USING MACHINE LEARNING AND THE HOUGH TRANSFORM TO SEARCH FOR GRAVITATIONAL WAVES DUE TO R-MODE EMISSION BY ISOLATED NEUTRON STARS

By

ANDREW L. MILLER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2019

ACKNOWLEDGMENTS

I have way too many people to thank for making this joint degree work.

First I thank my supervisors, Bernard Whiting and Pia Astone, for the tons of work they did to make this joint degree happen. I am very grateful to have met both of them. In detail, I thank Bernard Whiting, who has been a great adviser and friend. He was always been open to new ideas, and I greatly enjoyed these years working with him. And of course, I liked all the Italian, German and Syrian food and drinks we had together. I also thank Pia Astone, who always made time to help me. No matter the problem I had, she had the solution, and she was always so willing to help. It didn't matter how stupid my problem was, she always listened. I will greatly miss talking to her and seeing her all the time.

I thank Cristiano Palomba, who was always there to bounce ideas off of and for following my work so closely, and Paola Leaci, who helped me greatly when I tried to first use computing clusters such as CNAF to submit jobs and for always having my back. Of course I thank Sergio Frasca, for knowing the answer to every question and for the Italian snacks he has given me over the past five years. Also I thank Guido Mueller, for being the best graduate student coordinator, approving this joint degree, having a great sense of humor, and suggesting the solution to support my last semester at University of Florida.

I thank Guenakh Mitselmakher, for approving the funds to support me during my last semester at the University of Florida, and Ettore Majorana, for organizing my exchange with the Institute of Cosmic Ray Research in Tokyo and always finding funds to support my attendance at Virgo meetings and conferences, and Sabrina D'Antonio, for debugging my codes and always being so welcoming of me in Rome. I also thank Anna de Grossi, for helping Pia and me navigate the bureaucracy of an Italian university. I am very thankful for all of the secretaries and administrators who approved this agreement at every level. Finally I thank my parents and my aunt, whose continued support and almost daily Skype calls got me through my Ph.D.

TABLE OF CONTENTS

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

USING MACHINE LEARNING AND THE HOUGH TRANSFORM TO SEARCH FOR GRAVITATIONAL WAVES DUE TO R-MODE EMISSION BY ISOLATED NEUTRON STARS

By

Andrew L. Miller

December 2019

Chair: Bernard F. Whiting
Major: Physics

We present two new methods to search for gravitational waves (GWs) from isolated neutron stars lasting $O(hours - days)$, and the results of three searches of real LIGO/Virgo data. One method is a generalization of a method to detect continuous waves (CWs) that looks for signals that follow a specific model; the other uses machine learning to detect signals that deviate from this model. In two of our searches, we tried to find a remnant of the first-ever binary neutron star (BNS) merger, GW170817; in the third search, we looked everywhere in the sky for any kind of long duration transient signal in LIGO's third observing run. For the first method developed, the Generalized Frequency Hough, we present empirical and theoretical estimates of its sensitivity, an extensive follow-up procedure of GW candidates, the parameter space to which it is sensitive, and computational constraints when using it. For the second method, an application of convolutional neural networks, we demonstrate that its sensitivity is comparable to that of first method but is much faster to apply and can find signals that the first method cannot. Estimates of rates and sensitivities in future detectors, as well as implications of a GW detection from a neutron star, are also discussed. Additionally we explore the prospects of using a two-dimensional Fast Fourier Transform to improve the quality of time/frequency images to be used by either method.

Presentiamo due nuovi metodi per la ricerca di onde gravitazionali da stelle di neutroni isolate, di durata dell' ordine di ore o giorni.Uno dei due metodi è adatto alle ricerca di segnali che seguano una legge di potenza, l'altro, basato su procedure di machine learning, può invece

cercare segnali non modellizzati. Le sensibilità dei due metodi vengono qui confrontate e si mostrerà come siano entrambi adatti alla ricerca degli stessi segnali. Entrambi i metodi sono stati usati in una ricerca reale, qui presentata, fatta al tempo dell' evento GW170817, per cercare un possibile segnale dovuto all' oggetto rimasto dopo la fusione delle due stelle di neutroni. Notiamo che questi metodi possono entrambi essere utilizzati anche per la ricerca di segnali prodotti da esplosione di supernova. In questo lavoro presentiamo anche stime del tasso di fusione di stelle di neutroni e ne discutiamo le prospettive per i rivelatori gravitazionali di prossima generazione. Discuteremo inoltre le implicazioni della scoperta di onde gravitazionali da stelle di neutroni. Infine presenteremo una procedura, basata su trasformate di Fourier a due dimensioni, che abbiamo studiato nell'ottica di migliorare la qualità delle immagini tempo/frequenza, che sono utilizzate come dato di ingresso da entrambi i metodi

CHAPTER 1
INTRODUCTION

In 1915 Einstein proposed a new formulation of gravity, General relativity (GR), to address issues that Newtonian gravity could not explain, e.g. the perihelion of Mercury, and to make the theory of gravity compatible with relativity. In GR, there is no longer a force due to gravity; rather, mass curves the space and time around it, and this curvature tells other matter how to move. Einstein's insights have provided the basis for the field of GW physics, as well as for a great understanding of how the universe works. GR has surpassed all tests in the weak-field and now parts of the strong-field regime after observations of binary black hole (BBH) and BNS mergers by LIGO/Virgo (*1*, *2*). GWs are a direct consequence of Einstein's theory of gravity, and can be thought of as very tiny ripples in space-time that propagate from their source in all directions. This dissertation focused on developing methods to pull out these signals deeply embedded in the noise of our detector.

In Chapter 2, we detail the theory behind GWs; in Chapter 3, we explain the physics of neutron stars and begin to apply the theory of GWs to these objects. The past few years have seen significant progress in GW physics and astronomy, so we discuss some implications of the two most popular discoveries in Chapter 4. We describe our detectors and the various types of noise we encounter in Chapter 5. Chapter 6 deals with possible implications of a detection of an isolated neutron star for the equation of state and the rates of neutron star mergers. Afterwards in Chapter 7, techniques in data analysis and corrections to the data are elucidated. We explain a method to detect continuous gravitational wave signals from isolated neutron stars in Chapter 8 , and show its generalization to power-law signals in Chapter 9. We search for a remnant of binary neutron star merger GW170817 in Chapter 10 using the Generalized Frequency Hough. A technique used to enhance the quality of images is shown in Chapter 11. We then go over in detail different types of machine learning algorithms in Chapter 12 and characterize their sensitivities and robustness in Chapter 13. In Chapter 14 we use the machine learning algorithms in a proof-of-concept search for the same remnant of GW170817, but we

cover a different parameter space than our previous search. Afterwards in Chapter 15, we present a preliminary analysis on the data from LIGO's third observing run, the C00 calibration. Finally we draw some conclusions and make plans for future work in Chapter 16.

CHAPTER 2
THE ORIGIN OF GRAVITATIONAL WAVES

2.1 Introduction

This chapter focuses on developing the framework with which we analyze and think about gravitational waves from relativistic objects in the universe. We discuss in Section 2.2 the development of Einstein's field equations (EFEs). Then, in Section 2.3 we describe how GWs result naturally from linearizing the EFEs, as well as the energetics, polarizations, and detectability of GWs. Finally in Section 2.4, we categorize the four major sources of GWs. This chapter is heavily based on the information presented in (*3*).

2.2 Einstein Field Equations

To formulate a theory of gravity, one needs to have an equation of motion, that is, a model for the dynamics of massive and massless particles. These equations of motion are also called field equations, when treating gravity as a field. The key insight that Einstein made was that the curvature of space-time was related to the mass/density, meaning the equations should be of this form (*3*):

$$K_{\mu\nu} = \kappa T_{\mu\nu} \tag{2–1}$$

where $\kappa = \frac{8\pi G}{c^4}$ (fixed from requiring that EFE reproduce Newton's equations for small velocities), G is the gravitational constant, c is the speed of light, $T_{\mu\nu}$ is the energy-momentum tensor, and $K_{\mu\nu}$ is the tensor that is some nonlinear combination of the space and time coordinates themselves and their derivatives. We describe space and time in the so-called metric coefficients $g_{\mu\nu}$, and the derivatives of the metric are expressed in terms of the Ricci tensor $R_{\mu\nu}$, and the Ricci scalar R. The metric describes the bending of space and time due to a particular mass distribution. We define the Ricci scalar and tensor as derivatives of the metric coefficients:

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\alpha} + \Gamma^\alpha_{\mu\nu}\Gamma^\beta_{\alpha\beta} - \Gamma^\beta_{\mu\alpha}\Gamma^\alpha_{\nu\beta}, \tag{2–2}$$

where:

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\alpha} \left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu} \right) \tag{2–3}$$

and:

$$\partial_\alpha = \frac{\partial}{\partial x^\alpha}. \tag{2–4}$$

The Ricci scalar is:

$$R = g^{\mu\nu} R_{\mu\nu}. \tag{2–5}$$

$T_{\mu\nu}$ has these components:

- T_{00}: energy/volume
- T_{0i}: $c \times$ momentum/volume of source particles in the x^i-direction
- T_{ij}: internal stresses of source, related to pressure

The form of the equation has certain requirements:

1. The slow-velocity limit (Newtonian) implies that the derivatives of $g_{\mu\nu}$ should have no higher than linear terms in second order derivatives of $g_{\mu\nu}$.
2. $T_{\mu\nu}$ is symmetric, so $K_{\mu\nu}$ should be as well.

Under these conditions, the field equations take the form (*3*):

$$K_{\mu\nu} = aR_{\mu\nu} + bRg_{\mu\nu} + \lambda g_{\mu\nu} \tag{2–6}$$

λ can be chosen to be 0 or not depending on the cosmological model. When we allow $\lambda = 0$, we fix that $K_{\mu\nu}$ must be linear in second-order derivatives of $g_{\mu\nu}$.

In Newtonian physics, we have conservation of energy and momentum for a perfect fluid. The same is true in general relativity as long as we can neglect the gravitational field. When

we cannot, as in this case when deriving the EFE, the "conservation" equation represents the equation of motion of the matter in the presence of a gravitational field. This conservation condition is expressed by:

$$\nabla_\mu T^{\mu\nu} = 0 \tag{2–7}$$

Equation 2–7 places tight constraints on the form of EFE. By applying this equation and some mathematical identities that are coordinate-independent, we can determine that $b = 1/2$. When taking the limit as velocities become small (where Newton's laws apply), it can be seen that $a = -1$, so we have:

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

2.3 Gravitational Wave Solutions in the Linearized Regime

To see where GWs appear in GR, we begin with the following two claims: (1) assume we are very far from the source such that the overall metric can be described by the metric of special relativity $\eta_{\mu\nu}$, and (2) a GW will cause a small perturbation to this metric $h_{\mu\nu}$. This can be written as.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{2–9}$$

where $|h_{\mu\nu}| \ll 1$.

Substituting this metric into the EFE involves a lot of algebra, the highlights of which are the following:

1. All terms $O(h_{\mu\nu}^2)$ and higher are neglected.
2. A trace-reversed perturbation $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h$ is actually used to simply the resulting algebra.

3. A gauge condition resulting from the invariance of $h_{\mu\nu}$ under an infinitesimal transformation of the coordinates ($x^\mu \to x^\mu + \xi^\mu$) is applied to simply the final equation, in an analogous way as is done in electromagnetism with the vector potential A_μ.
    - Note: gauge conditions are allowed because the transformed perturbation $\bar{h}_{\mu\nu}$ and $h_{\mu\nu}$ both satisfy the EFE, so nothing about the dynamics of the system has changed. It's similar to adding a constant to an equation for a particle's position in classical mechanics: it doesn't affect the acceleration terms.

At this point, we arrive at an equation for $\bar{h}_{\mu\nu}$ that looks similar to a wave equation:

$$\Box^2 \bar{h}^{\mu\nu} = -\frac{16\pi G}{c^4} T^{\mu\nu}. \tag{2–10}$$

where $\Box^2 = \partial_\mu \partial^\mu$. This equation holds so long as the gauge condition is satisfied:

$$\partial_\mu \bar{h}^{\mu\nu} = 0 \tag{2–11}$$

There is however still some residual gauge freedom, so we also fix:

$$\Box^2 \xi^\mu = 0. \tag{2–12}$$

Solving Equation 2–10 in the absence of $T_{\mu\nu}$ gives wave solutions:

$$\bar{h}^{\mu\nu} = A^{\mu\nu} e^{ik_\rho x^\rho}, \tag{2–13}$$

where $A^{\mu\nu}$ is a constant/complex/symmetric tensor. In order for Equation 2–10 to be satisfied by Equation 2–13, k_ρ must be null, so $k_\rho k^\rho = 0$ implying the waves travel at the speed of light. If we write $k_\rho = (\frac{\omega}{c}, \vec{k})$, we see immediately that $\omega = ck$, the identical dispersion relation for light.

One can also solve Equation 2–10 with $T^{\mu\nu}$ by using a Green's function, and it can be shown in this way that $\bar{h}^{\mu\nu}$ is related to the volume integral over $T^{\mu\nu}$ at the retarded time:

$$\bar{h}^{kl}(ct,\vec{x}) = \frac{4G}{c^4}\int \frac{T^{kl}(ct-|\vec{x}-\vec{y}|,\vec{y})}{|\vec{x}-\vec{y}|}d^3\vec{y}. \tag{2–14}$$

Here, $\vec{y}$ is the position of the source, and $\vec{x}$ is the field point at which one measures $\bar{h}^{\mu\nu}$. We have been able to eliminate the T^{00} and T^{0l} terms because of the conservation of energy/momentum tensor.

Because we are far from the source of GWs, a Taylor expansion of the $|\vec{x}-\vec{y}|$ piece in Equation 2–14 can be done:

$$|\vec{x}-\vec{y}| = r - \vec{y}\cdot\hat{n} + ... \tag{2–15}$$

where r is the magnitude of the distance between us and the star, and $\hat{n}$ is the normal vector pointing from us to the source. We then take only the leading term:

$$\bar{h}^{kl}(ct,\vec{x}) = \frac{4G}{c^4}\frac{1}{r}\int T^{kl}(ct-r+\vec{y}\cdot\hat{n},\vec{y})d^3\vec{y}. \tag{2–16}$$

Moreover, we can also expand T^{kl} at $r=\infty$ where each term in the expansion corresponds to a certain multipole moment. The expansion of T^{kl} corresponds to an expansion in powers of v/c. By neglecting higher order moments, we are implicitly assuming that the speed of the source is small, and that the metric is not at all influenced by the speed of source.

We can see this more clearly if we take the Fourier Transform of T^{kl}:

$$T^{kl} = \int \frac{d^4k}{(2\pi)^4}T^{kl}(\omega,\vec{k})e^{-i\omega(t-r/c+\vec{y}\cdot\hat{n}/c)+i\vec{k}\cdot\vec{y}}. \tag{2–17}$$

Upon expanding the exponential piece, we have increasing powers of v/c:

$$e^{-i\omega(t-r/c+\vec{y}\cdot\hat{n}/c)+i\vec{k}\cdot\vec{y}} = e^{-i\omega(t-r/c)}\left[1 - i\frac{\omega}{c}y^i n^i + \frac{1}{2}\left(-i\frac{\omega}{c}\right)^2 y^i y^j n^i n^j + ...\right] \tag{2–18}$$

This expansion in the Fourier space is equivalent to expanding T^{kl} in position space:

$$T^{kl} = T^{kl} + y^i n^i \partial^0 T^{kl} + \frac{1}{2c^2} y^i y^j n^i n^j (\partial^0)^2 T^{kl} + ... \tag{2–19}$$

In this way, different spatial integrals over the various terms in Equation 2–18 represent different moments. By expanding T^{kl}, we arrive at:

$$\bar{h}^{kl}(ct, \vec{x}) = \frac{4G}{c^4} \left(\frac{1}{r} S^{kl} + \frac{1}{c} n_m \dot{S}^{kl,m} + ... \right), \tag{2–20}$$

where:

$$S^{ij}(t) = \int d^3x T^{ij}(t, \vec{x}) \tag{2–21}$$

$$S^{ij,k}(t) = \int d^3x T^{ij}(t, \vec{x}) x^k. \tag{2–22}$$

It is better to write the T^{ij} components of the energy momentum tensor in terms of the energy density T^{00} and the momenta densities T^{0i}. Defining the energies:

$$M = \frac{1}{c^2} \int d^3x T^{00}(t, \vec{x}) \tag{2–23}$$

$$M^i = \frac{1}{c^2} \int d^3 T^{00}(t, \vec{x}) x^i \tag{2–24}$$

$$M^{ij} = \frac{1}{c^2} \int d^3 T^{00}(t, \vec{x}) x^i x^j \tag{2–25}$$

...

and the momenta:

$$P^i = \frac{1}{c} \int d^3x T^{0i}(t, \vec{x}) \tag{2–26}$$

$$P^{i,j} = \frac{1}{c} \int d^3x T^{0i}(t, \vec{x}) x^j \tag{2–27}$$

...

and using energy/momentum conservation laws, we can find:

$$S^{ij} = \frac{1}{2}\ddot{M}^{ij} \tag{2–28}$$

$$\dot{S}^{ij,k} = \frac{1}{6}\dddot{M}^{ijk} + \frac{1}{3}\left(\ddot{P}^{i,jk} + \ddot{P}^{j,ik} - 2\ddot{P}^{k,ij}\right). \tag{2–29}$$

Equation 2–28 is the time-derivative of the quadrupole moment; the first and second terms in Equation 2–29 are the octupole moment and the current quadrupole moment, respectively.

The leading order term to the emission of gravitational radiation is the the second time derivative of the quadrupole term. Neglecting all higher order moments, the equation for the amplitude tensor of GW emission reads:

$$\bar{h}^{ij}(ct, \vec{x}) = -\frac{2G}{c^6 r}\ddot{Q}^{ij}(ct - r), \tag{2–30}$$

where $Q^{ij} = M^{ij} - \delta^{ij}M_{kk}$ is the quadrupole moment tensor. We will analyze these moments in a later chapter when we discuss GWs from isolated neutron stars and from r-modes.

Note that this approximation breaks down when the source's motion is relativistic because curvature and velocity are no longer independent; however, this is not so much a problem for GWs that come from an isolated neutron star, since we don't use post-Newtonian theory, an expansion of the energy due to gravitational waves in powers of velocity of the source, to construct the waveform.

2.3.1 Gravitational Wave Polarizations

According to Equation 2–13, the $A^{\mu\nu}$ term has 16 terms, however with an appropriate choice of gauge, there will only be two independent terms. $A^{\mu\nu}$ is symmetric ($A^{\mu\nu} = A^{\nu\mu}$) because $T^{\mu\nu}$ is, which means 16 independent terms become 10 independent ones. But, $A^{\mu\nu}$ also satisfies the gauge condition: $A^{\mu\nu}k_\mu = 0$, so this reduces the number of degrees of freedom to 6.

From here, it is best to work with a plane wave propagating in the z direction, so $k_\rho = (ck, 0, 0, k)$. Solving Equation 2–11 for ξ^ν with plane wave solutions, and doing some

algebra, all terms vanish except for 2 independent terms in $A^{\mu\nu}$: A^{11}, A^{21}, A^{12}, and A^{22}. To be traceless, $A^{11} = -A^{22} = A_+$ and symmetric, $A^{12} = A^{21} = A_\times$. We are now in the transverse-traceless (TT) gauge, a consequence of removing gauge degrees of freedom from $A^{\mu\nu}$. Of course, the transformation to the TT gauge works for a wave propagating in an arbitrary direction as well.

2.3.2 Effect of Gravitational Waves on Point Particles

To understand how a GW with two polarizations will affect the motion of a particle, we must consider a ring of particles through which it passes. We begin with the geodesic equation:

$$\frac{d^2x^\mu}{d\tau} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0, \tag{2–31}$$

where x^μ is the position four-vector and τ is the proper time (or in general any parameter by which one can parameterize the path of a particle).

Equation 2–31 is derived in a metric-independent way by using the variational principle, which states that freely falling test particles will follow the path in a curved space-time such that the action is extremized. In classical mechanics, this corresponds to the fact that a particle will always follow the shortest distance between two points.

If we only considered one point particle, we would find that the particle stays at rest regardless of the passing GW, because in the TT gauge, the coordinates are attached to individual particles. Therefore, we need to understand the difference in distance between two or more particles as the GW passes.

Let us imagine that we have two particles on the x-axis, at coordinate distance $L/2$ from the origin. If we calculate their distance as a GW perturbation comes through from the z-direction, $k_\rho = (\omega, 0, 0, \omega)$ with a single polarization:

$$
\begin{aligned}
ds^2 &= g_{\mu\nu} dx^\mu dx^\nu && \text{(2–32)} \\
L_0^2 &= (\eta_{11} + h_{11}) L^2 && \text{(2–33)} \\
&= (1 - h_+ \cos\omega t) L^2 && \text{(2–34)} \\
L_0 &\approx L \left(1 - \frac{1}{2} h_+ \cos\omega t\right). && \text{(2–35)}
\end{aligned}
$$

Here we have simply used Equation 2–13 to eliminate the A_+ term in favor of h_+ and taken the real part of 2–13.

If we consider two particles on the y-axis separated by L, we similarly get:

$$
L_0 \approx L \left(1 + \frac{1}{2} h_+ \cos\omega t\right). \qquad \text{(2–36)}
$$

The difference in sign of the h_+ term results in an out-of-phase oscillation on the y-axis with respect to the x-axis.

$h_\times$ is related to h_+ by a rotation of $\pi/4$ radians, so the motion of the test particles will also differ by this amount.

Based on Equation 2–35, we can see that the proper distance between two particles changes by an extremely small small distance $\frac{1}{2} h_+$ and oscillates harmonically as the GW passes through. We show the effect of a GW on a ring of point particles in Figure 2-1.

2.4 Sources of Gravitational Waves

We will describe here the four major sources of gravitational radiation to which ground-based interferometers are sensitive.

2.4.1 Compact Binaries

The only sources of GWs that have been detected so far are compact binaries, composed of either two black holes, two neutron stars or one black hole and one neutron star. In a binary system, the two objects orbit each other at a frequency Ω, and due to the emission of GWs, they lose orbital energy and angular momentum, causing them to get closer to each other

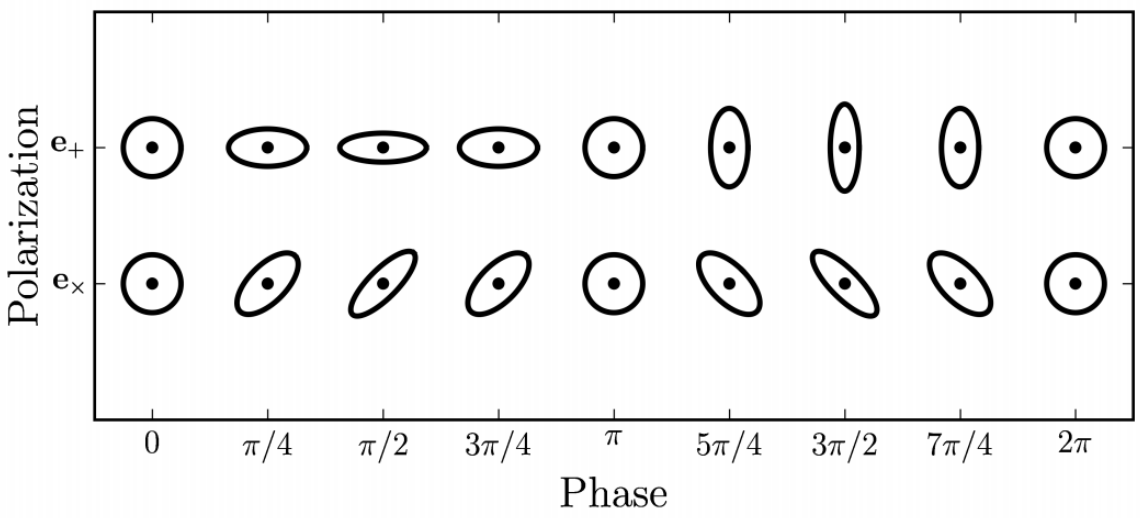


Figure 2-1: The effect of a plus/cross polarized GW on a ring of point particles. The wave is traveling perpendicular to the page.

and orbit more quickly. Eventually, the two objects will collide in a huge explosion of GWs, forming a remnant whose properties depend on the masses, spins, tidal deformabilities, etc. of the two original objects. The GW emitted from a compact binary is modeled in three distinct phases: (1) the inspiral, (2) the merger and (3) the ringdown. In the inspiral phase, the phase evolution of the binary is modeled using post-Newtonian expansion, which expresses the phase as a function of powers of v/c.

Next, the merger phase is modeled by numerically solving the equations of general relativity for a binary, since the two-body problem is not analytically solvable.

For black holes only, the ringdown is modeled using perturbation theory, and has an analytic solution: quasi-normal modes. The black hole essentially "rings" like a bell as it emits the remainder of its GW energy as it settles down into a stable state and is geometrically symmetric, so it no longer emits GWs.

Binaries are the most understood source of GWs because their waveforms can be exactly described in these three phases, though their rates were uncertain. When we know the

waveform, we can use the most sensitive detection filtering technique to pull out this signal from the data called matched filtering, discussed later.

2.4.2 Core Collapse Supernovae Explosions

It is possible that non-axisymmetric explosions of stars within our galaxy will emit detectable GWs. The rate of such events around 1-10 per century (*4*), and the last known event has occurred about 100 years ago (*5*), so there is anticipation that another will occur soon.

Signals from supernovae are much more complex than from binaries, since neutrino interactions, hydrodynamics, general relativity, shocks, and fluid instabilities all play a role in determining the time/frequency behavior of a GW. In fact, the GWs do not come just from an non-spherical structure, but from an asymmetric emission of neutrinos, from a non-spherical structure, or forces on the fluid caused by the pressure of the neutrinos. Many models of these interactions exist (*6*, *7*), though the searches for supernovae typically do not assume a signal morphology (*8*, *9*). To claim a detection, these searches assign a loudness criteria to extremely short duration events, and use all available detectors to look for coincident events.

As an example of a mechanism to produce anisotropic neutrino emission, convective motion occurs due to the fact that negative lepton and entropy gradients are being formed in the outer parts of the proto-neutron star (*6*). This situation may give rise to dynamical turnovers of material heated by neutrinos, which would cause large deviations from a spherically symmetric explosion (*10*). Specifically near the neutrino-sphere, neutrinos can be convectively transported, leading to anisotropic emission, caused by fluctuations in temperature and density in this region.

2.4.3 Stochastic Background

An incoherent superposition of all GW events in the universe will give rise to a background of events, characterized mostly by BBH mergers in the LIGO frequency bands. We also expect a stochastic GW background from the waves that were generated in the early universe as well, analogous to the Cosmic Microwave Background we observe electromagnetically. Possible

sources, detectable with future generation instruments, include waves from inflation, phase transitions, and cosmic strings. Each of these sources provides a unique contribution to the stochastic GW background (*11*).

2.4.4 Neutron Stars

Low mass x-ray binaries (LMXBs) are binary systems where one object is a neutron star or black hole that accretes matter from its companion, e.g. Scorpius X-1 (*12*). LMXBs can emit continuous (monochromatic) gravitational waves because this accretion should have spun up the rate of rotation of the neutron star to be much higher (> 1000 Hz) (*13*) than what we currently observe, which is $O(100)$ Hz (*14*). Something is limiting the spin-up of these systems: it is therefore possible that the system radiates away angular momentum in the form of gravitational waves that balances the accretion torque (*15*, *16*). The amplitude of the GW that balances these torques increases with the amount of observed X-ray flux independent of the distance the source is from us, so these kinds of binaries are very promising sources of GWs (*17*).

