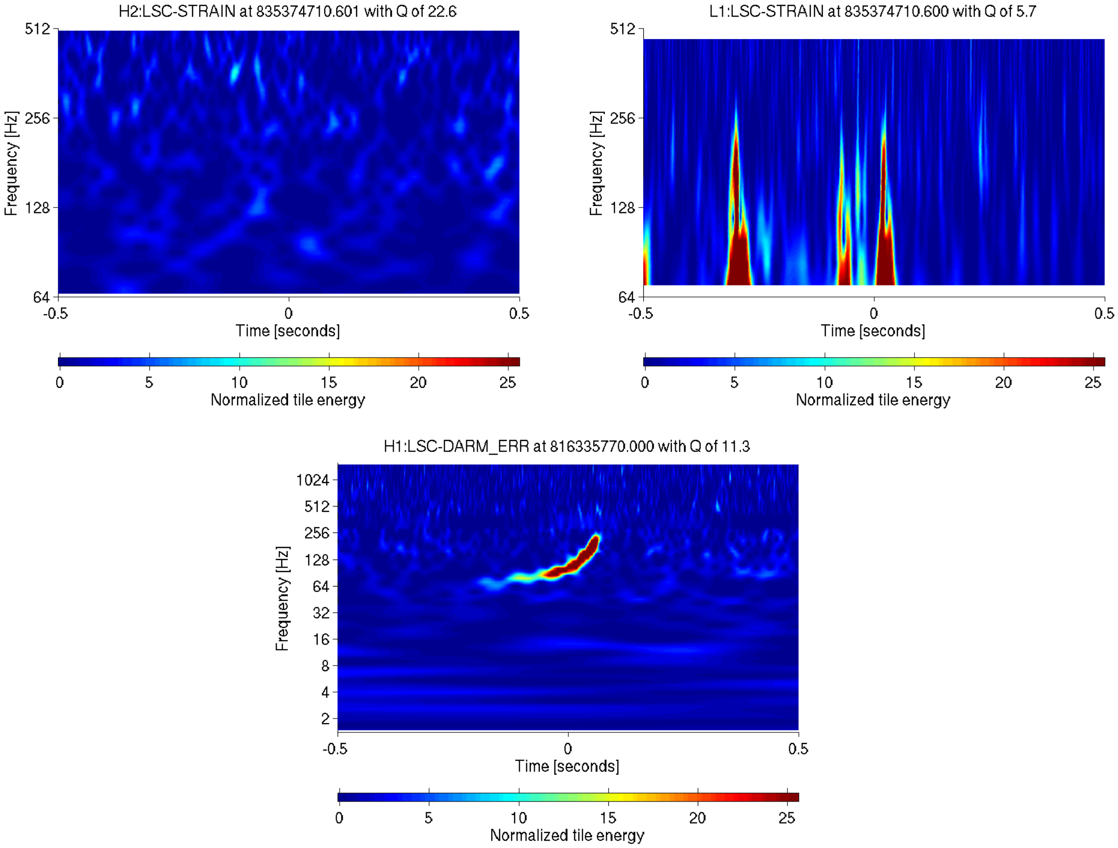

Figure 10.1: Comparison of Triggers from Loudest Event to an Injected Signal
A "Qscan" of the triggers in the loudest event (top) in this search after category 2 and 3 DQ flag vetoes have been applied and of an signal injected into the data (bottom). The Qscans show time-frequency content in the form of power in a Sine-Gaussian basis.

to do the following: Assess the coherence between the signals recorded by each individual detector operating at the time of the event. Verify the robustness of the trigger against small changes in the pipeline (i.e., changes in the adjacent Fourier transform boundaries or changes in the calibration of the data). Check the robustness across pipelines by employing other search techniques to analyze the same data (i.e., CBC pipelines using different templates or pipelines designed to search for unmodeled bursts). Finally, we will check for coincidence with external searches for gamma-ray bursts, optical transients, or neutrino events. (This last test is for information only, as a genuine GW event might or might not be accompanied by other signals.)

We examine the distribution of in-time triggers after each category of DQ flag veto is applied. In case there is a statistically significant outlier before category 2 or categories 2 and 3 DQ flag vetoes are applied, we carry out a follow-up exercise to see if the veto that eventually rejected the event was rightfully applied. There are two reasons that this could be important. First, a very strong GW from within the Milky Way could cause an instrumental saturation of the sort that we use as a veto; this kind of problem would be easy to diagnose if it were to occur, since the signal would be strong enough for us to see in the moments leading up to the signal-induced saturation. Second, we want to guard against false dismissal of a candidate by other kinds of vetoes, which can have nonnegligible dead-time associated with them. Some of our vetoes are associated with recognizable forms of false signals; we check to be sure that a vetoed loud event looks like that kind of false signal, and not like a genuine coalescence signal. In this search, there was a single statistically significant outlier in the distribution of events before the application of categories 2 and 3 DQ flag vetoes. The follow-up exercise confirmed that the category 3 DQ flag that vetoed that event was correctly applied, as it occurred during a time that was especially glitchy and was recommended to be removed from analysis as the data was taken.

10.2 Loudest Events

At the end of our pipeline we are left with a set of coincident triggers that are potential detection candidates. The cumulative distribution of events above an IFARc versus IFARc for the different observation times is shown in figure 10.2; the H1H2L1 distribution is also shown in figure 7.10. This figure shows that the loudest candidates in all three observation times were consistent with the estimated background and thus were likely accidental coincidences. Thus, the search yielded no detection candidates. The results of the reanalysis were consistent with the original analysis and did not produce any plausible GW signals. We report an upper limit on the rate of coalescences in section 10.3.

As an exercise to prepare for future detections, we carried the loudest several events (such as the three loudest events that appear in each of the histograms in figure 10.2) through the follow-up procedure.

Even though we know our background is underestimated for H1H2 coincident triggers, we re-viewed the two loudest H1H2 candidates using the detection checklist. In both of those cases, the waveforms from the two interferometers failed to match each other in detail, thus ruling them out as GW events.

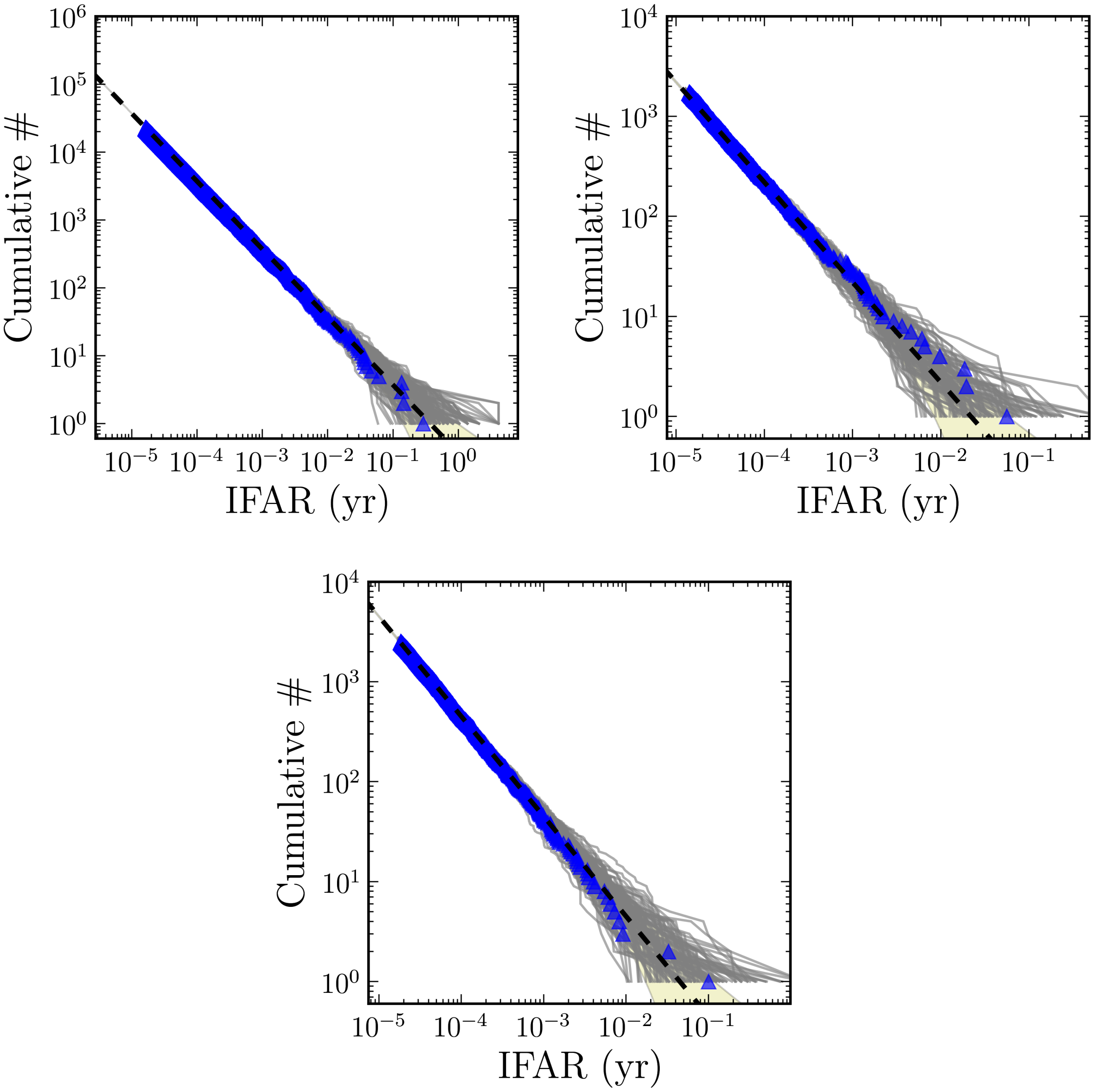

Figure 10.2: IFAR Loudest Events
The cumulative distribution of events above a threshold IFAR, for in-time coincident events, shown as blue triangles, from all coincidence categories for the observation times H1H2L1 (top left), H1L1 (top right), and H2L1 (bottom right). The expected background (by definition) is shown as a dashed black line. The 100 experimental trials that make up our background are also plotted individually as the solid grey lines. The yellow band shows the $N^{1/2}$ one-standard deviation errors on the average background estimate.

Table 10.1: Detailed Results of the BNS Upper Limit Calculation

BNS	$D_{\text{horizon}} = 30\,\text{Mpc}$		
Coincidence Time	H1H2L1	H1L1	H2L1
Cumulative Luminosity (L_{10})	250	230	120
Calibration Error	21%	3.9%	16%
Monte Carlo Error	5.4%	16%	13%
Waveform Error	26%	11%	20%
Galaxy Distance Error	14%	13%	6.1%
Galaxy Magnitude Error	17%	17%	16%
Λ [Eq. 9.2]	0.30	0.41	0.72
Marginalized Upper Limit $\left(\text{yr}^{-1} L_{10}^{-1}\right)$	3.9×10^{-2}		

Summary of the search for BNS systems. The horizon distance, as defined in Eq. 5.38, for the 4 km detectors and averaged over the search. The cumulative luminosity is rounded to two significant figures. The errors in this table are listed as logarithmic errors in the luminosity multiplier based on the cited sources of error.

Table 10.2: Detailed Results of the BBH Upper Limit Calculation

BBH	$D_{\text{horizon}} = 80\,\text{Mpc}$		
Coincidence Time	H1H2L1	H1L1	H2L1
Cumulative Luminosity (L_{10})	4900	4900	1800
Calibration Error	23%	36%	25%
Monte Carlo Error	3.2%	11%	9.9%
Waveform Error	27%	34%	27%
Galaxy Distance Error	18%	18%	20%
Galaxy Magnitude Error	16%	16%	17%
Λ [See (9.2)]	0.59	1.2	1.4
Marginalized Upper Limit $\left(\text{yr}^{-1} L_{10}^{-1}\right)$	2.5×10^{-3}		

Summary of the search for BBH systems. The rows of this table are defined as in Table 10.1.

10.3 Coalescence Rate Upper Limits

In the absence of detection, we set upper limits on the rate of CBCs per unit L_{10}, for several canonical binary masses, as a function of the total mass of the compact binary system, and as a function of the black hole mass for BHNS systems.

For each mass range of interest, we calculate the rate upper limit at 90% confidence level (CL) using the loudest event formalism [133, 130], described in chapter 9. We derive a Bayesian posterior distribution for the rate as described in chapter 9 and reference [130]. The cumulative luminosity for this search, calculated as described in chapter 9, can be found in Table 10.1, 10.2, and 10.3 for the BNS, BBH, and BHNS canonical mass CBC systems, respectively.

Table 10.3: Detailed Results of the BHNS Upper Limit Calculation

BHNS	$D_{\text{horizon}} = 50\,\text{Mpc}$		
Coincidence Time	H1H2L1	H1L1	H2L1
Cumulative Luminosity (L_{10})	990	980	390
Calibration Error	22%	47%	25%
Monte Carlo Error	3.6%	9.7%	11%
Waveform Error	26%	40%	27%
Galaxy Distance Error	17%	16%	17%
Galaxy Magnitude Error	17%	18%	18%
Λ [See (9.2)]	0.45	1.0	1.1
Marginalized Upper Limit $\left(\text{yr}^{-1} L_{10}^{-1}\right)$	1.1×10^{-2}		

Summary of the search for BHNS systems. The rows of this table are defined as in Table 10.1.

10.3.1 Upper Limits Neglecting Spin

We apply the upper limit calculation to three canonical binary masses as well as calculating the upper limit as two functions of mass. Our three canonical binary masses are BNS ($m_1 = m_2 = (1.35 \pm 0.04)$ $M_\odot$), BBH ($m_1 = m_2 = (5 \pm 1)$ $M_\odot$), and BHNS ($m_1 = (5 \pm 1)$ $M_\odot$, $m_2 = (1.35 \pm 0.04)$ $M_\odot$). We use Gaussian distributions in component mass centered on these masses, with standard deviations following the ± symbols. The neutron star mass distribution was chosen according to an analysis of 50 known pulsars in binary systems at the time (reference [140]). However, the mass distribution for black holes is still uncertain, therefore we have chosen a canonical value to represent stellar mass black holes, which allows us to compare results across different science runs.

We combine the results of this search from the three different observation times in a Bayesian manner, described in section 9.1, and the results from previous science runs [21, 22] are incorporated in a similar way.

Assuming that spin is not important in these systems, we calculate upper limits on the rate of binary coalescences using our injection families that neglect spin (section 6.2). There are a number of uncertainties that affect the upper limit, including systematic errors associated with detector calibration, simulation waveforms, Monte Carlo statistics, and galaxy catalog distances and magnitudes [131]. We marginalize over these as described in section 9.4 and obtain upper limits on the rate of binary coalescences of

$$\mathcal{R}_{90\%,\mathrm{BNS}} = 3.9 \times 10^{-2}\,\mathrm{yr}^{-1}{\mathrm{L}_{10}}^{-1}\ , \tag{10.1a}$$

$$\mathcal{R}_{90\%,\mathrm{BBH}} = 2.5 \times 10^{-3}\,\mathrm{yr}^{-1}{\mathrm{L}_{10}}^{-1}\ , \tag{10.1b}$$

$$\mathcal{R}_{90\%,\mathrm{BHNS}} = 1.1 \times 10^{-2}\,\mathrm{yr}^{-1}{\mathrm{L}_{10}}^{-1}\ . \tag{10.1c}$$

We also calculate upper limits for two additional cases: as a function of the total mass of the binary, with a uniform distribution in the mass ratio $q = m1/m2$, and as a function of the mass of the black hole in a BHNS system, holding fixed the mass of the neutron star at $m_{\mathrm{NS}} = 1.35$ $M_\odot$(figure 10.3).

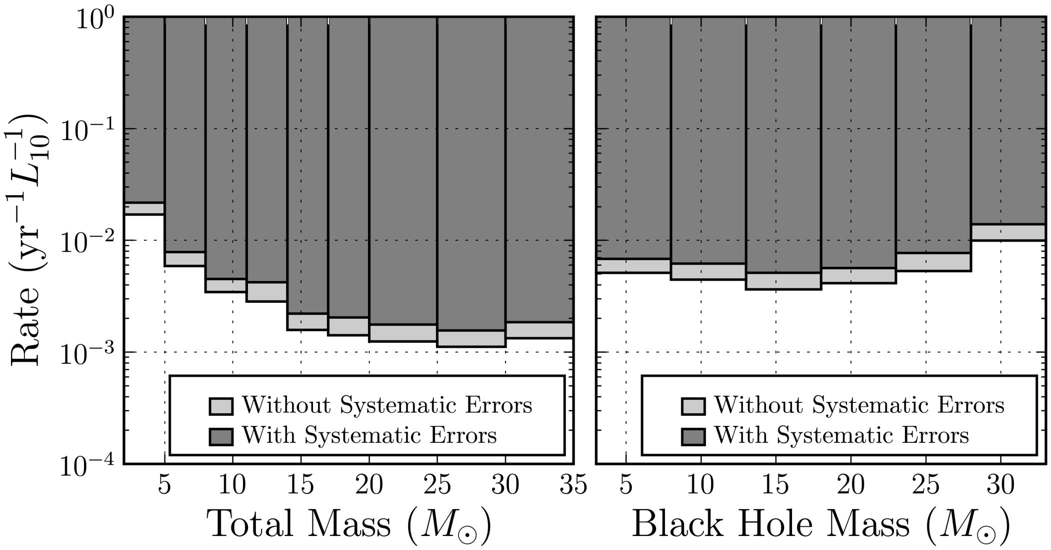

Figure 10.3: Upper Limits versus Mass
Upper limits on the binary coalescence rate per year and per L_{10} as a function of total mass of the binary system with a uniform distribution in the mass ratio (left) and as a function of the mass of a black hole in a BHNS system with a neutron star mass of $m_{\mathrm{NS}} = 1.35\ M_{\odot}$(right). The darker area shows the excluded region after marginalization over the estimated systematic errors. The lighter area shows the additional region that would have been excluded if the systematic errors of our most sensitive observation time had been ignored.

10.3.2 Upper Limits Including Spin

In the previous section, we reported upper limits on the rate of mergers for different classes of objects using injection waveforms generated assuming nonspinning objects. We can also evaluate the upper limits using injection waveforms that take into account the effects of spinning bodies.

Since the maximum possible rotational angular momentum S for a black hole of mass m is Gm^2/c, it is useful to describe the spin of a compact object in terms of the dimensionless spin parameter $\hat{a} = (cS) / \left(Gm^2\right)$. The distribution of black hole spin magnitudes within the range $0 \leq \hat{a} \leq 1$, as well as their orientations relative to binary orbits, is not well constrained by observations. To illustrate the possible effects of BH spins on our sensitivity to BBH and BHNS signals, we provide an example calculation using a set of injections using the SpinTaylor approximant, which calculates waveforms for systems whose component objects can include significant spin angular momentum. The spin magnitude $\hat{a}$ for this set of injections is uniformly distributed between 0 and 1 (section 6.2). However, assuming a canonical mass and uniform density for a component object, astrophysical observations of neutron stars show typical angular momenta corresponding to $\hat{a} \ll 1$ [141]. In addition, the spin effects are found to be weak for the frequency range of interest for LIGO [16] so the BNS upper limits in section 10.3.1 are valid even though we have ignored the effects of spin.