Isolated neutron stars are expected to emit GWs due to a small deformation on the surface of the star, or due to some fundamental oscillation mode (r-modes, for example, discussed later). A detection of a GW from a neutron star will give us direct insight into the bulk motion of the star, something that is impossible with electromagnetic radiation from pulsars. Moreover, this deformation or oscillation mode is related to equation of state (EoS) parameters, which would allow us to at least rule out some EoS. The rest of this thesis will be devoted to detecting GWs from isolated neutron stars.

CHAPTER 3
THE ASTROPHYSICS OF NEUTRON STARS

3.1 Introduction

Neutron stars are formed in stellar collapses when the progenitor star's mass is between 8-25 M_{sun} (*18*). Despite our ability to observe neutron stars, we still do not know a lot about the physics that governs their behavior and their emissions of GWs. We describe some defining physical aspects of neutron stars, including their magnetic fields and fast rotation rates, in Section 3.2. Then we analyze the types of GWs we expect from deformed neutron stars in Section 3.3. In Section 3.4, we describe a likely source of GWs from isolated, old neutron stars: continuous waves (CWs). Next, we describe a type of GW that has recently received a lot of attention in the LIGO/Virgo communities in Section 3.5: transient continuous waves (tCWs). We analyze additional possible GWs due to perturbations throughout the neutron star, so-called r-modes, in Section 3.6. We conclude this chapter by discussing the implications of a GW detection for the neutron star EoS in Section 3.8.

3.2 Neutron Stars

There have been only about 2000 neutron stars detected using electromagnetic radiation in our galaxy, but $O(10^8)$ are expected to exist (*19*). The neutron stars we have detected are commonly called pulsars, because of how regularly they shine light at us. They can be used as very precise clocks, and are in fact used as "arms" in GW experiments (*20*). The radius of a neutron star is $O(10$ km), but its mass is on the order of a solar mass, implying a very large density. The fastest spinning neutron star we have seen rotates at 716 Hz (*21*), which, together with high densities, make these compact objects a relativistic laboratory. They also have magnetic fields on $O(10^{10})$ T.

The neutron star is unsurprisingly composed mostly of neutrons: during stellar collapse the force of gravity is so strong that the outward pressure of electrons and protons repelling off each other is not enough to sustain the star (as it would be in the case of white dwarfs). The neutron star is supported by the pressure between all the neutrons that have been created due

to gravity forcing electrons and protons into such close proximity that they combine and form neutrons (*22*). The pressure that the neutrons exert on each other is enough to stop further collapse of the star to a black hole.

In the following subsections, we describe how the magnetic field strength and the maximum rotational frequency of a neutron star can be estimated.

3.2.1 Magnetic Field Strength

The strength of the magnetic field in a neutron star can be understood intuitively in the following way: astronomers can measure stellar magnetic fields by looking at the spectra that come from a star. Typically, the spectra show dips in energy at certain frequencies that correspond to certain elements being present in the star's atmosphere. If the star has a magnetic field, an absorption line is split into evenly spaced lines, which is due to the Zeeman effect: a magnetic field splits the energy levels in an atom. By measuring the spacing in the spectra, one can determine the magnetic field of a star. These fields are typically weak, $O(0.1)$ T maximum.

We can use this information to determine the magnetic field of a neutron star formed after stellar collapse. If we imagine a star that has the radius of our sun $R_s = 7 \times 10^5$ km, a mass of $10\ M_{sun}$, and a magnetic field of 0.1 T, we can calculate the magnetic flux $\phi = \int \vec{B} \cdot d\vec{A}$ through a ring that has the radius of the star before and after the collapse, and equate the two (since energy is conserved):

$$B_{ns} = B_s \left(\frac{R_s}{R_{ns}} \right)^2 \approx 10^9 \text{ T}. \tag{3–1}$$

3.2.2 Break-up Frequency

It is straightforward to compute the rotation frequency at which a neutron star is no longer stable, the so-called break-up frequency f_{bu}. If we assume that a point particle is rotating on the surface of the neutron star, then the only two forces acting on the particle (in the rotating frame) are gravity and the centrifugal force. Equating the two gives:

$$f_{bu} = \frac{1}{2\pi}\sqrt{\frac{GM}{R^3}} \approx 1833 \text{ Hz} \left(\frac{M}{M_{sun}}\right)^{1/2} \left(\frac{R}{10 \text{ km}}\right)^{-3/2}, \tag{3–2}$$

where M is the mass of the star and R is its radius. Moreover, we know that a neutron star has a maximum mass of around $3M_{sun}$ (otherwise, the neutron degeneracy pressure cannot sustain the inward force due to gravity and collapse to a black hole would occur), so we can place an upper bound on the maximum rotation frequency of the neutron star, ~ 3000 Hz.

This is a very simple calculation: clearly the material of which a neutron star is composed will also affect f_{bu}, meaning that the EoS will play a role. However, the calculation is meant to give an idea of what kind of GWs to expect from a rotating neutron star: recalling that the GW frequency is twice the rotational frequency, we expect GWs with frequencies no higher than ~ 6000 Hz.

This calculation raises another question: if the break-up frequency is ~ 3000 Hz, why have we not seen any pulsars that rotate at this frequency? The answer is not known, but some combination of the emission of GWs and electromagnetic radiation may be responsible for this restriction on the maximum spin (*23*).

Additionally, if the neutron star is accreting matter from a companion star, the rate at which it is accumulating mass will impact the maximum frequency allowed, since there will be an additional torque applied to the neutron star, spinning it up (*24*, *25*).

3.2.3 Indirectly Measuring an Upper Limit Magnetic Field of a Neutron Star

We can calculate the power radiated due to an arbitrary object in terms of the second time derivative of the magnetic dipole moment $\ddot{m}$, which is (*26*):

$$P_{mag} = \frac{dE}{dt} = \frac{\mu_0 |\ddot{m}|^2}{6\pi c^3}, \tag{3–3}$$

where μ_0 is the permeability constant. Now, we assume that a neutron star acts like a magnetic dipole, and that this dipole is inclined at an angle α with respect to the central axis of the star. The dipole rotates with angular velocity Ω, meaning $m = m_0 e^{i\Omega t}$. Taking two

time derivatives, and recalling that the magnetic moment of a uniformly magnetized sphere is $m = BR^3/\mu_0$:

$$P_{mag} = \frac{\Omega^4 B^2 R^6 \sin^2 \alpha}{6\pi c^3}. \tag{3–4}$$

We also know the time derivative of the rotational energy of the star is given by:

$$P_{rot} = \frac{d}{dt}\left(\frac{1}{2} I \Omega^2\right) = I\Omega\dot{\Omega}. \tag{3–5}$$

We can now attempt to bound the magnetic field of the neutron star by assuming that all the rotational energy is converted into magnetic dipole radiation, so we equate P_{mag} and P_{rot}, and write $\Omega = 1/P$ and $\dot{\Omega} = -\frac{\dot{P}}{P^2}$ we see that:

$$B = \sqrt{\frac{6\pi c^3 I}{R^6 \sin^2 \alpha} \frac{\dot{\Omega}}{\Omega^3}} = \sqrt{\frac{6\pi I c^3}{R^6 \sin^2 \alpha} P\dot{P}}. \tag{3–6}$$

The magnetic field is proportional to the spin period P and its derivative $\dot{P}$. So all one has to do is measure how quickly the pulsar is rotating and spinning down, which is done through measurements of how often we receive light pulses over time.

3.3 Gravitational Waves from Deformed, Rotating Neutron Stars

Typically a neutron star is modeled as an ellipsoid, a rotating rigid body with three semi-axes a, b, c of different lengths. To determine the strength of GWs, we need to calculate $\bar{h}^{ij}$ in Equation 2–30, and for this, we must determine the moment of inertia tensor Q^{ij}. Since we do not expect the neutron stars to be moving quickly (since their break-up frequency is mildly relativistic), we can write

$$T^{00} \approx \rho c^2, \tag{3–7}$$

where ρ is the density of the neutron star, which we assume to be uniform over the star volume in these calculations (this is not true in practice). Now we need to compute the principle moments of inertia of an ellipsoid. From classical mechanics, we know the moments of inertia of a rigid body are:

$$Q^{ij} = \int_V d^3x\rho(\vec{x})(r^2\delta^{ij} - x^i x^j), \tag{3–8}$$

where we have removed the trace for use in the quadrupole formula 2–30, and that a coordinate system can be chosen such that the moment of inertia tensor is diagonal. The components are:

$$I_1 = \int_V d^3x\rho(y^2 + z^2) = \frac{M}{5}(b^2 + c^2) \tag{3–9}$$

$$I_2 = \int_V d^3x\rho(x^2 + z^2) = \frac{M}{5}(a^2 + c^2) \tag{3–10}$$

$$I_3 = \int_V d^3x\rho(x^2 + y^2) = \frac{M}{5}(a^2 + b^2) \tag{3–11}$$

Let us now allow the rigid body to rotate, which we will describe the rotation as about the z-axis by a rate ω with the rotation matrix R:

$$R = \begin{bmatrix} \cos\omega t & \sin\omega t & 0 \\ -\sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

From now on primed coordinates refer to the rotating frame x',y',z', while unprimed refer to the inertial frame. So Q'^{ij} is time-independent, while Q^{ij} is time-dependent. They are related through tensor products:

$$Q = R^T Q' R \tag{3–12}$$

so

$$Q_{11} = I_1 \cos^2 \omega t + I_2 \sin^2 \omega t = I_1 + I_2 + \frac{I_1 - I_2}{2} \cos 2\omega t \quad (3\text{–}13)$$

$$Q_{12} = \frac{I_1 - I_2}{2} \sin 2\omega_t \quad (3\text{–}14)$$

$$Q_{22} = I_1 \sin^2 \omega t + I_2 \cos^2 \omega t = I_1 + I_2 - \frac{I_1 - I_2}{2} \cos 2\omega t \quad (3\text{–}15)$$

$$Q_{33} = I_3 \quad (3\text{–}16)$$

$$Q_{13} = Q_{23} = 0. \quad (3\text{–}17)$$

From the moments of inertia in the inertial frame, we see that GWs are radiated at twice the rotational frequency of the neutron star, $\omega_{gw} = 2\omega$. Now, we simply need to take the time-derivative of the appropriate components of Q^{ij}.

Using Equation 2–30, seeing that the moments of inertia have the same pre-factor $\frac{I_1 - I_2}{2}$, and exchanging angular GW frequency for frequency ($\omega = 2\pi f$), we can deduce that the pre-factor is:

$$h_0 = \frac{4\pi^2 G}{c^4} \frac{(I_1 - I_2)}{d} f(t)^2 = \frac{4\pi^2 G}{c^4} \frac{I_3 \epsilon}{d} f(t)^2, \quad (3\text{–}18)$$

where $\epsilon = \frac{I_1 - I_2}{I_3}$ is the level of asymmetry on the star and d is the star's distance from us.

The amplitude h_0 is modulated depending on the inclination angle of the observer with respect to the source of the radiation. This modulation affects h_+ and $h_\times$ differently.

We can now derive the power due to GW emission, which will be useful in Section 3.5 when we parameterize our analysis methods and search in terms of braking indices:

$$\frac{dE}{dt} = -L_{gw} = -\frac{G}{5c^5} < \dddot{Q}_{ij} \dddot{Q}^{ij} > \quad (3\text{–}19)$$

$$Q_{ij} = \frac{I_1 - I_2}{2} \begin{bmatrix} \cos 2\omega t & \sin 2\omega t & 0 \\ \sin 2\omega t & -\cos 2\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3\text{–}20)$$

Taking three time derivatives and substituting in I_3 and ϵ:

$$\dddot{Q}_{ij} = 4\omega^3 I_3 \epsilon \begin{bmatrix} \sin 2\omega t & -\cos 2\omega t & 0 \\ -\cos 2\omega t & -\sin 2\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3–21}$$

Plugging this into Equation 3–19 and time-averaging over the resulting sum of $\sin^2 \omega t$ and $\cos^2 \omega t$ terms:

$$\frac{dE}{dt} = \dot{E} = -\frac{32G}{5c^5}\omega^6 I_3^2 \epsilon^2. \tag{3–22}$$

Equating 3–22 with the rotational energy $E_{rot} = \frac{1}{2} I_3 \omega^2$ of a neutron star, we can obtain the spindown as a function of the frequency:

$$\dot{\omega} = -\frac{32G}{5c^5} I_3 \epsilon^2 \omega^5. \tag{3–23}$$

This is the power law behavior that we have developed a method to search for.

3.4 Continuous Gravitational Waves

CWs are quasi-infinite, quasi-monochromatic signals that we expect from small deformations on the surface of a quickly rotating neutron star. One can imagine a perfect sphere with a small "mountain" or "crater": it is this deformation that causes a time-varying quadrupole moment and hence the emission of CWs. Because the deformation is small, the signal is typically even smaller than those we expect from coalescing binaries. However, the signal is quasi-infinite, so the greatest limitation on sensitivity is how much data we have. This is because by integrating for longer times, the signal-to-noise ratio increases with the square root of the observation time. For an unchanging frequency f_0, we can see this clearly by looking at Equation 7–13 for a continuous wave signal, but the same result applies even when the source has a spindown or is Doppler modulated due to the earth/source relative motion. The amplitude $h(t)$ of a CW signal is a sinusoid $h(t) \sim \cos(2\pi f_0 t)$. The Fourier Transform of a sinusoid is a delta function, meaning the ideal matched filter is also a delta function.

In CW searches, one year of data are analyzed in small chunks of time, typically 1000-8000 s, because of the intrinsic spindown of the source and the modulation of the signal due to the Doppler shift. To elaborate further, a Fast Fourier Transform (FFT) length of the entire run is ideal for a sinusoidal signal, since all the power can be confined to the smallest frequency bin possible. However, even after correcting for the Doppler modulation of the earth, the source can still have an intrinsic spindown (as it must, because the decrease in rotational energy of the neutron star is what powers the GW emission). Therefore, unless one knows the intrinsic spindown a priori, one has to choose FFT lengths such that the signal will not spin out of one frequency bin during T_{FFT}, the time of the FFT. Therefore, CW searches use FFT lengths of $O(hours - days)$ depending on the computation power available for the search. The all-sky Frequency Hough (FH) pipeline can only afford $T_{FFT} = 8192$ s up to 128 Hz, because the number of points in the sky increases with the square of the frequency and because the variation due to the Doppler motion at these frequencies must remain within one frequency bin; however other searches such as Einstein-at-Home (*27*) can take day-long FFTs because they source the huge computation load to individuals' computers (*28*).

The model for a CW signal is:

$$f = f_0 + \dot{f}(t - t_0) + \frac{\ddot{f}}{2}(t - t_0)^2 + \dots \tag{3–24}$$

where f_0 is the frequency at time t_0, $\dot{f}$ is the spindown and $\ddot{f}$ is the second-order spindown.

Since the frequency does not change very much, the amplitude of CW signals is effectively constant, as shown in Equation 3–18. However, earth's rotation induces a Doppler modulation, and there are source-independent relativistic effects that must be corrected for– more information is in Section 7.3. At the detector, the signal is a time-varying function:

$$h(t) = F_+ h_+(t) + F_\times h_\times(t), \tag{3–25}$$

where $F_{+,\times}$ are the beam pattern functions given by (*29*):

$$F_+ = \sin\zeta[a(t)\cos 2\psi + b(t)\sin 2\psi] \tag{3–26}$$

$$F_\times = \sin\zeta[b(t)\cos 2\psi - a(t)\sin 2\psi], \tag{3–27}$$

where ψ is the polarization angle of the GW, ζ is the angle between the interferometers' arms, and $a(t)$ and $b(t)$ are written explicitly in (*29*) but depend on the detector site location, source position, the frequency of earth's rotation, and a reference phase at $t = 0$. Additionally,

$$h_+(t) = h_{0,+}\cos(2\pi f_0 t) \tag{3–28}$$

$$h_\times(t) = h_{0,\times}\sin(2\pi f_0 t), \tag{3–29}$$

and:

$$h_{0,+} = h_0\frac{1+\cos^2\iota}{2} \tag{3–30}$$

$$h_{0,\times} = h_0\cos\iota. \tag{3–31}$$

where ι is the inclination angle between us and the neutron star, and h_0 is given by Equation 3–18.

3.5 Transient Continuous Gravitational Waves

Transient continuous gravitational waves (tCWs), or long duration transients, are signals that come from isolated neutron stars and last O(hours-days). They are "long" with respect to burst and CBC signals, but they are "short" with respect to CWs. This portion of the parameter space only began recently due to a lack of promising sources (*30*–*33*). For these signals we assume a power-law model for the time/frequency evolution of the GW signal:

$$\dot{f} = -kf^n \tag{3–32}$$

where n is the braking index, and k is a proportionality constant that contains the physics of the emitted radiation, defined to be positive. This radiation is modeled as coming from the rotational energy of the neutron star; hence as the energy is emitted, the neutron star spins down. Integrating Equation 3–32, we get:

$$f(t) = \frac{f_0}{\left(1 + k(n-1)f_0^{n-1}(t - t_0)\right)^{\frac{1}{n-1}}} \tag{3–33}$$

For all braking indices except $n = 7$ (r-modes), we model the amplitude change as a function of time and frequency as in Equation 3–18 (*29*). This amplitude evolution as a function of time explicitly is (*34*):

$$h(t) = \frac{4\pi^2 G I_{zz}}{c^4}\frac{\epsilon}{d}f_0^2\left(1 + \frac{t}{\tau}\right)^{\frac{2}{1-n}} \tag{3–34}$$

where $\tau = [(n-1)kf_0^{n-1}]^{-1}$ is the spindown time scale, ϵ is the ellipticity (degree of deformation) of the neutron star, and I_{zz} is the z-component of the star's moment of inertia.

For r-modes ($n = 7$), we use the following relation (*35*, *36*):

$$h(t) = \sqrt{\frac{2048}{375}}\frac{GJMR^3}{c^5 d}\alpha f(t)^3 \tag{3–35}$$

where α is the saturation amplitude, the amplitude at which non-linear effects stop the growth of the r-mode and $J = 1.635 \times 10^{-2}$ is a dimensionless angular momentum for polytropic EoS models. R-modes will be discussed in Section 3.6.

3.6 Gravitational Waves from R-modes

A promising source of GWs from neutron stars comes from toroidal perturbations throughout the star, so-called r-modes. These modes are analogous to Rossby waves on earth, and have a Coriolis restoring force. In Subsection 3.6.1 we delve into the physics of r-modes: how r-modes come about from a rotating object, and the kinds of GWs they emit.

3.6.1 Velocity and Density Perturbations

r-mode oscillations are caused by small changes in the velocity and density of fluid matter in the star. To compute the form of these perturbations, we write the multipoles of Equation 2–20 as (*37*, *38*):

$$h_{\mu\nu} = \frac{1}{r}\sum_{l=2}^{\infty}\sum_{m=-l}^{m=l}\partial_0^l[D_{lm}T_{\mu\nu}^{lm} + J_{lm}(T_{\mu\nu})^{lm}] \tag{3–36}$$

where D_{lm} is the mass quadrupole moment and J_{lm} is the current quadrupole moment.

We can then use this expression for $h_{\mu\nu}$, coupled with an expression for the power of the GW, to write:

$$\partial_0 E_{gw} = \frac{1}{32\pi}\sum_{l=2}^{\infty}\sum_{m=-l}^{m=l} < |\partial_0^{l+1}D_{lm}|^2 + |\partial_0^{l+1}J_{lm}|^2 >, \tag{3–37}$$

and knowing that we are dealing with waves of the form $e^{i\omega t}$, it follows that:

$$\partial_0 E_{gw} = -\omega_r \sum_{l\geq 2} N_l \omega_i^{2l+1}(|\delta D_{lm}|^2 + |\delta J_{lm}|^2), \tag{3–38}$$

where $\omega_r = \omega_i + m\Omega$. Here ω_r is the rotational frequency of the star in the rotating frame, and ω_i is the rotational frequency in the inertial frame. They are simply off by a factor of $m\Omega$, but it is this factor that makes all the difference, as will be discussed in a following section 3.6.4. Additionally, we can express the mass and current quadruples in terms of these velocity and density perturbations δv and $\delta\rho$ respectively (*37*):

$$\delta D_{lm} = \int \delta\rho r^l Y_{lm}^* dV \tag{3–39}$$

$$\delta J_{lm} = \frac{2}{c}\sqrt{\frac{l}{l+1}}\int r^l(\rho\delta v + \delta\rho\Omega)Y_{lm}^{*B}dV, \tag{3–40}$$

where the Y_{lm} are the standard spherical harmonics and Y_{lm}^B are the magnetic multipoles, defined in (*39*).

Writing explicitly the r-mode eigenfunction ξ:

$$\frac{\xi}{r} = \sum_{l,m} \left(0, \frac{D_{lm}}{\sin\theta}\partial_\phi, -D_{lm}\partial_\theta\right) Y^B_{lm} e^{i\omega_r t}, \tag{3–41}$$

the velocity ($\delta_v = \partial_0 \xi$) and density perturbations equations can be derived (*38*, *40*):

$$\delta v = \alpha R\Omega \left(\frac{r}{R}\right)^l Y^B_{ll} e^{i\omega_r t} \tag{3–42}$$

$$\delta\rho = \alpha R^2\Omega^2 \rho \frac{\partial\rho}{\partial P}\left[\frac{2l}{2l+1}\sqrt{\frac{l}{l+1}}\left(\frac{r}{R}\right)^{l+1} + \delta\psi\right] Y_{l+1,l} e^{i\omega t} \tag{3–43}$$

3.6.2 Frequency Evolution of R-modes

To begin, we treat the neutron star as a rotating fluid using Euler's equation:

$$\rho\left[\frac{\partial\vec{v}}{\partial t} + \vec{v}\cdot(\nabla\vec{v})\right] + 2\rho(\vec{\Omega}\times\vec{v}) + \rho\vec{\Omega}\times(\vec{\Omega}\times\vec{r}) = -\nabla P - \rho\nabla\Phi, \tag{3–44}$$

where v is the velocity of the fluid, P is the pressure of the fluid and Φ is the gravitational potential.

We now consider small toroidal displacements (no perturbation in the radial direction) of the fluid star, as well as perturbations in velocity, density, pressure, and gravitational potential of the star. And we decompose the small toroidal displacement vector $\vec{\xi}$ into harmonics of the azimuthal angle, which can be done so long as the object is azimuthally symmetric.

This vector can be written as $\vec{\xi_0} e^{i(m\phi - \omega_r t)}$, where we have decomposed this perturbation into harmonics m with respect to the azimuthal angle, and assumed periodic time dependence.

We define the displacement vector ξ in terms of the velocity perturbation:

$$\delta\vec{v} = \partial_0\vec{\xi}, \tag{3–45}$$

where δv is the perturbation in fluid velocity. In this section δX refers to the perturbation of quantity X. In this case we want to see how the inertial toroidal frequencies in the inertial and rotating frames are related, which can be accomplished through the derivative transformation:

$$\frac{d\vec{\xi_i}}{dt} = \partial_t \xi_r + \vec{v} \cdot \nabla \xi_r \tag{3–46}$$

$$-\omega_i = -\omega_r + R\Omega \frac{1}{R} \partial_\phi \vec{\xi_r} \tag{3–47}$$

$$\omega_i = \omega_r - m\Omega \tag{3–48}$$

Applying these perturbations and linearizing Equation 3–44, we obtain the following (*37*):

$$\partial_0 \delta\vec{v} + 2\Omega \times \delta\vec{v} = -\omega_r^2 \vec{\xi} + 2i\omega_r \Omega \times \vec{\xi} = \frac{\delta\rho}{\rho^2}\nabla p - \frac{1}{\rho}\nabla\delta p - \nabla\delta\Phi. \tag{3–49}$$

Taking the curl of Equation 3–49, analyzing the radial part, neglecting terms of $O(\Omega^2)$ and knowing that the radial velocity is much smaller than the horizontal velocity, we get:

$$\partial_t(\vec{\nabla} \times \vec{v})_r + 2(\vec{v}_h \cdot \nabla_h)\vec{\Omega}_r + 2\Omega_r \cdot (\nabla_h \vec{v}_h) = 0. \tag{3–50}$$

Other necessary equation include the Poisson equation for the perturbed gravitational potential, an EoS, and the conservation of mass equation.

Considering only linear order terms in Ω, so dropping $\vec{v}_h \cdot \nabla_h \propto \Omega^3$, and assuming no spindown of the star ($\partial_t \Omega = 0$), we get in the rotating frame:

$$\frac{d}{dt}\left[\vec{\nabla} \times \vec{v} + 2\Omega\right] = 0. \tag{3–51}$$

Writing $v = i\omega_r \xi$, and using toroidal eigenvectors written in terms of the quadrupole moment:

$$\frac{\xi}{r} = \left(0, D_{lm}\frac{1}{\sin\theta}\partial_\phi Y_l^m, -D_{lm}\partial_\theta Y_l^m\right) e^{i\omega t}, \tag{3–52}$$

we get:

$$\nabla \times \xi = T_{lm}\frac{l(l+1)}{r^2} Y_l^m, \tag{3–53}$$

where the T_{lm} are function of r. We can then obtain (*37*):

$$\omega_{r-mode} = \frac{2m\Omega}{l(l+1)}. \tag{3–54}$$

Converting this frequency to the inertial frame, we get:

$$\omega_i = \omega_r - m\Omega = \frac{2m\Omega}{l(l+1)} - m\Omega \tag{3–55}$$

Now, requiring $l = m = 2$, we obtain that $\omega_i = -\frac{4}{3}\Omega$, which slightly differs from the GW frequency being twice the rotational frequency from a small deformation derived above.

3.6.3 Current Quadrupole Radiation

It is not the mass quadrupole that radiates most strongly for r-modes, but the current quadrupole moment, the second term in equation 2–29. The angular frequency of an r-mode, to $O(\Omega^3)$, is given by not neglecting the scalar product $\vec{v} \cdot \nabla$ term in Equation 3–50:

$$\omega_{r-mode} = \frac{2m\Omega}{l(l+1)} \left(1 - \eta \frac{R^3\Omega^2}{GM}\right), \tag{3–56}$$

where the constant $\eta \sim 0.7$. Because the Ω^2 factor is weighed down by M and η, the Ω term actually drives the frequency evolution more quickly than the Ω^2 term. This implies that the spindown of a neutron star due to r-mode emission will be due to the current-quadrupole principally.

3.6.4 CFS Instability

We study GWs from r-modes because of their ability to be driven unstable by GW radiation. Due to the difference in frames, a mode that is prograde in the rotating frame may be retrograde in the inertial frame, and as GWs are radiated away, angular momentum is actually subtracted from the already negative angular momentum of the star, thus causing a 'runaway' effect whereby the amplitude of the mode grows until some saturation mechanism terminates the growth. The instability condition is derived in (*37*).

To better understand this instability condition, we start with the linearized Euler Equation 3–49. In the lowest order approximation, the potential can be expressed in terms of the density,

the pressure perturbation and the potential perturbation. Using this, we can integrate the energy due to the perturbations (*41*):

$$E = \frac{1}{2}\int d^3x\rho|\delta v|^2 + \left(\frac{\delta P}{\rho} - \delta\phi\right)\delta\rho \tag{3–57}$$

Using expressions for the perturbations found in (*35*) and taking a time derivative, we find:

$$\frac{dE}{dt} = -\int [2\eta\delta\sigma^{ab}\sigma_{ab} + \zeta|\delta\sigma|^2]d^3x - \omega_i(\omega_i + m\Omega)\sum_{l>=2} N_l\omega_i^{2l}(|\delta D_{l,m}|^2 + |\delta J_{l,m}|^2) \tag{3–58}$$

where η is the shear viscosity of the medium through which the r-modes propagate and ζ is the bulk viscosity.

The first term tends to be negative since viscosity damps (or decreases) the energy of the r-mode over time. However, the second term may change sign, depending on if piece outside the sum is positive or negative. So the instability condition, that is, the way to ensure that the perturbation energy increases over time, is:

$$\omega_i(\omega_i + m\Omega) < 0 \tag{3–59}$$

We assume that only $l = m$ r-modes exist in neutron stars. Using Equation 3–55, we obtain:

$$\omega_i = -\frac{(l-1)(l+2)}{l+1}\Omega \tag{3–60}$$

So the instability condition Equation 3–59 is:

$$\omega_i(\omega_i + m\Omega) = -2\frac{(l-1)(l+2)}{(l+1)^2}\Omega^2, \tag{3–61}$$

which is negative definite for $l >= 2$, hence forcing the second term in Equation 3–58 to contribute positive energy over time, causing a growth of the r-mode.

3.7 Cowling Approximation

The Cowling Approximation (*42*) has been hinted at in the above r-mode formalism, but specifically refers to neglecting the perturbations of the metric due to a source's relativistic velocity. Many studies have explored the impacts of this approximation in the context of different oscillation modes of a star (*43*, *44*). Essentially, using enough nodes in the eigenfunctions of whichever mode one considers improves the approximation. Calculations of gravitational waves due to r-modes have also been done within this framework (*35*). While a linear relationship between the r-mode frequency and rotational frequency of the star is a good approximation of the r-mode frequency, the coefficient of proportionality depends strongly on the relativistic parameter $GM/(c^2R)$, which is $O(0.1)$ for neutron stars (*45*).

For the work presented here, the frequency of the r-mode is assumed to be produced only from $l = m = 2$ mode, where higher order modes are neglected. This is certainly an over-simplification of the gravitational wave emission, especially if the r-mode grows over time and hydrodynamic effects become important (*35*). Our methods can therefore not probe deeply the causes of r-mode evolution, but can discover the dominant behavior.

3.8 Equation of State

A central problem of neutron star physics is determining the EoS. General relativity does not provide this for us: we can solve Einstein's equations to get the Tolman-Oppenheimer-Volkoff (TOV) equations (*46*), and then only assume an EoS $p = p(\rho, ...)$ to solve the TOV equations. Therefore, we need to rely on observations and assumptions about the neutron star structure when trying to derive the EoS. GW observations of an isolated neutron star will allow us to constrain its mass and radius, thereby giving constraints on the tidal deformability of the neutron star, the moment of inertia, and the ellipticity. This is because we can measure the rotational frequency and spindown of the star, and in the case of signals predominately considered here, the braking index. Along with an estimation of the amplitude of the GW signal, and an electromagnetic counterpart to constrain the distance, we can deduce separately what the moment of inertia and the ellipticity of the neutron star are (*47*)

CHAPTER 4
THE DAWN OF MULTI-MESSENGER ASTRONOMY

4.1 Introduction

The era of GW astronomy began about 4 years ago with the detection of the first BBH merger (*2*). From subsequent detections in LIGO/Virgo's first and second observing runs of more BBHs and one BNS (*48*), physicists have been finally able to use GW to probe interesting questions about the universe in the strong-field regime. In Section 4.2, we describe the first ever detection of GWs by LIGO, as well as how this detection constrained speed of the graviton. Afterwards we discuss in Section 4.3 the first BNS observed by LIGO/Virgo (*1*) that allowed for one of the most extensive observation campaigns in scientific history (*49*).

4.2 GW150914

The first ever GW detection was made on 14 September 2015 by the two LIGO detectors (*2*). A BBH merger spanned the frequency range 35 to 250 Hz, and peaked in gravitational-wave strain at 1×10^{-21}, an extremely powerful event. From this merger, the LIGO/Virgo collaborations showed that general relativity works in the strong field regime, and that BBHs do exist in nature, though alternative theories to explain the specific nature of this object are possible (*50*). The implications of this merger also include higher-mass stellar black holes than previously thought, meaning that models of stellar evolution need to be changed. Moreover, we could establish some confidence in merger rates of BBHs. We discuss one interesting implication of GW150914: the limit we can place on the speed of gravity.

In theories about massive gravity, the speed of the graviton is less than the speed of light, meaning different modes (frequencies) travel at different speeds, given by:

$$(\hbar\omega)^2 = (\hbar ck)^2 + (m_g c^2)^2 \tag{4–1}$$

where m_g is the mass of the graviton and the term $(m_g c^2)^2$ is very small compared to the others. If the waves propagate at the speed of light, we have:

$$\omega^2 = c^2 k^2 \tag{4–2}$$

$$\omega \; d\omega = c^2 k \; dk \tag{4–3}$$

$$\frac{\omega}{k}\frac{d\omega}{dk} = c^2. \tag{4–4}$$

The speed of massive gravity, specifically the speed of a wave packet (the group velocity), is:

$$\frac{d\omega}{dk} = v_g = \frac{c^2}{\omega/k} \approx c\left[1 - \frac{1}{2}\left(\frac{m_g c^2}{\hbar\omega}\right)^2\right]. \tag{4–5}$$

The duration of the signal observed in GW150914 was about $\Delta t = 0.1$ s, from $f_{min} = 30$ Hz to $f_{max} = 150$ Hz. We can use this information to place an upper bound on the mass of gravity by assuming that the two modes f_{min} and f_{max} were delayed by the length of the signal Δt. The time of arrival of a mode is:

$$t = \frac{D}{v_g} \approx \frac{D}{c}\left[1 + \frac{1}{2}\left(\frac{m_g c^2}{\hbar\omega}\right)^2\right]. \tag{4–6}$$

where D is the distance to the source. Writing an equation using f_{min} and f_{max} and subtracting them, defining $\Delta t = t_{max} - t_{min}$ (where the subscripts min and max refer to the frequencies), we arrive at an upper bound on the mass of the graviton:

$$m_g^2 c^4 \leq 8\pi^2\hbar^2\left(\frac{c\Delta t}{D}\right)\left(\frac{f_{max}^2 f_{min}^2}{f_{min}^2 - f_{max}^2}\right) \tag{4–7}$$

$$m_g c^2 \leq 1 \times 10^{-22} \text{ eV}. \tag{4–8}$$

This is a much more stringent constraint than what previous experiments have given by about two orders of magnitude (*51*).

4.3 GW170817

On 17 August 2017, LIGO and Virgo saw a binary neutron star merger 40 Mpc away (*1*). What followed was one of the most extensive astronomical campaigns in history to see this merger in every lens of the electromagnetic spectrum. We were able to see this merger in every frequency band, which gave us insight into the remnant of the merger, as well as the physics of neutron stars. Searches for a possible remnant of this merger will be presented in chapters 10 and 14.

4.3.1 Constraints on Difference between the Speeds of Light and Gravity

The GWs and the gamma ray burst within $\Delta t = 1.7$ s of each other (*52*), which allows us to constrain how much faster the GW could have been travelling than light. So we can place constraints on how much the speeds of light and gravity differ with the following logic:

$$\begin{aligned} \Delta v &= v_{gw} - v_{em} && \text{(4–9)} \\ &= \frac{D}{t_{gw}} - \frac{D}{t_{em}} && \text{(4–10)} \\ &= D\left(\frac{t_{em} - t_{gw}}{t_{em} t_{gw}}\right) && \text{(4–11)} \\ \Delta v &= v_{em}\left(\frac{\Delta t}{t_{gw}}\right). && \text{(4–12)} \end{aligned}$$

To zeroth order in Equation 4–6, assuming a massless gravity, $t_{gw} \approx D/v_{em}$; therefore:

$$\frac{\Delta v}{v_{em}} \approx v_{em}\frac{\Delta t}{D}. \qquad \text{(4–13)}$$

In the paper (*53*), the authors first assume that the GWs and EM waves were emitted at the same time, and attribute $\Delta t = 1.74 \pm 0.05$ s entirely to faster GWs. This provides an upper, positive bound to the difference in the two speeds, where the authors use $\Delta t = 1.79$ s and $D = 26$ Mpc to be conservative. Then, the authors consider that the two waves were emitted at times differing by more than 1.74 ± 0.05 s, allowing then the EM waves to make up some part of the travel. They consider $\Delta t = 10$ s, and arrive at:

$$-3.74 \times 10^{-15} \leq \frac{\Delta v}{v_{em}} \leq 6.67 \times 10^{-16}. \tag{4–14}$$

4.3.2 Constraints on Equation of State

LIGO/Virgo observations of the masses of the neutron stars can begin to constrain the possible EoS of the neutron star (*1*). In a typical matched filtering search, the phase evolution of the signal is required to be exactly known, and this EoS information enters as one of the higher post-Newtonian corrections to the phase evolution, in a parameter called the tidal deformability Λ. Λ is higher for less compact (so higher radius) stars, and smaller for more compact stars.

Assuming different priors for spins of the neutron star, LIGO/Virgo placed different constraints on the tidal deformability, and ruled out many equations of states that produced less compact neutron stars- see Figure 4-1. The reason that such priors are necessary is because LIGO/Virgo is not as sensitive to the individual masses as it is to the chirp mass, since there are degeneracies in the phase evolution of the signal.

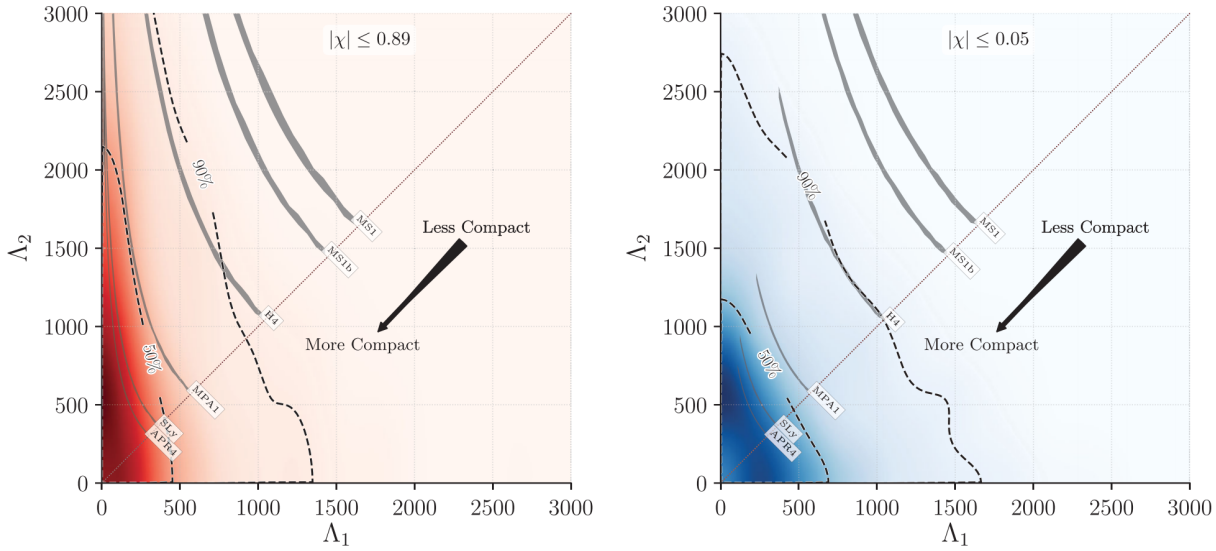


Figure 4-1: This figure, taken from (*1*) shows the constraints on the tidal deformability parameter Λ for low and high spin priors. The color is the probability density. Predictions for $\Lambda_{1,2}$ from different equations of state are drawn on the plot, and some equations of state that predict less compact starts are excluded at 50 and 90% confidence.

CHAPTER 5
GRAVITATIONAL WAVE DETECTORS

5.1 Introduction

No GW detection would have been possible without the ingenious designs for an instrument to measure extremely tiny perturbations in space-time. In this chapter we discuss the basics of the detector used in GW physics. Section 5.2 demonstrates the key principles that motivated the design of current GW detectors. Afterwards in Section 5.3, we describe many of the sources of noise which experimentalists needed to account for when designing such a sensitive instrument. Finally in Section 5.4, we explain some of the key components of the GW detectors.

5.2 Detection Principle

LIGO and Virgo employ interferometry to detect GW signals. An interferometer is a device that interferes two beams of light sent in perpendicular directions. A schematic of an interferometer is shown in Figure 5-1. A (quite powerful) laser serves as input to the experiment, which is then split into 2 perpendicular beams by a beam splitter. A simplified scheme of the working principle is that the beams travel through a cavity formed by the two mirrors, bounce off the mirrors, rejoin at the beam splitter, and are forwarded to the photodetector. What is measured is an interference pattern: if the lengths of the arms are exactly equal, and there is no noise, we expect that the two rays will constructively interfere. If the light rays are out of phase with each other completely, we expect destructive interference and a completely dark interference pattern.

GWs alter the position of the mirrors as discussed in Section 2.3.2. This movement is a function of the source properties that cause the GW. In essence, every black hole merger, neutron star, supernova, etc. will change the distance between the two mirrors in a slightly different way depending on the masses, spins, deformation, level of asymmetry, etc. A GW is therefore like a fingerprint, allowing us to confidently identify the source of such a mirror movement, if we are sure that noise does not interfere.

The GW therefore changes the arrival time of the light rays at the photodetector, and this causes different kinds of interference patterns, because the light rays are now out of phase.

Figure 5-1 is a much more simplified version of what LIGO/Virgo actually use because the noise and the precision needed in this experiment required the development of many advanced techniques. We describe many of the noise sources in LIGO/Virgo in this chapter. More details of the detector design are discussed in the Subsection 5.4.

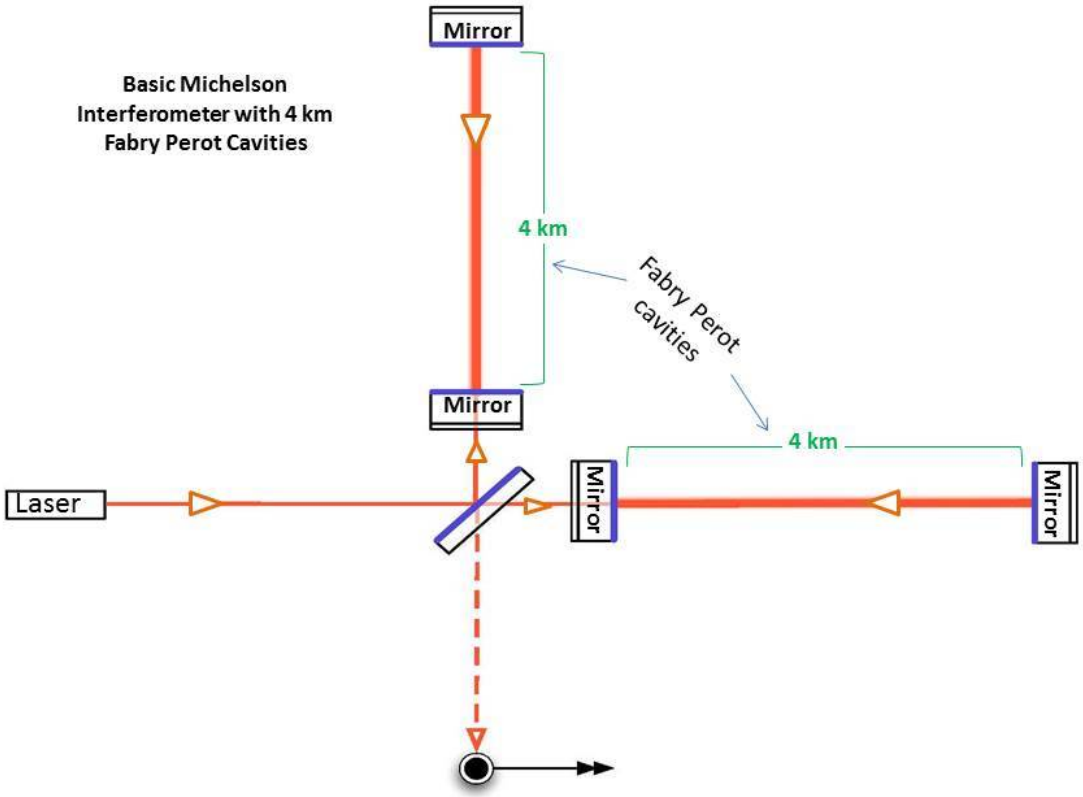


Figure 5-1: This figure shows a simplified version of a LIGO/Virgo interferometer. A laser beam is split into each of the arms, reflected off the mirrors, goes through $O(100)$ roundtrips in the Fabry-Perot (FP) cavity, and recombines at the beam splitter. The output is read by a photodiode, where the phase shift between the light rays in each arm is measured.

5.3 Noise Sources

5.3.1 Seismic Noise

Seismic noise is due to the shaking of the earth at many frequencies, and is a very strong source of noise. We need orders of magnitude of attenuation to have a shot of seeing GWs of any kind. To achieve a large attenuation, LIGO/Virgo uses a system of pendula that, for each pendulum in the cascade, reduces S_n by a factor of f_r^2/f^2, where f_r is the resonance frequency of the pendulum. To show this, we consider a simple pendulum with mass m hanging from a rod of length l whose support point can move in time, with a known function (following the derivation in (*54*)):

$$\vec{R}(t) = X(t)\vec{i} + Y(t)\vec{j}. \tag{5–1}$$

where $\vec{R}(t)$ is the overall displacement of the system, and $X(t)$ and $Y(t)$ describe the horizontal direction $\vec{i}$ and vertical direction $\vec{j}$ time-dependent displacements, respectively. The position of the mass is:

$$\vec{r} = \vec{R} + l(\sin\theta\vec{i} + \cos\theta\vec{j}). \tag{5–2}$$

where l is the length of the pendulum and the angle θ is from the vertical to the pendulum. These two quantities completely describe the system.

We need to write the Lagrangian and derive the equation of motion for the mass m suspended by the pendulum, which we will then use to determine the impact of different forces on the mass. We write the kinetic energy T and potential energy V as:

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(|\dot{\vec{R}}|^2 + l^2\dot{\theta}^2 + 2l\dot{\theta}(\dot{X}\cos\theta - \dot{Y}\sin\theta) \tag{5–3}$$

$$V = -m\vec{g}\cdot\vec{r} = -mgl\cos\theta - mgY. \tag{5–4}$$

where dots denote time derivatives. Now we can write the Lagrangian and then the equation of motion:

$$ml^2\ddot{\theta} + ml(g - \ddot{Y})\sin\theta + ml\ddot{X}\cos\theta = 0. \tag{5–5}$$

If there is no displacement ($X = Y = 0$), we recover the equation of motion for a simple pendulum.

We will focus on a small, periodic displacement in the x-direction: $X = X_0\cos(\omega t)$, where ω is the driving frequency (which could be the GW frequency or a seismic disturbance). The equation of motion takes the form, for small angles and $Y(t) = 0$:

$$l\ddot{\theta} + g\theta + \ddot{X} = 0 \tag{5–6}$$

$$l\ddot{\theta} + g\theta = \omega^2\cos\omega t. \tag{5–7}$$

Because the oscillation is driven, the pendulum will, after some transients in the beginning, oscillate at the same frequency as the support point:

$$\theta = \theta_0\cos\omega t. \tag{5–8}$$

Plugging θ into the equation of motion, we arrive at:

$$\theta_0 = \frac{X_0}{l}\frac{\omega^2}{\omega_0^2 - \omega^2}, \tag{5–9}$$

where $\omega_0^2 = g/l$ is the oscillation frequency of the unperturbed simple pendulum.

The new position of the mass in the x-direction is: $x = X + l\theta$, which comes from Equation 5–2 in the small-angle approximation. We now would like to compute the transfer function F, which relates the movement of the suspension point of the pendulum to the movement of the mass:

$$F(\omega) = \frac{x}{X} = 1 + \frac{l\theta}{X} = 1 + \frac{\omega^2}{\omega_0^2 - \omega^2} \tag{5–10}$$

$$F(\omega) = \frac{\omega_0^2}{\omega_0^2 - \omega^2} = F(f) = \frac{f_r^2}{f_r^2 - f^2}. \tag{5–11}$$

This transfer function represents an attenuation in displacement: if something attempts to move the suspension point, so long as the rate at which it tries to move the suspension point is not near or below the resonance frequency, the motion of the mass will be attenuated: if $f >> f_r$, we see that $F(f) \sim f_r^2/f^2$. If we keep the resonance frequency of the pendulum significantly lower that the GW frequency we would like to find, we can apply N pendulums to get $(f_r^2/f^2)^N$ levels of attenuation in the motion of the test mass.

5.3.2 Thermal Noise

Thermal noise is caused by Brownian/random motion of the atoms that compose the detector. It affects both the mirrors and the suspensions. We can express the power spectral density of the noise using the fluctuation-dissipation theorem:

$$S(\omega) = 4k_B T \text{ Re } Z(\omega), \tag{5–12}$$

where Z is the impedance, T is the temperature and k_B is the Boltzmann constant.

This theorem states that as the force varies in time/frequency (fluctuations), there will be a loss of mechanical energy of the system (dissipation). It is expressed above in the frequency domain.

In LIGO/Virgo, the thermal noise is calculated for the suspensions/mirrors by modeling these systems as damped harmonic oscillators with a dampening term proportional to the velocity of the atoms in motion:

$$\ddot{x} + \omega_0^2 x + \gamma \dot{x} = \frac{F(t)}{m}, \tag{5–13}$$

where γ is the damping factor, and $F(t)$ is the driving force. The displacement spectral density is given in terms of the fluctuation-dissipation theorem (*55*):

$$X(\omega) = \frac{1}{\omega |Z(\omega)|} (4kT \text{ Re } Z(\omega))^{1/2}. \tag{5–14}$$

This term is found by casting the equation of motion in the form:

$$F(\omega) = -i\omega Z(\omega)x(\omega), \tag{5–15}$$

For a damped harmonic oscillator, allowing $\gamma(\omega) \to \phi(\omega)$ (for generality), we can write by inspection:

$$Z(\omega) = \frac{im}{\omega}\left(\omega^2 - \omega_0^2 + i\omega^2\phi(\omega)\right) \tag{5–16}$$

The displacement noise then is:

$$X(\omega)^2 = \frac{kt\omega_0^2\phi(\omega)}{m\omega[(\omega_0^2 - \omega^2)^2 + \omega_0^4\phi^2(\omega)]} \tag{5–17}$$

where $\phi(\omega)$ is the loss angle of the oscillator whose fundamental mode is ω_0, and takes different forms for thermal noise at the mirror, suspensions, etc.

5.3.3 Shot Noise

Shot noise results from the fact that we are counting photons bouncing off each mirror to measure the phase shift of the light, and there is an uncertainty in the number of photons that arrive at the photodiode. To calculate the spectral density of this noise, we do the following: let N_γ be the number of photons that we would like to count at the photodetector in observation time T. The angular frequency of the laser light is ω_L. Then the power due to these photons at the photodetector is:

$$P = \frac{1}{T}N_\gamma\hbar\omega_L, \tag{5–18}$$

where $\hbar$ is the reduced Plank's constant. When we count a number of independent and discrete events, the probability of such photons hitting the mirrors follows a Poisson distribution, and in the limit that we count many (we count around $O(10^9)$), the uncertainty ΔN_γ is:

$$\Delta N_\gamma = \sqrt{N_\gamma}. \tag{5–19}$$

So the fluctuation is the observed power of the laser beam is (replacing N_γ by ΔN_γ in the above equation):

$$\Delta P = \frac{1}{T}\sqrt{N_\gamma}\hbar\omega_L \tag{5–20}$$

$$\Delta P = \left(\frac{\hbar\omega}{T}P\right)^{1/2}. \tag{5–21}$$

To determine P, we must write equations for the photon's electric field as it travels through each cavity and account for how it is affected by the passing of a GW. The electric field $E_{x,y}$ along each arm will have the form:

$$E_{x,y} = \mp\frac{1}{2}E_0 e^{-i\omega_L t + 2ik_L L_{x,y}}, \tag{5–22}$$

and the total field $E_{tot} = E_1 + E_2$ is what we will see at the photodetector. Its square modulus is

$$|E_{tot}|^2 = E_0^2 \sin^2[k_L(L_y - L_x)] = E_0^2 \sin^2\phi_0 \tag{5–23}$$

Knowing just this and that $P \propto |E_{tot}|^2$ and $E_0 \sim P_0$ we can compute the change in the power of the laser just due to noise:

$$(\Delta P)_{\text{shot}} = \left(\frac{\hbar\omega_L}{T}P_0\right)^{1/2} |\sin\phi_0| \tag{5–24}$$

The next step is to determine $(\Delta P)_{\text{GW}}$, which will affect the phase of E_x and E_y that recombine at the photodetector. To do this, we need to compute the arrival times of the photons from each arm when affected by the GW (in the transverse-traceless gauge).

Based on the analysis in Sections 2.3.1 and 2.3.2, the perturbation $h_{\mu\nu}$ has only h_+ and $h_\times$ components, so the metric is written as:

$$ds^2 = g_{\mu\nu}x^\mu x^\nu = -c^2dt^2 + (1+h_+)dx^2 + (1-h_+)dy^2 + 2h_\times dxdy + dz^2 \tag{5–25}$$

where $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$.

Consider a GW with $+$ polarization traveling in the z-direction:

$$h_+(t) = h_0 \cos \omega_{gw} t \tag{5–26}$$

For a photon, $ds^2 = 0$, and when it travels only in the x-direction, $dy^2 = dz^2 = 0$. The photon is affected by the passing GW, and we can express the distance traveled from the beam splitter to the mirror L_x as (the forward trip of the light):

$$L_x = c(t_1 - t_0) - \frac{c}{2}\int_{t_0}^{t_1} dt' h_+(t') \tag{5–27}$$

where t_0 is the time that the light is emitted and t_1 is the time that the light reaches the end of the arm of length L_x. For the return trip, the expression is the same, just with $t_1 \to t_2$ and $t_0 \to t_1$. t_2 is then the time that the photon returns back to its origin.

By summing the forward and backward parts:

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{1}{2}\int_{t_0}^{t_2} dt' h_+(t'). \tag{5–28}$$

In the case of no GWs, we just get the standard expression: $t_2 - t_0 = \frac{2L_x}{c}$. Since the correction h_+ is small, we can let $t_2 \to t_0 + \frac{2L_x}{c}$.

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{1}{2}\int_{t_0}^{t_0+2L_x/c} dt' h_0 \cos \omega_{gw} t. \tag{5–29}$$

We integrate this expression and get:

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{h_0 L_x}{c}\,\mathrm{sinc}(\omega_{gw} L_x/c)\cos(\omega_{gw}(t_0 + L_x/c)) \quad (5\text{–}30)$$

$$= \frac{2L_x}{c} + \frac{h_0 L_x}{c}\,\mathrm{sinc}(\omega_{gw} L_x/c) h_+(t_0 + L_x/c) \quad (5\text{–}31)$$

It is more useful to define $t_2 = t$ as the time which we observe the light at the photodetector, and then compute the corresponding t_0. This means that $h_+(t_0 + L_x/c) \to h_+(t - 2L_x/c + L_x/c)$ based on Equation 5–28 in which we neglect the h_+ integral:

$$t_0^x = t - \frac{2L_x}{c} - \frac{L_x}{c} h_+(t - L_x/c)\,\mathrm{sinc}(\omega_{gw} L_x/c) \quad (5\text{–}32)$$

and $x \to y$ for the other arm (with a sign flip of the third term due to the minus sign present in the dy^2 term in the metric).