Using the above injections, we obtain marginalized upper limits on the rate of binary coalescences of

$$\mathcal{R}_{90\%,\mathrm{BBH}} = 3.2 \times 10^{-3}\,\mathrm{yr}^{-1}{\mathrm{L}_{10}}^{-1}\ , \tag{10.2}$$

$$\mathcal{R}_{90\%,\mathrm{BHNS}} = 1.4 \times 10^{-2}\,\mathrm{yr}^{-1}{\mathrm{L}_{10}}^{-1}\ . \tag{10.3}$$

We can see that the effect of spin on the relatively low mass systems targeted in this search is rather small, modifying the average distance we can see by a factor of order ∼5%–10%, and thus our upper limit by a factor that is three times larger (since the volume we are sensitive to is proportional to the cube of the distance) of order ∼15%–30%.

10.4 Summary of the Search

We can compare the above results to a prediction made for the upper limit on the rate of BNS coalescences from reference [131]. The example upper limit in that reference is computed assuming a horizon distance of 35 Mpc for the detector, an observation time of 1 yr, and a loudest event from the analysis with a combined SNR of 10. The result they predict, after marginalization over systemic errors is $7.3 \times 10^{-3}\,\mathrm{yr}^{-1}\,{\mathrm{L}_{10}}^{-1}$. If we adjust the parameters above to match the sensitivity (average horizon distance of 30 Mpc) and duration (0.37 yr of analyzed data) of this search, the predicted result becomes $3.1 \times 10^{-2}\,\mathrm{yr}^{-1}\,{\mathrm{L}_{10}}^{-1}$, which is very close to the upper limit we obtain in equation (10.1a).

Aside from possibly detecting GWs, this is what Initial LIGO was proposed to achieve and what it was built for. S5 was the first science run since the LIGO detectors reached their design sensitivity and also the first long science run LIGO has undertaken. The search for BNS GW signals, the flagship search, concluded with a result as sensitive as expected, a great achievement for LIGO and the LIGO Scientific Collaboration (LSC).

We can also extend this prediction to BBH systems by adjusting the predicted result by the factor $(M_{\mathrm{BNS}}/M_{\mathrm{BBH}})^{5/2}$, which comes from the mass dependence of the amplitude of a CBC GW in equation (2.43). This results in a predicted result for BBH systems of $1.2 \times 10^{-3}\,\mathrm{yr}^{-1}\,{\mathrm{L}_{10}}^{-1}$. This is further from the obtained result in equation (10.1b) than the BNS comparison, however this is expected due to the higher FAR for higher mass templates. This shows that the IFARc detection statistic prevents the contamination of the lower FAR for lower mass templates from the higher FAR from the higher mass templates.

Chapter 11

Future Improvements for Compact Binary Coalescence Searches

In previous chapters we focused on the procedures used for the Search for Low Mass CBCs in the First Year of LIGO's Fifth Science Run (S5) Data, referred to as "this search," including data analysis improvements that were made based on the collective experience of analyzing previous science runs' data. In this chapter, we focus on such improvements that are yet to be made, or in the process of being implemented within the LSC CBC Group that will further benefit future searches. These include improvements in separating signals from noise with better detection statistics (section 11.1), better estimating the background of a particular search (section 11.2), coherently combining the data from multiple detectors (section 11.3), cleaning the data before analysis to remove detector glitches (section 11.4), and lowering the latency of the searches to improve the scientific output (section 11.5).

11.1 Detection Statistic

The goal of developing new detection statistics is to get to the *ideal detector response to Gaussian noise*. We are already close to this ideal for the low mass search, and in particular the low mass portion of the low mass search. This is because the low mass waveforms a more broadband than the high mass waveforms, which makes the standard SNR calculation relatively insensitive to detector glitches. As was seen in this search's waveforms, including information about the background noise of the detectors allowed us to develop a detection statistic that prevents the contamination of the low mass portion of parameter space by the inclusion of the high mass portion.

The addition of the background information can be seen as a step toward defining a likelihood ratio for the detection statistic. The likelihood ratio Λ is the ratio of the probability that a particular trigger c was associated with a true signal $P(c|h)$ and the probability that a particular trigger was due to background noise alone $P(c|0)$ given by

$$\Lambda = \frac{P(c|h)}{P(c|0)} \ . \tag{11.1}$$

As we saw in section 9.3.2, the FARc detection statistic can be converted to a FAP (i.e., the probability of getting any background triggers louder than that trigger) using equation (7.6a).

The final step toward a likelihood ratio detection statistic is calculating $P(c|h)$ using the injected signals. This procedure is being demonstrated to gain additional sensitivity in the triggered search for CBC signals associated with short GRBs [142]. The reason for this is that including the $P(c|h)$ term introduces more information that we can use to distinguish noise triggers from signal triggers. However, there is a downside of introducing this term. As in calculating $P(c|0)$ where included information about how the pipeline responded to noise in the GW channel, for $P(c|h)$ we include information about how the pipeline responds to signals in the GW channel. However, there is a difference between the two calculations. For the former, there was a well defined way to characterize the noise in the detectors (i.e., time shifting the data before analysis). Whereas for the latter, since we do not know the astrophysical distribution of signals we end up making a choice for the distributions of signals. The resulting likelihood ratio detection statistic then depends more strongly on those assumptions of the signal distributions than the previous detection statistic.

A different approach toward an improved detection statistic is to use a multivariate classifier to rank order the triggers. Such a classifier develops a classification algorithm by taking in a large number of parameters (such as the many recovered parameters of the coincident triggers we record and information from auxiliary channels if the detectors). It does this for both background triggers and injection triggers and then looks for how the two classes of triggers separate in the multidimensional space. Different classifications of this sort are currently being tested using triggers from the analysis of this search. This technique can be powerful because it allows easy inclusion even more information beyond that included in the present detection statistic, ρ_{eff} (equation (5.40)), and the associated background and signal probability factors $P(c|0)$ and $P(c|h)$.

11.2 Background Estimation

One of the main questions we ask at the end of a search is "what is the probability that a particular trigger came from the background noise?" As described above, the FARc detection statistic has a simple answer to this question, however the current procedure of calculating the FARc is limited in the smallest FAP it can estimate, when there is only one background trigger with significance equal to or greater than the trigger in question. Naïvely, since we perform 100 time-shifted analyses to estimate the background, we would expect to be able to estimate a minimum FAP of 1%. Unfortunately this was under the assumption that we had a single experiment, while in fact the FARc calculation combines the results of nine different trigger categories, which are effectively different experiments. We then have to multiply the minimum FAP by this trials factor giving us a minimum FAP that we can estimate of 9%.

One way to get around this problem and lower the minimum FAP is to increase the number of time-shifts we perform. However there are two problems associated with this. First of all, we want to estimate the background for triggers with roughly the same combination of detector noise conditions. This means that we cannot shift the data by too large an amount, otherwise the noise characteristics from multiple detectors will no longer be overlapping. Since there are time correlations in the data from a single detector, we need to shift the data of multiple detectors by a minimum amount so we do not include those correlations in our background estimation. The combination of the two effects gives us a maximum number of time-shifts we can perform. Second, on taking into account practical considerations, as we increase the number of shifts, the amount of time it takes to analyze the data increases roughly by the same factor.

A different way to estimate the background is to build a model of the background that we could then use to extrapolate to low FAPs. One way to do this is to produce a joint probability distribution function (PDF) for the background starting from the PDFs from the original single-detector triggers. As stated in chapter 5, for each detector the PDF of the SNR squared can be modelled as a χ^2 distribution with two degrees of freedom for the Gaussian noise component. Since real data also contains a nonGaussian component that dominates in the tail of the distribution, we would add a Poisson distribution to capture these tails. To factor in the effect of signal-based vetoes, we would need to run those portions of the pipeline with and without the vetoes to see their effect in each detector. We would then multiply the PDF of each detector by the ratio of the two to get the PDF including vetoes. Finally, since we would want to rank order the triggers according to effective SNR, we would construct the theoretical distribution for the value of the χ^2 veto (i.e., χ^2 distribution with $2p-2$ degrees of freedom and noncentrality parameter given by that for glitches), and combine it with the SNR PDF to obtain an effective SNR PDF for single detectors. The joint PDF for coincident triggers could then be taken as the outer product of the single-detector distributions, normalized such that its integral gives a rate that matches the number of triggers in the time-shifted analyses divided by the sum of the time-shifted analyses observation times.

11.3 Coherent Analysis

Some of the more powerful vetoes we use in analyzing GW-detector data are the amplitude consistency vetoes between interferometers. These vetoes ensure that the strength of a trigger is consistent between different detectors. However, there is also phase information that can be compared. This additional information can be included by matched filtering a coherent data stream, as described in section 3.6 and references [143, 144].

11.4 Data Cleaning and Noise Regression

Currently, all of the loudest triggers produced by analyzing LIGO data are the result of glitches in the detectors. As described in section 6.8.1, we remove times during which we expect the detectors to be glitching. Currently, as the signal processing we use for templated CBC searches introduces artifacts that can last for a few seconds around a glitch, there are wide windows we must add to these times to ensure that triggers resulting from such glitches are removed. These windows can make up a significant portion of each time we wish to veto, so trying to reduce or eliminate these windows could help save additional analysis time. One way to do this is to switch over from vetoing such times to eliminating the glitches from the data processing. This could happen if instead of vetoing the required times, plus their windows, the required times' data was excised and replaced with the appropriately colored Gaussian noise such that the transients from the excision boundaries were minimized. These times could then be flagged as not involving real data but would allow the analysis of adjacent times.

Another option, which would be much better when possible, would be the regression of noise using information from the auxiliary data channels. For situations where the transfer function between the auxiliary data channel and the strain data channel are known, this would allow excess noise from these sources to be virtually eliminated from strain data and improving sensitivity. For channels with a time-varying coupling to the strain data that is recorded, adaptive filtering could be used to eliminate both excess noise as well as transients caused by an increase and then decrease of the coupling.

11.5 Low Latency Searches

Improving the latency of CBC searches for GWs would have a significant scientific impact. Currently, there is an inherent latency in the trigger generation of roughly 30 minutes, since the analysis proceeds in batch mode analyzing 2048 seconds of data at a time (chapter 6). This is a long enough time that any electromagnetic counterparts could have faded into the background by the time telescopes received an announcement of a likely GW. There are several techniques that are being utilized to reduce this latency. References [145, 146] take advantage of the chirp-like nature of CBC GW signals to break the signal up into different bands, which can be combined to generate a trigger with an SNR that is roughly the same as the current full band analysis. These techniques can reduce the latency of the search to a few seconds. Additionally, for BNS signals in Adv. LIGO, if only the lowest frequency bands are used, a trigger for a GW could be produced a few minutes *before* the arrival of the signal from the neutron stars' collision, alerting the LIGO detectors to remain online for the next few minutes, potentially even changing the response of the detectors to the high frequency portion of the inspiral by dynamically detuning the signal extraction optical cavity, and alerting electromagnetic telescopes where and when to look, resulting in multimessenger observations of the same event.

Chapter 12

Testing Effects of a Massive Graviton

The General Theory of Relativity has proven to be very predictive over the past century. As esoteric as it may seem, its effects can impact our every day lives, such as the General Relativistic corrections needed for GPS navigation to work. However, we know the predictions of General Relativity must break down at some small length scale since quantum effects will become important at the Planck scale.

General Relativity could also have flaws at large scales. It has also been observed that stars within galaxies and galaxies within galactic clusters orbit around the center of mass faster than they should based on the amount of visible matter [147, 148]. There are one of two conclusions that can be drawn from this, either there is more matter there than we can see, or General Relativity does not describe gravity at large distances.

Many people currently believe the former of these two options, positing the existence of dark matter to explain the missing mass. The latter option requires a modification to General Relativity at large distances, which could be accounted for in the form of a massive graviton.

In this chapter we will compare the analysis of Will [149] using 1.5PN waveforms to a new analysis using complete inspiral-merger-ringdown (IMR) waveforms. The rest of this chapter is as follows: we briefly summarize the effects of a massive graviton (section 12.1), give a dispersion relation for GWs in general (section 12.2) and specific to the analysis of GWs from CBC signals (section 12.3 and 12.4), use this dispersion relation to predict theoretical bounds for the mass of the graviton from future CBC GW observations (section 12.5), and then compare these bounds to current observational bounds and theoretical bounds from other techniques (section 12.7).

12.1 Effects of Massive Gravitons

Endowing the graviton with mass would have many affects on different physical processes. These include changing the dispersion relation for gravitational waves such that they no longer travel at the speed of light [149], increasing the energy loss due to gravitational waves of a given frequency [150], and causing the Newtonian gravitational potential to have a Yukawa form instead of being $1/r$ which can affect the orbits of planets [149] and gravitational lensing observations [151]. The rest of this chapter will focus on the first of these effects expect the final section (section 12.7) where we compare these bounds to the bounds from other methods.

12.2 Dispersion of Gravitational Waves

In reference [149], Will derived the propagation effects for two GWs traveling over cosmological distances, which we repeat below. Let us distinguish these GWs using the presence or absence of a “$'$” on the properties of the two waves. Let us also denote properties of either wave at the time of emission with an e subscript and at the time of arrival or measurement with an a subscript. The emitted energies of these GWs is then given as E_e and E'_e and the emission time difference as $\Delta t_e \equiv t'_e - t_e$. For clarity, we also recover the G and c terms for formulae in this chapter.

Let us assume the GWs propagate in a flat Friedman-Robertson-Walker homogeneous and isotropic spacetime, with background metric given as

$$ds^2 = -c^2 dt^2 + a^2(t)\left[dr^2 + r^2\left(d\theta^a + \sin^2\theta d\phi^2\right)\right] \, , \tag{12.1}$$

where $a(t)$ is the expansion factor of the universe.

Consider a graviton moving radially with mass m, 4-velocity $u^\alpha = dx^\alpha/d\tau$, and 4-momentum

$p^\alpha = mu^\alpha$. Let it be emitted at a distance $r = r_e$ and detected at $r = 0$. The energy-momentum relation is given as

$$m^2c^4 = -p^\alpha p^\beta g_{\alpha\beta} = E^2 - a^{-2}p_r^2c^2 \ , \tag{12.2}$$

where $E = p^0c$ is the energy of the particle, m is the mass of the particle, and p_r is the particles radial momentum. Combining this with the ratio $p^rc/E = p^r/p^0 = dr/dt$, after eliminating E we find

$$\frac{dr}{dt} = -\frac{c}{a}\left(1 + \frac{a^2m_g^2c^2}{p_r^2}\right)^{-1/2} \ . \tag{12.3}$$

Since the graviton will be following a geodesic, its 4-velocity will satisfy the geodesic equation $u^\alpha{}_{;\mu}u^\mu = 0$. Using the property of the metric that $g^{\mu\nu}{}_{;\alpha} = 0$, we also know that $u_{\alpha;\mu}u^\mu = 0$. Expanding the covariant derivative using a Christoffel symbol, this is then given by $u_{\alpha,\mu}u^\mu - u_\nu u^\mu \Gamma^\nu{}_{\alpha\mu} = 0$. Expanding the Christoffel symbol and using its symmetries, we can write this as $u_{\alpha,\mu}u^\mu - u^\mu u^\nu g_{\mu\nu,\alpha} = 0$. Specializing this to u_r and using the fact that only the time and radial components of the 4-velocity u^α are nonzero, and both $g_{0\mu}$ and $g_{r\mu}$ are independent of r, we then find $u_{\alpha,\mu}u^\mu = 0$, which shows u^r to be conserved. This then implies that p_r is a conserved component of the 4-momentum. We choose to evaluate it at the moment of emission of the GW, $p_r^2c^2 = a^2(t_e)\left(E_e^2 - m_g^2c^4\right)$.

Using the relation from quantum mechanics between a wave's frequency f and its energy $E = hf$, where h is Planck's constant, and the definition of the Compton wavelength of a particle $\lambda \equiv h/m$, we find that $m_gc^2/E_e = c/(\lambda_g f_e)$. Let us assume that the lowest frequency we are interested in is such that $E_e \gg m_gc^2$. Then the dispersion relation above can be expanded to second order in $c/(\lambda_g f_e)$ as

$$\frac{dr}{dt} = -\frac{c}{a}\left(1 - \frac{c^2a^2}{2a^2(t_e)\lambda_g^2 f_e^2}\right) \ , \tag{12.4a}$$

which after integration gives

$$r_e = \int_{t_e}^{t_a} \frac{c}{a(t)}dt - \frac{c^3}{2a^2(t_e)\lambda_g^2 f_e^2}\int_{t_e}^{t_a} a(t)dt \ . \tag{12.4b}$$

Let us apply this to the two GW emitted with energies E_e and E_e' with a time difference of

$\Delta t_e \equiv t'_e - t_e$. Application to the first wave yields

$$r_e = \int_{t_e}^{t_a} \frac{c}{a(t)} dt - \frac{c^3}{2a^2(t_e)\lambda_g^2 f_e^2} \int_{t_e}^{t_a} a(t) dt \ , \tag{12.5a}$$

while application to the second yields

$$r_e = \int_{t_e+\Delta t_e}^{t_a+\Delta t_a} \frac{c}{a(t)} dt - \frac{c^3}{2a^2(t_e)\lambda_g^2 f'^2_e} \int_{t_e+\Delta t_e}^{t_a+\Delta t_a} a(t) dt \ . \tag{12.5b}$$

Combining these two results gives us

$$\begin{aligned} \int_{t_a}^{t_a+\Delta t_a} \frac{c}{a(t)} dt - \int_{t_e}^{t_e+\Delta t_e} \frac{c}{a(t)} dt \quad &= \quad \frac{c^3}{2\lambda_g^2} \left(\frac{1}{a^2(t_e+\Delta t_e) f'^2_e} \int_{t_e+\Delta t_e}^{t_a+\Delta t_a} a(t) dt \right. \\ &\quad \left. - \frac{1}{a^2(t_e) f_e^2} \int_{t_e}^{t_a} a(t) dt \right) \ . \end{aligned} \tag{12.6}$$

This can be simplified even further in the case where the characteristic time it takes a to change is large compared to the emission time difference, characterized by $\Delta t_e \ll a/\dot{a}$, and where the time difference between t_e and t_a is large to the form

$$\Delta t_a = (1+Z) \left[\Delta t_e + \frac{Dc}{2\lambda_g^2} \left(\frac{1}{f'^2_e} - \frac{1}{f_e^2} \right) \right] \ , \tag{12.7}$$

where $Z \equiv a_0/a(t_e) - 1$ is the cosmological redshift, and D is a distance given by

$$D \equiv \frac{(1+Z)c}{a_0} \int_{t_e}^{t_a} a(t) dt \ , \tag{12.8}$$

where $a_0 \equiv a(t_a)$ is the present value of the expansion factor. D is different from standard cosmological distances, such as the luminosity distance $D_L \equiv a_0(1+Z)\chi_e$. For a matter-dominated spatially-flat universe, D and D_L are given by

$$D = \frac{2c}{5H_0}(1+Z) \left[1 - (1+Z)^{-5/2} \right] \ , \tag{12.9a}$$

$$D_L = \frac{2c}{H_0}(1+Z)\left[1-(1+Z)^{-1/2}\right] \ , \tag{12.9b}$$

where H_0 is the Hubble constant given by $c/H_0 \simeq 42$Gpc. Both cases reduce to the familiar Hubble law cZ/H_0 in the limit $Z \ll 1$. It should also be noted that the frequency upon arrival to the detector will have undergone a redshift such that $f_a = f_e/(1+Z)$.

12.3 Dispersion of Compact Binary Coalescence Gravitational Waveforms

This dispersion relation above, equation (12.4), tells us that GW of different frequencies will travel at different speeds. This can change the morphology of GW signals emitted from a single source that originally were emitted with a well-defined frequency evolution.