Now, we can compute the phase shift in the electric field due to the GW. The free propagation of the light in the cavity does not change the phase, and the reflection/transmission of the light only give additional $1/2$ factors in the amplitude, so:

$$\Delta\phi_{gw,x} = \omega_L t_0^x = \frac{\omega_L L_x}{c} h_0 \cos\omega_{gw} t\,\mathrm{sinc}(\omega_{gw} L_x/c). \quad (5\text{–}33)$$

where this phase shift is between the time observed t and the time that the beam left the origin t_0. The analysis for the y-direction is the same, except that $h_+ \to -h_+$ because of the metric, so:

$$\Delta\phi_{gw,x} = -\Delta\phi_{gw,y}. \quad (5\text{–}34)$$

At the photodetector, we see the difference of the phase shifts in x and y:

$$\Delta\phi_{mich} = \Delta\phi_{gw,x} - \Delta\phi_{gw,y} = 2\Delta\phi_{gw,x}. \quad (5\text{–}35)$$

The total phase shift is:

$$\phi_{tot} = \phi_0 + \Delta\phi_{mich} \tag{5–36}$$

and the power $P \sim |E_{tot}|^2$ is (in the spirit of Equation 5–23):

$$\begin{aligned} P &= P_0^2 \sin^2(\phi_0 + \phi_{gw,x}) && (5–37)\\ &= \frac{P_0}{2}\left(1 - \cos[2\phi_0 + \Delta\phi_{mich}(t)]\right). && (5–38) \end{aligned}$$

Now, we can calculate the change in power due to a GW signal by subtracting from P the contribution due to only noise $P_{noise} = P_0 \sin^2\phi_0$, and by allowing $\cos\phi_{mich} \sim 1$ and $\sin\phi_{mich} \sim \phi_{mich}$:

$$P - P_{noise} = (\Delta P)_{gw} = \frac{P_0}{2}\Delta\phi_{mich}|\sin 2\phi_0| \tag{5–39}$$

From here, we can compute the signal-to-noise ratio of the GW signal:

$$\begin{aligned} \left(\frac{S}{N}\right) &= \frac{\Delta P_{gw}}{\Delta P_{shot}} && (5–40)\\ &= \left(\left(\frac{P_0}{\hbar\omega_L}\right)^{1/2} |\cos\phi_0|\Delta\phi_{mich}\right) T^{1/2}. && (5–41) \end{aligned}$$

Comparing Equation 5–41 with Equation 7–13, we can read off S_n:

$$\frac{S_n}{h_0} = \left(\frac{\hbar\omega_L}{P_0}\right)^{1/2} \frac{1}{|\cos\phi_0|\Delta\phi_{mich}}. \tag{5–42}$$

To get a better idea of how S_n behaves, let us choose $\phi_0 = \pi/4$ so $\cos\phi_0 = 1/\sqrt{2}$, and let us analyze the simple case of $\omega_{gw}L_x/c \ll 1$, so we can Taylor expand Equation 5–35 and obtain:

$$\Delta\phi_{mich} = h_0\frac{2L\omega_L}{c} \tag{5–43}$$

$$S_n^{1/2} = \frac{1}{L}\sqrt{\frac{\hbar c^2}{2P_0\omega_L}}. \tag{5–44}$$

A key finding is that the shot noise gets better as more laser power is used. While true, another noise source, radiation pressure noise, increases with increasing laser power, as we will discuss in Subsection 5.3.4. But in Fabry-Perot (FP) cavities (see Section 5.4.2) , the types of cavities actually used in LIGO/Virgo, the shot noise is frequency dependent. We will discuss FP cavities in a later section.

5.3.4 Radiation Pressure Noise

Radiation pressure noise is due to the fact that a different number of discrete photons are hitting the mirrors at any given time, causing a time-varying, stochastic force to push the mirrors. The number of photons hitting the mirrors fluctuates as a function of time because the counting statistics are Poisson-distributed- see Equation 5–19). Therefore, we need to calculate the impact that this force F has on the sensitivity of the detector.

The photons are reflected, resulting in a doubling in the momentum transferred. Therefore the power that is incident on the mirror is: $P = 2Fc$. The noise results because the force on the mirrors is always changing slightly, meaning that we will use Equation 5–24 to describe this fluctuation:

$$\Delta F = 2\frac{\Delta P}{c} = 2\sqrt{\frac{\hbar\omega_L P}{c^2 T}} \tag{5–45}$$

where we have set $\phi_0 = \pi/2$.

We would now like to compute the power spectral density of the force. We begin by calculating the auto-correlation of the fluctuations in the force, ΔF. The fluctuations are independent since we are describing a random process, so we can write:

$$< \Delta F(t)\Delta F(t) >= \frac{1}{2T}S_F, \tag{5–46}$$

where S_F is the spectral density of the force. Plugging in Equation 5–45 into Equation 5–46, we get:

$$S_F^{1/2} = 2\sqrt{\frac{2\hbar\omega_L.P}{c^2}} \tag{5–47}$$

To relate the power spectral density of the force to the power spectral density of the displacement (or the strain), we need to specify the transfer function. Let's assume a simple system where this stochastic force is the only one acting on the mirror of mass M: $F = M\ddot{x}$. The transfer function will relate the power spectral density of the force to that of the displacement. We can evaluate the transfer function by solving for the relationship between F and x in Fourier space:

$$\tilde{F}(f) = -4\pi^2 M f^2 \tilde{x}(f) = T_{(F\to x)}\tilde{x}(f). \tag{5–48}$$

since derivatives in time become multiplications in frequency domain, and $T_{(F\to x)}(f)$ is the transfer function. Here f is the frequency at which the photons (or the GW) cause the mirror to oscillate. Then:

$$S_x^{1/2}(f) = \frac{2}{M(2\pi f)^2}\sqrt{\frac{2\hbar\omega_L P}{c^2}}. \tag{5–49}$$

To relate the power spectral density of the displacement of the mirrors to that of the strain, we can use $\Delta L = Lh = T_{x\to h}h$ for a simple interferometer, which means the transfer function is just L:

$$S_n^{1/2} = \frac{4}{ML(2\pi f)^2}\sqrt{\frac{2\hbar\omega_L P}{c^2}}. \tag{5–50}$$

5.4 Detector Design

5.4.1 Simple Michelson Interferometer

We can calculate the optimal length of a simple interferometer by maximizing the phase shift due to GWs 5–35. Fully expressing the sinc in Equation 5–33, we find:

$$\phi_{mic} = \frac{2\omega_L}{\omega_{gw}} \cos(\omega_{gw} t) \sin(\omega_{gw} L_x / c). \tag{5–51}$$

Maximizing with respect to L_x and writing $\omega_{gw} = 2\pi f$, we arrive t:

$$L_x = \frac{c}{4} \frac{1}{f} = 750 \text{ km } \left(\frac{100 \text{ Hz}}{f} \right). \tag{5–52}$$

5.4.2 The Fabry-Perot Interferometer

Because it is not feasible to build a detector of 750 km on earth, the collaborations have thought of a clever way to effectively increase the arm length of the detector using FP cavities. Essentially, light is reflected back and forth between the beam splitter and the end mirror many times before being read into the photodetector.

The difficulty in this approach lies in ensuring that the mirrors in the cavity are as close to perfect reflectors as possible. Moreover, it is best to keep the cavities in resonance, since this ensures that the electric field that is shot into the cavity (and therefore the number of photons in the cavity) remains large. This means that various beams within the cavity are all constructively interfering with each other.

The sensitivity (the amount of phase variation we are able to see) is enhanced in a FP cavity by a factor $\frac{2}{\pi}\mathcal{F}$ relative to a Michelson interferometer, where $\mathcal{F}$ is the Finesse, a quantity that is primarily related to how well the mirrors reflect light. This factor can be computed by using the reflected electric field amplitude at both mirrors, originally coming from working through equations that relate incoming, transmitted and reflected electric fields to each other for a certain number n of roundtrips (*55*).

CHAPTER 6
PROSPECTS OF DETECTING NEUTRON STARS: RATES AND IMPLICATIONS

6.1 Introduction

Isolated neutron stars are formed as a result of BNS mergers or supernovae, so it is interesting to discuss the rates at which these events are expected to happen. Astrophysically, we can gain some insight into the rates of mergers by observing how often we see gamma ray bursts (GRBs), which could be indicative of mergers, as in the case of GW170817, or stellar collapse. However, in this chapter we will focus on what GW observations alone can tell us. In Section 6.2, we explain how event rates are calculated theoretically and empirically based on GW data. We then describe ways to optimize our sensitivity to long duration transient GW signals in Section 6.3. In Section 6.4, we discuss what we can learn from a hypothetical detection of a power-law GW signal. Finally, we calculate constraints that can be placed on the rates of BNS mergers accessible to current and future detectors in Section 6.5.

6.2 Merger Rates: What is Known

In a 12 month period (20 May 2018- 19 May 2019), an average of 1.25 GRBs per day above threshold, and 0.25 GRBs per day subthreshold (*56*), were observed by Fermi-GBM and Swift-BAT (*57*, *58*). It is possible that some of these GRBs are due to mergers of neutron stars.

There are different ways to estimate the rates of neutron star coalescences. Using one gravitational wave observation (GW170817) and an observation time equal to the duration of LIGO's second observing run, LIGO/Virgo inferred a merger rate of 1540^{+3200}_{-1200} Gpc^{-3} yr^{-1} (*1*) by applying a Bayesian method (*59*). The method is as follows: many injections are done uniformly throughout the parameter space, and if the injection is recovered (above a certain SNR) with a false alarm rate lower than a given threshold, then a signal in this portion of the parameter space is detected. By spreading out injections across the source parameter space and at different redshifts, the space-time volume $< VT >$ to which the detector is sensitive is estimated. Then, given a certain confidence level c, the rate R_c can be estimated (*59*):

$$R_c = \frac{-ln(1-c)}{< VT >} \tag{6–1}$$

More precisely, this space-time volume is (*60*, *61*):

$$< VT >= T \int_0^{z_{max}} \frac{dV_c}{dz} \frac{1}{1+z} p(\theta) \frac{N_{found}}{N_{inj}} d\theta dz \tag{6–2}$$

where $\frac{dV_c}{dz}$ is the differential co-moving volume of injections in the astrophysical parameter space, $p(\theta)$ is the distribution function for the astrophysical population of neutron stars from which injection source parameters are drawn , N_{found} is the number of found injections, N_{inj} is the total number of injections done, z is the redshift, T is the observation time and θ are all source parameters. The quantity $\int_0^{z_{max}} \frac{dV_c}{dz} \frac{1}{1+z} dz$ is in fact the astrophysical volume in which the injections were done.

Other ways to estimate the rate of BNS mergers include using just GRBs as indicators of mergers, or assuming a proportionality between merger rates and stellar birth rates in spiral galaxies nearby. The latter birth rate is estimated from the measured blue luminosity, which can be done because of the binaries' short lifetimes in Milky Way-like galaxies, and slow rate of change in galaxies' star formations rates (*62*, *63*). Moreover, observations of pulsars with the same spin period allow a calculation of the scale factor (*64*, *65*). The scale factor relates the number of pulsars identifiable in a galaxy to those potentially found by pulsar surveys (whose search volume is smaller than the galaxy), meaning it is a measure of how many similar pulsars to ones we have already found in BNS systems exist in a galaxy. By dividing the scale factor by the estimated lifetimes of binaries, the merger rate is obtained, taking into account all detected BNS systems. Corrections to the scale factor can be made because (1) the size of the detectable emission cone of pulsars is uncertain, (2) the pulsars are just too faint to be seen by any pulsar survey, or (3) the orbital periods are so short that the pulses are not distinguishable due to high orbital frequency (*66*).

6.3 Optimizing Sensitivity

Since our interest will be in transient CW signals, we will have to deal with very high spindowns, meaning that they progress to lower frequencies (and hence improved detector sensitivity) very quickly. However, higher frequency signals have higher amplitudes, as seen in Equation 3–34, but if a signal has a lower spindown, longer T_{FFT} can be used, improving sensitivity (see Equation 9.6). This means that for different combinations of $f, \dot{f}, T_{obs}$ and T_{FFT} (parameters defined in 3.5), the sensitivity will change. We therefore compute a figure of merit for the sensitivity for various combinations of these parameters, showing in which portions of the parameter space it is better to analyze short and stronger signals versus longer and weaker signals.

The figure of merit is:

$$S(f) \propto \frac{(T_{FFT}T_{obs})^{1/4}}{\sqrt{S_n}} f^2, \tag{6–3}$$

which is computed for every frequency that the signal passes through during T_{obs}.

For braking index $n = 5$ and $T_{obs} = 10^5$ s (~ 1 day) and for a range of initial frequencies and spindowns, we plot the optimal T_{FFT} and the amount of time skipped if the GW begins at $t = 0$, using the noise power spectral density from O2 Livingston data. Signals at higher frequencies/spindowns tend to require shorter T_{FFT}, which means that it may actually be better to exclude some data so that we can use longer T_{FFT} in order to optimize sensitivity for this analysis. This is because we can use longer T_{FFT} for signals with lower spindowns. This is shown in Figure 6-1.

The optimal T_{FFT} and amount of time to skip are also both sensitive to the braking index. We display this in Figure 6-2.

6.4 Implications of a Detection

If a fixed power law signal is detected by our methods, described in chapters 9, 12 and 13, we will have an estimation of f_0, k and n. From these parameters, and from an estimation of the distance d due to an electromagnetic counterpart, we can break the degeneracy between I

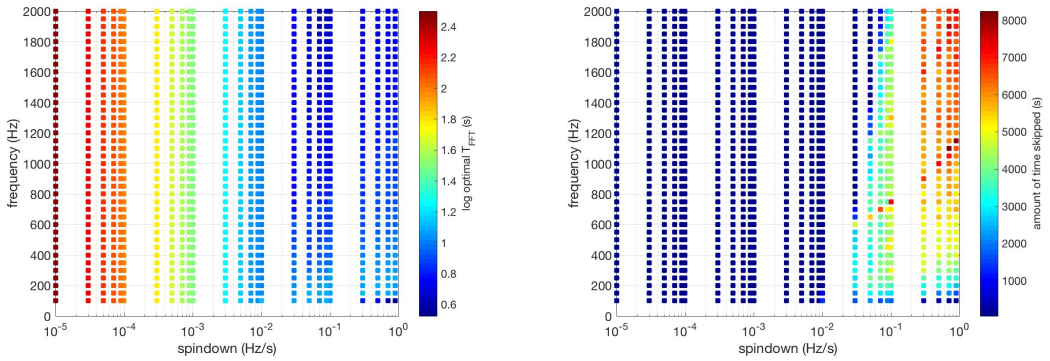


Figure 6-1: For a wide range of frequencies and spindowns, we plot the optimal T_{FFT} in seconds (left) and how much time to skip in the analysis after a NS is born (right), with the criteria that the sensitivity in Equation 6–3 is maximized.

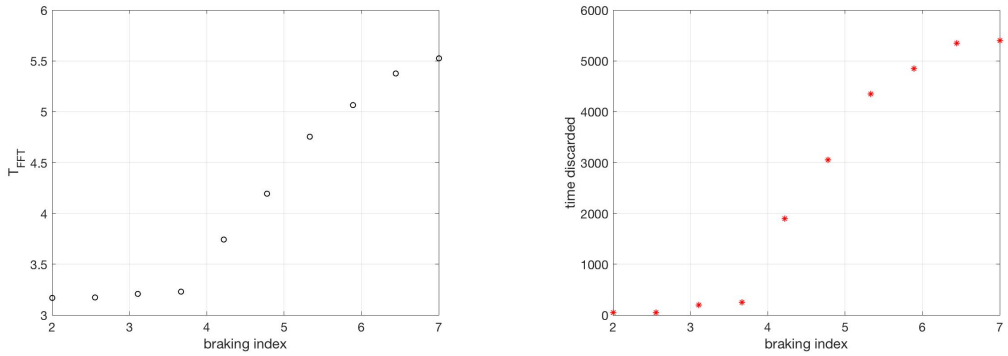


Figure 6-2: We vary the braking index for a fixed $f_0 = 2000$ Hz, $\dot{f}_0 = -10^{-1}$ Hz/s, $T_{obs} = 10^5$ s. Left: optimal T_{FFT}; right: amount of time skipped, both as a function of braking index.

and ϵ in equations 3–23 and 3–33 if we assume that all the energy emitted by the pulsar goes into GWs. This section follows closely the analysis presented in (*47*).

For the case $n = 5$, we can arrive at:

$$\epsilon = \frac{5c}{4\pi^2} \frac{\sqrt{1 + \frac{t}{\tau}}}{h_0 d} \frac{\dot{f}_0}{f_0^3} \tag{6–4}$$

and the z-component of the moment of inertia is:

$$I_3 = \frac{c^3}{5G} \frac{h_0^2 d^2}{(1 + \frac{t}{\tau})} \frac{f_0}{\dot{f}_0} \tag{6–5}$$

Moreover, we can also deduce an ellipticity-independent quantity from equations 3–22 and 3–23:

$$\frac{\dot{E}}{\omega\dot{\omega}} = I_3 = 4\pi \int \rho(r) r^4 dr \tag{6–6}$$

$\dot{E}$ can be estimated by fitting waveforms to the time-series data with different amplitudes h_0 for given recovered parameters of $f_0, \dot{f}_0$, and n (under the assumption that all energy emitted goes into GWs) (*12*, *67*):

$$\dot{E} = \frac{2\pi^2 c^3}{5G} d^2 f^2 h_0^2 \tag{6–7}$$

In terms of the GW frequency and spindown f and $\dot{f}$:

$$\frac{\dot{E}}{f\dot{f}} = 4\pi^2 I_3 = 16\pi^3 \int \rho(r) r^4 dr \tag{6–8}$$

With these parameters, we can begin to deduce what the moment of inertia of the neutron star is, giving us insight into its equation of state.

6.5 Rates and Possible Detections of Newly Born Neutron Stars in Current and Future Detectors

The existing gravitational wave detectors have been able to see BBHs and BNSs out to O(Gpc) and O(Mpc) distances, respectively. However, due to longer arms and a triangular design (*68*, *69*), future detectors, specifically the US-based Cosmic Explorer (CE) and Europe-based Einstein Telescope (ET) will be able to see black hole binaries out to redshifts of $z \sim 20$. Specifically the event rates inferred by GW detections in advanced LIGO/Virgo imply hundreds of events per year for future detectors.

It is more instructive now to consider a theoretically calculated rate, co-moving volume, etc. as opposed to one based on injections described in Section 6.2. Essentially, the rate (number of detections per time) $\dot{N}$ is related to the source density $\dot{\mathcal{N}}$ and the probability that the SNR ρ of the signal is above a threshold (*70*–*72*):

$$\dot{N} = 4\pi\dot{\mathscr{N}} \int_0^\infty dr r^2 P(\rho^2 > \rho_{det}^2) \tag{6–9}$$

We need to relate the probability to the flux of gravitational waves (energy per unit area per unit time) from a single source, which is (*3*):

$$F_{GW} = \frac{c^3}{16\pi G} < \dot{h}_+^2(t) + \dot{h}_\times^2(t) > \tag{6–10}$$

When assuming isotropic emission, we can find E_{GW} over the entire time of GW emission:

$$E_{GW} = 4\pi r^2 T F_{GW} \tag{6–11}$$

Defining $h_{rss}^2 = \int_{-\infty}^{\infty} df |\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2$, computing a time average over the duration T of the GW, and taking the Fourier transform of Equation 6–10, we arrive at:

$$E_{GW} = \frac{\pi^2 c^3}{G} r^2 f_0^2 h_{rss}^2 \tag{6–12}$$

Now that we have a relationship between the strain h_{rss} and E_{GW}, we can directly relate the SNR ρ to E_{GW}. For a narrowband signal, ρ is simply related to h_{rss} as (*70*):

$$\rho^2 = 2 \int_{-\infty}^{\infty} df \frac{|F_+ \tilde{h}_+(f) + F_\times \tilde{h}_\times(f)|^2}{S(f)} \tag{6–13}$$

We expand the square, and drop second order terms in h because for elliptical polarizations, h_+ and $h_\times$ are orthogonal, and for linear polarizations, $h_\times = 0$. Since we are considering a narrowband signal for a short duration, the beam pattern functions can be seen as being constant in time. Knowing this, we arrive at:

$$\rho^2 = \Theta^2 \frac{h_{rss}^2}{S(f)} \tag{6–14}$$

where Θ^2 is the combination of the detector beam pattern functions and source polarization and inclination angles. Now, we can use Equation 6–13 to relate ρ and E_{GW}:

$$\rho^2 = \frac{\Theta^2}{\alpha}\frac{G}{\pi^2 c^3}\frac{E_{GW}}{S(f)r^2 f^2} \tag{6–15}$$

where $\alpha = 1, 8/15, 2/5$ for isotropic, linear or circularly polarized emission.

Using Equation 6–15, we can rewrite the probability term in Equation 6–9 as:

$$P(\rho^2 > \rho^2_{det}) = P(\Theta^2 > r^2/r_0^2) \tag{6–16}$$

where:

$$r_0^2 = \frac{G}{\alpha\pi^2 c^3}\frac{E_{GW}}{S(f)f^2\rho^2_{det}} \tag{6–17}$$

Thus the detection rate is:

$$\dot{N}_{det} = \frac{4}{3}\pi r_0^3 \dot{\mathscr{N}}\left[3\int_0^\infty dx x^2 P(\Theta^2 > x^2)\right] \tag{6–18}$$

Defining the effective detection range R_{eff} as the radius of a sphere with volume V such that the detection rate is $\dot{\mathscr{N}}V$:

$$R_{eff} = r_0\left[3\int_0^\infty dx x^2 P(\Theta^2 > x^2)\right]^{1/3} \tag{6–19}$$

and evaluating the integral numerically (there is little difference among the value of this integral for different polarizations), we arrive at:

$$R_{eff} \approx \left(\frac{G}{2\pi^2 c^3}\frac{E_{gw}}{S(f)f^2\rho^2_{det}}\right)^{1/2} \tag{6–20}$$

and the rate is:

$$\dot{N} = \frac{4}{3}\pi R^3_{eff}\dot{\mathscr{N}} \tag{6–21}$$

We plot the strain curves of current and future detectors in Figure 6-3 (we have taken the noise spectral densities from (*73*)) and the effective distance we can see, for $\rho_{det} = 8$ and BNS

masses of 1.4M_{sun} in Figure 6-4. $\dot{N}$ is determined by how many signals we actually see (and in the case of no detections, how many injections in the space-time volume $< VT >$ we recover).

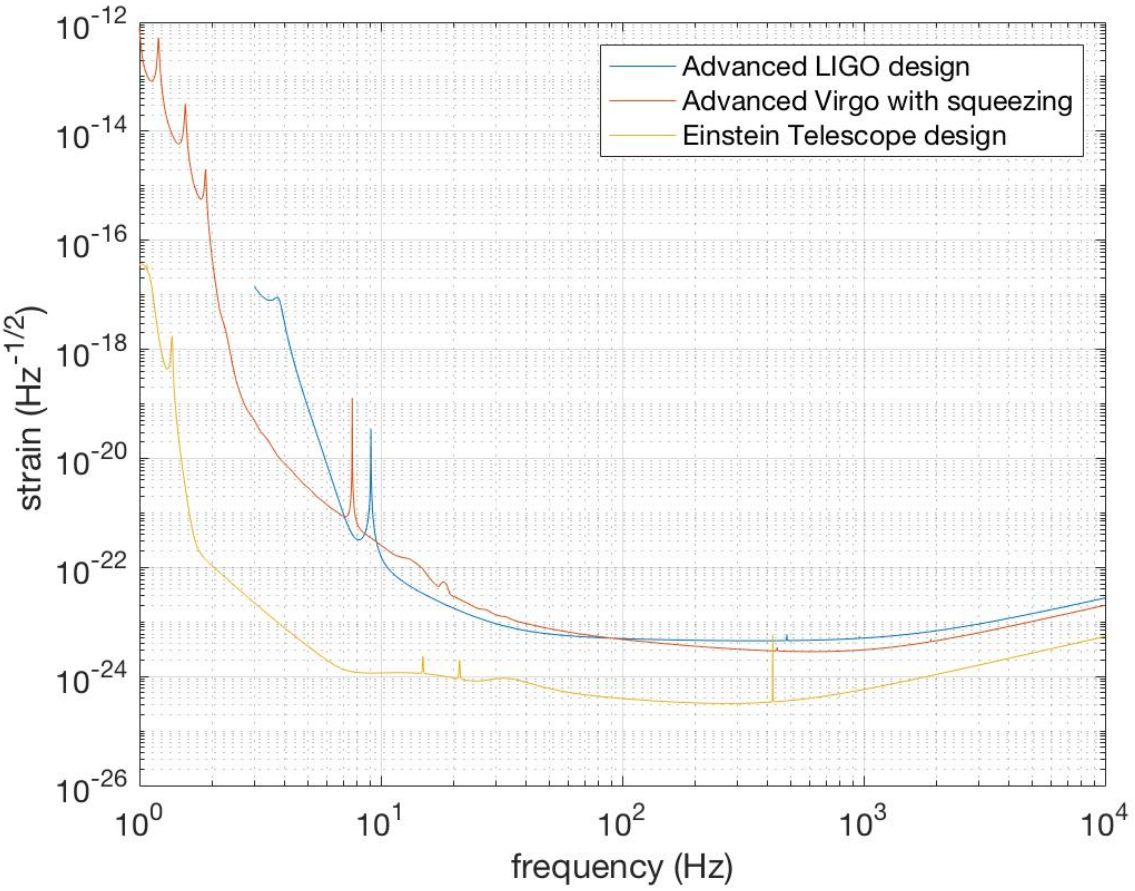


Figure 6-3: Strain curves for advanced LIGO at design sensitivity, advanced Virgo with squeezing employed and Einstein Telescope at design sensitivity. Einstein Telescope will benefit from an order of magnitude increase in sensitivity relative to LIGO and Virgo.

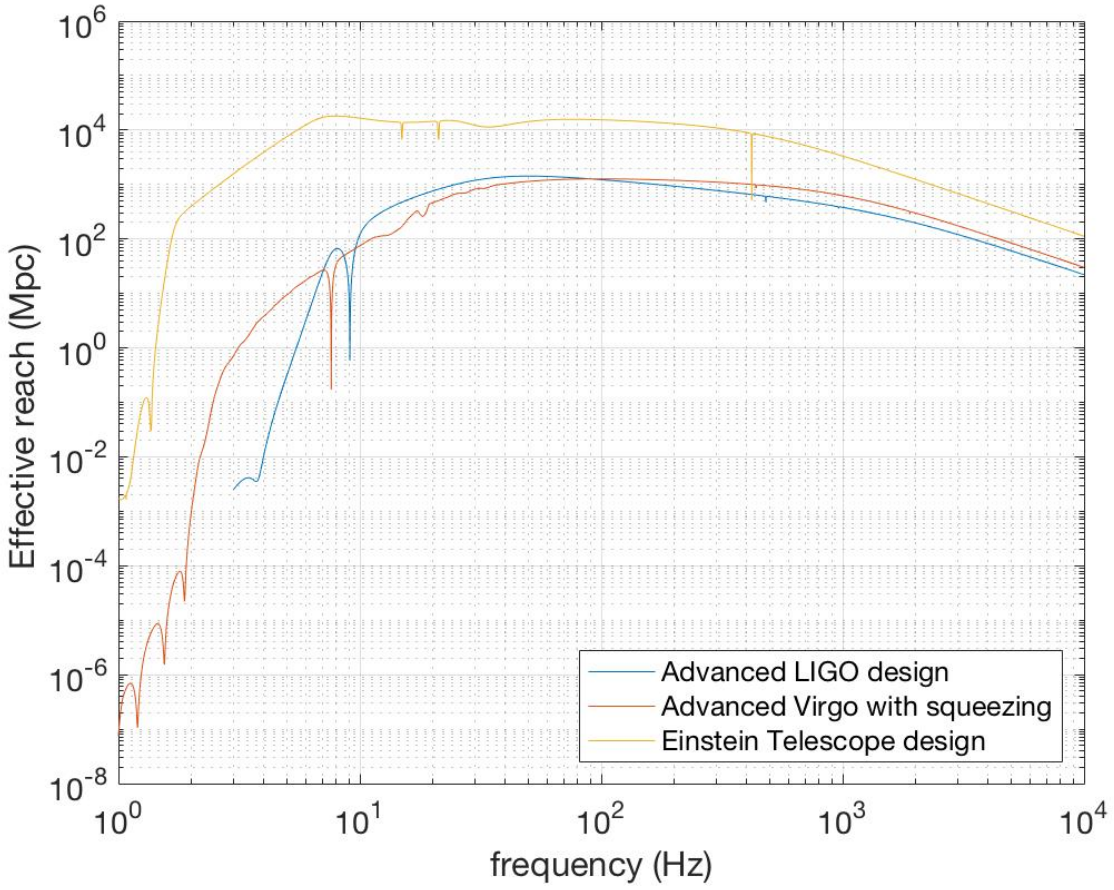


Figure 6-4: The effective reach of a search for binary neutron stars is plotted as a function of detection frequency for different detectors. Parameters: $m_1 = m_2 = 1.4M_{sun}$, $E_{GW} = 10^{-2}M_{sun}c^2$,$\rho_{det} = 8$

CHAPTER 7
DATA ANALYSIS

7.1 Introduction

In this chapter we will discuss various data analysis techniques used in GW searches. We begin in Section 7.2 by describing the most sensitive technique, matched filtering, and why it is difficult to apply it in CW searches. In Section 7.3, we discuss the ways in which GWs are influenced as they propagate from their source to the earth, which result in many time shifts that must be accounted for when running a real search.

7.2 Ideal Detection Technique: Matched Filtering

Let $d(t) = h(t) + n(t)$, where $d(t)$ is the data that we take, $h(t)$ is the signal we would like to find, and $n(t)$ is the noise we do not want. We would like to filter the data, i.e. multiply it, by some optimally chosen function $K(t)$. We define the output of the data multiplied by some filter function $K(t)$ over time as:

$$\hat{d} = \int_{-\infty}^{+\infty} dt d(t) K(t). \tag{7–1}$$

The signal power S is the expected value of $\hat{d}$ if a signal is present in $d(t)$. We calculate this as:

$$\begin{aligned} S &= \int_{-\infty}^{+\infty} dt < d(t) > K(t) && (7–2)\\ &= \int_{-\infty}^{+\infty} dt < h(t) > K(t) && (7–3)\\ &= \int_{-\infty}^{+\infty} df < \tilde{h}(f) > K(t), && (7–4) \end{aligned}$$

where $< .. >$ is an ensemble average, and then we determine the noise to be:

$$N^2 = \sigma^2 =< d(t)^2 > - < d(t) >^2 \quad (7\text{–}5)$$

$$= < d(t)^2 > \quad (7\text{–}6)$$

$$= \int_{-\infty}^{+\infty} dt dt' K(t) K(t') < n(t) n(t') > \quad (7\text{–}7)$$

$$= \int_{-\infty}^{+\infty} df df' \tilde{K}(f') \tilde{K}(f) e^{2\pi i (ft - f't')} < \tilde{n}(f) \tilde{n}(f') > . \quad (7\text{–}8)$$

$< n(t) = 0 >$ because the noise is assumed to be random. Understanding that different Fourier components of the noise are uncorrelated, we only have a contribution to the integral when $f = f'$:

$$N^2 = \int_{-\infty}^{+\infty} df \frac{1}{2} S_n(f) |\tilde{K}(f)|^2, \quad (7\text{–}9)$$

where $S_n(f)$ is the spectral power density of the noise, units of Hz^{-1}. $S_n(f)$ is measured and depends on the different types of noise sources that are present at each frequency.

To find this optimal filter function, we can rewrite S/N into inner products between the noise power spectral density and the signal. When viewed in this way, the max S/N is when the "signal vector" is parallel to the "noise vector". Essentially, this means that the ideal filter function is the $h(t)$ signal itself:

$$\left(\frac{S}{N}\right)^2 = 4 \int_0^{\infty} df \frac{|\tilde{h}(f)|^2}{S_n(f)}. \quad (7\text{–}10)$$

For a CW signal, taking the Fourier transform of Equation 3–25 (*55*):

$$\tilde{h}(f) = \frac{1}{2} \delta(f - f_0)[F_+ h_{0,+} + F_\times h_{0,\times}], \quad (7\text{–}11)$$

Accounting for the fact that the observation time is fixed at T, the Dirac Delta function must be replaced by a regularized Dirac Delta function (*55*):

$$\left(\frac{S}{N}\right)^2 = |F_+ h_{0,+} + F_\times h_{0,\times}|^2 \int_0^\infty df \frac{\delta(f-f_0)\delta(0)}{S_n(f)} \quad (7\text{–}12)$$

$$= = |F_+ h_{0,+} + F_\times h_{0,\times}^2|^2 \frac{T}{S_n(f_0)} \quad (7\text{–}13)$$

We see that for CW signals, the SNR increases with the square root of the observation time. Here, we remind that $h_{0,+}$ and $h_{0,\times}$ are functions of just f_0 and time, so they can be pulled out of the integral in Equation 7–13. The beam pattern functions are also narrowly peaked around the sidereal rotational frequency of the earth, and can therefore be pulled out of the integral as well.

Matched filtering is very computationally expensive because one needs to try many combinations of parameters to obtain the best match for a potential signal. Moreover, it relies on the fact that one can fully trust the model. Unlike in BBH and BNS systems, where the phase evolution is known, the physics governing the GW emission of isolated neutron stars is unknown.

7.3 Time Delays

The GW that arrives at our detectors has been altered by a variety of general relativistic factors that must be accounted for. In CW data analysis, what we wish to do is transform the data we receive on earth (a non-inertial reference frame with many time delays) to the solar system barycenter (SSB) frame, which is assumed inertial relative to the source of GWs.

We can express the overall time delay in terms of three separate delays (*55*):

$$t_b = t + \Delta_R + \Delta_S + \Delta_E, \quad (7\text{–}14)$$

where t_b is the time in the SSB frame, t is the time we measure at the detector, Δ_R is the Roemer delay, Δ_E is the Einstein delay, and Δ_S is the Shapiro delay. We explain these delays in the following subsections.

7.3.1 Roemer Delay

The Roemer delay (*74*) is a correction that at first order results because the earth/sun distance changes, which affects our transformation from t to t_b (*55*):

$$\Delta_R = \frac{\mathbf{r_{os}} \cdot \hat{\mathbf{n}}}{c}, \tag{7–15}$$

where $\mathbf{r_{os}} = \mathbf{r_{oe}} + \mathbf{r_{es}} + \mathbf{r_{sb}}$ is the changing observer-SSB distance due to changes in distance between observer/earth due to earth's rotation about its axis ($\mathbf{r_{oe}}$), between earth/sun due to earth's revolution around the sun ($\mathbf{r_{es}}$), and between sun/SSB due to the sun's revolution around the SSB ($\mathbf{r_{sb}}$). $\hat{\mathbf{n}}$ is a unit vector pointing towards the source from the SSB.

7.3.2 Shapiro Delay

The Shapiro delay (*75*) refers to the fact that the gravitational field of the solar system causes light or GWs to travel along its geodesics, affecting the arrival times (*55*). This is a weak-field effect, so we can start with the Newtonian metric:

$$ds^2 = -[1 - 2\phi(\vec{x})]c^2dt^2 + (1 - 2\phi(\vec{x}))d\vec{x}^2 \tag{7–16}$$

where $\phi = -GM/(rc^2)$ is the Newtonian potential, and we assume that GWs propagate at the speed of light, so $ds^2 = 0$. Keeping only terms $O(\phi)$, we obtain:

$$\begin{aligned} cdt &= \pm\sqrt{\frac{1 - 2\phi(\vec{x})}{1 + 2\phi(\vec{x})}} && \text{(7–17)} \\ &\approx \pm\frac{1 - \phi(\vec{x})}{1 + \phi(\vec{x})} \times \left(\frac{1 - \phi(\vec{x})}{1 - \phi(\vec{x})}\right) && \text{(7–18)} \\ cdt &\approx \pm[1 - 2\phi(\vec{x})] && \text{(7–19)} \end{aligned}$$

Let us define t_e as the time of emission, and t_{obs} as the time we observe the GWs, and $\vec{r}_p$ as the position of the pulsar and $\vec{r}_{obs}$ as the position of the observer at t_{obs}:

$$c(t_{obs} - t_e) = \int_{\vec{r}_{obs}}^{\vec{r}_p} (1 - 2\phi(\vec{x}))d\vec{x} \tag{7–20}$$

$$= |\vec{r}_p - \vec{r}_{obs}| - 2\int_{\vec{r}_{obs}}^{\vec{r}_p} \phi(\vec{x})d\vec{x} \tag{7–21}$$

$$= |\vec{r}_p - \vec{r}_b| + (\vec{r}_b - \vec{r}_{obs}) \cdot \vec{n} - 2\int_{\vec{r}_{obs}}^{\vec{r}_p} \phi(\vec{x})d\vec{x} \tag{7–22}$$

We have made a clever transformation of the coordinate distance, which is valid because the distance between the SSB and the pulsar is much greater than the distance between the barycenter and observer. The only reason for this transformation is to express the delay in terms of the Roemer delay and identify the Shapiro delay piece:

$$t_{obs} \approx \left(t_e + \frac{1}{c}|\vec{r}_p - \vec{r}_b|\right) + \frac{1}{c}(\vec{r}_b - \vec{r}_{obs}) \cdot \vec{n} - \frac{2}{c}\int_{\vec{r}_{obs}}^{\vec{r}_p} \phi(\vec{x})d\vec{x} \tag{7–23}$$

We can identify the piece in (...) as t_{SSB}, the time that the signal would arrive if there were no general relativistic effects. The second term represents the Roemer delay, and the third term is the Shapiro delay, which depends on the mass and distance/geometry to the source:

$$\Delta_S = -\frac{2}{c}\int_{\vec{r}_{obs}}^{\vec{r}_p} \phi(\vec{x})d\vec{x} \tag{7–24}$$

The biggest contribution to Δ_S is the sun, so we can write:

$$\Delta_S = -\frac{2GM}{c^3}\int_0^D \frac{d\rho}{r} \tag{7–25}$$

where D is the distance between the sun of mass M and the source, and we are integrating over the earth/source distance ρ. This is shown schematically in Figure 7-1. By the Pythagorean theorem, we can write:

$$r^2 = (d_{ES} + \rho\cos\theta)^2 + \rho\sin\theta^2 \tag{7–26}$$

where θ is the angle between earth and the source, respectively, and d_{ES} is the earth/sun distance, on average 1 AU.

By combining equations 7–25 and 7–26, we find the Shapiro delay to be:

$$\Delta_S = \frac{2GM}{c^3}\left[\ln\left(\frac{2D}{d_{es}}\right) - \ln(1+\cos\theta)\right] > 0 \tag{7–27}$$

The first term is a constant offset in the arrival time of the signal, dependent on the earth and the sun distances/mass, and the distance of the sun from the source. The second term is only based on the relative source/earth location. The Shapiro offset is always a delay, hence why $\Delta_S > 0$

7.3.3 Einstein Delay

The Einstein delay (*76*) is caused by the fact that the clock we use on earth to measure the arrival time of the GWs measures its own proper time, and needs to be converted to the coordinate time of the SSB. Starting from the same metric Equation 7–16, writing $ds^2 = c^2 d\tau^2$, and dividing everything by $c^2 dt^2$, we see:

$$\frac{d\tau}{dt} = \left[(1+2\phi) - (1-2\phi)\frac{v^2}{c^2}\right]^{1/2} \tag{7–28}$$

where $\phi = \phi(\vec{x}_{obs})$ and $v = d\vec{x}_{obs}/dt$ is the velocity of the earth around the sun (and its rotational velocity, but this is a small correction). Taylor expanding and keeping terms $O(v)$ and $O(\phi)$:

$$\frac{d\tau}{dt} = 1 + \phi - \frac{v^2}{2c^2} \tag{7–29}$$

ϕ is the gravitational redshift; the $v^2/(2c^2)$ term is the transverse Doppler shift. Integrating and rewriting so that we express everything in the coordinate time:

$$t \approx \tau + \int^t \left(\frac{v^2}{2c^2} - \phi\right) dt \tag{7–30}$$

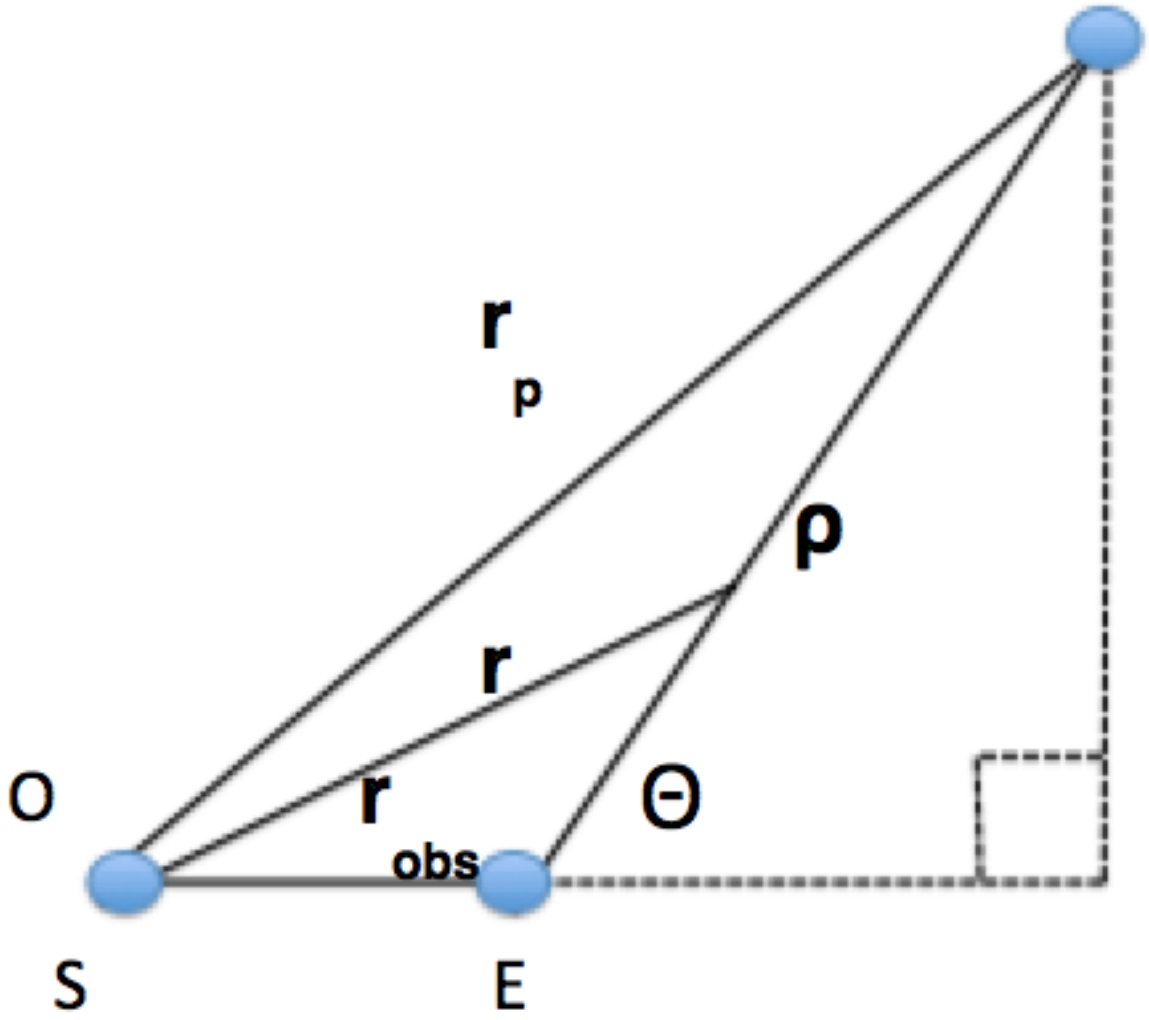


Figure 7-1: Diagram showing the geometry of the source relative to the earth and sun, used in our calculation of the Shapiro delay. All bold quantities in this figure refer to vectors.

where both ϕ and v are time-dependent, and the second term is the Einstein delay:

$$\Delta_E = \int^t \left(\frac{v^2}{2c^2} - \phi \right) dt \tag{7–31}$$

CHAPTER 8
THE ORIGINAL FREQUENCY HOUGH

8.1 Introduction

The FH is an example of a hierarchical method, and is shown in Figure 8-1. Hierarchical methods reduce computational time in exchange for a small loss in sensitivity, which can be made up for by analyzing data for a long time ($O(years)$). The data are divided into small time segments which are analyzed coherently, then combined incoherently. "Coherent" analysis refers to performing FFTs on individual data segments and retaining the signal's phase information: FFTs transform the strain $h(t)$ to the frequency plane of the detector, so that a periodic signal would show up as a spike in the frequency plane. Each FFT is actually interlaced by 50% to reduce the known effect of truncating an infinite time series. The data are smoothed by applying a flat-cosine window function.

We discuss how we create a database of FFTs of different lengths in Section 8.2. In Section 8.3, we describe the creation of thresholded time-frequency maps, which are the input to our method, the FH Transform, examined in detail in Section 8.4. Finally we give a theoretical estimate of the FH's sensitivity in Section 8.5.

8.2 Construction of Short Fast Fourier Transform Databases

The first step in the method involves constructing a so-called short Fast Fourier Transform Database (SFDB), which is essentially the raw strain data Fourier transformed with one of four T_{FFT} that cover slightly different frequency ranges. There are currently 4 databases: one up to 128 Hz with $T_{FFT} = 8192$ s, one up to 512 Hz with $T_{FFT} = 4096$ s, one up to 1024 Hz with $T_{FFT} = 2048s$ and finally one up to 2048 Hz with $T_{FFT} = 1024$ s. The databases are constructed in this way to confine the frequency evolution due to earth's rotation to one frequency bin. We can estimate the maximum T_{FFT} that can be used as a function of the frequency with the following logic: we know the Doppler shift equation is:

$$f(t) = f_0 \left(1 + \frac{\vec{v} \cdot \hat{n}}{c}\right) \tag{8–1}$$

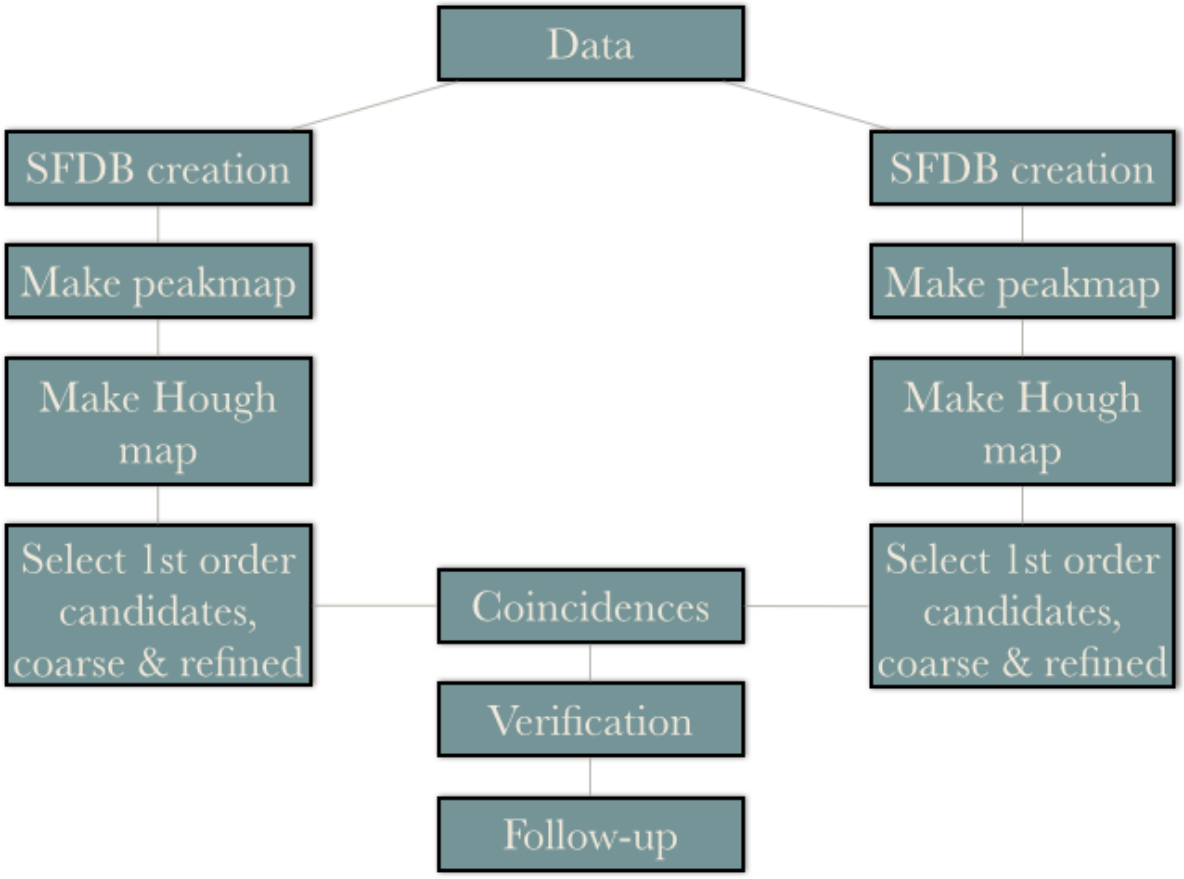


Figure 8-1: Scheme of the Frequency Hough hierarchical method.

It is the change of the velocity of the detector with respect to the source due to earth's rotation and orbit around the sun that leads us to consider the maximum T_{FFT} we can use. We imagine that in time Δt, the velocity changes by Δv. Hence the change in frequency is:

$$\Delta f = f_0 \frac{\Delta v}{c} \tag{8–2}$$

If the earth turns an angle $\Delta\theta$ in the integration time T, then we have $\Delta\theta = \Omega T$. We need now to relate this to Δv, so we can imagine that the earth is moving at a certain speed $v_1 = R\Omega\cos\theta$ and that a nearby part of the earth is moving at a speed $v_2 = R\Omega\cos(\theta + \Delta\theta)$, where R is the average radius of the earth. Then:

$$v_2 - v_1 = \Delta v = R\Omega(\cos(\theta + \Delta\theta) - \cos\theta) \tag{8–3}$$

Assuming $T \ll 86400$ s (one day) so that $\Delta\theta \ll 1$, we can expand the first term with a trigonometric identity. Neglecting $O(\Delta\theta^2)$ terms, we arrive at:

$$\Delta v = R_E\Omega\Delta\theta = R_E\Omega^2 T \tag{8–4}$$

where R_E is the distance to the rotation axis at the latitude θ (where the detector is at the initial time). Writing $\Omega = 2\pi/T_E$, we arrive at the maximum coherence time T_{max}:

$$T_{max} = \frac{T_E}{2\pi}\sqrt{\frac{c}{R_E f_0}} \approx 1.1 \times 10^4 \sqrt{\frac{100\ \text{Hz}}{f_0}} \tag{8–5}$$

where T_{max} is the maximum time one can analyze without having to do a Doppler (sky) correction and T_E is the rotational period of Earth. We can see that lower frequencies permit longer T_{FFT}, and vice-versa.

We can repeat this order-of-magnitude analysis for the orbital motion of the earth around the sun, and show that its effect is actually less important than that of the rotation around the earth.

For each FFT in the SFDB, the average noise is estimated based on an auto-regressive mean that tracks the variation in the noise without being sensitive to peaks in the spectrum. Here, auto-regressive means that the estimation for the noise in the spectrum is based not just on the frequency at which the estimate is done, but also the frequencies nearby. In this way, large "glitches" can be taken out of the data, as they will be literally seen as spikes. But a (transient) continuous gravitational wave signal will be weak, probably not so high above the noise, so the contribution due to a signal will not be included in the spectrum estimation. See (*77*) for more details.

A visual representation of the calculation of the auto-regressive spectrum and selection of peaks is shown in Figure 8-2.

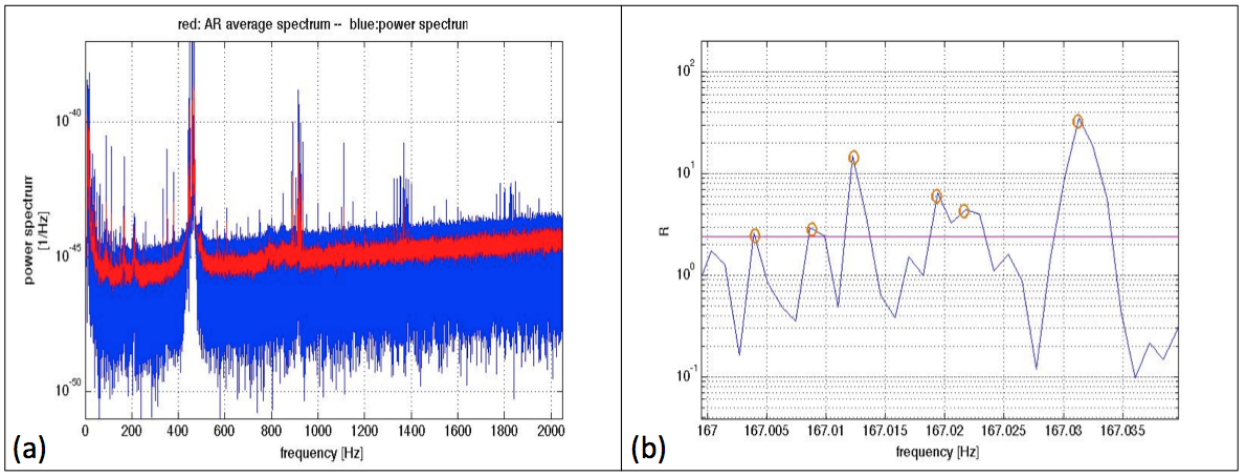


Figure 8-2: (a) The auto-regressive and power spectrum are shown and (b) This is a zoomed in version of a peakmap showing the selection of just local maxima in time and frequency.

8.3 Peakmap Creation

We now would like to create a time/frequency representation, but one that is useful for our adaptation of the Hough transform. To do this, we take each FFT in the database and divide it by an estimation of the auto-regressive spectrum. This is called "whitening" the data. At this point, we have essentially normalized power, and we apply a threshold of 2.5 to exclude likely sources of noise (*77*).

In more details, we calculate the power spectral density S_i (the periodogram) for each FFT:

$$S_i = \mid FFT \mid^2 \tag{8–6}$$

$$FFT = \frac{A}{\sigma} \tag{8–7}$$

A is the power at a particular frequency and σ is the standard deviation of the noise.

Then the ratio between the power spectral density and the auto-regressive spectrum S_i^{AR} is calculated:

$$R_i(j) = \frac{S_i}{S_i^{AR}} \tag{8–8}$$

This ratio is computed across all frequency bins j for the i^{th} FFT, and is usually around 1 unless a peak at some frequency in the spectrum is present (see Figure 8-2 for a visual representation of this). We compare the ratios to a threshold value defined based on the number of total candidates we want, therefore the threshold must be chosen carefully as it affects the sensitivity of the search. A peak occurs is defined by a particular FFT's start time i and frequency bin j. The amplitude of the peak does not factor into this analysis, however we only select peaks that are local maxima: this reduces the computational time and decreases sensitivity to glitches (huge spikes in the spectra due to noise) in exchange for a small loss in sensitivity.