With this in mind, let us analyze what happens to the inspiral waveform we calculated in chapter 2, where the orbital phase evolution $\Phi(t)$ is given in equation (2.36b). Taking the time derivative of the phase evolution $2\Phi(t)$, we find the frequency evolution $\pi f(t) = d\Phi/dt$ to be

$$f(t) = \left(\frac{GM}{c^3\pi^2}\right)^{1/2}\left[\frac{5c^9}{256G^3M^3\eta\left(\frac{405GM}{16c^3\eta}+t_{\rm ISCO}-t\right)}\right]^{3/8} \ , \tag{12.10}$$

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 we have defined $t_{\rm ISCO}$ to be the time when the frequency is that corresponding to the ISCO $f_{\rm ISCO} = c^3/(G6^{3/2}\pi M)$.

Let us consider equation (12.7) for two waves emitted at $t = t_{\rm ISCO} - \Delta t$ and $t' = t_{\rm ISCO}$ and find the distance D at which they will arrive at the same time (i.e., $\Delta t_a = 0$). Plugging these two times into equation (12.10), using the frequencies in equation (12.7), and solving for D where $\Delta t_a = 0$, we find

$$D = \frac{2\lambda_g^2\Delta t}{\frac{\pi^2}{5^{3/4}M}\left[256\eta M^3\left(\frac{405M}{16\eta}+\Delta t\right)\right]^{3/4} - 216\pi^2M^2} \ . \tag{12.11}$$

Taking the limit of this as $\Delta t \to 0$, we find the distance at which the end of the waveform starts to

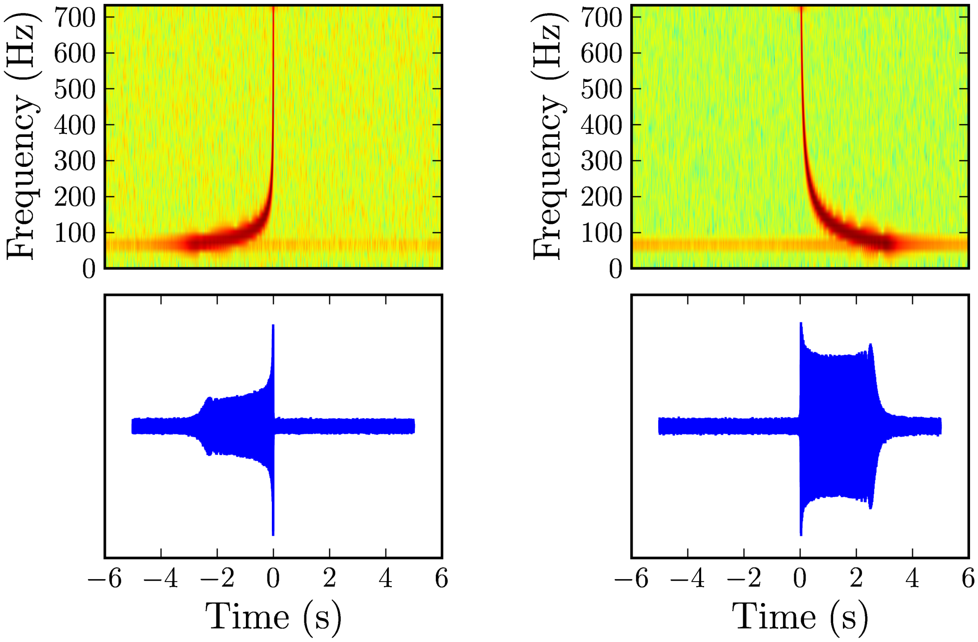

Figure 12.1: Waveform Inversion
Plots of inspiral waveforms at $D \ll D_{\text{inv}}$ (left) and $D \gg D_{\text{inv}}$ (right). The upper plots show spectrograms of the waveforms below. This makes it easier to see what frequency information arrives at what time.

become inverted (D_{inv}) to be

$$D_{\text{inv}} = \frac{5\lambda_g^2}{16\pi^2\eta M} \, . \tag{12.12}$$

Different waveforms are plotted in figure 12.1 showing how the morphology of the GW wave changes as it propagates due to dispersion effects of a massive graviton. The end of the waveform starts becoming inverted at $D \sim D_{\text{inv}}$. At $D \gg D_{\text{inv}}$, the information at the end of the waveform (i.e., the higher frequency portion of the waveform) can be seen to arrive before information that was emitted earlier (i.e., the lower frequency portion of the waveform).

12.4 Recovery of Dispersed Compact Binary Coalescence Gravitational Waveforms

In chapter 2 we found the inspiral portion of CBC waveforms as a function of time. Converted to the form found in [149], these are given as

$$h(t) = A(t)e^{-i\Phi(t)} \ , \tag{12.13a}$$

$$\Phi(t) = \Phi_0 + 2\pi \textstyle\int_{t_c}^{t} f(t)dt \ , \tag{12.13b}$$

$$A(t) = \tfrac{\mathcal{M}}{D_L}\left(\tfrac{5\mathcal{M}}{t_c - t}\right)^{1/4} \ . \tag{12.13c}$$

When searching for these signals in GW detector data, as discussed in chapter 2, it is convenient to convert these waveforms to the frequency domain using the SPA (equation (2.41)). Converting the SPA phase (equation (2.45)) to the form found in [149], $\Psi(f)$ is given as

$$\Psi(f) = 2\pi \int_{f}^{f_c} (t_c - t)df' + 2\pi f t_c + \phi_0 \ , \tag{12.14}$$

where the subscript c denotes a value at the time of coalescence. These waveforms are used to search for signals in the GW detector data with the technique known as matched filtering (discussed in chapter 5) where an inner product is defined as

$$(h_1|h_2) = 4\Re\left[\int_0^\infty \frac{\tilde{h}_1^* \tilde{h}_2}{S_n(f)}df\right] \ , \tag{12.15}$$

where $\Re[\cdots]$ denotes the real part of $[\cdots]$, $\tilde{h}^*$ denotes the complex conjugate of $\tilde{h}$, and $S_n(f)$ is the power spectral density of the particular detector.

Since these are the waveforms that are used for the detection of signals, all of the f's above are measured frequencies, which as we have stated above, are related to the emitted frequencies by $f = f_e/(1+Z)$. We are now in a position to see how the massive graviton propagation effects change a signal of the form given in equation (2.41). The term which is modified by these effects is the integral term in the phase $\Psi(f)$ (equation (12.14)). Applying equation (12.7) to this term we find it changes to be

$$\begin{aligned}\int_f^{f_c}(t_c - t)df' &= \int_{f_e}^{f_{ec}}(t_{ec}-t_e)df'_e + \int_{f_e}^{f_{ec}}\frac{D}{2\lambda_g^2 f_e'^2}df'_e - \int_{f_e}^{f_{ec}}\frac{D}{2\lambda_g^2 f_{ec}^2}df'_e \\ &= \int_{f_e}^{f_{ec}}(t_{ec}-t_e)df'_e - \frac{D}{2\lambda_g^2 f_e} + \frac{D}{\lambda_g^2 f_{ec}} - \frac{D}{2\lambda_g^2 f_{ec}^2}f_e\ . \qquad (12.16)\end{aligned}$$

These additional terms can be seen to have different effects on the waveform. The last term has an f_e^1 dependence, which is the same dependence as the time offset term in equation (12.14) and can be interpreted as a change in the arrival time. The second-to-last term has an f_e^0 dependence, which is the same dependence as the phase offset term in equation (12.14) and can be interpreted as a change in the phase at coalescence. The other additional term (the second term) has an f_e^{-1} dependence, which is the same as the 1.0PN phase term. This is the term that will change the morphology of the waveform. Its effect on equation (2.45) is to add an additional $-\beta f^{-1}$ term, where $\beta \equiv \frac{\pi D}{\lambda_g^2(1+Z)}$. To 1.5PN order, $\Psi(f)$ is then given as

$$\begin{aligned}\Psi(f) &= 2\pi f t_c + \phi_0 - \frac{3}{128\eta}\left[(\pi M f)^{-5/3} + \left(\frac{3715}{756}+\frac{55}{9}\eta\right)(\pi M f)^{-1}\right. \\ &\quad \left. -16\pi\,(\pi M f)^{-2/3}\right] - \beta f^{-1} \qquad (12.17)\end{aligned}$$

12.5 Using the Dispersion of Gravitational Waves to Bound the Mass of the Graviton

In reference [149], Will used the above formalism with restricted inspiral waveforms to 1.5PN order in the phase to predict bounds on the mass of the graviton from the detection of signals from Adv. LIGO and LISA. Will's analysis used the rms error of associated massive graviton phase term from parameter estimation techniques. Specifically, the rms error of a parameter θ^α is given by

$$\Delta\theta^\alpha \equiv \sqrt{\left\langle(\theta^\alpha - \langle\theta^\alpha\rangle)^2\right\rangle} = \sqrt{\Sigma^{\alpha\alpha}}\ , \qquad (12.18)$$

where Σ is the inverse of the Fischer Information matrix Λ given by

$$\Lambda_{\alpha\beta} \equiv \left(\frac{\partial h}{\partial \theta^{\alpha}} | \frac{\partial h}{\partial \theta^{\beta}} \right) . \tag{12.19}$$

For this analysis, we will compare the results of Will's analysis [149] using 1.5PN waveforms to a similar analysis using new IMR waveforms. The IMR waveforms we use are frequency domain phenomenological waveforms whose phenomenological parameters have been tuned using Numerical Relativity waveforms with symmetric mass ratios $0.25 \geq \eta \geq 0.16$ [152, 153]. These waveforms are given by

$$\tilde{h}(f) \equiv A_{\mathrm{eff}}(f) e^{i\Psi_{\mathrm{eff}}(f)} , \tag{12.20}$$

where the effective amplitude $A_{\mathrm{eff}}(f)$ and phase Ψ_{eff} are given by

$$A_{\mathrm{eff}}(f) \equiv C \begin{cases} (f/f_1)^{-7/6} & \text{if } f < f_1 \\ (f/f_1)^{-2/3} & \text{if } f_1 \leq f < f_2 \\ w\mathcal{L}(f, f_2, \sigma) & \text{if } f_2 \leq f < f_3 \\ 0 & \text{if } f \geq f_3 \, , \end{cases} \tag{12.21a}$$

$$\Psi_{\mathrm{eff}}(f) \equiv 2\pi f t_0 + \phi_0 + \sum_{k=0}^{6} \psi_k f^{(k-5)/3} , \tag{12.21b}$$

where $C = \frac{M^{5/6} f_1^{-7/6}}{D_L \pi^{2/3}} \sqrt{\frac{5\eta}{24}}$ is the amplitude parameter for optimally oriented and located sources, $w = \frac{\pi\sigma}{2} \left(\frac{f_2}{f_1} \right)^{-2/3}$ is a normalization such that A_{eff} is continuous through f_2, $\mathcal{L}(f, f_2, \sigma) = \frac{1}{2\pi} \frac{\sigma}{(f-f_2)^2 + \sigma^2/4}$ is a Lorentzian function centered on f_2 with width σ, t_0 is the arrival time, and ϕ_0 is the phase offset. For these waveforms, f_1 corresponds to the where the PN amplitude evolution ends (i.e., where the inspiral ends and the merger begins), f_2 corresponds to where the merger ends and the ringdown begins, and f_3 is the frequency at which we cut off the waveform. The ringdown portion of the waveform is modelled as a Lorentzian which agrees with the quasi-normal mode ringing of a perturbed black hole from black hole perturbation theory.

The phenomenological amplitude parameters $\alpha_k = \{f_1, f_2, f_3, \sigma\}$ are given by

$$\alpha_k = \frac{a_k\eta^2 + b_k\eta + c_k}{\pi\mathcal{M}}\eta^{3/5} , \tag{12.22a}$$

where a_k, b_k, and c_k are given in Table 12.1. The phenomenological phase parameters ψ_k are given by

$$\psi_k = \frac{x_k\eta^2 + y_k\eta + z_k}{(\pi\mathcal{M})^{(5-k)/3}}\eta^{-k/5} , \tag{12.22b}$$

where x_k, y_k, and z_k are given in Table 12.1. The phenomenological phase parameters do not match the PN terms because of the choice of the authors in references [152, 153]. They allowed these terms to vary while trying to maximize the "faithfulness" and "effectualness" of the phenomenological waveforms to recovering hybrid PN-Numerical Relativity waveforms.

Table 12.1: Phenomenological Coefficients

Parameter	a_k	b_k	c_k
f_1	6.6389 $\times 10^{-1}$	-1.0321 $\times 10^{-1}$	1.0979 $\times 10^{-1}$
f_2	1.3278	-2.0642 $\times 10^{-1}$	2.1957 $\times 10^{-1}$
σ	1.1383	-1.7700 $\times 10^{-1}$	4.6834 $\times 10^{-2}$
f_3	1.7086	-2.6592 $\times 10^{-1}$	2.8236 $\times 10^{-1}$
Parameter	x_k	y_k	z_k
ψ_0	-1.5829 $\times 10^{-1}$	8.7016 $\times 10^{-2}$	-3.3382 $\times 10^{-2}$
ψ_2	3.2967 $\times 10^{1}$	-1.9000 $\times 10^{1}$	2.1345
ψ_3	-3.0849 $\times 10^{2}$	1.8211 $\times 10^{2}$	-2.1727 $\times 10^{1}$
ψ_4	1.1525 $\times 10^{3}$	-7.1477 $\times 10^{2}$	9.9692 $\times 10^{1}$
ψ_6	1.2057 $\times 10^{3}$	-8.4233 $\times 10^{2}$	1.8046 $\times 10^{2}$

Unitless coefficients describing the amplitude and phase of the phenomenological waveforms from reference [152] to be used in equations (12.22). There is no ψ_1 term as the 0.5PN phasing term is 0.

The amplitude C given above is the amplitude one would find for an optimally oriented and optimally located source. Averaging over random angles for the two sky location angles, the polarization angle, and the inclination angle, there is an overall factor of 2/5 we should apply to this. For this reason we redefine $C = \frac{2M^{5/6}f_1^{-7/6}}{5D_L\pi^{2/3}}$ from here on.

Applying equation (12.16) to this waveform results in a measured phase evolution of

$$\Psi_{\text{eff}}(f) \equiv 2\pi f t_0 + \phi_0 + \sum_{k=0}^{6}\psi_k f^{(k-5)/3} - \beta f^{-1} , \tag{12.23}$$

where t_0 is the measured arrival time which has absorbed the massive graviton term with the f^1 dependence and ϕ_0 is the measured phase offset which has absorbed the massive graviton term with the f^0 dependence.

In producing this result, we have made an assumption which we must check. This is that, at the order to which we the phase evolution of the binary, energy loss due to massive graviton effects are negligible. Since we use phenomenological terms up to the equivalent 3.0PN order and ignore massive graviton effects of order $(r/\lambda_g)^2$ when computing the energy loss at the source, this implies we require $r^2\lambda_g^{-2}v^{-6} \ll 1$. Combining this with $v^2 \sim M/r$ and $v = 2\pi rf$, we find the frequencies this holds for are $f > \lambda_g^{-3/5}M^{-2/5}/(2\pi)$, which is $f > 2.6 \times 10^{-3}(M/M_\odot)^{-2/5}$Hz assuming the current bound on $\lambda_g > 10^{12}$ km. For the systems we consider, this is always satisfied.

The parameters we use in the Fischer Matrix calculation are $\ln C$, ϕ_0, t_0, $\ln \mathcal{M}$, $\ln \eta$, and β. The partial derivatives of $\tilde{h}$ with respect to these are given by

$$\frac{\partial \tilde{h}(f)}{\partial \ln C} = \tilde{h}(f) \ , \tag{12.24a}$$

$$\frac{\partial \tilde{h}(f)}{\partial \phi_0} = -i\tilde{h}(f) \ , \tag{12.24b}$$

$$\frac{\partial \tilde{h}(f)}{\partial t_0} = 2\pi i\tilde{h}(f) \ , \tag{12.24c}$$

$$
\begin{aligned}
\frac{\partial \tilde{h}(f)}{\partial \ln \mathcal{M}} \quad &= \quad C e^{i\Psi_{\rm eff}} \begin{cases} \frac{7}{6}\frac{\partial \ln f_1}{\partial \ln \mathcal{M}} \left(\frac{f}{f_1}\right)^{-7/6} & \text{if } f < f_1 \\ \frac{2}{3}\frac{\partial \ln f_1}{\partial \ln \mathcal{M}} \left(\frac{f}{f_1}\right)^{-2/3} & \text{if } f < f_1 \\ w\mathcal{L}(f, f_2, \sigma)\left(\frac{\partial \ln w}{\partial \ln \mathcal{M}} + \frac{\partial \ln \mathcal{L}(f,f_2,\sigma)}{\partial \ln \mathcal{M}}\right) & \text{if } f_2 \le f < f_3 \\ 0 & \text{if } f \ge f_3 \end{cases} \\
&+ \quad iCe^{i\Psi_{\rm eff}} \sum_{k=0}^{6} \left[\psi_k \frac{\partial \ln \psi_k}{\partial \ln \mathcal{M}} f^{(k-5)/3}\right] \begin{cases} (f/f_1)^{-7/6} & \text{if } f < f_1 \\ (f/f_1)^{-2/3} & \text{if } f_1 \le f < f_2 \\ w\mathcal{L}(f, f_2, \sigma) & \text{if } f_2 \le f < f_3 \\ 0 & \text{if } f \ge f_3 \;, \end{cases}
\end{aligned} \tag{12.24d}
$$

$$
\begin{aligned}
\frac{\partial \tilde{h}(f)}{\partial \ln \eta} \quad &= \quad C e^{i\Psi_{\rm eff}} \begin{cases} \frac{7}{6}\frac{\partial \ln f_1}{\partial \ln \eta} \left(\frac{f}{f_1}\right)^{-7/6} & \text{if } f < f_1 \\ \frac{2}{3}\frac{\partial \ln f_1}{\partial \ln \eta} \left(\frac{f}{f_1}\right)^{-2/3} & \text{if } f < f_1 \\ w\mathcal{L}(f, f_2, \sigma)\left(\frac{\partial \ln w}{\partial \ln \eta} + \frac{\partial \ln \mathcal{L}(f,f_2,\sigma)}{\partial \ln \eta}\right) & \text{if } f_2 \le f < f_3 \\ 0 & \text{if } f \ge f_3 \end{cases} \\
&+ \quad iCe^{i\Psi_{\rm eff}} \sum_{k=0}^{6} \left[\psi_k \frac{\partial \ln \psi_k}{\partial \ln \eta} f^{(k-5)/3}\right] \begin{cases} (f/f_1)^{-7/6} & \text{if } f < f_1 \\ (f/f_1)^{-2/3} & \text{if } f_1 \le f < f_2 \\ w\mathcal{L}(f, f_2, \sigma) & \text{if } f_2 \le f < f_3 \\ 0 & \text{if } f \ge f_3 \;, \end{cases}
\end{aligned} \tag{12.24e}
$$

$$
\frac{\partial \tilde{h}(f)}{\partial \beta} = -if^{-1}\tilde{h}(f) \;. \tag{12.24f}
$$

When taking the inner products in calculating Λ, we find it is helpful to introduce the definition ofof an integral $I_a^b(p,q)$ given by

$$I_a^b(p,q) \equiv 4\int_a^b \frac{f^{-p/3}}{((f-f_2)^2+\sigma^2/4)^q\, g_\alpha(f)}df\ , \tag{12.25}$$

where $g_\alpha(f)$ is the frequency dependence of the detectors PSD. The PSD can then be written as $S_n(f) = S_0 g_\alpha(f)$. In this notation, the SNR can be written as

$$\begin{aligned} \rho &\equiv (h|h)^{1/2} \\ &= \frac{C}{\sqrt{S_0}}\left[f_1^{7/3} I_0^{f_1}(7,0) + f_1^{4/3} I_{f_1}^{f_2}(4,0) + \frac{w^2\sigma^2}{4\pi^2} I_{f_2}^{f_3}(0,2)\right]^{1/2}\ . \end{aligned} \tag{12.26}$$

In the calculation of the inverse of the Fischer Matrix, we find all of the terms to be proportional to $\sqrt{S_0}/C$. Let us then define Δ as $\Delta\beta \equiv \Delta^{1/2}\sqrt{S_0}/C$. In setting an upper limit on the mass of the graviton, we view $\Delta\beta$ as the upper bound on β which results in a lower bound on λ_g given by

$$\lambda_g > \left(\frac{D}{(1+Z)D_L}\right)^{1/2}\left(\frac{\pi^2 C\mathcal{M}}{S_0^{1/2}\Delta^{1/2}}\right)^{1/2}\ . \tag{12.27}$$

In Table 12.2 we summarize the bounds on λ_g we obtain from the complete IMR waveforms and compare them to the results from Will [149]. This table uses two different noise models, one for signals which Adv. LIGO will be sensitive to, and one for signals LISA will be sensitive to. The noise model for Adv. LIGO is given by $S_0 = 3\times 10^{-48}$ and $g_\alpha(f) = ((f/f_n)^{-4} + 2 + 2(f/f_n)^2)/5$ where $f_n = 70$Hz. Since Adv. LIGO is sensitive down to $f_0 = 40$Hz, this is where we start the integration instead of at 0Hz. The noise model for LISA is given by $S_0 = 4.2\times 10^{-41}$ and $g_\alpha(f) = (10^{1/2}(f/f_n)^{-14/3} + 1 + (f/f_n)^2/1000 + 313.5(f/f_n)^{-(6.398+3.518\log_{10} x)})$ where $f_n = 10^{-3}$Hz. For LISA, we start the integration at the lower frequency corresponding to a $T = 1$yr integration time, such that $f_0 = (5\mathcal{M}T^{-1}/256)^{3/8}/(\pi\mathcal{M})$.