We create the Hough map from the peakmap; therefore, the selection of peaks is a critical step in our analysis. Even if a peak is due to noise and is above the threshold, it will appear in the peakmap, and if a signal does not have a peak in the peakmap, it is lost. We can write the probability that a peak is due to noise as $p_0 = P(\theta; 0)$ and $p_\lambda = P(\theta; \lambda)$, if a peak is due to signal, where θ is the threshold and λ is the spectral amplitude of the signal, defined as:

$$\lambda = \frac{4|\tilde{h}(f)|^2}{T_{FFT} S_n}, \tag{8–9}$$

where $h(f) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} dt$ is the Fourier transform of $h(t)$, the GW signal, and S_n is the spectral noise density, which changes as a function of frequency. In Gaussian noise, the probability distribution of the peakmap is exponential, so we can calculate p_0. The probability of having a peak in a frequency bin between two values in the spectrum, x and $x + dx$ is: $p(x) = e^{-x} dx$. The probability that this peak is a local maximum, that is, the two values nearby are not peaks, is $(1 - p(x))^2$. Thus,

$$p_0 = \int_\theta^\infty e^{-x}(1 - e^{-x})^2 dx = e^{-\theta} - e^{-2\theta} + \frac{1}{3} e^{-3\theta} \tag{8–10}$$

The probability of finding a peak due to a signal p_λ is calculated based on knowing that the probability density of the spectrum is a non-central χ^2 distribution with 2 degrees of freedom and non-centrality parameter λ. This probability density $p(x;\lambda)$ can be written as:

$$p(x;\lambda) = e^{(-x\lambda/2)} I_0(\sqrt{2x\lambda}) \tag{8–11}$$

where I_0 is the modified zeroth order Bessel function whose mean value $\bar{x} = 1 + \lambda/2$. Then, applying the small signal approximation, $\lambda \ll x$, we can expand I_0 and the exponential term to get a simpler expression for $p(x;\lambda)$:

$$p(x;\lambda) \approx e^{-x}\left(1 - \frac{\lambda}{2} + \frac{\lambda}{2}x\right) \tag{8–12}$$

Again we can compute the probability of selecting a local maximum:

$$p_\lambda = \int_\theta^\infty p(x;\lambda) \int_0^x p(x';\lambda) dx' dx \tag{8–13}$$

Doing the integration, and keeping terms up to $O(\lambda)$, we arrive at:

$$p_\lambda = p_0 + \frac{\lambda}{2}\theta(e^{-\theta} - 2e^{-2\theta} + e^{-3\theta}) \tag{8–14}$$

Knowing these two probabilities, we can discuss now how we determine the threshold used in the peakmap.

The threshold is chosen based on the maximization of the average value of our detection statistic. The FH uses a simple detection statistic:

$$CR = \frac{y - \bar{z}}{\sigma} = \frac{y - Np_0}{\sqrt{Np_0(1-p_0)}}. \tag{8–15}$$

The Hough map is binomial because each peak is either a noise peak or a signal peak, so we can use the expressions for the mean and standard deviation of a binomial distribution. y is the number of peaks at in the Hough map, $\bar{z}$ is the average number count of the noise, and σ is the standard deviation of the noise. It is upon maximizing the expectation value of the critical

ratio that we can arrive at determining an adequate threshold. Assuming a signal is present, we can write that the average value of the number count will be $y = Np_\lambda$. Thus (78):

$$CR = \frac{N(p_\lambda - p_0)}{\sqrt{Np_0(1-p_0)}} = \sqrt{N}\psi(\theta;\lambda). \tag{8–16}$$

By plotting $\psi(\theta;\lambda)$ to see at which value ψ is maximal, we can determine the appropriate threshold. Actually, plotting this function shows that the ideal threshold is $\theta = 2$, but we use $\theta = 2.5$ to reduce computational time at a small cost in sensitivity.

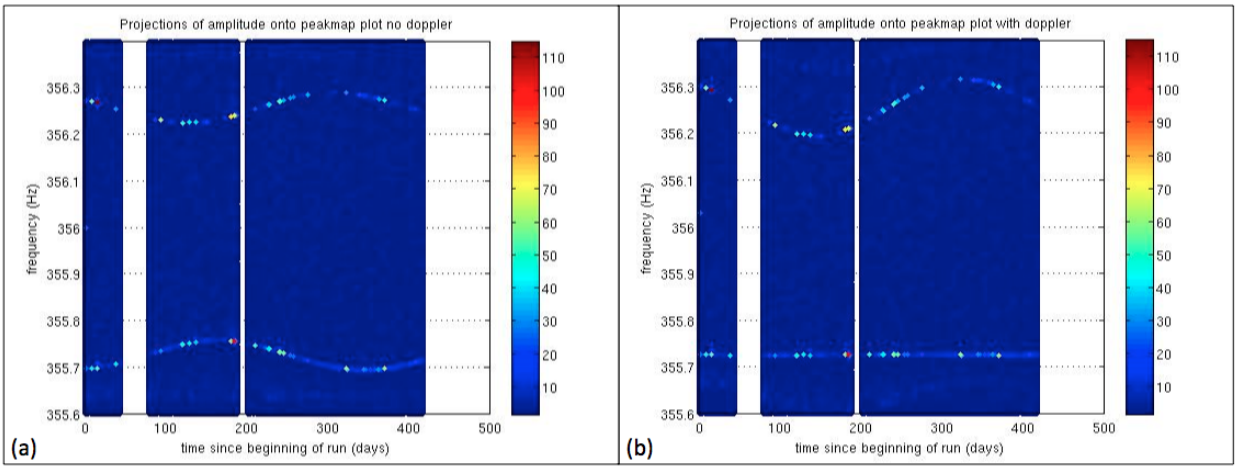


Figure 8-3: (a) Pre-Doppler corrected peakmap for a signal at $f = 355.725$ Hz. The signal is visible at a sinusoidal curve, and a glitch is present at a higher frequency. (b) Doppler corrected peakmap. Signal is visible as horizontal line, glitch has had its power spread in multiple frequency bins.

8.4 The Frequency Hough Transform

The FH is a clever implementation of the Hough transform, a pattern recognition technique originally used to find tracks in particle physics experiments. This method maps the points (the peaks) in the time/frequency domain of the detector to the frequency/spindown domain of the source. This method requires that the signal's time/frequency evolution be linear, so a Taylor series expansion is used as shown in 3–24, where the second order spindown $\ddot{f}$ is typically neglected. The method works as follows: for each position in the sky, the peakmap is corrected for the Doppler shift. We see how powerful this Doppler correction can

be to distinguish a glitch and signal in Figure 8-3 Then, for each peak in the time/frequency plane, a line is drawn for different choices of $\dot{f}$ at different times t:

$$\dot{f} = -\frac{f_0}{(t - t_0)} + \frac{f}{(t - t_0)}. \tag{8–17}$$

From this information, the values of f_0 (the y-intercept) are calculated, and then histogrammed in the $f_0/\dot{f}$ plane- see Figure 8-4. A candidate signal will have a large number of peaks that are traced back to a particular f_0.

In practice, the FH employs an adaptive procedure that weights less noisy areas of the detector and accounts for the differing detector sensitivity across the sky (*78*).

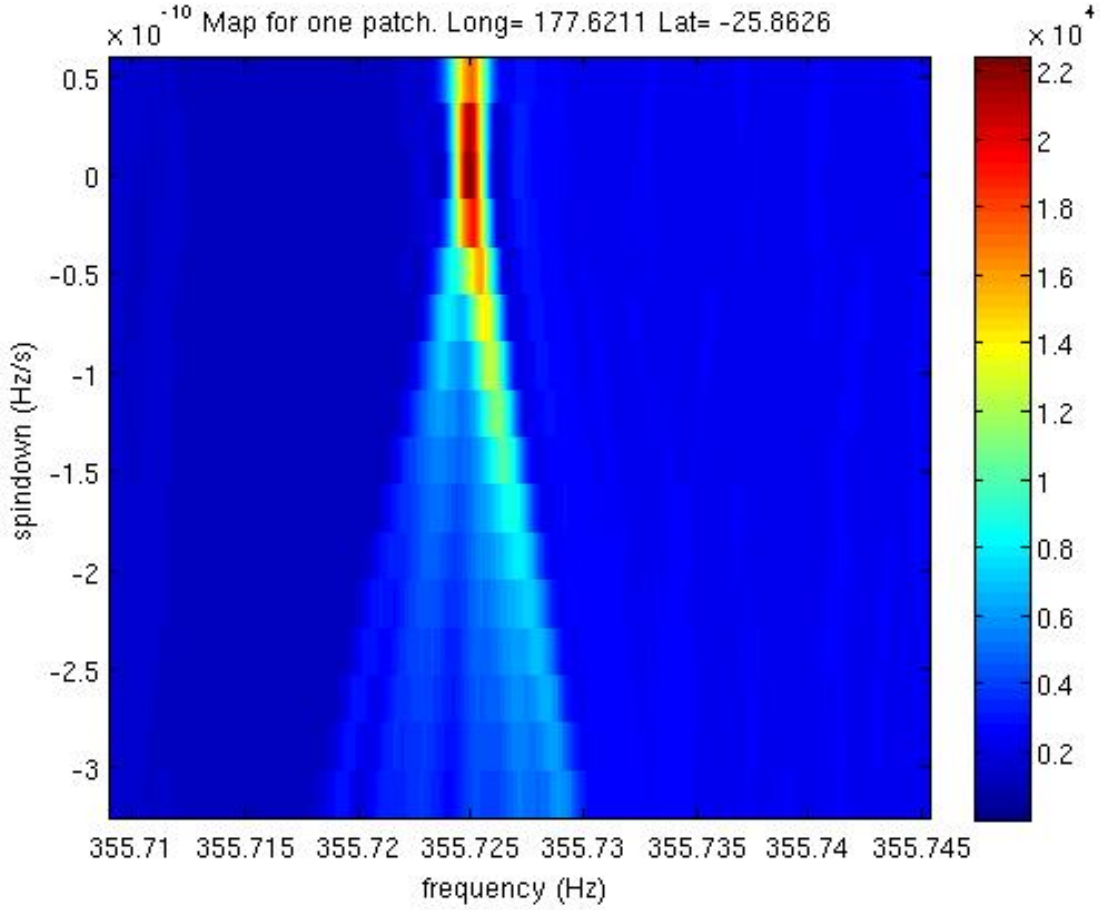


Figure 8-4: This is a zoomed in version of a Hough map created for a signal at $f = 355.725Hz$, $\dot{f} = -2.69 \times 10^{-18}$, longitude $= 178.0718°$, and latitude $= -26.1775°$. The color bar represents the number of FFTs that captured a peak at a particular frequency/spin-down.

8.5 Theoretical Sensitivity Estimate

Sensitivity is defined as the minimum signal amplitude that a method can detect at a certain confidence level (typically, 50%, 90% or 95% is used) in many experiments (usually the percentage of 100 signals at a given frequency/amplitude is sufficient). It can be calculated independently of any search; hence, it provides a good benchmark when computing upper limits in the real search-one's sensitivity cannot be better than the theoretical estimation!

The crucial detail is that the probability distribution of the Hough map, though binomial, is assumed to be Gaussian:

$$P_n(\theta_{thr};\lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\mu)^2}{2\sigma^2}} \tag{8–18}$$

This assumption works well if the number of FFTs is large, and the threshold on the peakmap is such that the probability of selecting a peak is not too close to 0 or 1 (see equations 8–10 and 8–14. Moreover, it is clear that this assumption can break down in frequency bands that have many disturbances, which occur frequently in the detectors. The probability distribution of the Hough map is therefore not completely Gaussian.

Integrating this probability distribution for signals that have surpassed our threshold number count n_{thr} (which can be related to the critical ratio by Equation 8–16):

$$P_{n>n_{thr}}(\lambda) = \int_{n_{thr}}^{\infty} P_n(\theta_{thr};\lambda)dn = \Gamma \tag{8–19}$$

where Γ is the confidence level at which one claims a detection. Then:

$$\operatorname{erfc}\left(\frac{n_{thr} - Np_\lambda}{\sqrt{2Np_\lambda(1-p_\lambda)}}\right) = 2\Gamma \tag{8–20}$$

Now, we must also include the false alarm probability into our analysis. We perform a similar integral, expect that $\lambda = 0$ so the probability distributions we use are p_0:

$$\int_{n_{thr}}^{\infty} P_n(\theta_{thr};0)dn = P^*_{fa} \tag{8–21}$$

therefore:

$$n_{thr}(N, \theta_{thr}, P^*_{fa}) = Np_0 + \sqrt{2Np_0(1-p_0)}\mathrm{erfc}^{-1}(2P^*_{fa}) \tag{8–22}$$

Now, we can write $P^*_{fa} = N_{cand}/N$, since we only select the candidates with the highest number count, but it turns out that P_{fa} is so small that the term that is linked to it can be neglected, since it is compared to a term with p_0, the probability that there are noise peaks.

We combine equations 8–20 and 8–22, along with the expressions for p_0 and p_1, equations 8–10 and 8–14, in the small signal approximation. Essentially, the equation can be put into a form in which it is solvable as a quadratic equation, and then of the two possible solutions, the one that gives a physically meaningful interpretation is chosen. In this case, the solution is the one in which λ increases as $\Gamma > 0.5$, since we expect stronger signals to be more easily and consistently detectable. We arrive at:

$$\lambda_{min} \approx \frac{2}{\theta_{thr}}\sqrt{\frac{p_0(1-p_0)}{Np_1^2}\left(CR_{thr} - \sqrt{2}\mathrm{erfc}^{-1}2\Gamma\right)} \tag{8–23}$$

where $p_1 = e^{-\theta_{thr}} - 2e^{-2\theta_{thr}} + e^{-3\theta_{thr}}$.

Now, we must relate this quantity to $h_{0,min}$, the minimum signal amplitude that we can detect. In order to do this, we first write an expression for the GW signal:

$$\begin{aligned} h(t) &= F_+ h_{0+}\cos\left(\phi(t)\right) + F_\times h_{0\times}\sin\left(\phi(t)\right) \\ h_{0+} &= h_0\frac{1+\cos^2\iota}{2} \\ h_{0\times} &= h_0\cos\iota \end{aligned} \tag{8–24}$$

where $\phi(t) = 2\pi f_0 t$ (since the spindown is so small and is confined to once frequency, we can neglect the spindown and assume a constant frequency signal).

Then, we perform the exact Fourier transform of a sinusoidal signal, and account for the fact that when we take the Fourier transform with a finite length (as we do in practice), we should apply a window function. We obtain:

$$\tilde{h}(f) \approx T_{FFT}\frac{(F_+A_+ - iF_\times A_\times)}{2}\frac{\sin(\pi(f-f_k)T_{FFT})}{\pi(f-f_k)T_{FFT}} \tag{8–25}$$

where A_+ and $A_\times$ are the detector response functions that depend on the angle between the star rotation axis and the line of sight, given by:

$$\begin{aligned} A_+ =& a_0 + a_{1c}\cos\Omega_\oplus t + a_{1s}\sin\Omega_\oplus t + a_{2c}\cos 2\Omega_\oplus t + \\ & a_{2s}\sin 2\Omega_\oplus t \end{aligned} \tag{8–26}$$

$$\begin{aligned} A_\times =& b_{1c}\cos\Omega_\oplus t + b_{1s}\sin\Omega_\oplus t + b_{2c}\cos 2\Omega_\oplus t + \\ & b_{2s}\sin 2\Omega_\oplus t \end{aligned} \tag{8–27}$$

where $\Omega_\oplus$ is the sidereal angular frequency of the earth, whose coefficients a_0, a_{1c}, etc. depend on the relative source/detector positions, and are given in (*79*).

Since we want to know our sensitivity at any sky location, frequency, etc., we take the average over the two beam pattern functions and the two amplitude functions. This involves averaging over the sky position, the inclination angle, the wave polarization angle, and the frequency of the signal. After performing these averages, relying heavily on equations B15-B19 in (*78*) we arrive at:

$$h_{0,min} \approx \frac{4.02}{N^{1/4}\theta_{thr}^{1/2}}\sqrt{\frac{S_n(f)}{T_{FFT}}}\left(\frac{p_0(1-p_0)}{p_1}\right)^{1/4}\sqrt{CR_{thr} - \sqrt{2}\mathrm{erfc}^{-1}(2\Gamma)} \tag{8–28}$$

This is the minimum amplitude at which the FrequencyHough can detect a signal under the assumption of Gaussian noise.

CHAPTER 9
THE GENERALIZED FREQUENCY HOUGH

9.1 Introduction

In order to search for tCWs (described in 3.5), we had to generalize the FH to be able to see power-law signals, or develop an entirely new method. In the end we were able to adapt the existing FH Transform to search for a different type of signal. This chapter describes the mechanics of this new method, as well as the different problems that we encountered when using the Generalized Frequency Hough (GFH).

The main reason that this method works is shown in Section 9.2. We define how we can claim a detection of a GW, i.e. what our detection statistic is, in Section 9.3. Afterwards in Section 9.4 we discuss the general procedure to follow up interesting candidates that we may receive in a search. The construction of the grids in the parameter space that we would use to run a search, expressed in terms of braking indices, spindowns and initial frequencies, is demonstrated in Section 9.5. We then estimate the theoretical and empirical sensitivity of our method in Sections 9.6 and 9.7 respectively, ensuring that they are in reasonable agreement.

9.2 The Coordinate Transformation

The FH works well if the signal's frequency evolution is linear; however, when we consider young neutron stars formed after a merger or in a supernova, the frequency evolution is highly nonlinear. We assume a power-law model for the frequency evolution of a GW signal as in 3–33. However, this equation is not linear and therefore we cannot apply the FH. But if we change coordinates (*31*):

$$x = \frac{1}{f^{n-1}}; x_0 = \frac{1}{f_0^{n-1}} \tag{9–1}$$

Under this transformation, Equation 3–33 becomes the equation of a line:

$$x = x_0 + (n-1)k(t - t_0) \tag{9–2}$$

Points in the $(t - t_0,\, x)$ plane are mapped to lines in the $(x_0,\, k)$ plane:

$$k = -\frac{x_0}{(n-1)(t-t_0)} + \frac{x}{(n-1)(t-t_0)} \tag{9–3}$$

We now have an analogous equation to equations 3–24.

We show in the left panel of Figure 9-1 a peakmap created with an injected r-mode signal. In the right panel of Figure 9-1, we see a peakmap in the transformed space. It is this (Doppler-corrected) map that is the input to the GFH. To create these r-mode signals, we simply simulate their expected phase evolution basd on equation 3–33, because performing numerical simulations and identifying the fluid motions as r-modes is very difficult (*80*).

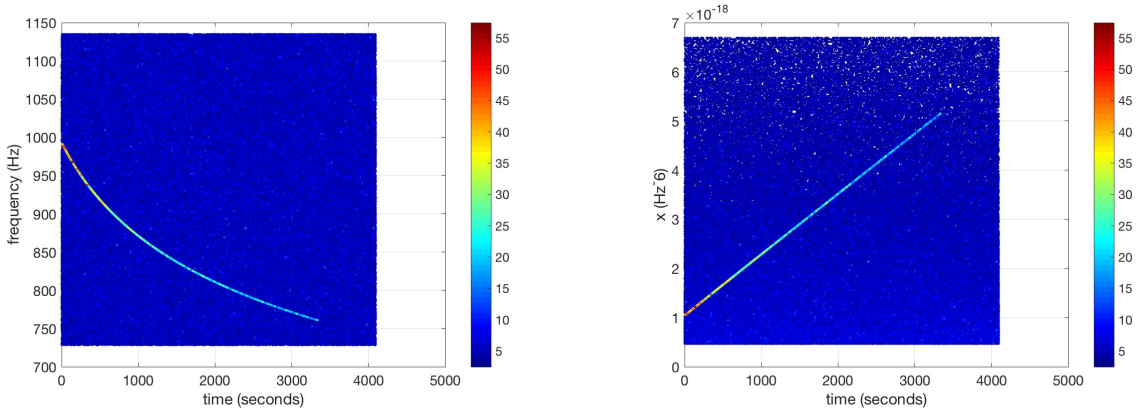


Figure 9-1: A peakmap of an injected r-mode ($n = 7$) signal in white noise with $h_0 = 1.5 \times 10^{-22}$, $f_0 = 993.25$ Hz, $\dot{f}_0 = -1.99 \times 10^{-1}$ Hz/s before and after the change of coordinates. The color represents normalized power.

Notice that the peaks are denser near the low values of x (which correspond to higher frequencies). This effect is also seen in the FH map; it occurs because in this new space, the noise is not uniform. The change of coordinates takes points that are equally spaced in frequency and concentrates them at higher frequencies (lower x values) and spreads them out at lower frequencies (high x values). So when the GFH is performed, more points in t/x are mapped to lower x_0 values. The non-uniformity of the noise can affect the sensitivity; therefore, we construct peakmaps in a larger frequency band than we plan to actually analyze (20-25$\%$ larger). This is by far the simplest way to deal with the non-linearity of the noise in this new space at the cost of a very small increase in computational cost.

In Figure 9-2, we show the output of the GFH: a histogram in x_0/k space. The number count corresponds to the number of FFTs in which a signal with parameters x_0 and k was found. The lines intersect at the true parameters of the injected signal, and the highest number count refers to the most likely parameters of the signal.

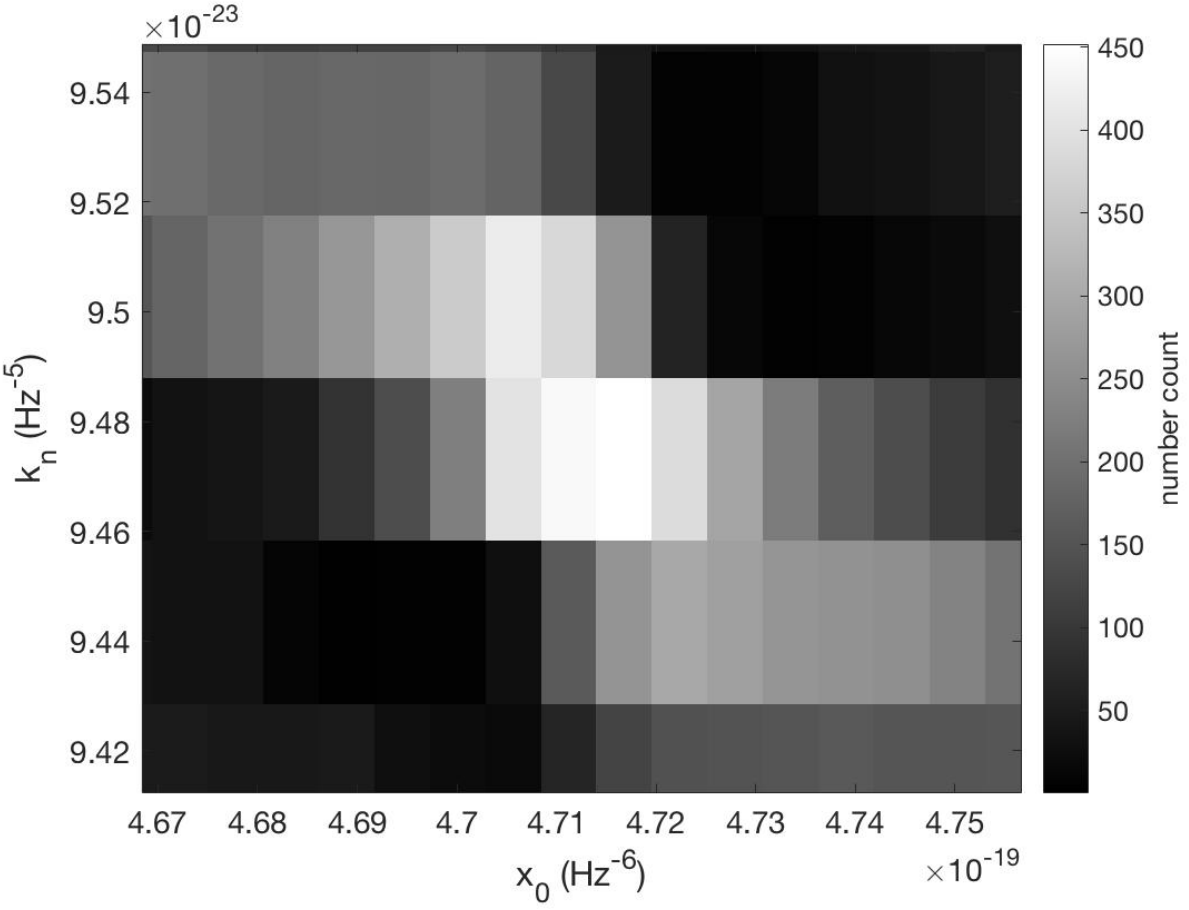


Figure 9-2: Hough map clearly showing the injected signal recovered, with the parameters given in the caption of Figure 9-1. Lower values of x_0 correspond to higher frequencies, and vice-versa. The color corresponds to the number of lines that hit each x_0/k bin. The reference time t_0 is taken to be at the beginning of the time segment analyzed.

9.3 Detection Statistic and Candidates

In both the original and generalized FHs, we evaluate a detection statistic called the critical ratio to determine if a candidate is significant:

$$CR = \frac{y - \mu}{\sigma} \tag{9–4}$$

where y is the number count in the Hough plane in a particular bin (a potential candidate), μ and σ are the average and standard deviation of the number counts in the map due to noise.

In each of the maps, we select significant candidates. Each candidate will have an x_0, k, n, longitude and latitude. From x_0 and k, we can find f_0 and $\dot{f}_0$ with the inverse transformation of Equation 9–1. Candidates are chosen in each "square" of the Hough map, whose size (in bins) is determined by the number of candidates we wish to select in each map. We employ this way of selecting candidates because we want to cover the parameter space as uniformly as possible, and so that we are not blinded by major disturbances, nor the non-uniformity of the noise. Because the noise is nonuniform, one will always see higher number counts at lower values of x_0, even in pure Gaussian noise. So we cannot simply select the candidates with the highest number counts; we must evaluate the critical ratio in each (almost independent) subsection of the Hough map.

Once the analysis has been performed on all detectors and a set of candidates has been produced form each one, we then perform coincidences, which simply find if candidates selected in each detector are close to each other in the parameter space. We define "close" in the parameter space as being within a certain number of x_0 and k bins. For the candidates in each detector, the distance between them in the parameter space is calculated in the following way, as was done in (*78*):

$$d = \sqrt{\left(\frac{x_{0,2} - x_{0,1}}{\delta x_0}\right)^2 + \left(\frac{k_2 - k_1}{\delta k}\right)^2} \tag{9–5}$$

where δx_0 is the size of the bins in x_0, and δk is the varying size of the bins in k (see 9.5). Note that in general we should also include error in the sky location, but for now we are only considering a search where the position of the source is fixed.

9.4 Follow-up

We now employ a follow-up procedure for significant candidates returned by the search. For each candidate, we have an estimation of its f_0, $\dot{f}_0$, longitude, latitude, and n. We can therefore reconstruct the frequency evolution of the signal $f(t)$, and more importantly the

phase evolution. We make a heterodyne correction $h'(t) = h(t)e^{j\phi_{sd}(t)}$, where the phase ϕ_{sd} is (*81*):

$$\phi_{sd}(t) = 2\pi \cdot \int_{t_0}^{T_{obs}} \frac{f_0}{\left(1 + k(n-1)f_0^{n-1}(t' - t_0)\right)^{\frac{1}{n-1}}} dt' + \phi_{t_0} \tag{9–6}$$

We are essentially removing all phase variation for a GW signal in the data; hence, even a rapidly varying signal will be monochromatic. In practice, however, this correction will not be perfect, so there will be a residual spin-up or a spindown.

This issue can be addressed in a follow-up search by increasing T_{FFT} by a factor of about 10, but is generally limited by the amount of follow-ups one can afford. Moreover, the frequency variations that CW searches can handle are orders of magnitude smaller than for tCW searches. Therefore, we have to quantify how much we can increase T_{FFT} as a function of how wrong we are in our correction. Additionally, because of 3–33, errors in our returned f_0 and $\dot{f}_0$ parameters will affect the new frequency evolution differently.

In general, we have found that if we would like to increase the T_{FFT} by a factor of 10, we would need to consider a grid around the candidate's parameters over-resolved by a factor of 10 in each parameter, determined by requiring that when using the new FFT length, a candidate would not spin out of one frequency bin during the observation time. This would mean that one candidate would require 100 separate follow-ups. See Figure 9-3.

To actually perform the follow-up, it is easiest to use the Band-Sampled Data (BSD) framework (*81*). The framework allows us to use any T_{FFT} to correct for the signal's phase evolution, and is very flexible and easy to use. BSDs allow for both frequency and time bands to be extracted when doing injections, cleaning, etc. They are constructed from SFDBs by taking the inverse FFT of the data and constructing an analytic representation of the signal.

An example: imagine that we run a search using the generalized FH and recover a candidate at $n = 7$ (r-mode) with $f_0 = 795$ Hz and $\dot{f}_0 = -0.25$ Hz/s using an initial $T_{FFT,0} = 2$ s. The signal would appear in the peakmap (before the Generalized FH is applied) as in the left-hand panel of Figure 9-4, which is an injection done within the BSD framework.

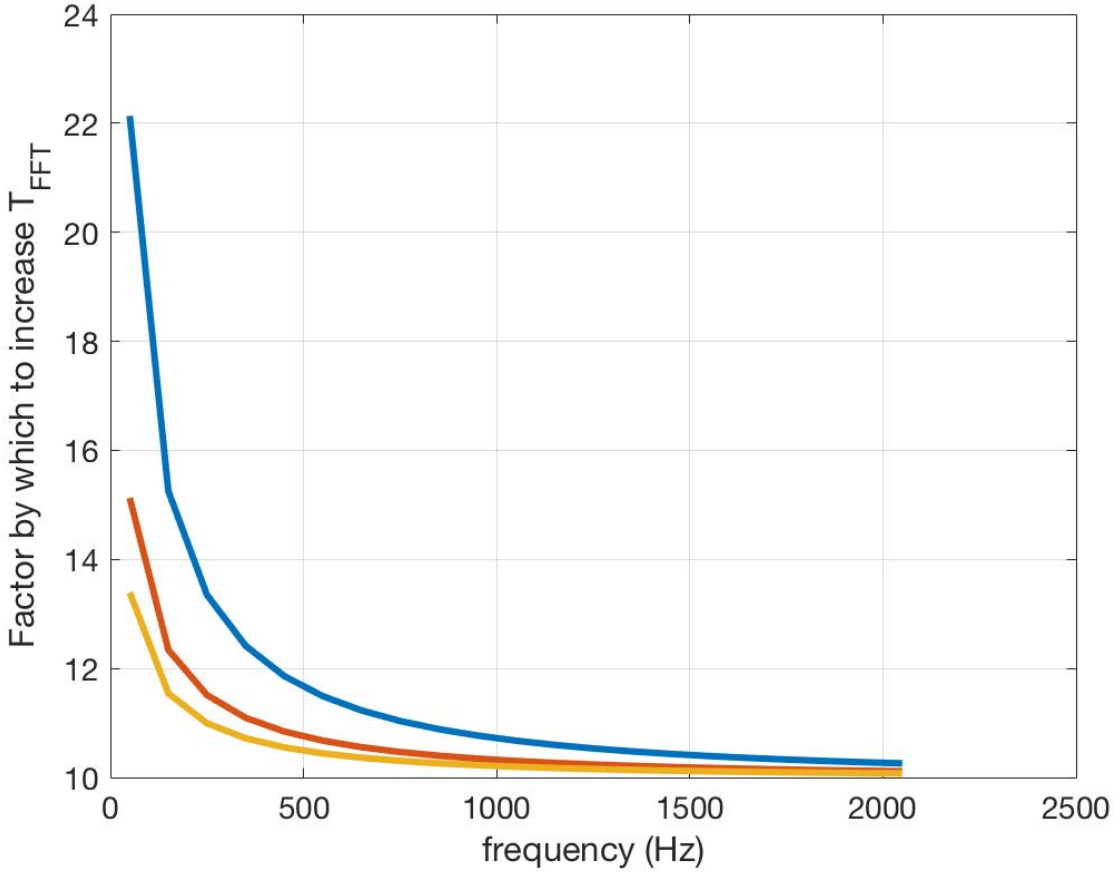


Figure 9-3: We chose the original $T_{FFT,0} = 2$ s, $\dot{f}_0 = 0.25$ Hz/s. We show on the y-axis the factor by which we can increase T_{FFT} during the follow-up as a function of frequency such that a signal would still be confined to one new frequency bin if we use an over-resolution factor of 10 in both x_0 and k. Blue, red, yellow $\rightarrow n = 3, 5, 7$ (also top, middle, bottom)

Knowing these parameters, we can then correct the signal's phase evolution and use longer FFTs, since the signal's residual spindown will be lower. In the right panel, by applying Equation 9–6, we obtain a monochromatic signal, where we have used $T_{FFT} = 2T_{FFT,0}$. After correcting for the phase, we then apply the original FH (*78*) because any residual spindown remaining (when the phase correction is not perfect) could be small enough such that the time/frequency behavior is linear. We have done tests by correcting for the wrong frequency, spindown and/or braking index, and have found that if the parameters are off by around one bin at the most, we can still recover the signal after coincidences for certain amplitudes.

After performing the original FH, we select candidates in this Hough map, where a candidate is defined by having a frequency, spindown and sky location, and do coincidences between the two detectors. If there is no coincidence, the candidate is vetoed; otherwise, we use the candidates' recovered parameters to correct the peakmap that was fed to the original FH (the right panel of Figure 9-4). This correction requires a simple shifting of bins, which is done for the Doppler correction. After this second correction, we project this peakmap onto the frequency axis and calculate a new critical ratio. The left panel of Figure 9-5 shows, for one detector, the projection of the peakmap created with $T_{FFT,0}$; the right panel shows the projection of the peakmap made with $T_{FFT} = 2T_{FFT,0}$, that has been created with the follow-up procedure.

This process is done for each detector individually. We then weight the critical ratio by each detector's sensitivity around the candidate frequency, and compute an average critical ratio. If this weighted average critical ratio is greater than that computed on the original $T_{FFT,0}$ peakmaps by a factor of $T_{FFT}^{1/4}$, then we have a candidate that requires more analysis; otherwise, the candidate is vetoed.

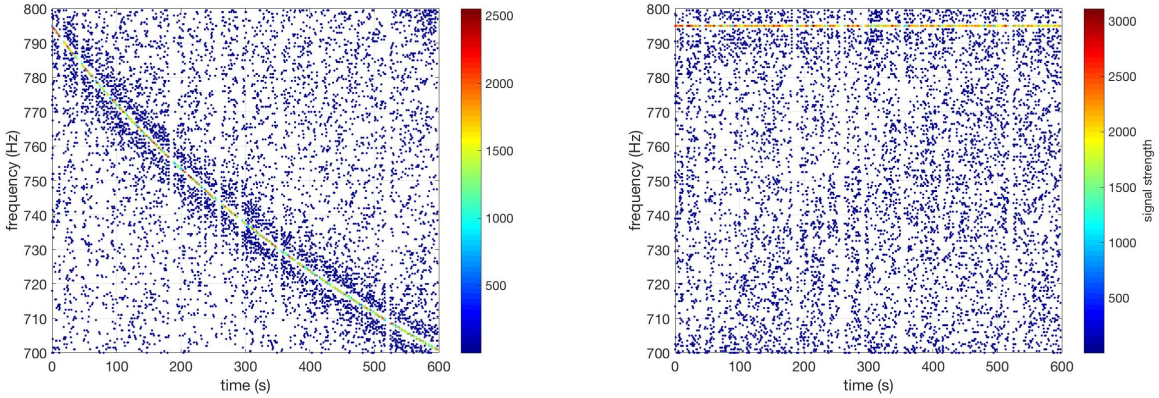


Figure 9-4: Left: peakmap of an injected r-mode ($n = 7$) signal in white noise with $h_0 = 1 \times 10^{-22}$, $f_0 = 795$ Hz, $\dot{f}_0 = -2.5 \times 10^{-1}$ Hz/s. Right: peakmap after phase evolution correction; a monochromatic signal is seen at $f_0 = 795$ Hz. The peakmap on the left was constructed with $T_{FFT} = 2$ s; on the right, as is the case of the follow-up, we used $T_{FFT} = 4$ s, twice as large as in the original analysis. The signal was injected and its phase corrected using the BSD framework.

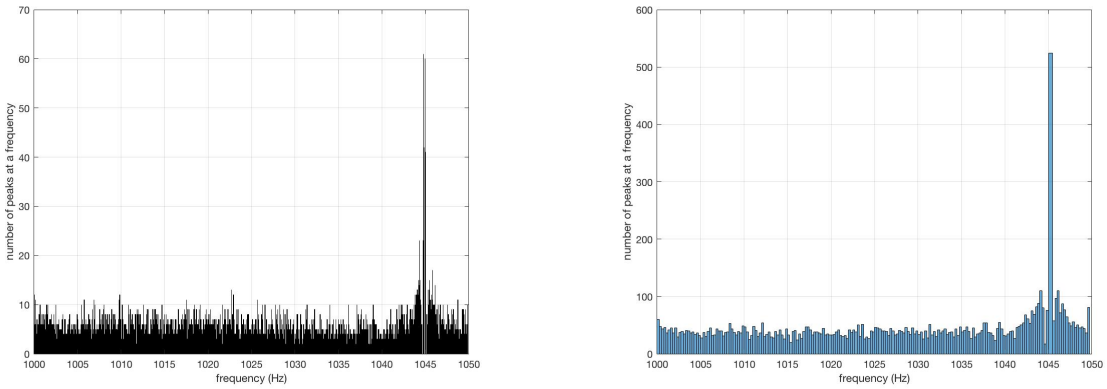


Figure 9-5: A signal was simulated with $f_0 = 1045$ Hz,$\dot{f}_0 = -0.03125$ Hz/s, $n = 7$, and $h_0 = 8 \times 10^{-23}$ in white noise. Left: we performed the Generalized FH, obtained candidates, and corrected the phase evolution in original peakmap $T_{FFT,0} = 4$ s, then projected that peakmap onto the frequency axis. Right: we then corrected for the phase evolution of this signal while increasing the T_{FFT} by a factor of 2, ran the original FH, and used the best candidate's parameters to shift the bins in the peakmap as we do for the Doppler shift. In one detector, the critical ratio is increased by about 20%; when using both detectors, the weighted average critical ratio is increased by about 22%, which roughly equals $T_{FFT}^{1/4}$

9.5 Grids in the Parameter Space

Each parameter we use in the GFH is discretized in some way. In the following sections, we describe how we construct the grid in each of the parameters. None of the steps are uniform.

9.5.1 Grid on x_0

The grid in x_0 is obtained by taking the derivative of Equation 9–1:

$$dx_0 = (n-1)\frac{\delta f}{f^n} \tag{9–7}$$

The step dx_0 changes as a function of the frequency analyzed, as well as the $T_{FFT} = 1/\delta f$ chosen. However, having a nonuniform grid in x_0 is computationally very heavy, especially at higher frequencies, since we cannot fill every bin simultaneously as we can do if the grid is uniform. Therefore, we over-resolve the grid by taking $f = f_{max}$. This results in more bins, and a splitting of the signal across some x_0 bins in some cases. However, it makes the analysis

computationally feasible. The computation time increase is negligible since we fill all x_0 bins simultaneously when constructing the Hough map.

9.5.2 Grid on k

We must also construct a grid in k, since the Hough map is constructed across all times for a fixed k. An expression for the step can be derived analytically if we take the log_{10} of Equation 3–32:

$$\dot{f} = -kf^n \tag{9–8}$$

$$log\ |\dot{f}| = n\ log\ f + log\ k \tag{9–9}$$

Equation 3–32 forms lines in this space, where different braking indices correspond to lines with different slopes. If we consider a transformation $f \to f + \delta f$ and $k \to k + dk$, we can find dk such that the spindown remains constant when moving one frequency bin $\delta f = 1/T_{FFT}$. After solving equations 9–8 and 9–9 simultaneously, we find:

$$dk = k\left(\left(1 + \frac{\delta f}{f_{0,max}}\right)^{-n} - 1\right) \tag{9–10}$$

Since $\delta f \ll f_{0,max}$, we can Taylor expand to better understand the behavior of this grid:

$$dk \approx -nk\frac{\delta f}{f_{0,max}} \tag{9–11}$$

The grid depends on n and k, so it is nonuniform and changes each time we perform the Hough.

The minimum and maximum values for k are related to the maximum and minimum spindowns we wish to analyze, respectively:

$$k_{min} = \frac{\dot{f}_{0,min}}{f^n_{max}} \tag{9–12}$$

$$k_{max} = \frac{\dot{f}_{0,max}}{f^n_{min}} \tag{9–13}$$

To understand the properties of the k grid, we plot in Figure 9-6 how the range in k and the step size dk change as a function of braking index:

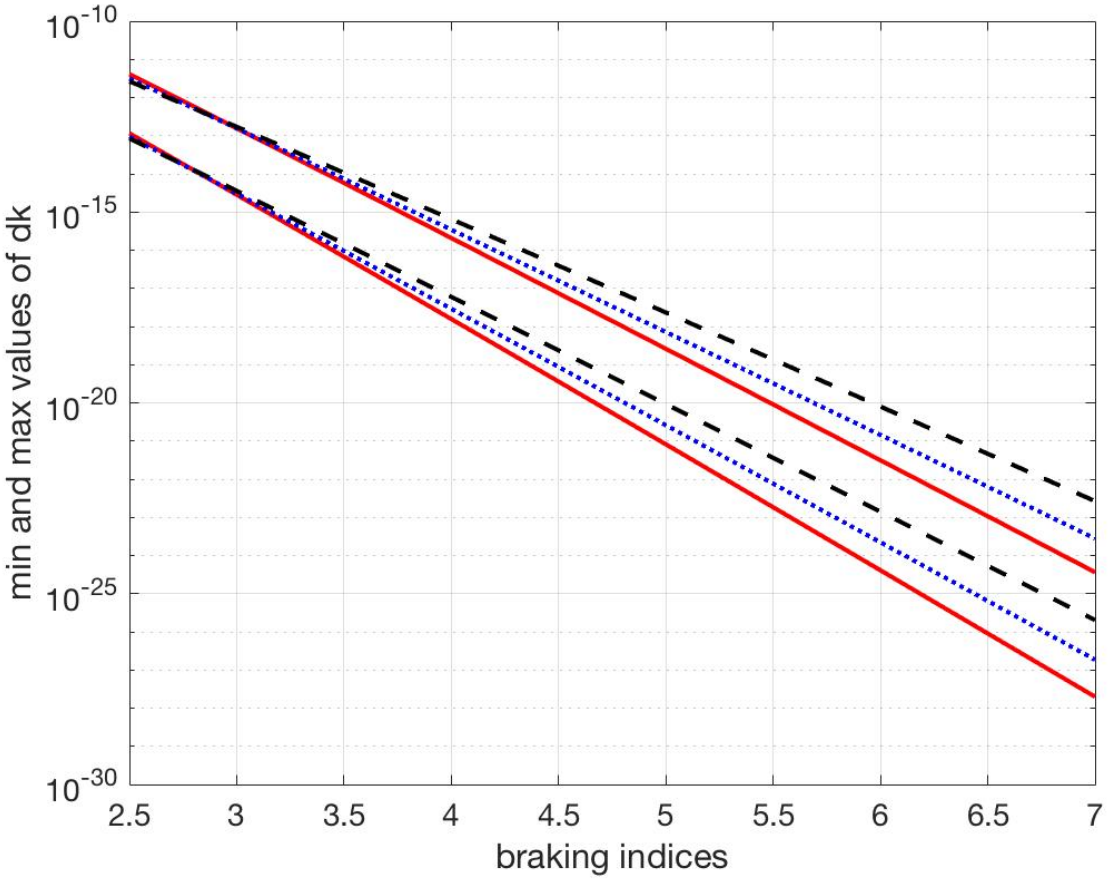


Figure 9-6: The ranges of the step dk for different T_{FFT}. Red, blue, and black correspond to $T_{FFT} = 2, 4, 8$ s (lowest, middle, highest curves) for frequency bands of [1000 2400] Hz, [600 1400] Hz, and [350 800] Hz respectively. $T_{obs} = 5000$ s and $\Delta t = 1 \times 10^6$ s. The upper (lower) curves correspond to the maximum (minimum) values for dk for each choice of T_{FFT}/frequency band.

It is clear that dk decreases with braking index, since the spindowns are smaller. However, the range $(k_{max} - k_{min})/dk$ decreases with increasing braking index, which means that we should expect fewer points in the grid on k at higher frequencies. This is shown in Figure 9-7.

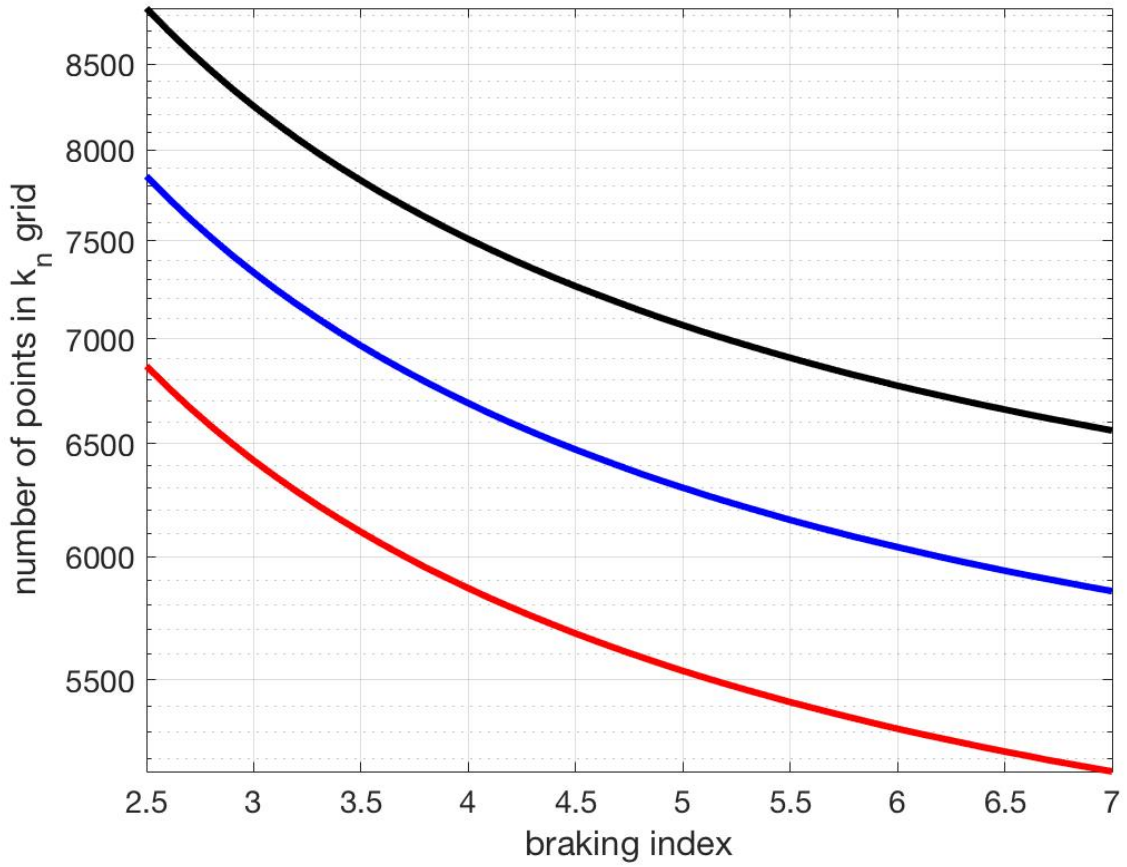


Figure 9-7: The number of points in the grid on k as a function of braking index. Less points are required to capture the variation in spindown at higher braking indices. Red, blue, and black lines correspond to $T_{FFT} = 2, 4, 8$ s for frequency bands of [1000 2400] Hz, [600 1400] Hz, and [350 800] Hz respectively.

9.5.3 Grid on n

The search is model-dependent, and yet the power law models for neutron stars are quite uncertain. Therefore we search across different braking indices, which can correspond to either a combination of emission mechanisms- electromagnetic and GW emission from a neutron star – or a braking index that is very slowly varying in time. We construct a grid of different braking indices between $n = 2.5$ and $n = 7$ with a step that is calculated so that the frequency

variation stepping from $n_1 = n$ to $n_2 = n + dn$ for the signal duration is confined to one frequency bin:

$$f_1(t) = \frac{f_0}{\left(1 + k(n_1 - 1) f_0^{n_1 - 1}(t - t_0)\right)^{\frac{1}{n_1 - 1}}} \quad (9\text{–}14)$$

$$f_2(t) = \frac{f_0}{\left(1 + k(n_2 - 1) f_0^{n_2 - 1}(t - t_0)\right)^{\frac{1}{n_2 - 1}}} \quad (9\text{–}15)$$

$$\Delta f = f_2(t) - f_1(t) \leq \delta f = \frac{1}{T_{FFT}} \quad (9\text{–}16)$$

We empirically find the value of dn at each braking index n, k and f_0 such that for the duration of the analysis T_{obs}, the frequency variation remains within one frequency bin. δf is fixed by our choice of T_{FFT}, which depends on the maximum spindown we wish to analyze. Therefore, the grid depends on the FFT length, the braking index, the spindown range and frequency range.

We provide here some plots describing the behavior of the grid on n. In Figure 9-8, we show how the number of points changes in the grid as a function of the braking index. At higher braking indices, the magnitude of the spindown decreases quicker than at lower braking indices, meaning that fewer points in the grid n are required to cover the parameter space.

Additionally, the number of points in this grid changes as a function of T_{FFT} for two reasons: (1) the frequency bin size becomes smaller with increasing T_{FFT} and (2) the frequency dependence as a function of time becomes much weaker as the spindowns we analyze decrease (since $\dot{f} = 1/T_{FFT}^2$ and T_{FFT} is increasing), see Figure 9-9.

9.6 Theoretical Estimate of Sensitivity

We derive an analytic expression for the sensitivity of a semi-coherent search for long transient periodic signals, such as those emitted by a newborn magnetar, assuming the analysis is done with the generalized FH transform. The expression is in fact a generalization of the sensitivity computed for standard continuous wave signals: 9–22.

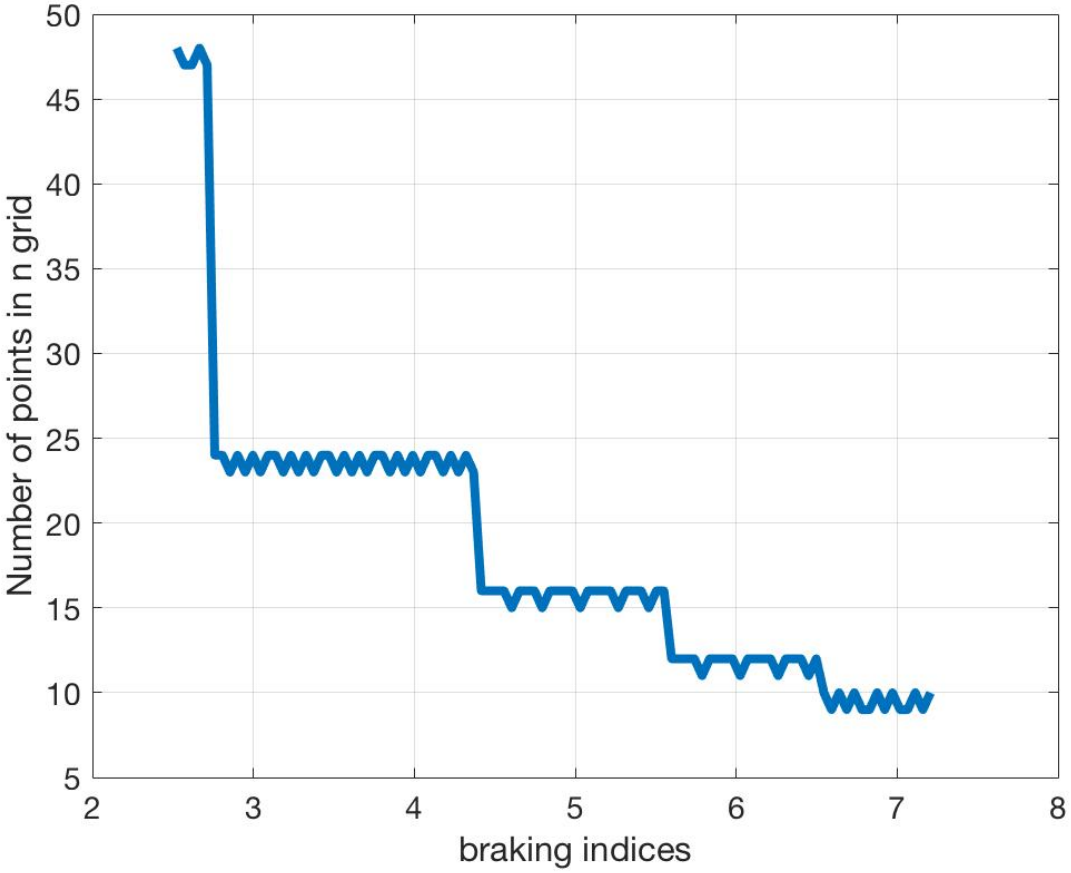


Figure 9-8: The plot shows the number of points in the grid on n as a function of the braking index. Parameters used in this plot: $T_{FFT} = 2$ s, $f_{0,max} = 2000$ Hz, $\dot{f}_{0,max} = 1/T_{FFT}^2$ $n = [2.5, 7]$, histogrammed into 100 n bins of size $\delta n \sim 4 \times 10^{-2}$.

As in the previous sensitivity estimation, we assume a period signal but that can vary (rapidly or not) in time. Recall that our semi-coherent analysis requires that a signal not spin out of a frequency bin during T_{FFT}. So, in each T_{FFT}, the frequency and therefore amplitude of the signal is constant.

In order to compute an average sensitivity we first introduce the quantity

$$\Lambda = \frac{T_{FFT}}{2T_{obs}} \sum_{i=1}^{i=N} \lambda(f(t_i)) \tag{9–17}$$

where $\lambda(f(t_i))$ is:

$$\lambda_i = \lambda(f(t_i)) = \frac{4|\tilde{h}(f(t_i))|^2}{T_{FFT} S_n(f(t_i))} \tag{9–18}$$

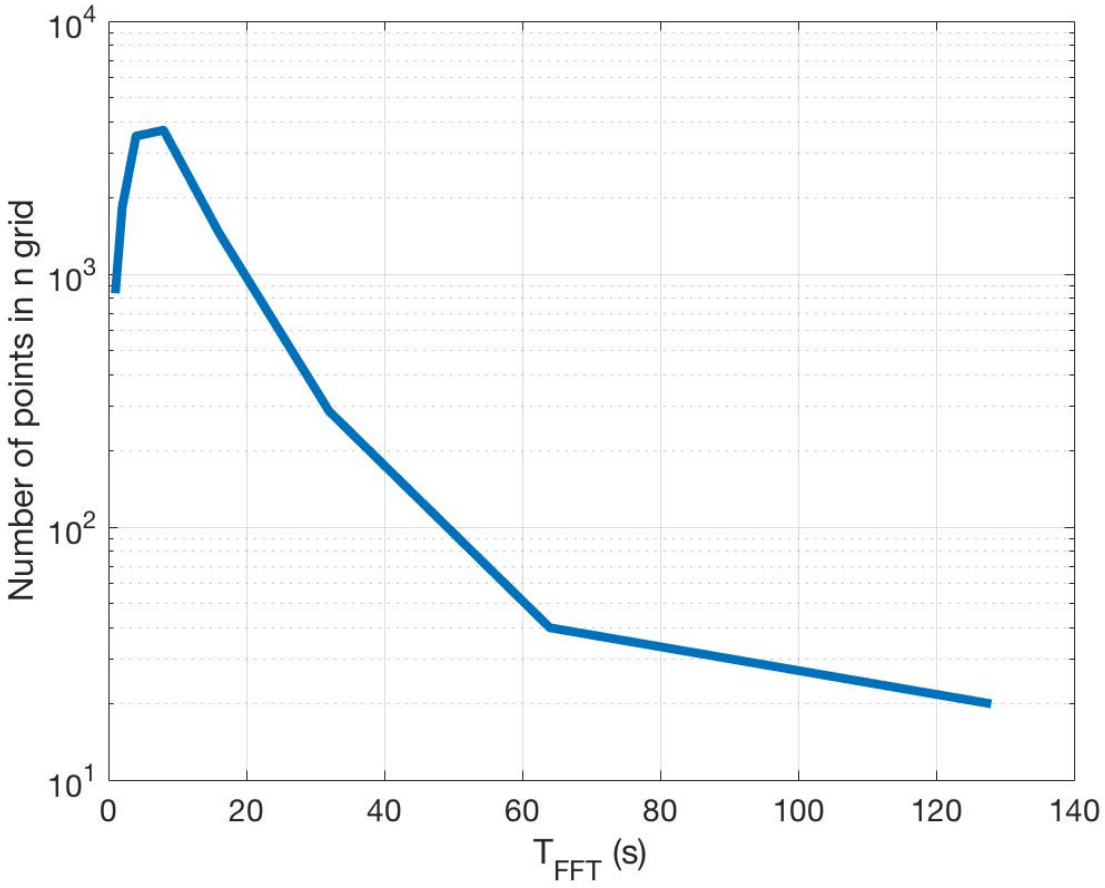


Figure 9-9: This plot shows the number of points in the grid on n as a function of T_{FFT} for a fixed observation time $T_{obs} = 10^5$ s and maximum frequency $f_{0,max} = 2000$ Hz. Note that for longer T_{FFT}, the spindowns we are analyzing are smaller. We see a sharp decay in the number of points in the grid because the frequency is not varying so strongly, and therefore Equation 3–33 reduces to Equation 3–24.

where $\tilde{h}(f(t_i))$ is the Fourier transform of the GW signal, $N = T_{obs}/T_{FFT}$ is the number of FFTs used, and $S_n(f(t_i))$ is the detector noise power spectrum. This quantity is essentially the same as that quoted in 8–9, except that instead of having one contribution due to the strength of the signal in the observation time, we have many different contributions, each within one T_{FFT}.