We see from table 12.2 and figure 12.2 that for some signals, we can obtain better bounds from the PN inspiral-only waveform than from the complete IMR waveform. If we repeat the above analysis for the phenomenological waveform keeping only the inspiral portion with the $f^{-7/6}$

Table 12.2: Massive Graviton Bounds

m_1 ($M_\odot$)	m_2 ($M_\odot$)	Detector	Distance (Mpc)	Inspiral Bound (km)	Complete Waveform Bound (km)
10	10	Adv. LIGO	1500	6.0×10^{12}	4.8×10^{12}
50	50	Adv. LIGO	3000	2.8×10^{12}	4.8×10^{13}
10^4	10^4	LISA	3000	7.4×10^{15}	4.3×10^{15}
10^5	10^5	LISA	3000	2.3×10^{16}	1.1×10^{16}
10^6	10^6	LISA	3000	5.4×10^{16}	2.1×10^{17}
10^7	10^7	LISA	3000	6.9×10^{16}	5.1×10^{17}
10^8	10^8	LISA	3000	5.8×10^{16}	9.7×10^{17}
10^9	10^9	LISA	3000	2.6×10^{16}	1.1×10^{18}

The theoretical bounds on λ_g that could be placed based on a GW detection from different mass CBC signals. The column "Inspiral Waveform" refers to the bounds one could get from just looking at the inspiral portion of the waveform up to the ISCO frequency. Where available, these agree quantitatively with Will [149]. The column "Full Waveform" refers to the bounds one could get from looking at the complete IMR waveform which has been characterized by Ajith [152] and Ajith et al. [153].

amplitude frequency dependence and setting the cutoff frequency to be f_{ISCO}, we find that the PN waveform always beats the phenomenological waveform. This is due to the differences between the phenomenological phasing parameters and the PN phasing parameters Ajith et al. noted in [153]. The above observation implies the parameters of the PN waveform are less degenerate than the parameters of the phenomenological waveform.

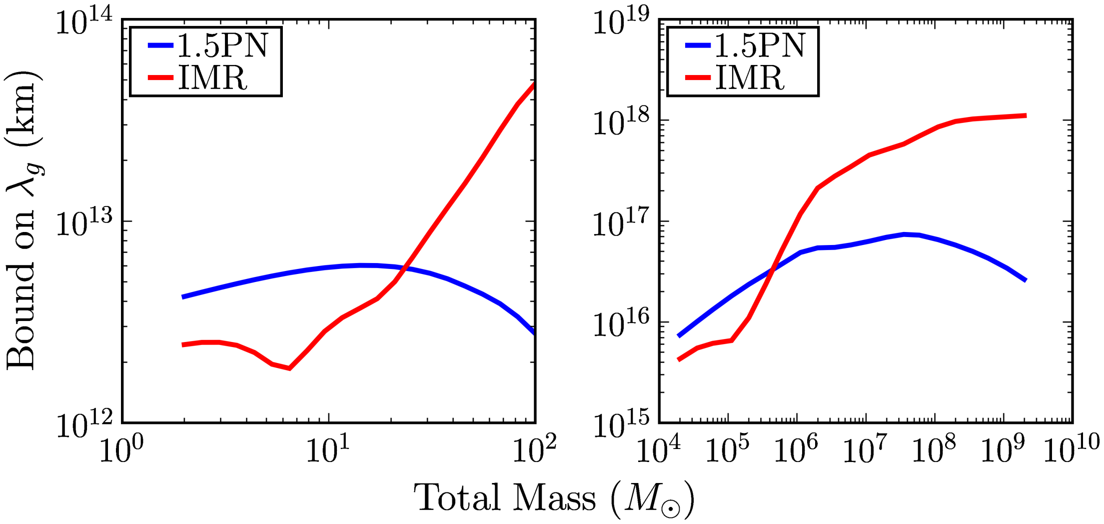

Figure 12.2: Massive Graviton Bounds

Plots of theoretical bounds on λ_g that could be placed based on a GW detection from Adv. LIGO mass systems (left) and LISA mass systems (right). For each mass of the Adv. LIGO systems, the bound was calculated using lesser of the distance corresponding to an SNR of 10 for the 1.5PN waveform and a distance of 3 Gpc and a minimum frequency of 10 Hz. For each mass of the LISA systems, the bound was calculated using a distance of 3 Gpc and a bandwidth corresponding to 1 year of integration time. The 1.5PN curves correspond to the bounds one could get from just looking at the 1.5PN waveform up to the ISCO frequency, as was done in Will [149]. The IMR curves correspond to the bounds one could get from looking at the complete IMR waveform which has been characterized by Ajith [152] and Ajith et al. [153].

12.6 Beyond the Fischer Matrix Approach

The approach that has been taken by Will [149] and here to constrain the mass of the graviton is the Fischer information matrix approach. This approximation is only valid in the large SNR limit and does not include known issues involved in real data analysis pipelines, which search for inspiral signals, such as recovery of a signal with incorrect parameters and separating the amplitude correlations between distance, mass, and sky location. A full simulation is required to accurately determine the achievable bounds on the mass of the graviton with one, or even several, observed signals.

12.7 Other Bounds on the Graviton's Mass

Let us compare these results to the results available from other physical effects of a massive graviton. The first effect we shall look at is changes in planetary motion due to a Yukawa gravitational potential. If this were the case, Kepler's third law would be violated since the gravitational force would no longer follow an inverse-square law. This leads to a bound on the mass of the graviton in the form of $\lambda_g > (\frac{1-a_p^2}{6\eta_p})^{1/2}$ where a_p is the semimajor axis of planet p, η_p is planet p's bound on the η parameter given by Talmadge et al. [154]. Both λ_g and a_p are given in astronomical units. The most stringent bound of this form comes from the orbit of Mars which limits the mass of the graviton such that $\lambda_g > 2.8 \times 10^{12}$ km [149].

The next bound on the mass of the graviton is from binary pulsar observations. If the graviton had mass, the orbits of binary pulsars would decay at a slightly faster rate than predicted by General Relativity, due to additional energy loss from the leading order massive graviton terms in the power radiated. Combining the observations of PSR B1913+16 and PSR B1534+12, the 90% confidence bound on the mass of the graviton is $\lambda_g > 1.6 \times 10^{10}$ km [150].

The final bound on the mass of the graviton we shall look at can be obtained by studying gravitational lensing data. Assuming a specific distribution of dark matter, lensing data can be studied looking for an effect which would change the distribution of the variance of the power spectrum $\gamma^2(\theta)$ smoothed over a filter of radius θ. When comparing the predictions of a $1/r$ potential versus the Yukawa potential, Choudury et al. [151] find the bound on the mass of the graviton to be $\lambda_g > 3 \times 10^{21}$ km [151]. However, this upper limit is less robust than the others above as one must assume a specific distribution of dark matter before calculating the bound.

Chapter 13

Toward Understanding Black Hole Merger Dynamics

The stunning breakthroughs in Numerical Relativity over the past few years, starting with those by Pretorius [155] in spring 2005, have provided greater insight into the realm of strongly curved spacetime and dynamical gravity. Non-spinning binary black hole merger simulations are quickly becoming the norm, and numerical relativists have started looking at more interesting situations, including the orbit and merger of spinning binary black holes. Surprising results from this sort of situation are exemplified in the *extreme-kick configuration* in which two identical, spinning black holes are initially in a (quasi-)circular orbit, with oppositely directed spins lying in the orbital plane (figure 13.1).

As Campanelli, Lousto, Zlochower, and Merritt [156, 157] (CLZM) discovered and Healy et al. [158] helped flesh out, of all equal mass, quasi-circular initial configurations, this one has the

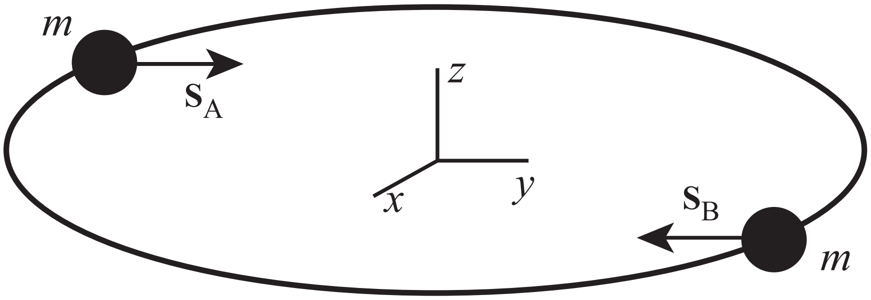

Figure 13.1: Extreme-kick Configuration
Extreme-kick configuration for a black-hole binary: Identical holes, A and B with masses $m = M/2$ move in a circular orbit with their spin angular momenta $\boldsymbol{S}_A$ and $\boldsymbol{S}_B$ antialigned and lying in the orbital plane.

largest kick speed for the final black hole. Not only that, it also exhibits intriguing orbital motions.

During the inspiral phase, as the holes circle each other, they bob up and down (in the z direction of figure 13.1), sinusoidally and synchronously. After merger the combined hole gets kicked up or down with a final speed that depends on the orbital phase at merger (relative to the spin directions). This bobbing-then-kick, as deduced by CLZM from numerical simulations, is graphed quantitatively in figure 13.2.

Qualitatively, Pretorius [159] has offered a lovely physical explanation for the holes' bobbing (figure 13.2) in this configuration: In figure 13.3, taken from his paper, we see snapshots of the holes at four phases in their orbital motion. In each snapshot, each hole's spin drags space into motion (drags inertial frames) in the direction depicted by gray, semicircular arrows. In phase B, hole 1 drags space and thence hole 2 into the sheet of paper (or computer screen); and hole 2 drags space and thence hole 1 also inward. In phase D each hole drags the other outward. This picture agrees in phasing and semiquantitatively in amplitude with the bobbing observed in the simulations (figure 13.2).

However, momentum conservation dictates that, when the holes are moving upward together with momentum $p^z_A + p^z_B$, there must be some equal and opposite downward momentum in their gravitational field (in the curved spacetime surrounding them); and when the holes are moving downward, there must be an equal and opposite upward field momentum. Our aim is to quantitatively answer the following questions: How is this field momentum distributed? What are the details of the momentum flow between field and holes? And to what extent are other momentum-flow processes responsible for the motion shown in figure 13.2?

The outline of this chapter is as follows: section 13.1 gives a brief overview of the Landau-Lifshitz formalism, which can be used to calculate the energy and momentum content of the gravitational field, section 13.2 sets up the calculation of momentum consveration by splitting the total momentum into its respective pieces, section 13.3 calculates each of the pieces of the momentum both for the general binary and in the extreme-kick configuration, and section 13.4 combines these results to show that the momentum of the objects balanced by the field momentum.

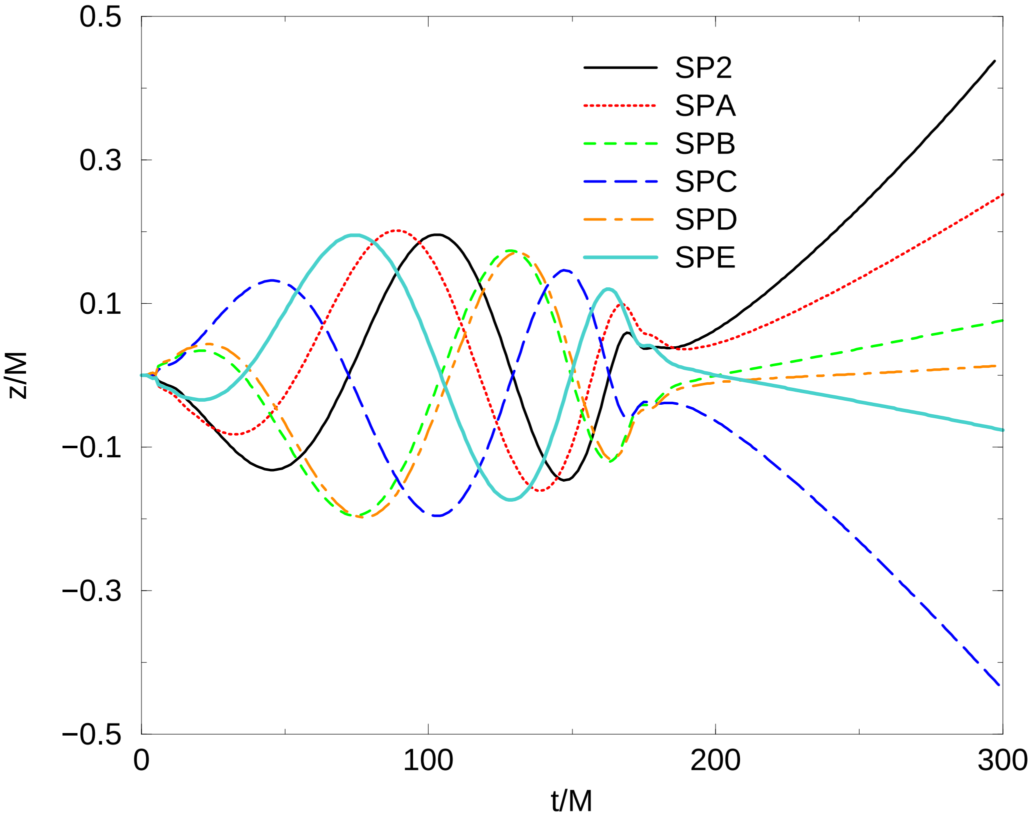

Figure 13.2: Bobbing Motion in Extreme-kick Configuration
Bobbing and kick of binary black holes in the extreme-kick configuration of figure 13.1, as simulated by Campanelli, Lousto, Zlochower, and Merritt (CLZM) [157]. Plotted vertically (as a function of time horizontally) is the identical height z of the two black holes, and then transitioning through merger (presumably at $t/M \sim 170$), the height of the merged hole, above the initial orbital plane. This height versus time is shown for six different initial configurations, each leading to a different orbital phase at merger. In all six configurations, the initial holes' spins are half the maximum allowed, $a/m = 0.5$. The height z and time t are those of the "punctures" that represent the holes' centers in the CLZM computations, as defined in their computational coordinate system, which becomes Lorentz at large radii. These z and t are measured in units of the system's total mass $M \simeq 2m$.

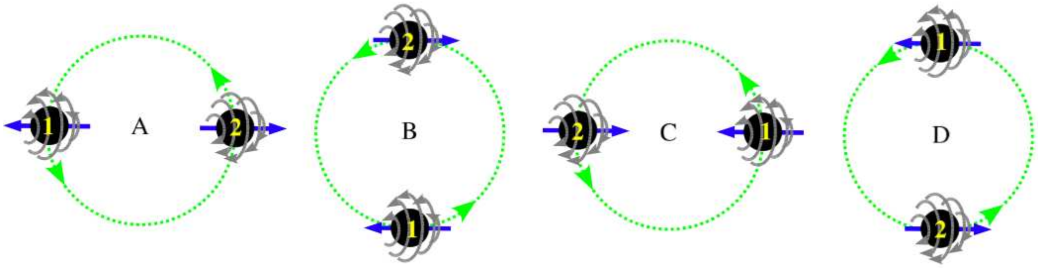

Figure 13.3: Pretorius' Explanation
Pretorius' physical explanation for the holes' bobbing in the extreme-kick configuration, in which the effect is attributed qualitatively to frame-dragging.

13.1 The Landau-Lifshitz Formalism in Brief

Here we give a detailed analysis of momentum flow in generic compact binary systems. We begin in this section with a brief review of the Landau-Lifshitz (LL) formulation of general relativity as a nonlinear field theory in flat spacetime [160].

The Landau-Lifshitz formulation [160] starts by choosing an arbitrary coordinate system that is asymptotically Lorentz, to this an auxiliary flat spacetime is added by asserting that the chosen ("preferred") coordinates are globally Lorentz in the auxiliary spacetime (so in them the auxiliary metric has components $\mathrm{diag}(-1,1,1,1)$). In this formulation, gravity is described by the physical metric density

$$\mathfrak{g}^{\mu\nu} = \sqrt{-g} g^{\mu\nu} \ , \tag{13.1}$$

where g is the determinant of the covariant components of the physical metric, and $g^{\mu\nu}$ are the contravariant components of the physical metric.

The EFEs are then reformulated as a nonlinear field theory in the space of the flat, auxiliary metric. Making use of the superpotential

$$H^{\mu\alpha\nu\beta} \equiv \mathfrak{g}^{\mu\nu}\mathfrak{g}^{\alpha\beta} - \mathfrak{g}^{\mu\alpha}\mathfrak{g}^{\nu\beta} \ , \tag{13.2}$$

the EFEs are given by

$$H^{\mu\alpha\nu\beta}{}_{,\alpha\beta} = 16\pi\tau^{\mu\nu} \ , \tag{13.3}$$

where $\tau^{\mu\nu} = (-g)(T^{\mu\nu} + t^{\mu\nu}_{\mathrm{LL}})$ is the total effective stress-energy tensor, indices after the comma denote partial derivatives (covariant derivatives with respect to the flat auxiliary metric), and the Landau-Lifshitz pseudotensor $t^{\mu\nu}_{\mathrm{LL}}$ (actually a real tensor in the auxiliary flat spacetime) is given by equation (100.7) of LL [160] or equivalently equation (20.22) of MTW [1]. By virtue of the symmetries of the superpotential (which are the same as those of the Riemann tensor), the field equations in the form (13.3) imply the differential conservation law for 4-momentum

$$\tau^{\mu\nu}{}_{,\nu} = 0 \ , \tag{13.4}$$

which is equivalent to $T^{\mu\nu}{}_{;\nu} = 0$ (where the semicolon denotes a covariant derivative with respect to the physical metric).

It is shown in LL and in MTW that the total 4-momentum of any isolated system (as measured gravitationally in the asymptotically flat region far from the system) is

$$p^{\mu}_{\rm tot} = \frac{1}{16\pi} \oint_{\mathcal{S}} H^{\mu\alpha 0j}{}_{,\alpha} d\Sigma_j \ , \tag{13.5}$$

where $d\Sigma_j$ is the surface-area element defined using the flat auxiliary metric, and the integral is over an arbitrarily large closed surface $\mathcal{S}$ surrounding the system. Differentiating this, we find

$$\frac{dp^{\mu}_{\rm tot}}{dt} = \frac{1}{16\pi} \oint_{\mathcal{S}} H^{\mu\alpha 0j}{}_{,\alpha 0} d\Sigma_j \ . \tag{13.6}$$

Let us look at $H^{\mu\alpha 0j}{}_{,\alpha 0} = H^{\mu\alpha\nu j}{}_{,\alpha\nu} - H^{\mu\alpha kj}{}_{,\alpha k}$. The first term is $-16\pi\tau^{\mu j}$ by virtue of the field equations (13.3) and the antisymmetry of the superpotential on its last two indices (13.2). That same antisymmetry on the second term $-H^{\mu\alpha kj}{}_{,\alpha k}$ permits us to write it as the curl of a 3-vector field, whose surface integral vanishes by virtue of Stokes' theorem. Combining these two results we find the total 4-momentum satisfies the standard conservation law

$$\frac{dp^{\mu}_{\rm tot}}{dt} = - \oint_{\mathcal{S}} \tau^{\mu j} d\Sigma_j \ . \tag{13.7}$$

This proof has also been given in LL and MTW where it relied on an assumption that the interior of $\mathcal{S}$ be simply connected, i.e., that it not contain any black holes.

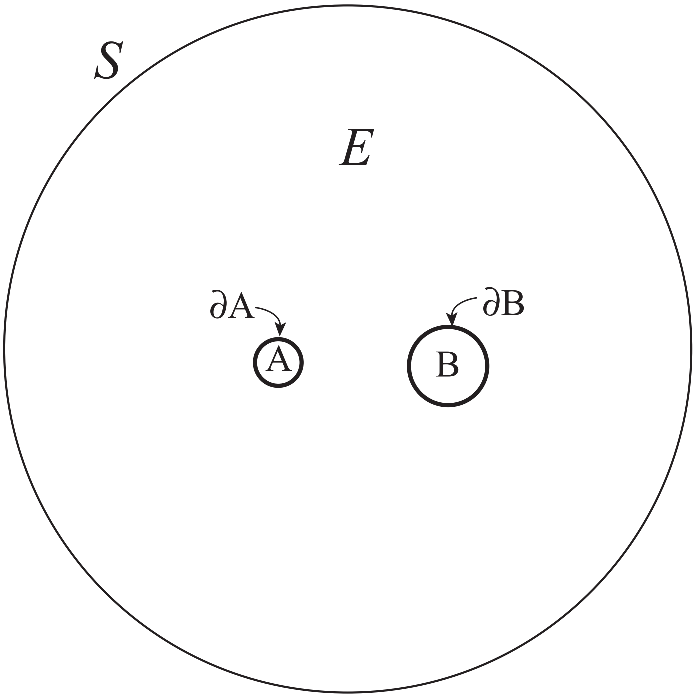

Figure 13.4: Division of Space
The regions of space around and inside a compact binary system.

13.2 Momentum Conservation for a Fully Nonlinear Compact Binary

We now apply this LL formalism to a binary system made of black holes and/or neutron stars; see figure 13.4. We denote the binary's two bodies by the letters A and B, and the regions of space inside them by these same letters, and their surfaces by ∂A and ∂B. For a black hole, ∂A could be the hole's absolute event horizon or its apparent horizon, whichever one wishes. For a neutron star, ∂A will be the star's physical surface. We denote by $\mathcal{E}$ the region outside both bodies, but inside the arbitrarily large surface $\mathcal{S}$ where the system's total momentum is computed.

By applying Gauss's theorem to equation (13.5) for the binary's total 4-momentum and using the EFE (13.3), we obtain an expression for the binary's total 4-momentum as a sum over contributions from each of the bodies and from the gravitational field in the region $\mathcal{E}$ outside them:

$$p^{\mu}_{\rm tot} = p^{\mu}_{A} + p^{\mu}_{B} + p^{\mu}_{\rm field} \ . \tag{13.8a}$$

Here

$$p_A^\mu \equiv \frac{1}{16\pi} \oint_{\partial A} H^{\mu\alpha 0j}{}_{,\alpha} d\Sigma_j \tag{13.8b}$$

is the 4-momentum of body A and similarly for body B, and

$$p_{\rm field}^\mu \equiv \int_{\mathcal{E}} \tau^{0\mu} d^3x \tag{13.8c}$$

is the gravitational field's 4-momentum in the surrounding space.

If either of the bodies has a simply connected interior (is a star rather than a black hole), then we can use Gauss's theorem and the EFEs (13.3) to convert the surface integral (13.8b) for the body's 4-momentum into a volume integral over the body's interior:

$$p_A^\mu = \int_A \tau^{0\mu} d^3x \ . \tag{13.8d}$$

By an obvious extension of the argument we used to derive equation (13.7) for the rate of change of the binary's total 4-momentum, we can deduce from equation (13.8b) the corresponding equation for the rate of change of the 4-momentum of body A:

$$\frac{dp_A^\mu}{dt} = -\oint_{\partial A} (\tau^{\mu k} - \tau^{\mu 0} v_A^k) d\Sigma_k \ . \tag{13.9}$$

Here the second term arises from the motion of the boundary of body A with coordinate velocity $v_A^k = dx^k_{A\ {\rm cm}}/dt$. Equation (13.9) describes the flow of field 4-momentum into and out of body A.

We shall use equations (13.8), (13.7), and (13.9), specialized to linear momentum (index μ made spatial) as foundations for our study of momentum flow in compact binaries.

The actual values of the body and field 4-momenta, computed in the above ways, will depend on the arbitrary coordinate system that we chose, in which to make the auxiliary metric be diag$(-1, 1, 1, 1)$ and in which to perform the above computations. This is a "gauge dependence," which we will fix by our choice of coordinates. In the remainder of this chapter we shall choose

Harmonic coordinates, so the gravitational field satisfies the Harmonic gauge condition

$$\mathfrak{g}^{\alpha\beta}{}_{,\beta} = 0 \, , \tag{13.10}$$

and we shall specialize the above equations to the 1.5 post-Newtonian approximation and use them to study momentum flow during the inspiral phase of generic compact binaries.

13.3 Post-Newtonian Momentum Flow in Generic Compact Binaries

13.3.1 Field Momentum Outside the Bodies

In Harmonic gauge at leading post-Newtonian order, the Landau-Lifshitz formalism gives for the density of field momentum

$$\tau^{0j}\boldsymbol{e}_j = -\frac{\boldsymbol{g}\times\boldsymbol{H}}{4\pi} + \frac{3}{4\pi}\dot{U}_N\boldsymbol{g} \tag{13.11}$$

(equation (4.1a) of [161]). Here, to the accuracy we need, $\boldsymbol{g}$ is the Newtonian gravitational acceleration field (the *gravitoelectric field*), $\boldsymbol{H}$ is the gravitational analog of the magnetic field (the *gravitomagnetic field*), U_N is the Newtonian potential and the dot denotes differentiation with respect to time, and $\boldsymbol{e}_j$ is the jth basis vector of the flat-spacetime field theory that we are using.

This field momentum can be split into several pieces, only one of which will flow back and forth between the field and the bobbing holes. Each of the components in equation (13.11) (i.e., $\boldsymbol{g}$, $\boldsymbol{H}$, and U_N) have contributions from body A and body B, where

$$\boldsymbol{g}_A = -\frac{m}{r_A^2}\boldsymbol{n}_A \, , \tag{13.12a}$$

$$U_{N,A} = \frac{m}{r_A} \, , \tag{13.12b}$$

(equations (2.5) and (6.1) of [161]). In addition, $\boldsymbol{H}$ can be split into terms arising from either velocity ($\boldsymbol{H}^{\rm velo}$) or spin ($\boldsymbol{H}^{\rm spin}$) contributions of a particular body:

$$\boldsymbol{H}_A^{\rm velo} = \frac{4m_A(\boldsymbol{n}_A\times\boldsymbol{v}_A)}{r_A^2}\,, \tag{13.12c}$$

$$\boldsymbol{H}_A^{\rm spin} = -2\frac{(3\boldsymbol{n}_A\cdot\boldsymbol{S}_A)\boldsymbol{n}_A-\boldsymbol{S}_A}{r_A^3}\,, \tag{13.12d}$$

(equations (2.5) and (6.1) of [161]), with $\boldsymbol{n}_A$ the unit radial vector pointing from the center of body A to the field point, r_A the distance from the center of body A to the field point, $\boldsymbol{v}_A$ the vectorial velocity of body A, and $\boldsymbol{S}_A$ the vectorial angular momentum of body A. The fields for body B are the same as equations (13.12), but with each subscript A replaced by a B.

The full field momentum density is written most concisely as

$$\tau^{0j} = \tau^{0j}_{\rm spin} + \tau^{0j}_{\rm velo}, \tag{13.13a}$$

where $\tau^{0j}_{\rm spin}$ and $\tau^{0j}_{\rm velo}$ are the terms that depend on the spins and the velocities, respectively. These terms are given by

$$\begin{aligned}\tau^{0j}_{\rm spin}\boldsymbol{e}_j &= \frac{m_B}{2\pi r_A^3 r_B^2}\left[3(\boldsymbol{S}_A\cdot\boldsymbol{n}_A)(\boldsymbol{n}_A\times\boldsymbol{n}_B)-(\boldsymbol{S}_A\times\boldsymbol{n}_B)\right]\\ &\quad -\frac{1}{2\pi}\frac{m_A}{r_A^5}(\boldsymbol{S}_A\times\boldsymbol{n}_A)+(A\leftrightarrow B),\end{aligned} \tag{13.13b}$$

and

$$\begin{aligned}\tau^{j0}_{\rm velo}\boldsymbol{e}_j &= \frac{m_A}{4\pi r_A^2}\left\{\frac{m_B[4(\boldsymbol{n}_B\cdot\boldsymbol{v}_A)\boldsymbol{n}_A-4(\boldsymbol{n}_A\cdot\boldsymbol{n}_B)\boldsymbol{v}_A]}{r_B^2}\right.\\ &\quad\left.-\frac{3m_B(\boldsymbol{n}_A\cdot\boldsymbol{v}_A)\boldsymbol{n}_B}{r_B^2}+\frac{m_A[(\boldsymbol{n}_A\cdot\boldsymbol{v}_A)\boldsymbol{n}_A-4\boldsymbol{v}_A]}{r_A^2}\right\}\\ &\quad +(A\leftrightarrow B),\end{aligned} \tag{13.13c}$$

where $(A\leftrightarrow B)$ means the same expression with labels A and B interchanged.

We are only interested in that portion of the field momentum that is induced by the holes' spins, since this is the portion that must flow back and forth between the field and the bobbing holes in order to conserve total momentum. As can be seen from equations (13.12), this portion arises from one hole's gravitoelectric field $\boldsymbol{g}$ coupling to the spin-induced part of the other hole's gravitomagnetic field $\boldsymbol{H}^{\rm spin}$

$$\delta\tau^{0j}\boldsymbol{e}_j = -\frac{\boldsymbol{g}_A \times \boldsymbol{H}_B^{\rm spin}}{4\pi} - \frac{\boldsymbol{g}_B \times \boldsymbol{H}_A^{\rm spin}}{4\pi} \; . \tag{13.14}$$

Combining equations (13.14) and (13.12) we obtain for the binary's density of field momentum (that portion which must flow during bobbing)

$$\delta\tau^{0j}\boldsymbol{e}_j = \frac{m_B}{2\pi r_A^3 r_B^2}\left[3(\boldsymbol{S}_A\cdot\boldsymbol{n}_A)(\boldsymbol{n}_A\times\boldsymbol{n}_B) - (\boldsymbol{S}_A\times\boldsymbol{n}_B)\right] + (A\leftrightarrow B) \; . \tag{13.15}$$

In order to integrate this over the region $\mathcal{E}$, we find it convenient to rewrite the bodies' gravitoelectric (13.12a) and gravitomagnetic fields (13.12d) as

$$g_K^j = m_K\left(\frac{1}{r_K}\right)_{,j} \; , \quad H_K^j = -2S_K^i\left(\frac{1}{r_K}\right)_{,ij} \; , \tag{13.16}$$

where K is A or B and where, as before, r_K is the (flat-space) distance of the field point from the center of mass of body K. Inserting equations (13.16) into expression (13.14) and manipulating the derivatives, we obtain the following expression for the field momentum density:

$$\delta\tau^{0j} = -\frac{1}{2\pi}\epsilon_{jpl}\left[S_A^q m_B\left(\frac{1}{r_A}\right)_{,q}\left(\frac{1}{r_B}\right)_{,l}\right]_{,p} + (A\leftrightarrow B) \; . \tag{13.17}$$

Notice that this expression for the momentum density is the curl of a vector field; or, equally well, it can be viewed as the divergence of a tensor field.

The total spin-induced, flowing field momentum is the integral of expression (13.17) over the exterior region $\mathcal{E}$ (cf., figure 13.4). Using Gauss's law, that volume integral can be converted into the following integral over the boundary of $\mathcal{E}$

$$\delta p^j_{\rm field} = -\frac{1}{2\pi}\epsilon_{jpl}S_A^q m_B\int_{\partial\mathcal{E}}\left(\frac{1}{r_A}\right)_{,q}\left(\frac{1}{r_B}\right)_{,l} d\Sigma_p + (A\leftrightarrow B) \; . \tag{13.18}$$

The boundary of $\mathcal{E}$ has three components: the surface $\mathcal{S}$ far from the binary on which we compute the binary's total momentum, and the surfaces ∂A and ∂B of bodies A and B. The integral over $\mathcal{S}$ vanishes because the integrand is $\propto 1/r^4$ and the surface area is $\propto r^2$ and $\mathcal{S}$ is arbitrarily far from the binary, $r \to \infty$. When integrating over the bodies' surfaces, we shall flip the direction of the vectorial surface element so it points out of the bodies (into $\mathcal{E}$), thereby picking up a minus sign and bringing equation (13.18) into the form

$$\begin{aligned} \delta p^j_{\text{field}} &= \frac{1}{2\pi}\epsilon_{jpl} S^q_A m_B \left[\int_{\partial\mathcal{A}} \left(\frac{1}{r_A}\right)_{,q} \left(\frac{1}{r_B}\right)_{,l} d\Sigma_p \right. \\ &\quad \left. + \int_{\partial\mathcal{B}} \left(\frac{1}{r_A}\right)_{,q} \left(\frac{1}{r_B}\right)_{,l} d\Sigma_p \right] + (A \leftrightarrow B)\,. \end{aligned} \tag{13.19}$$

We presume (as is required by the PN approximation) that the bodies' separation is large compared to their radii. Then on ∂A, we can write $(1/r_A)_{,q} = -n^q_A/r^2_A$ and $(1/r_B)_{,l} = n^l_{AB}/r^2_{AB}$, where n_A is the unit vector pointing away from the center of mass of body A, n^l_{AB} is the unit vector pointing from the center of mass of body B toward the center of mass of body A, and r_{AB} is the (flat-spacetime) distance between the two bodies' centers of mass. The first integral in equation (13.19) then becomes $n_{AB}/r^2_B \int_{\partial A} n^q_A/r^2_A d\Sigma_p$. For simplicity we take the surface of integration to be a sphere immediately above the physical surface of body A and ignore the tiny contribution from the region between that sphere and the physical surface. On this sphere, we write $d\Sigma_p = r^2_A n^p_A d\Omega_A$, where $d\Omega_A$ is the solid angle element, and we then carry out the angular integral using the relation $\int_{\partial A} n^q_A n^p_A d\Omega_A = (4\pi/3)\delta_{qp}$. Thereby we obtain for the first integral in (13.19) $(4\pi/3)\delta_{qp} n_{AB}/r^2_B$ independent of the radius r_A of the sphere of integration. (If the body is not spherical, the contribution from the tiny volume between our spherical integration surface and the physical surface will be negligible.) Evaluating the second integral in equation (13.19) in the same way, and carrying out straightforward manipulations, we obtain for the external field momentum

$$\delta \boldsymbol{p}_{\text{field}} = \frac{2}{3r^2_{AB}}(m_B \boldsymbol{S}_A - m_A \boldsymbol{S}_B) \times \boldsymbol{n}_{AB}\ . \tag{13.20}$$

13.3.1.1 Field Momentum Specialized to the Extreme-Kick Configuration

For the extreme-kick configuration, which has $m_A = m_B = m$ and $\boldsymbol{S}_B = -\boldsymbol{S}_A$, the field momentum (13.20) becomes

$$\delta \boldsymbol{p}_{\text{field}} = \frac{4}{3}\frac{m}{r_{AB}^2}\boldsymbol{S}_A \times \boldsymbol{n}_{AB} \ . \tag{13.21}$$

Figure 13.5 shows the z-component (perpendicular to the orbital plane) of the field-momentum density $\delta\tau^{0z}$, as measured in the orbital plane at four different moments in the binary's orbital evolution. Only that part of the momentum that flows during bobbing (equation (13.15)) is pictured. Red depicts momentum density flowing out of the paper ($+z$ direction), and blue, into the paper. The yellow arrows show the holes' vectorial spins $\boldsymbol{S}$, and the arrowed circle is the binary's orbital trajectory. In the top-left and bottom-right frames, the black holes are momentarily stationary at the top and bottom of their bobbing. Nevertheless, the momentum density has a nontrivial distribution. In the top-right and bottom-left frames, the black holes are moving downward and upward, respectively, with maximum speed. In both cases, the field-momentum density between the two holes flows in the same direction as the bobbing, whereas the momentum surrounding the binary is in the opposite direction and larger. This leads to net momentum conservation for the binary, as discussed in section 13.4.

It is worth noting that the four figures, going counterclockwise from the top-left, are taken a quarter period apart in orbital phase. The first and third differ by half an orbital period (as do the second and fourth); and, consequently, the momentum patterns of each pair are identical, but signs are reversed (red exchanged with blue, as dictated by the symmetry of the configuration). This feature is responsible for the sinusoidal bobbing.

13.3.2 Centers of Mass and Equation of Motion for the Binary's Compact Bodies

Restrict attention, temporarily, to a body that is a star rather than a black hole, and temporarily omit the subscript K that identifies which body. Then, following the standard procedure in special

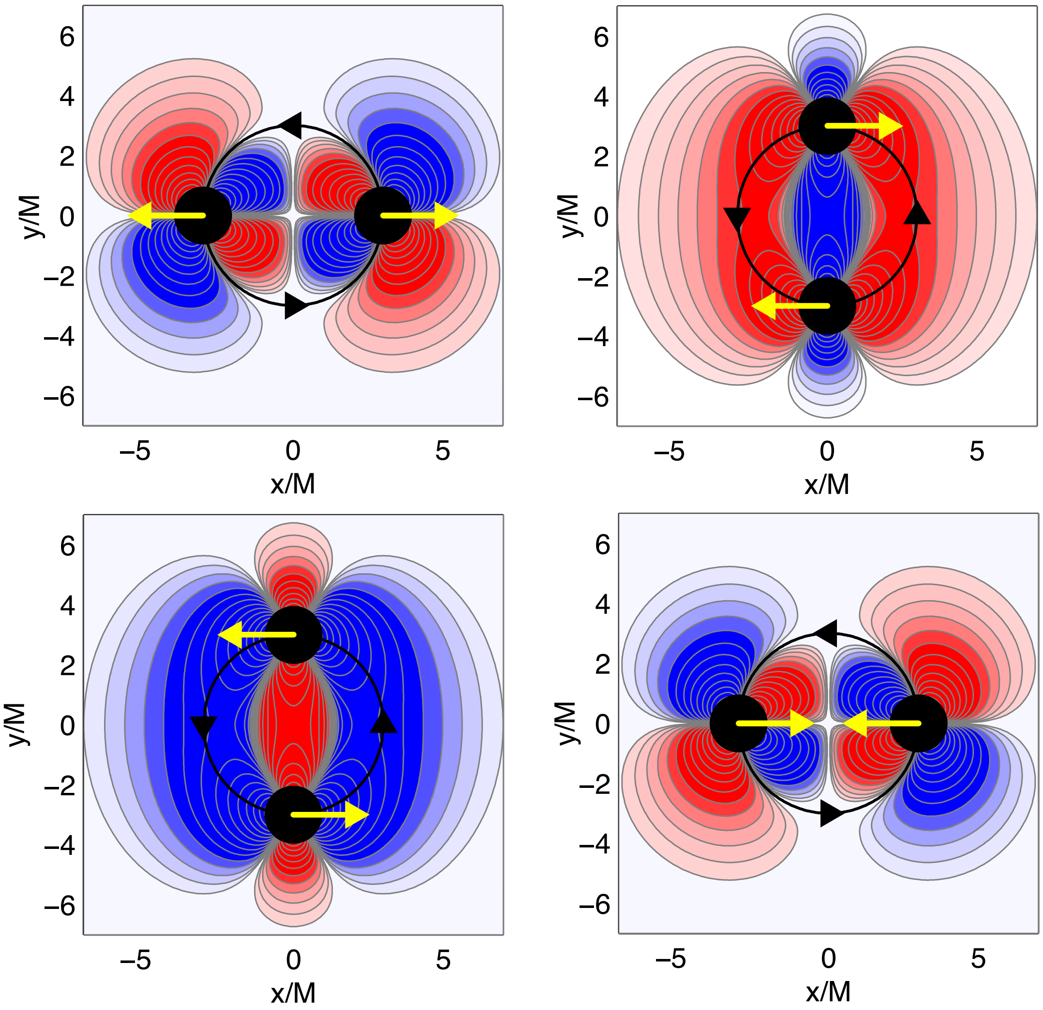

Figure 13.5: Field Momentum Density

The four pictures show the z-component of field-momentum density $\delta\tau^{0z}$ in the orbital plane at four different times, a quarter orbit apart. Red represents positive momentum density (coming out of the paper), and blue, negative (going into the paper). Only the piece of momentum density $\delta\tau^{0z}$ that flows during bobbing (equation (13.15)) is depicted. The yellow arrows are the black holes' vectorial spins; the large, black arrowed circle shows the orbital path of the two holes. In the top left picture, one sees the density of momentum when the black holes are at the top of their bob (maximum z) and momentarily stationary. The gravitational-field momentum is zero, but the momentum density itself shows rich structure. A quarter orbit later, in the top right, the holes are moving downward (into the paper) at top speed. The momentum between the black holes (blue region) flows into the paper with them, while surrounding momentum (red region) flows out of the paper ($+z$-direction). A half orbit after the first picture, in the lower left, the holes are momentarily at rest at the bottom of their bob (minimum z), the net field momentum is zero, and the momentum distribution is opposite that in the first picture (as one would expect during sinusoidal bobbing). Similarly, three quarters of the way through the orbit, in the lower right, the holes have reached their maximum upward speed, and the momentum distribution is identical to the second picture, but with the opposite sign.

relativity (e.g., Box 5.6 of MTW [1]), we define the star's center-of-mass world line to be that set of events $x^\mu_{\rm cm}$ satisfying the covariant field-theory-in-flat-spacetime relationship

$$S^{\alpha\beta} p_\beta = 0 \; . \tag{13.22}$$

Here $p^\beta = \int \tau^{0\beta} d^3x$ is the body's 4-momentum and

$$S^{\alpha\beta} \equiv \int [(x^\alpha - x^\alpha_{\rm cm})\tau^{\beta 0} - (x^\beta - x^\beta_{\rm cm})\tau^{\alpha 0}] d^3x \tag{13.23}$$

is the body's tensorial angular momentum. Here the integrals extend over the star's interior, and because the star's momentum is changing, we take the time component of $x^\mu_{\rm cm}$ to be the same as the time at which the integral is performed, $x^0_{\rm cm} = x^0$. (If the momentum were not changing, this restriction would be unnecessary; cf. Box 5.6 of MTW.)

In a reference frame where the body moves with ordinary velocity $v^j = p^j/p^0$, equation (13.22) says $S^{i0} = S^{ij} v_j$. We wish to rewrite this in a more illuminating form, accurate to first order in the velocity $\boldsymbol{v}$. At that accuracy, we can evaluate S^{ij} in the body's rest frame, obtaining $S^{ij} = \epsilon^{ijk} S_k$ where S_k is the body's spin angular momentum

$$S_k = \int \epsilon_{klm} (x^l - x^l_{\rm cm}) \tau^{m0} d^3x \; . \tag{13.24}$$

Using definition (13.23) of S^{i0} with $x^0_{\rm cm} = x^0$, our definition (13.22) of the center of mass then takes the concrete form

$$m\boldsymbol{x}_{\rm cm} = \int \boldsymbol{x} \tau^{00} d^3x - \boldsymbol{v} \times \boldsymbol{S} \; . \tag{13.25}$$

Here on the left side we have replaced $p^0 = \int \tau^{00} d^3x$ by its value in the body's rest frame, which is the mass m, since the two differ by amounts quadratic in $\boldsymbol{v}$.

Notice that, *when computed in the body's rest frame so* $\boldsymbol{v} = 0$*, the center of mass is* $m\boldsymbol{x}_{\rm cm} = \int \boldsymbol{x}\tau^{00} d^3x$*, but when computed in any frame moving slowly with respect to the rest frame, this expression must be corrected by the term* $-\boldsymbol{v} \times \boldsymbol{S}$. This is called the *physical spin supplementary* *condition* [162].

In our Harmonic coordinate system and at the 1.5PN order of our analysis, the dominant, time-time component of the EFEs (13.3) reduces to $\eta^{\mu\nu}\mathfrak{g}^{00}{}_{,\mu\nu} = 16\pi\tau^{00}$. The type of analysis carried out in section 19.1 of MTW [1] then reveals that *in the star's rest frame, the monopolar part of its* $\mathfrak{g}^{00}$ *is centered on the location* $\boldsymbol{x}_{\rm cm}$*; or, equivalently, when one expands the star's* $\mathfrak{g}^{00}$ *around* $\boldsymbol{x}_{\rm cm}$ *in its own rest frame, there is no dipolar* $1/r^2$ *term (no mass dipole moment). This well-known result (e.g., [163, 164]) can be used as an alternative definition of* $\boldsymbol{x}_{\rm cm}$*—a definition that works for black holes as well as for stars.*

Using this monopolar-field-centered definition of $\boldsymbol{x}_{\rm cm}$, Thorne and Hartle [164] have employed matched asymptotic expansions (valid for black holes) to derive the equations of motion for a system of compact bodies (e.g., a compact binary) (their equations (4.10) and (4.11)). For a compact binary, the spin-induced contributions to these equations of motion at 1.5PN order are (equation (4.11c) of Thorne and Hartle)

$$\begin{aligned} m_A \frac{d\delta\boldsymbol{v}_A}{dt} &= \frac{m_A}{r_{AB}^3}\left[6\boldsymbol{n}_{AB}(\boldsymbol{S}_B \times \boldsymbol{n}_{AB}\cdot\boldsymbol{v}_{AB}) + 4\boldsymbol{S}_B \times \boldsymbol{v}_{AB}\right. \\ &\qquad \left. -6(\boldsymbol{S}_B \times \boldsymbol{n}_{AB})(\boldsymbol{v}_{AB}\cdot\boldsymbol{n}_{AB})\right] \\ &\quad + \frac{m_B}{r_{AB}^3}\left[6\boldsymbol{n}_{AB}(\boldsymbol{S}_A \times \boldsymbol{n}_{AB}\cdot\boldsymbol{v}_{AB}) + 3\boldsymbol{S}_A \times \boldsymbol{v}_{AB}\right. \\ &\qquad \left. -3(\boldsymbol{S}_A \times \boldsymbol{n}_{AB})(\boldsymbol{v}_{AB}\cdot\boldsymbol{n}_{AB})\right] \ . \end{aligned} \tag{13.26}$$

Here

$$\boldsymbol{v}_A \equiv \frac{d\boldsymbol{x}_{{\rm cm}\,A}}{dt} \ , \quad \delta\boldsymbol{v}_A = \frac{d\delta\boldsymbol{x}_{{\rm cm}\,A}}{dt} \ , \quad \boldsymbol{v}_{AB} = \boldsymbol{v}_A - \boldsymbol{v}_B \tag{13.27}$$

are the velocity of (the center of mass of) body A, the spin-induced perturbation of that velocity, and the relative velocity of bodies A and B. The first two lines of equation (13.26) are due to frame dragging by the other body (body B); the last two lines are a force due to the coupling of body A's spin to B's spacetime curvature.

13.3.2.1 Equations of Motion Specialized to the Extreme-Kick Configuration

As in the previous section, we now specialize the above discussion to the extreme-kick configuration. In this configuration we are trying to explain the bobbing motion of the bodies as they orbit each other (i.e., the motion of the bodies perpendicular to the standard orbital plane). As noted above for the equal mass case, the symmetries of this configuration are such that $\boldsymbol{v}_A = -\boldsymbol{v}_B$, $\boldsymbol{S}_A = -\boldsymbol{S}_B$, and $\boldsymbol{v}_{AB} \cdot \boldsymbol{n}_{AB} = 0$. Since we are only interested in the portion of the equations of motion perpendicular to the orbital plane, the necessary portions of the equations of motion we shall look at are:

$$\left(\frac{d^2 \delta \boldsymbol{x}_{\mathrm{cm}\,A}}{dt^2}\right)_{\mathrm{FD}} = \frac{1}{r_{AB}^3} 4\boldsymbol{S}_B \times \boldsymbol{v}_{AB} \; , \tag{13.28a}$$

$$\left(\frac{d^2 \delta \boldsymbol{x}_{\mathrm{cm}\,A}}{dt^2}\right)_{\mathrm{SC}} = \frac{1}{r_{AB}^3} 3\boldsymbol{S}_A \times \boldsymbol{v}_{AB} \; , \tag{13.28b}$$

where the subscript FD refers to the frame dragging piece, while the subscript SC refers to the spin-curvature coupling piece. The sum of these two effects gives

$$\left(\frac{d^2 \delta \boldsymbol{x}_A}{dt^2}\right)_{\text{spin effects}} = -\frac{2}{r_{AB}^3} \boldsymbol{S}_A \times \boldsymbol{v}_A \; , \tag{13.29}$$

where we have used the fact that $\boldsymbol{S}_A = -\boldsymbol{S}_B$ and $\boldsymbol{v}_{AB} = \boldsymbol{v}_A - \boldsymbol{v}_B = 2\boldsymbol{v}_A$. We get the same acceleration of hole B by replacing all subscript A's by subscript B's because $\boldsymbol{S}_B = -\boldsymbol{S}_A$, $\boldsymbol{v}_B = -\boldsymbol{v}_A$.

We can easily integrate this equation in time by noting that the spin precesses much more slowly than the orbital motion so $\boldsymbol{S}_A$ can be approximated as constant, and noting that $\boldsymbol{v}_A$ rotates with angular velocity $\Omega = \sqrt{M/r_{AB}^3} = \sqrt{2m/r_{AB}^3}$, where M is the total mass and m is the mass of each hole. The result, after one integration, can be written as

$$m\delta \boldsymbol{v}_A = m\delta \boldsymbol{v}_B = -\frac{m}{r_{AB}^2} \boldsymbol{S}_A \times \boldsymbol{n}_{AB} \; . \tag{13.30}$$

These pieces of the total momentum we will call the *kinetic* momentum.

13.3.3 The Momenta of the Binary's Bodies

Let us now turn our attention to the total momentum contained by the bodies. This is done for two cases, first for stars where we can integrate over the interior, and then for black holes where we must convert the volume integrals to surface integrals. We shall find that both methods yield the same result. We initially omit the star's label A or B for ease of notation.

13.3.3.1 The Momentum of a Star

For a star we can derive an expression for the momentum $p^j = \int \tau^{0j} d^3x$ (with the integral over the star's interior) in terms of the star's velocity $v^j = dx^j_{\rm cm}/dt$ by differentiating the center-of-mass equation (13.25) with respect to time. To allow for the possibility that the mass might change with time, we set $m = \int \tau^{00} d^3x$ before doing the differentiation; i.e., we differentiate

$$\boldsymbol{x}_{\rm cm} \int_A \tau^{00} d^3x = \int_A \boldsymbol{x} \tau^{00} d^3x - \boldsymbol{v} \times \boldsymbol{S} \, . \tag{13.31}$$

Using $\tau^{00}{}_{,0} = -\tau^{0k}{}_{,k}$ and Gauss's theorem, we bring the left side into the form $\boldsymbol{v} \int_A \tau_{00} d^3x - \boldsymbol{x}_{\rm cm} \int_{\partial A} (\tau^{0j} - \tau^{00} v^j) d\Sigma_j$. The last term arises from the motion of the surface of the star through space with velocity $\boldsymbol{v}$. Manipulating the time derivative of the integral on the right side of equation (13.31) in this same way, we bring it into the form $\int_A \tau^{0j} d^3x - \int_{\partial A} x^j (\tau^{0k} - \tau^{00} v^k) d\Sigma_k = p^j - \int_{\partial A} x^j (\tau^{0k} - \tau^{00} v^k) d\Sigma_k$, where p^j is the star's momentum. Inserting these expressions for the left side and the right-side integral into equation (13.31), noting that the star's spin angular momentum evolves (due to precession) far more slowly than its velocity, denoting the time derivative of its velocity by $d\boldsymbol{v}/dt = \boldsymbol{a}$ (acceleration), solving for $\boldsymbol{p}$, and restoring subscript As, we obtain

$$\boldsymbol{p}_A = m_A \boldsymbol{v}_A + \int_{\partial A} (\boldsymbol{x} - \boldsymbol{x}_{{\rm cm}\,A})(\tau^{0k} - \tau^{00} v^k_A) d\Sigma_k + \boldsymbol{a}_A \times \boldsymbol{S}_A \, . \tag{13.32}$$

Although we have derived this equation for a star, it must be true also for a black hole. The reason is that all the quantities that appear in it are definable without any need for integrating over the body's interior, and all are expressible in terms of the binary's masses and spins and its bodies' vectorial separation, in manners that are insensitive to whether the bodies are stars or holes. To illustrate this statement, in section 13.3.3.2 we deduce (13.32) for a black hole, restricting ourselves to spin-induced portion of the momentum that is being exchanged with the field, $\delta \boldsymbol{p}_A$.

It is this $\delta \boldsymbol{p}_A$ that interests us. Because the spin has no influence on τ^{00} at the relevant order (which is $\delta\tau^{0k} \sim gH$ and $\delta\tau^{00} \sim g^2$ where g and H are the gravitoelectric and gravitomagnetic fields), equation (13.32) implies that

$$\delta \boldsymbol{p}_A = m_A \delta \boldsymbol{v}_A + \int_{\partial A} (\boldsymbol{x} - \boldsymbol{x}_{\mathrm{cm}\,A}) \delta\tau^{0k} d\Sigma_k + \boldsymbol{a}_A \times \boldsymbol{S}_A \ . \tag{13.33}$$

The acceleration $\boldsymbol{a}_A$ of body A is, at the order needed, just the gravitoelectric field of body B at the location of A, $\boldsymbol{a}_A = -(m_B/r_{AB}^2)\boldsymbol{n}_{AB}$. Performing the surface integral on a sphere just above the body's physical surface we can write $\boldsymbol{x} - \boldsymbol{x}_{\mathrm{cm}\,A} = \boldsymbol{n}_A r_A$ and $d\Sigma_k = r_A^2 n_A^k d\Omega_A$. Inserting these into equation (13.33), we obtain

$$\delta \boldsymbol{p}_A = \underbrace{m_A \delta \boldsymbol{v}_A}_{\substack{\text{kinetic}\\\text{term}}} \underbrace{+ \int_{\partial A} r_A^3 \delta\tau^{0k} \boldsymbol{n}_A n_A^k d\Omega_A}_{\text{surface term}} \underbrace{+ \frac{m_B}{r_{AB}^2} \boldsymbol{S}_A \times \boldsymbol{n}_{AB}}_{\text{SSC term}} \ . \tag{13.34}$$

Here "SSC term" refers to the "spin supplementary condition" required to get the correct, physical center of mass; see text following equation (13.25). In the surface term, the field momentum density $\delta\tau^{0k}$ is given by equation (13.17). The second term ($A \leftrightarrow B$) is smaller than the first by M/r_{AB} and thus is negligible. Inserting the first term into the integral, using $(1/r_A)_{,qp} = (3n_A^q n_A^p - \delta_{qp})/r_A^3$ and $(1/r_B)_{,l} = -n_B^l/r_B^2$, and $\int n_A^j n_A^l d\Omega_A = \frac{4\pi}{3}\delta_{jl}$, we bring equation (13.34) into the form

$$\begin{aligned} \delta \boldsymbol{p}_A &= \underbrace{m_A \delta \boldsymbol{v}_A}_{\substack{\text{kinetic}\\\text{term}}} \underbrace{- \frac{2}{3}\frac{m_B}{r_{AB}^2} \boldsymbol{S}_A \times \boldsymbol{n}_{AB}}_{\text{surface term}} \underbrace{+ \frac{m_B}{r_{AB}^2} \boldsymbol{S}_A \times \boldsymbol{n}_{AB}}_{\text{SSC term}} \\ &= m_A \delta \boldsymbol{v}_A + \frac{1}{3}\frac{m_B}{r_{AB}^2} \boldsymbol{S}_A \times \boldsymbol{n}_{AB} \ . \end{aligned} \tag{13.35}$$

13.3.3.2 The Momentum of a Black Hole

In section 13.3.3.1 we derived expression (13.32) for the momentum of a body in a binary, assuming the body is a star so we could do volume integrals, and we then asserted that this expression is also valid for black holes. The spin-induced portion of this expression that gets exchanged with the field as the body moves is given by equation (13.33), which reduces to (13.35). In this section we shall sketch a derivation of equation (13.35) directly from the surface-integral definition (13.8b) of a black hole's momentum,

$$\delta p^j_A = \frac{1}{16\pi} \int_{\partial A} \delta H^{j\alpha 0k}{}_{,\alpha} d\Sigma_k \ . \tag{13.36}$$

To evaluate this surface integral up to desired 1.5PN-order accuracy turns out to require some 2.5PN fields. Qualitatively, this can be anticipated because the superpotential we use in the surface integral is *sourced* by the spin-orbit piece of field momentum, and therefore necessarily a nonleading PN term. One can see this more clearly by expanding $\delta H^{j\alpha 0k}{}_{,\alpha}$ in terms of the metric density and using the symmetries of the superpotential H (which are the same as the Riemann tensor). In general, the momentum is given by

$$\delta p^j_A = -\frac{1}{16\pi} \int_{\partial A} (\mathfrak{g}^{jk}\mathfrak{g}^{\alpha 0} - \mathfrak{g}^{j\alpha}\mathfrak{g}^{0k})_{,\alpha} d\Sigma_k \ . \tag{13.37}$$

In Harmonic gauge, however, $\mathfrak{g}^{\alpha\beta}{}_{,\beta} = 0$, and the spatial metric is flat until 2PN order while the time-space components are of 1.5PN order. As a result, the terms at lowest and next-to-lowest PN order are contained within two terms,

$$\delta p^j_A = \frac{1}{16\pi} \int_{\partial A} (\mathfrak{g}^{j0}{}_{,k} + \mathfrak{g}^{jk}{}_{,0}) d\Sigma_k \ . \tag{13.38}$$

In this expression, the momentum arises from linear terms involving the metric density, instead of quadratic ones. As a result, one must keep pieces of the metric perturbation that are of higher PN accuracy. (Note: if we were to evaluate the time derivative of $\delta \boldsymbol{p}_A$ using the surface integral (13.9), we would not face such a delicacy; the integrand there is quadratic and requires only 1.5PN fields for its evaluation.)

To find the momentum in terms of the standard post-Newtonian parameters, we resort to a standard way that the metric perturbations are written in recent post-Newtonian literature, e.g., by Blanchet, Faye, and Ponsot [165]:

$$\begin{aligned} g_{00} &= -1+2V-2V^2+8\hat{X}\ , & (13.39\text{a})\\ g_{i0} &= -4V_i-8\hat{R}_i\ , & (13.39\text{b})\\ g_{ij} &= \delta_{ij}(1+2V+2V^2)+4\hat{W}_{ij}\ , & (13.39\text{c})\\ \sqrt{-g} &= 1+2V+4V^2+2\hat{W}_{kk}\ . & (13.39\text{d}) \end{aligned}$$

For spinning systems, we adopt the notation of Tagoshi, Ohashi, and Owen [166] where $\mathcal{O}(m,n)$ means to order c^m for nonspinning terms and χc^n for terms involving a single spin χ. (Here $\chi = |\boldsymbol{S}|/m^2$ is the body's dimensionless spin.) In this notation, terms we are interested in are of the order $\mathcal{O}(3,6)$, while the above post-Newtonian potentials have been obtained up to the following orders [166, 167]:

$$\begin{gathered} V=\mathcal{O}(2,5),\quad V_j=\mathcal{O}(3,4),\\ \hat{W}_{jk}=\mathcal{O}(4,5),\quad \hat{R}_j=\mathcal{O}(5,6). \end{gathered} \tag{13.40}$$

In terms of these post-Newtonian potentials, V, V_i, $\hat{R}_i$, $\hat{X}$ and $\hat{W}_{ij}$, the perturbed metric density is

$$\begin{aligned} \mathfrak{g}^{00} &= -1-4V-2\left(\hat{W}_{kk}+4V^2\right)+\mathcal{O}(6,7), & (13.41\text{a})\\ \mathfrak{g}^{0i} &= -4V_i-8\left(\hat{R}_i+VV_i\right)+\mathcal{O}(6,7), & (13.41\text{b})\\ \mathfrak{g}^{ij} &= \delta_{ij}-4\left(\hat{W}_{ij}-\frac{1}{2}\delta_{ij}\hat{W}_{kk}\right)+\mathcal{O}(6,7). & (13.41\text{c}) \end{aligned}$$

As a consequence, equation (13.38) is given by

$$\begin{aligned} \delta p^j_A &= \frac{1}{16\pi}\int_{\partial A}\left\{\left[-4V_j - 8\left(\hat{R}_{j(S)} + V_{(m)}V_{j(S)}\right)\right]_{,k}\right. \\ &\quad \left. -4\left[\hat{W}_{jk(S)} - \frac{1}{2}\delta_{jk}\hat{W}_{ii(S)}\right]_{,0}\right\}d\Sigma_k\,, \end{aligned} \tag{13.42}$$

where a subscript (S) means keep only the parts of those potentials proportional to the spins of the bodies, and a subscript (M) involves monopolar pieces of the potential without spins (proportional to the masses of the bodies). Terms without a subscript have both pieces.

Tagoshi, Ohashi and Owen express the potentials $V_{(M)}$, V_j, $\hat{R}_{j(S)}$ and $\hat{W}_{jk(S)}$ in terms of the bodies' masses, vectorial velocities, vectorial spins, and vectorial separations, and distance to the field-point location (their equations (A1a), (A1d), (A1f), and (A1g)). While the full equations are quite lengthy, the portions that generate momentum flow—those involving the coupling of the mass of one body to the spin of the other—are somewhat simpler. For convenience, we give these portions of the equations below, rewritten in our notation, with the typos noted by Faye, Blanchet, and Bounanno [167] corrected:

$$V_{(M)} = \frac{m_A}{r_A} + (A \leftrightarrow B)\,, \tag{13.43a}$$

$$\begin{aligned} V_j &= \frac{m_A v^j_A}{r_A} + \epsilon_{jkl}S^k_A\left\{n^l_A\left[-\frac{3m_B}{2r^2_A r_{AB}} - \frac{m_B(\boldsymbol{n}_A\cdot\boldsymbol{n}_{AB})}{4r_A r^2_{AB}}\right] + n^l_{AB}\frac{3m_B}{4r_A r_{AB}}\right\} \\ &\quad +(A \leftrightarrow B)\,, \end{aligned} \tag{13.43b}$$

$$\hat{W}_{jk(S)} = \frac{n^i_A}{r^2_A}\left[\frac{1}{2}\left(\epsilon_{ijl}S^l_A v^k_A + \epsilon_{ikl}S^l_A v^j_A\right) - \delta_{jk}\epsilon_{ilm}v^l_A S^m_A\right] + (A \leftrightarrow B)\,, \tag{13.43c}$$

$$\begin{aligned} \hat{R}_{j(S)} &= \epsilon_{jkl}S^k_A\left[n^l_A\left(-\frac{m_B}{2r^2_A r_{AB}} + \frac{m_B}{r_{AB}s^2}\right) + n^l_{AB}\left(-\frac{m_B}{2r_A r^2_{AB}} + \frac{m_B}{2r^2_{AB}r_B} + \frac{m_B}{r_A s^2}\right)\right. \\ &\quad \left. +n^l_B\left(\frac{m_B}{r_A s^2} + \frac{m_B}{r_{AB}s^2}\right)\right] \\ &\quad +n^j_A\epsilon_{ikl}S^l_A\left[n^i_A(n^k_{AB} + n^k_B)\left(\frac{m_B}{r_A s^2} + \frac{2m_B}{s^3}\right) - 2n^i_{AB}n^k_B\frac{m_B}{s^3}\right] \\ &\quad +n^j_{AB}\epsilon_{ikl}S^l_A\left[-2n^i_A n^k_b\frac{m_B}{s^3} + (n^i_A + n^i_B)n^k_{AB}\left(-\frac{m_B}{r_{AB}s^2} - \frac{2m_B}{s^3}\right)\right] \\ &\quad +(A \leftrightarrow B)\,. \end{aligned} \tag{13.43d}$$

Here, as before, m_A, $\boldsymbol{v}_A$, and $\boldsymbol{S}_A$ are the mass, velocity, and spin angular-momentum of object A; r_A is the separation of body A from a point in space and r_{AB} is the separation of the two objects; and $\boldsymbol{n}_A$ and $\boldsymbol{n}_{AB}$ are unit vectors pointing along r_A and r_{AB}, respectively. A new quantity, $s = r_A + r_B + r_{AB}$, has been introduced, in addition. Inserting these expressions into equation (13.42) gives us $\delta H^{j\alpha 0k}{}_{,\alpha}$, and the momentum of body A is then found by performing a surface integral over A's surface. The surface integrals are computed under the same assumptions as in section 13.3.1; namely the separation of the bodies is much larger than their radii, and each surface of integration is a sphere immediately above a body's surface. When they are computed, they give the same result as equation (13.35),

$$\delta \boldsymbol{p}_A = m_A \delta \boldsymbol{v}_A + \frac{1}{3}\frac{m_B}{r_{AB}^2}\boldsymbol{S}_A \times \boldsymbol{n}_{AB} \; . \tag{13.44}$$

As before, the momentum for body B is given by exchanging A and B.

As a consistency check, we can evaluate the system's total momentum by doing a surface integral at infinity:

$$\delta p^j_{\text{tot}} = \frac{1}{16\pi}\oint_{\mathcal{S}} \delta H^{j\alpha 0k}{}_{,\alpha} d\Sigma_k \; . \tag{13.45}$$

The quantity $\delta H^{j\alpha 0k}{}_{,\alpha}$ is exactly the same as above, from which one can find

$$\delta \boldsymbol{p}_{\text{tot}} = m_A \delta \boldsymbol{v}_A + \frac{m_B}{r_{AB}^2}\boldsymbol{S}_A \times \boldsymbol{n}_{AB} + (A \leftrightarrow B) \; . \tag{13.46}$$

This, combined with the fact that

$$\delta \boldsymbol{p}_{\text{tot}} = \delta \boldsymbol{p}_A + \delta \boldsymbol{p}_B + \delta \boldsymbol{p}_{\text{field}} \; , \tag{13.47}$$

as well as equation (13.44), gives

$$\delta \boldsymbol{p}_{\text{field}} = \frac{2m_B}{3r_{AB}^2}\boldsymbol{S}_A \times \boldsymbol{n}_{AB} + (A \leftrightarrow B), \tag{13.48}$$

as found in section 13.3.1.

13.4 Momentum Conservation

The total spin-induced momentum perturbation, $\delta \boldsymbol{p}_{\rm tot} = \delta \boldsymbol{p}_A + \delta \boldsymbol{p}_B + \delta \boldsymbol{p}_{\rm field}$ (equations (13.35) and (13.20)) is

$$\delta \boldsymbol{p}_{\rm tot} = m_A \delta \boldsymbol{v}_A + m_B \delta \boldsymbol{v}_B + \frac{1}{r_{AB}^2} (m_B \boldsymbol{S}_A - m_A \boldsymbol{S}_B) \times \boldsymbol{n}_{AB} \ . \tag{13.49}$$

Momentum conservation requires that the time derivative of this $\delta \boldsymbol{p}_{\rm tot}$ vanish. The time derivative of the kinetic terms can be read off the equation of motion (13.26):

$$\begin{aligned} & m_A \frac{d\delta \boldsymbol{v}_A}{dt} + m_B \frac{d\delta \boldsymbol{v}_B}{dt} \\ & = -(m_B \boldsymbol{S}_A - m_A \boldsymbol{S}_B) \times [\boldsymbol{v}_{AB} - 3(\boldsymbol{n}_{AB} \cdot \boldsymbol{v}_{AB}) \boldsymbol{n}_{AB}] \ . \end{aligned} \tag{13.50}$$

By inserting $\boldsymbol{n}_{AB} = (\boldsymbol{x}_{{\rm cm}\, A} - \boldsymbol{x}_{{\rm cm}\, B})/r_{AB}$ into the second term of equation (13.49) and differentiating with respect to time, we obtain the negative of expression (13.50). Therefore,

$$d\delta \boldsymbol{p}_{\rm tot}/dt = 0 \ ; \tag{13.51}$$

i.e., as the binary's evolution drives spin-induced momentum back and forth between the bodies and the field, the total momentum remains conserved, as it must.

Interestingly, during the summation of momentum terms, one finds that the surface terms in $\delta \boldsymbol{p}_A + \delta \boldsymbol{p}_B$ have exactly cancelled the field momentum $\delta \boldsymbol{p}_{\rm field}$, leaving the total momentum as the sum of the bodies' kinetic term and their SSC term.

13.4.1 Momentum Conservation for Extreme-Kick Configuration

Let us examine in detail how momentum conservation is achieved in the presence of the bodies' bobbing. Our detailed analysis (above) breaks each object's momentum perturbation $\delta \boldsymbol{p}_{A,B}$ into three terms, the *kinetic momentum* $m\delta \boldsymbol{v}_{A,B}$, a term due to the *SSC*, and a *surface* integral term

Table 13.1: Pieces of Total Momentum

Body	Kinetic		SSC	Surface	Total
	Frame-Dragging	Spin-Curvature			
$\boldsymbol{p}_A$	$-4\boldsymbol{a}_B \times \boldsymbol{S}_B$	$3\boldsymbol{a}_A \times \boldsymbol{S}_A$	$\boldsymbol{a}_A \times \boldsymbol{S}_A$	$-\frac{2}{3}\boldsymbol{a}_A \times \boldsymbol{S}_A$	$-\frac{2}{3}\boldsymbol{a}_A \times \boldsymbol{S}_A$
$\boldsymbol{p}_B$	$-4\boldsymbol{a}_A \times \boldsymbol{S}_A$	$3\boldsymbol{a}_B \times \boldsymbol{S}_B$	$\boldsymbol{a}_B \times \boldsymbol{S}_B$	$-\frac{2}{3}\boldsymbol{a}_B \times \boldsymbol{S}_B$	$-\frac{2}{3}\boldsymbol{a}_B \times \boldsymbol{S}_B$
$\boldsymbol{p}_{\rm field}$	$\frac{2}{3}(\boldsymbol{a}_A \times \boldsymbol{S}_A + \boldsymbol{a}_B \times \boldsymbol{S}_B)$				

Spin-dependent, time-varying pieces of body and field momenta at 1.5PN order, for the extreme-kick binary (circular orbit with spins antialigned and in the orbital plane). The body momenta are broken down into kinetic, SSC and surface terms and are expressed in terms of the bodies' spins $\boldsymbol{S}_{A,B}$ and Newtonian-order gravitational accelerations $\boldsymbol{a}_{A,B} = -m_{B,A}\boldsymbol{n}_{AB,BA}/r_{AB}^2$. See equations ((13.26)) and ((13.35)) for a similar decomposition in a generic binary.

(see table 13.1). The total kinetic momentum

$$m\delta\boldsymbol{v}_A + m\delta\boldsymbol{v}_B = -\left(\boldsymbol{a}_A \times \boldsymbol{S}_A + \boldsymbol{a}_B \times \boldsymbol{S}_B\right) \neq 0 \,, \tag{13.52}$$

is not conserved because of the noncancellation between the frame-dragging and spin-curvature coupling terms. The total body momentum $\delta\boldsymbol{p}_A + \delta\boldsymbol{p}_B$ is not conserved either; it sums up to 2/3 the total kinetic momentum:

$$\delta\boldsymbol{p}_A + \delta\boldsymbol{p}_B = -\frac{2}{3}(\boldsymbol{a}_A \times \boldsymbol{S}_A + \boldsymbol{a}_B \times \boldsymbol{S}_B) \neq 0 \,, \tag{13.53}$$

To achieve momentum conservation, there is a nonzero spin-dependent total field momentum distributed outside of the bodies, with

$$\delta\boldsymbol{p}_{\rm field} = \frac{2}{3}(\boldsymbol{a}_A \times \boldsymbol{S}_A + \boldsymbol{a}_B \times \boldsymbol{S}_B) \,. \tag{13.54}$$

Note that this total external field momentum is only $-2/3$ the spin-dependent total kinetic momentum—instead of the -1 that one might have naively expected.

13.5 Summary

We have studied generic binary compact objects and mometum flow that occurs between the bodies and the surrounding spacetime using the Landau-Lifshitz formalism. In this analysis, we have found that the momentum of the absolute motion of the binary system is balanced by the momentum contained in the surrounding gravitational field. These results were also specialized to the interesting case of spinning black holes in the extreme-kick configuration. We think this formalism holds great promise for understanding the dynamics of compact objects and anticipate additional insights and intuition gained from applying this sort of analysis to other interacting black hole environments.

Bibliography

[1] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Company, NY, 1973).

[2] K. S. Thorne, Reviews of Modern Physics **52**(2), 299 (1980).

[3] L. Blanchet, B. R. Iyer, C. M. Will, and A. G. Wiseman, Classical and Quantum Gravity **13**(4), 575 (1996).

[4] A. Buonanno and T. Damour, Physical Review D **59**(8), 084006 (1999).

[5] A. Buonanno and T. Damour, Physical Review D **62**(6), 064015 (2000).

[6] T. Damour, P. Jaranowski, and G. Schäfer, Physical Review D **62**(8), 084011 (2000).

[7] T. Damour, Physical Review D **64**(12), 124013 (2001).

[8] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, Physical Review D **63**(4), 044023 (2001).

[9] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, Physical Review D **57**(2), 885 (1998).

[10] A. Buonanno, Y. Chen, and M. Vallisneri, Physical Review D **67**(10), 104025 (2003), erratum-ibid. 74 (2006) 029904(E).

[11] L. E. Kidder, Physical Review D **52**(2), 821 (1995).

[12] L. Blanchet, T. Damour, B. R. Iyer, C. M. Will, and A. G. Wiseman, Physical Review Letters **74**(18), 3515 (1995).

[13] L. Blanchet, Physical Review D **54**(2), 1417 (1996).

[14] L. Blanchet, G. Faye, B. R. Iyer, and B. Joguet, Physical Review D **65**(6), 061501 (2002).

[15] L. E. Kidder, C. M. Will, and A. G. Wiseman, Physical Review D **47**(10), R4183 (1993).

[16] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Physical Review D **49**(12), 6274 (1994).

[17] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **69**(12), 122001 (2004).

[18] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **72**(8), 082001, 082001 (2005).

[19] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **72**(8), 082002, 082002 (2005).

[20] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, J. Agresti, P. Ajith, B. Allen, J. Allen, R. Amin, S. B. Anderson, *et al.* (LIGO Scientific Collaboration), Physical Review D **73**(6), 062001, 062001 (2006).

[21] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **77**(6), 062002, 062002 (2008).

[22] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **78**(4), 042002, 042002 (2008).

[23] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration and K. Hurley), The Astrophysical Journal **681**(2), 1419 (2008).

[24] A. J. Weinstein (for the LIGO Scientific Collaboration), Classical and Quantum Gravity **21**(5), S677 (2004).

[25] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **72**(4), 042002 (2005).

[26] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, J. Agresti, B. Allen, J. Allen, R. Amin, S. B. Anderson, W. G. Anderson, *et al.* (LIGO Scientific Collaboration), Physical Review D **72**(6), 062001 (2005).

[27] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, J. Agresti, P. Ajith, B. Allen, J. Allen, R. Amin, S. B. Anderson, *et al.* (LIGO Scientific Collaboration and TAMA Collaboration), Physical Review D **72**(12), 122004 (2005).

[28] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, J. Allen, R. Amin, S. B. Anderson, W. G. Anderson, *et al.* (LIGO Scientific Collaboration), Classical and Quantum Gravity **23**(8), S29 (2006).

[29] L. Baggio, M. Bignotto, M. Bonaldi, M. Cerdonio, M. D. Rosa, P. Falferi, S. Fattori, P. Fortini, G. Giusfredi, M. Inguscio, *et al.* (LIGO Scientific Collaboration and AURIGA Collaboration), Classical and Quantum Gravity **25**(9), 095004 (2008).

[30] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **76**(6), 062003 (2007).

[31] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Classical and Quantum Gravity **24**(22), 5343 (2007).

[32] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **77**(6), 062004 (2008).

[33] B. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain, *et al.* (LIGO Scientific Collaboration), Classical and Quantum Gravity **25**(24), 245008 (2008).

[34] B. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review Letters **101**(21), 211102 (2008).

[35] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **69**(8), 082004 (2004).

[36] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, J. Agresti, B. Allen, J. Allen, R. Amin, S. B. Anderson, W. G. Anderson, *et al.* (LIGO Scientific Collaboration), Physical Review D **72**(10), 102004 (2005).

[37] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review Letters **94**(18), 181103 (2005).

[38] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **76**(8), 082001 (2007).

[39] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **76**(4), 042001 (2007).

[40] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **77**(2), 022001 (2008).

[41] B. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. Amin, D. P. Anderson, S. B. Anderson, W. G. Anderson, *et al.* (LIGO Scientific Collaboration), Physical Review D **79**(2), 022001 (2009).

[42] B. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain, *et al.* (LIGO Scientific Collaboration), The Astrophysical Journal Letters **683**(1), L45 (2008).

[43] B. P. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. S. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review Letters **102**(11), 111102 (2009).

[44] B. Abbott, R. Abbott, R. Adhikari, A. Ageev, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, H. Armandula, *et al.* (LIGO Scientific Collaboration), Physical Review D **69**(12), 122004 (2004).

[45] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, J. Allen, R. Amin, S. B. Anderson, W. G. Anderson, *et al.* (LIGO Scientific Collaboration), Physical Review Letters **95**(22), 221101 (2005).

[46] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Araya, *et al.* (LIGO Scientific Collaboration), The Astrophysical Journal **659**(2), 918 (2007).

[47] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration), Physical Review D **76**(8), 082003 (2007).

[48] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, *et al.* (LIGO Scientific Collaboration and ALLEGRO Collaboration), Physical Review D **76**(2), 022001 (2007).

[49] H. Collins, *Gravity's Shadow: The Search for Gravitational Waves* (The University of Chicago Press, Chicago, 2004).

[50] J. Weber, Physical Review Letters **17**(24), 1228 (1966).

[51] J. Weber, Physical Review Letters **22**(24), 1320 (1969).

[52] J. Weber, Physical Review Letters **25**(3), 180 (1970).

[53] B. L. Brown, A. P. Mills, and J. A. Tyson, Physical Review D **26**(6), 1209 (1982).

[54] E. Mauceli, Z. K. Geng, W. O. Hamilton, W. W. Johnson, S. Merkowitz, A. Morse, B. Price, and N. Solomonson, Physical Review D **54**(2), 1264 (1996).

[55] G. A. Prodi, L. Conti, R. Mezzena, S. Vitale, L. Taffarello, J.-P. Zendri, L. Baggio, M. Cerdonio, A. Colombo, V. Crivelli Visconti, *et al.*, in E. Coccia, G. Veneziano, and P. G., eds., *GRAVITATIONAL WAVES - Proceedings of the Second Edoardo Amaldi Conference* (World Scientific, Singapore, 1999), pp. 148–158.

[56] P. Astone, M. Bassan, P. Bonifazi, P. Carelli, M. G. Castellano, G. Cavallari, E. Coccia, C. Cosmelli, V. Fafone, S. Frasca, *et al.*, Physical Review D **47**(2), 362 (1993).

[57] P. Astone, M. Bassan, P. Bonifazi, P. Carelli, E. Coccia, C. Cosmelli, V. Fafone, S. Frasca, A. Marini, G. Mazzitelli, *et al.*, Astroparticle Physics **7**(3), 231 (1997).

[58] D. G. Blair, E. N. Ivanov, M. E. Tobar, P. J. Turner, F. van Kann, and I. S. Heng, Physical Review Letters **74**(11), 1908 (1995).

[59] R. L. Forward, General Relativity and Gravitation **2**(2), 149 (1971).

[60] A. de Waard, Y. Benzaim, G. Frossati, L. Gottardi, H. van der Mark, J. Flokstra, M. Podt, M. Bassan, Y. Minenkov, A. Moleti, *et al.*, Classical and Quantum Gravity **22**(10), S215 (2005).

[61] O. D. Aguiar, L. A. Andrade, J. J. Barroso, F. Bortoli, L. A. Carneiro, P. J. Castro, C. A. Costa, K. M. F. Costa, J. C. N. de Araujo, A. U. de Lucena, *et al.*, Classical and Quantum Gravity **23**(8), S239 (2006).

[62] R. L. Forward, Physical Review D **17**(2), 379 (1978).

[63] R. Weiss, in *Lincoln Research Laboratory of Electronics Quarterly Progress Report* (1972), no. 105, pp. 54–76.

[64] K. Tsubono (TAMA Collaboration), in K. Tsubono, M.-K. Fujimoto, and K. Kuroda, eds., *Proceedings of the TAMA International Workshop on Gravitational Wave Detection (Saitama, 1996)* (Universal Academic Press, Tokyo, 1997), p. 183.

[65] K. Danzmann, in E. Coccia, G. Pizzella, and F. Ronga, eds., *First Edoardo Amaldi Conference on Gravitational Wave Experiments* (World Scientific, Singapore, 1995), p. 100.

[66] B. Caron, A. Dominjon, C. Drezen, R. Flaminio, X. Grave, F. Marion, L. Massonnet, C. Mehmel, R. Morand, B. Mours, *et al.*, Classical and Quantum Gravity **14**(6), 1461 (1997).

[67] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, *et al.*, Science **256**(5055), 325 (1992).

[68] *LISA homepage*, accessed April 1, 2009, http://lisa.nasa.gov/.

[69] S. Kawamura, M. Ando, T. Nakamura, K. Tsubono, T. Tanaka, I. Funaki, N. Seto, K. Numata, S. Sato, K. Ioka, *et al.*, Journal of Physics: Conference Series **122**, 012006 (2008).

[70] M. V. Sazhin, Soviet Astronomy **22**, 36 (1978).

[71] S. Detweiler, The Astrophysical Journal **234**(3), 1100 (1979).

[72] The Planck Collaboration, ArXiv Astrophysics e-prints (2006), arXiv:astro-ph/0604069.

[73] M. R. Nolta, J. Dunkley, R. S. Hill, G. Hinshaw, E. Komatsu, D. Larson, L. Page, D. N. Spergel, C. L. Bennett, B. Gold, *et al.*, The Astrophysical Journal Supplement Series **180**(2), 296 (2009).

[74] B. G. Keating, P. A. R. Ade, J. J. Bock, E. Hivon, W. L. Holzapfel, A. E. Lange, H. Nguyen, and K. W. Yoon, in S. Fineschi, ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (2003), vol. 4843 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 284–295.

[75] C. Pryke, P. Ade, J. Bock, M. Bowden, M. L. Brown, G. Cahill, P. G. Castro, S. Church, T. Culverhouse, R. Friedman, *et al.*, The Astrophysical Journal **692**(2), 1247 (2009).

[76] A. C. Taylor, New Astronomy Reviews **50**(11-12), 993 (2006).

[77] C. Bischoff, L. Hyatt, J. J. McMahon, G. W. Nixon, D. Samtleben, K. M. Smith, K. Vanderlinde, D. Barkats, P. Farese, T. Gaier, *et al.*, The Astrophysical Journal **684**(2), 771 (2008).

[78] D. Samtleben (for the QUIET collaboration), ArXiv e-prints (2008), 0806.4334.

[79] P. Oxley, P. A. Ade, C. Baccigalupi, P. deBernardis, H.-M. Cho, M. J. Devlin, S. Hanany, B. R. Johnson, T. Jones, A. T. Lee, *et al.*, in M. Strojnik, ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (2004), vol. 5543 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 320–331.

[80] B. P. Crill, P. A. R. Ade, E. S. Battistelli, S. Benton, R. Bihary, J. J. Bock, J. R. Bond, J. Brevik, S. Bryan, C. R. Contaldi, *et al.*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (2008), vol. 7010 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference.*

[81] D. Baumann, M. G. Jackson, P. Adshead, A. Amblard, A. Ashoorioon, N. Bartolo, R. Bean, M. Beltran, F. de Bernardis, S. Bird, *et al.*, ArXiv e-prints (2008), 0811.3919.

[82] J. Hough and S. Rowan, Living Reviews in Relativity **3**(3) (2000), http://www.livingreviews.org/lrr-2000-3.

[83] P. R. Saulson, Physical Review D **42**(8), 2437 (1990).

[84] D. Sigg (for the LIGO Scientific Collaboration), Classical and Quantum Gravity **25**(11), 114041 (2008).

[85] X. Siemens, B. Allen, J. Creighton, M. Hewitson, and M. Landry, Classical and Quantum Gravity **21**(20), S1723 (2004).

[86] Y. Gürsel and M. Tinto, Physical Review D **40**(12), 3884 (1989).

[87] O. Krause, S. M. Birkmann, T. Usuda, T. Hattori, M. Goto, G. H. Rieke, and K. A. Misselt, Science **320**(5880), 1195 (2008).

[88] J. L. Blum and J. M. Miller, in *Bulletin of the American Astronomical Society* (2009), vol. 41, p. 384.

[89] R. A. Hulse and J. H. Taylor, Astrophysical Journal **195**, L51 (1975).

[90] CBC Group, *Rate predictions for compact binary coalescence events* (2009), http://www.lsc-group.phys.uwm.edu/ligovirgo/cbc/protected/papers/rates/rates.pdf.

[91] V. Kalogera, C. Kim, D. R. Lorimer, M. Burgay, N. D'Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, *et al.*, The Astrophysical Journal Letters **601**(2), L179 (2004), http://stacks.iop.org/1538-4357/601/L179.

[92] V. Kalogera, C. Kim, D. R. Lorimer, M. Burgay, N. D'Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, *et al.*, The Astrophysical Journal Letters **614**(2), L137 (2004).

[93] C. Kim, V. Kalogera, D. R. Lorimer, , and T. White, The Astrophysical Journal **616**(2), 1109 (2004).

[94] C. Kim, V. Kalogera, and D. R. Lorimer, in *A life with stars* (2006).

[95] R. O'Shaughnessy, C. Kim, V. Kalogera, , and K. Belczynski, The Astrophysical Journal **672**(1), 479 (2008).

[96] R. O'Shaughnessy, C. Kim, T. Fragos, V. Kalogera, , and K. Belczynski, The Astrophysical Journal **633**(2), 1076 (2005).

[97] R. M. O'Leary, F. A. Rasio, J. M. Fregeau, N. Ivanova, , and R. O'Shaughnessy, The Astrophysical Journal **637**(2), 937 (2006).

[98] R. M. O'Leary, R. O'Shaughnessy, and F. A. Rasio, Physical Review D **76**(6), 061504 (2007).

[99] R. M. O'Leary, B. Kocsis, and A. Loeb, ArXiv e-prints (2008), 0807.2638.

[100] M. C. Miller and V. M. Lauburg, The Astrophysical Journal **692**(1), 917 (2009).

[101] M. Freitag, M. A. Gürkan, and F. A. Rasio, Monthly Notices of the Royal Astronomical Society **368**, 141 (2006).

[102] E. Salpeter, The Astrophysical Journal **121**, 161 (1955).

[103] M. J. Benacquista, Living Reviews in Relativity **9**(2) (2006), http://www.livingreviews.org/lrr-2006-2.

[104] S. Sigurdsson and E. S. Phinney, The Astrophysical Journal **415**, 631 (1993).

[105] S. Sigurdsson and L. Hernquist, Nature **364**, 423 (1993).

[106] S. R. Kulkarni, P. Hut, and S. McMillan, Nature **364**, 421 (1993).

[107] D. Guetta and L. Stella, ArXiv e-prints (2008), 0811.0684.

[108] J. Goodman, The Astrophysical Journal **308**, L47 (1986).

[109] B. Paczynski, The Astrophysical Journal **308**, L43 (1986).

[110] D. Eichler, M. Livio, T. Piran, and D. N. Schramm, Nature **340**, 126 (1989).

[111] R. Narayan, B. Paczynski, and T. Piran, The Astrophysical Journal **395**, L83 (1992).

[112] J. S. Bloom, S. Sigurdsson, and O. R. Pols, Monthly Notices of the Royal Astronomical Society **305**, 763 (1999).

[113] K. Belczynski, V. Kalogera, , and T. Bulik, The Astrophysical Journal **572**(1), 407 (2002).

[114] K. Belczynski, R. Perna, T. Bulik, V. Kalogera, N. Ivanova, and D. Lamb, in *36th COSPAR Scientific Assembly* (2006), vol. 36 of *COSPAR, Plenary Meeting*, p. 3470.

[115] D. J. Champion, D. R. Lorimer, M. A. McLaughlin, J. M. Cordes, Z. Arzoumanian, J. M. Weisberg, and J. H. Taylor, Monthly Notices of the Royal Astronomical Society **350**, L61 (2004).

[116] K. Belczynski, R. E. Taam, V. Kalogera, F. A. Rasio, , and T. Bulik, The Astrophysical Journal **662**(1), 504 (2007).

[117] C. Hopman, D. Guetta, E. Waxman, , and S. P. Zwart, The Astrophysical Journal Letters **643**(2), L91 (2006).

[118] D. B. Fox, D. A. Frail, P. A. Price, S. R. Kulkarni, E. Berger, T. Piran, A. M. Soderberg, S. B. Cenko, P. B. Cameron, A. Gal-Yam, *et al.*, Nature **437**, 845 (2005).

[119] A. M. Soderberg, E. Berger, M. Kasliwal, D. A. Frail, P. A. Price, B. P. Schmidt, S. R. Kulkarni, D. B. Fox, S. B. Cenko, A. Gal-Yam, *et al.*, The Astrophysical Journal **650**(1), 261 (2006).

[120] N. Weiner, *Extrapolation, interpolation, and smoothing of stationary time series with engineering applications* (M.I.T. Technology Press and John Wiley and Sons, Inc., New York, 1950).

[121] B. J. Owen, Physical Review D **53**(12), 6749 (1996).

[122] B. J. Owen and B. S. Sathyaprakash, Physical Review D **60**(2), 022002 (1999).

[123] T. Cokelaer, Physical Review D **76**(10), 102004 (2007).

[124] B. S. Sathyaprakash and S. V. Dhurandhar, Physical Review D **44**(12), 3819 (1991).

[125] B. S. Sathyaprakash, Physical Review D **50**(12), R7111 (1994).

[126] *LSC Algorithm Library software packages* LAL*,* LALWRAPPER*, and* LALAPPS, http://www.lsc-group.phys.uwm.edu/lal.

[127] K. S. Thorne, in S. W. Hawking and W. Israel, eds., *Three hundred years of gravitation* (Cambridge University Press, Cambridge, 1987), chap. 9, pp. 330–458.

[128] S. Droz, D. J. Knapp, E. Poisson, and B. J. Owen, Physical Review D **59**(12), 124016 (1999).

[129] A. S. Sengupta, J. A. Gupchup, and C. A. K. Robinson, In Preparation (2008).

[130] R. Biswas, P. R. Brady, J. D. E. Creighton, and S. Fairhurst, ArXiv e-prints (2007), 0710.0465.

[131] P. R. Brady and S. Fairhurst, Classical and Quantum Gravity **25**(10), 1050002 (2008).

[132] B. Allen, J. K. Blackburn, P. R. Brady, J. D. E. Creighton, T. Creighton, S. Droz, A. D. Gillespie, S. A. Hughes, S. Kawamura, T. T. Lyons, *et al.*, Physical Review Letters **83**(8), 1498 (1999).

[133] P. R. Brady, J. D. E. Creighton, and A. G. Wiseman, Classical and Quantum Gravity **21**(20), S1775 (2004).

[134] R. K. Kopparapu, C. Hanna, V. Kalogera, R. O'Shaughnessy, G. Gonzalez, P. R. Brady, and S. Fairhurst, The Astrophysical Journal **675**(2), 1459 (2008).

[135] E. S. Phinney, The Astrophysical Journal **380**, L17 (1991).

[136] R. O'Shaughnessy, K. Belczynski, , and V. Kalogera, The Astrophysical Journal **675**(1), 566 (2008).

[137] LIGO Scientific Collaboration, *Calibration of the LIGO detectors for S5*, Tech. Rep. In Preparation, LIGO Project.

[138] Y. Pan, A. Buonanno, J. G. Baker, J. Centrella, B. J. Kelly, S. T. McWilliams, F. Pretorius, and J. R. van Meter, Physical Review D (Particles, Fields, Gravitation, and Cosmology) **77**(2), 024014 (2008).

[139] R. Gouaty (for the LIGO Scientific Collaboration), Classical and Quantum Gravity **25**(18), 184006 (2008).

[140] S. E. Thorsett and D. Chakrabarty, The Astrophysical Journal **512**(1), 288 (1999).

[141] R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs, The Astronomical Journal **129**(4), 1993 (2005), http://www.atnf.csiro.au/research/pulsar/psrcat/.

[142] LIGO Scientific Collaboration and Virgo Collaboration, In Preparation (2009).

[143] S. Bose, A. Pai, and S. Dhurandhar, International Journal of Modern Physics D **9**(3), 325 (2000).

[144] A. Pai, S. Dhurandhar, and S. Bose, Physical Review D **64**(4), 042004 (2001).

[145] F. Marion (on behalf of the Virgo Collaboration), in J. Dumarchez and J. T. T. Vân, eds., *Gravitational waves and experimental gravity—Proceedings of the XXXVIIIth RECONTRES DE MORIOND* (The Gioi Publishers, 2004), pp. 145–150.

[146] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, In Preparation (2009).

[147] F. Zwicky, The Astrophysical Journal **86**, 217 (1937).

[148] K. C. Freeman, The Astrophysical Journal **160**, 811 (1970).

[149] C. M. Will, Physical Review D **57**(4), 2061 (1998).

[150] L. S. Finn and P. J. Sutton, Physical Review D **65**(4), 044022 (2002).

[151] S. R. Choudhury, G. C. Joshi, S. Mahajan, and B. H. J. McKellar, Astroparticle Physics **21**(5), 559 (2004).

[152] P. Ajith, Classical and Quantum Gravity **25**(11), 114033 (2008).

[153] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, A. M. Sintes, J. T. Whelan, B. Brügmann, P. Diener, N. Dorband, *et al.*, Physical Review D **77**(10), 104017 (2008).

[154] C. Talmadge, J. P. Berthias, R. W. Hellings, and E. M. Standish, Physical Review Letters **61**(10), 1159 (1988).

[155] F. Pretorius, Physical Review Letters **95**(12), 121101 (2005).

[156] M. Campanelli, C. Lousto, Y. Zlochower, , and D. Merritt, The Astrophysical Journal Letters **659**(1), L5 (2007).

[157] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, Physical Review Letters **98**(23), 231102 (2007).

[158] J. Healy, F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner, Physical Review Letters **102**(4), 041101 (2009).

[159] F. Pretorius, in M. Colpi, P. Casella, V. Gorini, U. Moschella, and A. Possenti, eds., *Physics of Relativistic Objects in Compact Binaries: From Birth to Coalescence* (Springer Verlag and Canopus Publishing Limited, 2009), vol. 359, pp. 305–369.

[160] L. D. Landau and E. M. Lifshitz, *Classical Theory of Fields* (Addison Wesley, Redding Mass., 1962), p. Sec. 100.

[161] J. D. Kaplan, D. A. Nichols, and K. S. Thorne, submitted to Physical Review D (2009).

[162] L. E. Kidder, Physical Review D **52**(2), 821 (1995).

[163] K. S. Thorne, Reviews of Modern Physics **52**(2), 299 (1980).

[164] K. S. Thorne and J. B. Hartle, Physical Review D **31**(8), 1815 (1985).

[165] L. Blanchet, G. Faye, and B. Ponsot, Physical Review D **58**(12), 124002 (1998).

[166] H. Tagoshi, A. Ohashi, and B. J. Owen, Physical Review D **63**(4), 044006 (2001).

[167] G. Faye, L. Blanchet, and A. Buonanno, Physical Review D **74**(10), 104033 (2006).