The quantity Λ is an average of the λ_i over the observation window and satisfies the condition that if $\lambda_i = \text{const}$ then $\Lambda = \text{const}$, i.e. we are back to the standard situation of CW signals in which the frequency and the amplitude do not significantly change over the

observation time. One can think of Λ as the signal amplitude decreasing over time (since the amplitude is a function of frequency), but the "strength" of the signal is also proportional to how long it lasts; hence, we have placed the summation in Equation 9–17. As the signal amplitude h_0 varies with frequency as f^2 we write

$$h_0(t_i) = \mathcal{A} f^2(t_i) = \mathcal{A}\mathcal{F}_i \tag{9–19}$$

where:

$$\mathcal{A} = \frac{4\pi^2 G I_{zz}}{c^4}\frac{\epsilon}{d} \tag{9–20}$$

, This $\mathcal{A}$ depends only the star ellipticity ϵ, moment of inertia and distance d. Combining equations

$$\Lambda \approx 4\frac{T_{FFT}}{2T_{obs}}\mathcal{A}^2\sum_i \frac{\mathcal{F}_i^2}{S_n(f_i)}\frac{2.4308}{25\pi}T_{FFT} \tag{9–21}$$

Following the same procedure given in (*78*), we obtain:

$$\mathcal{A}_{min} = \frac{4.02}{N^{1/4}\theta_{thr}^{1/2}}\sqrt{\frac{N}{T_{FFT}}}\left(\sum_i \frac{\mathcal{F}_i^2}{S_n(f_i)}\right)^{-1/2}\left(\frac{p_0(1-p_0)}{p_1^2}\right)^{1/4}\sqrt{\left(CR_{thr} - \sqrt{2}\mathrm{erfc}^{-1}(2\Gamma)\right)} \tag{9–22}$$

where θ_{thr} is the threshold for peak selection in the whitened spectra, S_n is the noise spectral density of the detector, p_0 is the probability of selecting a peak above the threshold θ_{thr} if the data contains only noise , $p_1 = e^{-\theta_{thr}}$ - $2e^{-2\theta_{thr}} + e^{-3\theta_{thr}}$, CR_{thr} is the threshold we use to select candidates in the final FH map, and Γ is the chosen confidence level.

The minimum detectable strain at a given confidence level can be obtained from Equation 9–22 using a suitable "frequency" (indeed $h_{0,min} = \mathcal{A}_{min} \cdot \mathrm{frequency}^2$). We use the initial frequency f_0. The sensitivity depends on the signal evolution through the ratio $\sum_i \frac{\mathcal{F}_i^2}{S_n(f_i)}$. To calculate the sensitivity we must fix n, f_0, $\dot{f}_0$ and T_{obs}. That is, for a given detector and given search parameters (T_{FFT}, θ_{thr}, CR_{thr}, Γ etc.) and for each specific signal model, we will have a different sensitivity. If $f_i = const$, which is a good approximation for a standard CW signal case, then $\mathcal{F}_i = f_0^2$ and as a consequence $\sum_i \frac{\mathcal{F}_i^2}{S_n(f_i)} = \frac{N f_0^4}{S_n(f)}$. Inserting this expression

in Equation 9–22, we recover the standard sensitivity expression in (*78*). If the emission of GWs is due to r-modes, the signal amplitude scales with f^3. In this case $\mathcal{F}_i = f^3(t_i)$ and $\mathcal{A} = 1.11 \cdot 10^{-9} I \alpha$. The distance reach of the search is easily obtained inverting Equation 9–22:

$$d_{max} = 5.72{\cdot}10^{-9} I_{38} \epsilon_{-3} \frac{T_{FFT}}{\sqrt{T_{obs}}} \left(\sum_i \frac{\mathcal{F}_i^2}{S_n(f_i)} \right)^{1/2} \left(\frac{p_0(1-p_0)}{N p_1^2} \right)^{-1/4} \sqrt{\frac{\theta_{thr}}{\left(CR_{thr} - \sqrt{2}\mathrm{erfc}^{-1}(2\Gamma)\right)}} \tag{9–23}$$

with I_{38} the star moment of inertia (with respect to the rotation axis) in units of $10^{38} kg \cdot m^2$ and ϵ_{-3} the star ellipticity in units of 10^{-3}. For r-mode emission Equation 9–23 must be modified replacing the numerical coefficient by $1.11 \cdot 10^{-12}$ and it should be noted that ϵ is in fact the mode amplitude α.

9.7 Empirical Estimate of Sensitivity

We empirically estimate the minimum amplitude detectable at $\Gamma = 90\%$ confidence at some initial frequencies for fixed braking indices of $n = 3, 5, 7$. We use 100 injections in Gaussian noise between f_0 and $f_0 + 10$ Hz and define a detection as when a candidate selected in a Hough map is within a distance of 3 x_0/k bins from the injection. In a real search, there is a coincidence step between candidates selected in different detectors, but for this empirical sensitivity estimation, we only do coincidences between a candidate found in one detector and the injection. The granularity (that is, the spacing in h_0) is 10% of the theoretical h_0 for the corresponding f_0. The minimum detectable amplitudes are shown in Figure 9-11 where the theoretical sensitivity curves for the same parameters are also plotted. In this plot, we have used varying estimations of CR_{thr} to reflect the fact that due to the flat noise curve we assumed, signals with the same amplitude in higher frequency will generate candidates with higher critical ratios because the amplitude of the GW signal is proportional to f^2- see Equation 3–34. The plot demonstrates a good agreement between experimental and theoretical results. Additionally, we show plots of detection efficiency as a function of amplitude for a

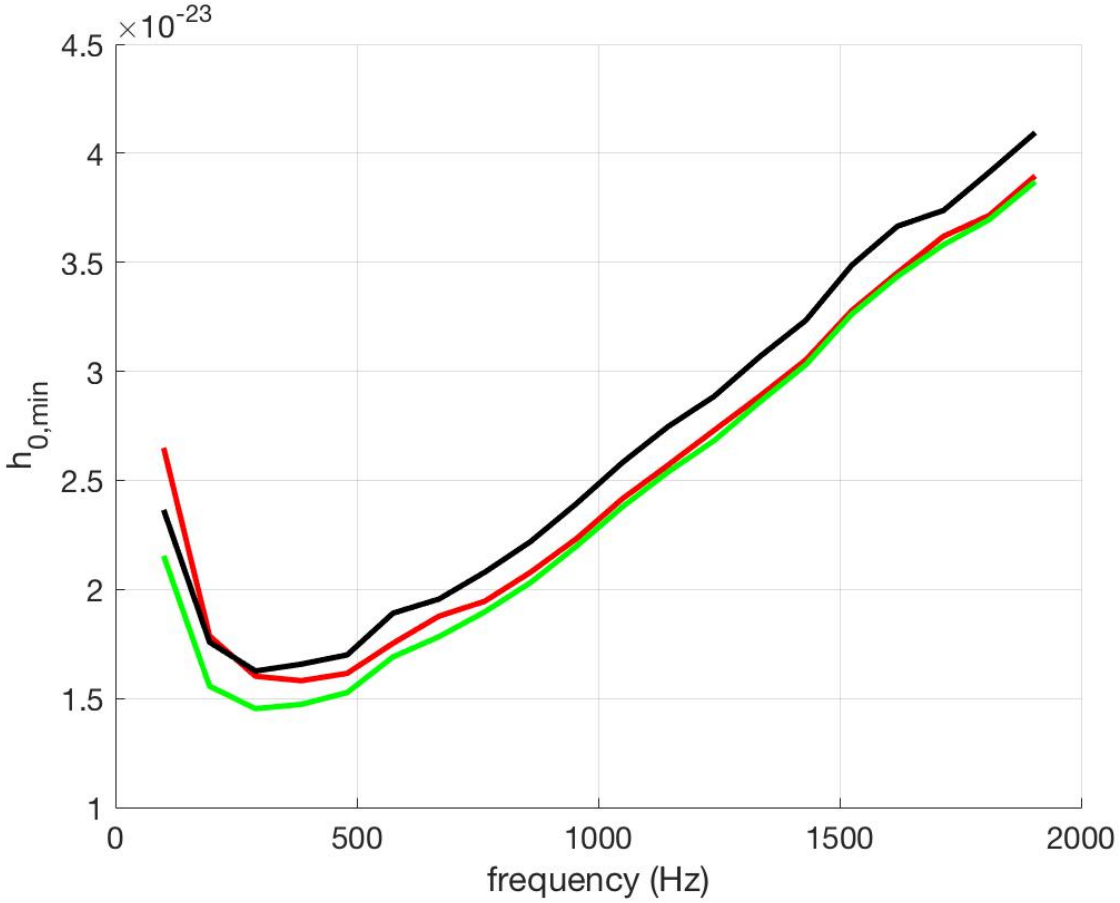


Figure 9-10: Theoretical $h_{0,min}$ for different initial frequencies f_0 for n=3, 5, 7 (red, green, black) where the noise distribution has been weighted with the O2 Livingston sensitivity curve for $\dot{f}_0 = 1/16$ Hz/s. There appears to be a dependence on the details of Equation 9–22.

few initial frequencies in Figure 9-12. The false alarm probability is $\sim 0.01\%$, but in general changes in each Hough map.

In this case the typical distance reach is of the order of 0.5-1 Mpc, obtained from the theoretical estimates shown in Figure 9-10 for O2 Livingston data, assuming a moment of inertia and star ellipticity. The moment of inertia could be larger than the canonical $10^{38}\ kg \cdot m^2$ for a neutron star born after a coalescence. The other parameter which may have some impact is the threshold for candidate selection CR_{thr}, which depends on how many candidate follow-ups we can afford. The distance reach is typically smaller for shorter T_{obs}.

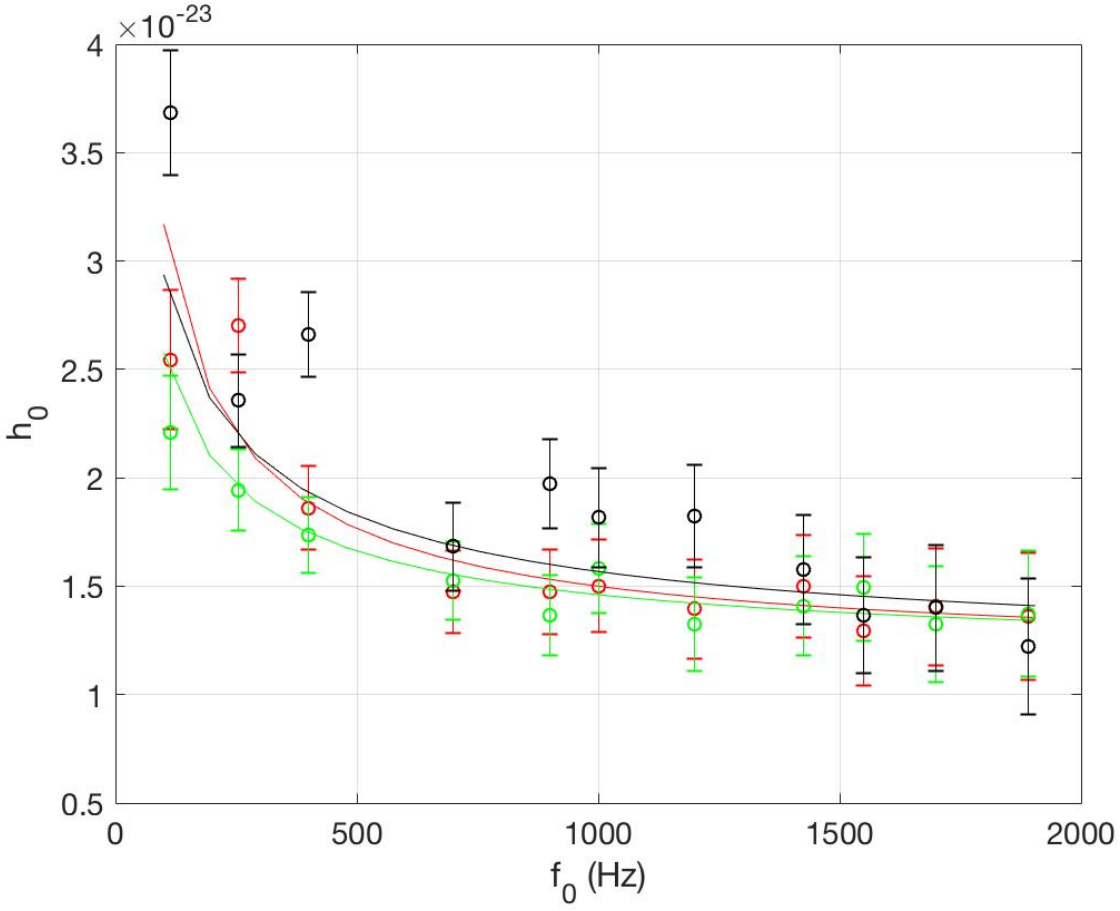


Figure 9-11: Minimum detectable amplitude $h_{0,min}$ as a function of initial frequency f_0 for braking indices $n = 3, 5, 7$ (red, green and black / bottom, middle, and top). The continuous curves are the theoretical sensitivity estimates; the curves marked with circles are the sensitivity estimates obtained from injections in white noise at a level consistent with real O2 Livingston data about 1 hour after GW170817 ($S_n = 7.94 \times 10^{-24} \frac{1}{\sqrt{Hz}}$); parameters: $T_{obs} = 5000$ s, $T_{FFT} = 4$ s, $\theta_{thr} = 2.5$, $\dot{f}_0 = 1/T_{FFT}^2$, $\Gamma = 0.9$. To compute the theoretical sensitivity estimates, we use a varying CR_{thr} reflective of the average critical ratio we recover from injections at 90% confidence, which changes because signals with higher frequencies with the same amplitudes as those at lower frequencies will be recovered with higher critical ratios.

It is worth noting that in Figure 9-10, the sensitivity depends heavily on the details of Equation 9–22: specifically, the decaying frequency is a strong function of the braking index, which can induce changes in the sensitivity curve at various frequencies. This is why at some frequencies, signals $n = 3$ or $n = 5$ have lower values of minimum detectable amplitude.

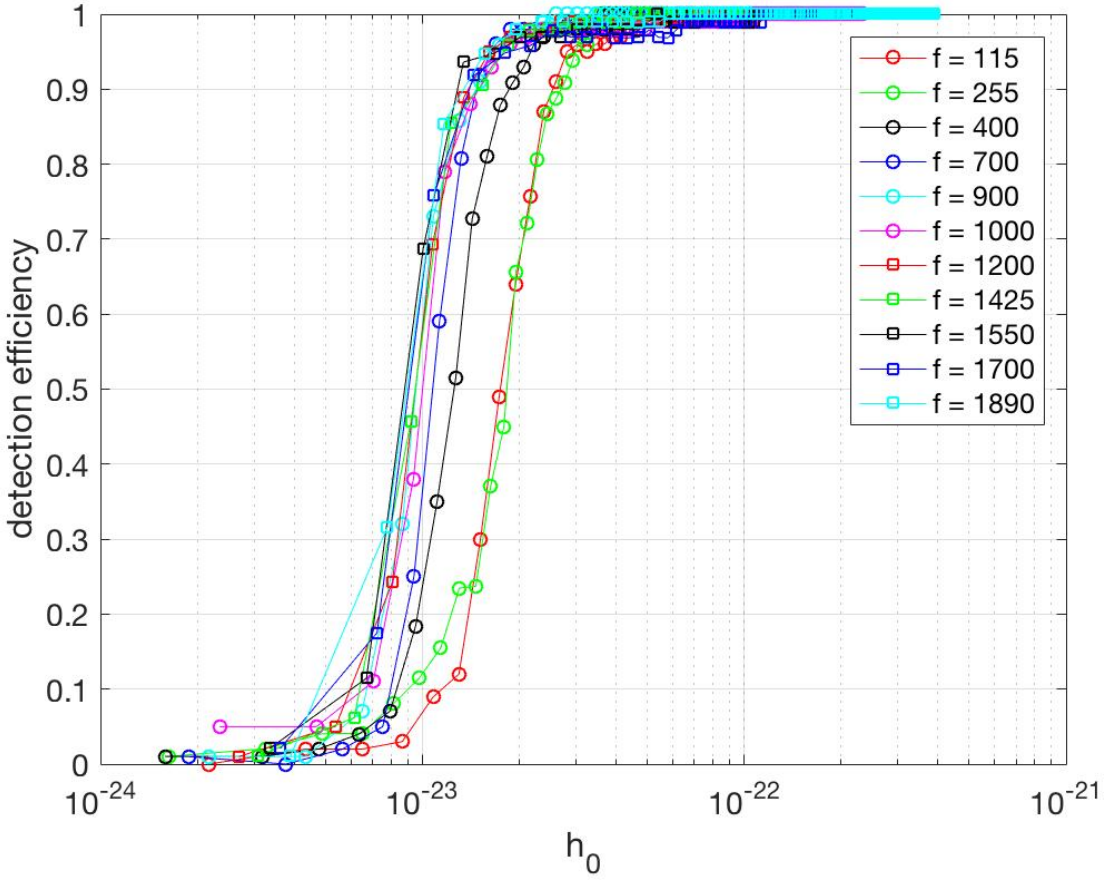
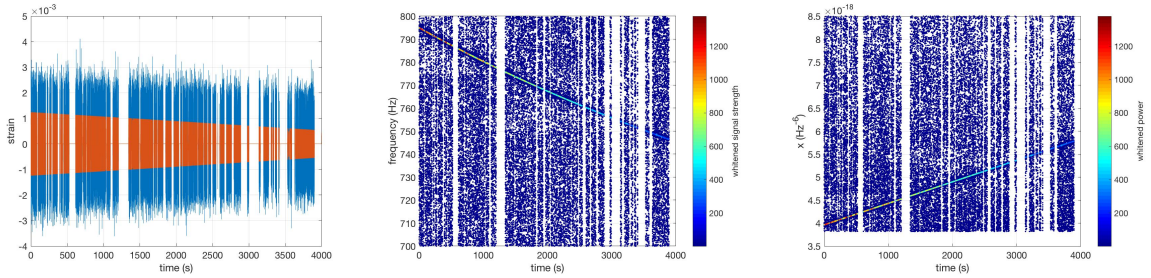


Figure 9-12: Detection efficiency plotted as a function of h_0 for $n = 5$ for $T_{obs} = 5000$ s and $T_{FFT} = 4$ s. Different colors correspond to different initial frequencies f_0. The curves seem to follow the sigmoid distribution, which is as expected. We theoretically estimate the false alarm probability to be order of 0.01% using equation 61 of (*78*); however the grids on x_0 and k change in each Hough map, so the false alarm probability also changes. Moving from left to right tends to correspond to decreasing f_0.

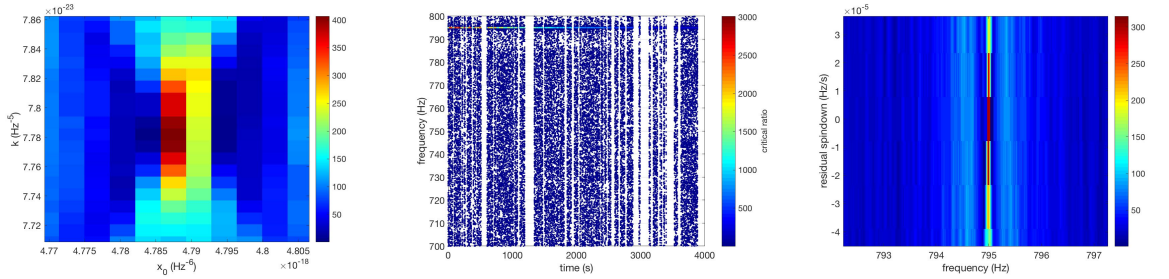
If we are able, thanks to observations in the EM band or to robust theoretical models, to restrict the possible range of the search parameters, f_0, k, n, we can in principle make a deeper search. For instance, by making a first rough search over a limited range of braking indexes we could make a second step searching over the braking index *residuals* (with respect to the initial rough grid) and this would allow us to increase T_{FFT} and then gain in sensitivity.



(a) Time series strain O2 Livingston data

(b) Time/frequency map made with $T_{FFT} = 8$ s.

(c) Transformation of 9-13(b) using equation 9–1



(d) Hough map showing injected signal's parameters

(e) Phase-corrected time/frequency map using 9-13(d)

(f) Hough map, phase-corrected signal's parameters

Figure 9-13: These plots illustrate the Generalized Frequency Hough pipeline, showing how we estimate the parameters of a GW signal from an isolated neutron star starting with the strain time series. Note that in Figure 9-13(d), a reference time equal to midway through the observation time ($t_0 \sim 2000$ s) was used, which corresponds to a particular.

CHAPTER 10
SEARCHING FOR A REMNANT OF GW170817

10.1 Introduction

GW170817 was an extremely exciting event due not only to the physics we learned from the merger, but also because of the unknown nature of the product of such a merger. Therefore the LIGO/Virgo collaborations wanted to determine the identity of the remnant using four separate pipelines. In Section 10.2, we motivate why we should search for a remnant of a BNS merger and what we could hope to learn from a GW detection. Afterwards in Section 10.3, we show the portion of the parameter space our method, the GFH, was sensitive to in this search. The next three sections, Sections 10.4, 10.5, and 10.6 deal with practical aspects of the search, that is, the selection of candidates after the GFH transform, the follow-up of potentially interesting signals, and the upper limits we were able to place on GWs originating from a remnant of GW170817 respectively. We finish the chapter with some discussion of future searches in Section 10.7.

10.2 Motivation

GW170817 was the first BNS merger ever observed by LIGO/Virgo. It has created many opportunities in astrophysics for deep study into the nature of neutron star mergers and neutron star physics. One of these aspects is answering the question, what remained after the merger? The mass of the two neutron stars, and the mass of the remnant, have quite wide uncertainties, so it is not clear what formed after the merger. There are essentially 4 options as to what could happen: (1) immediate collapse to a black hole, (2) formation of hypermassive neutron star (HMNS) that is supported by differential rotation lasting for $O(s)$, then uniform rotation, then collapsing to a black hole, (3) a supramassive neutron star (SMNS) that exists is sustained by normal rotation for $O(hours)$, then collapses to a black hole, and (4) a stable neutron star that would in theory, after allowing for some time for the remnant to stabilize, emit continuous GWs (*82*). In Figure 10-1 the possible remnant scenarios are shown

schematically as a function of the initial mass of the binary, as well as their associated time scales.

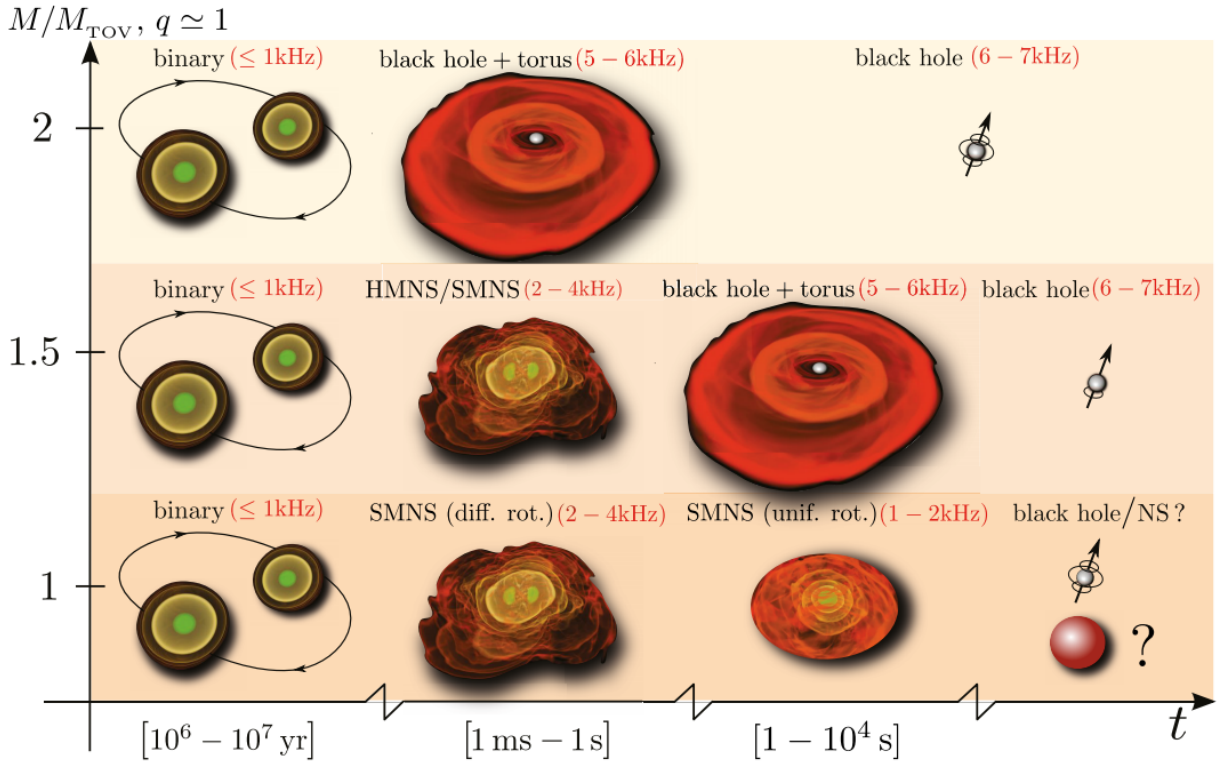


Figure 10-1: Diagram taking from (*82*) showing the stages of a binary system's evolution as a function of the total initial mass of the system. The total mass itself, and its relation to the EoS-dependent M_{TOV} affects the product of the merger, and each remnant has an associated lifetime.

It is possible that electromagnetic observations could give us information on the post-merger remnant; however, in reality it is often ambiguous. For example, the kilanova (r-process) model cannot fully explain the spectra that we see, so hybrid models have been considered: early EM emission is powered by radioactive decay of r-process elements, and emission at later times is powered by the spindown of a long-lived remnant (*83*). Additionally, it has been shown that some stiffer equations of state can produce magnetic fields consistent with the electromagnetic observations (whereby measuring the period and rate of change of period of a pulsar, we can calculate the magnetic field since $B \propto \sqrt{P\dot{P}}$- see Equation 3–6).

This relation comes from equating the rotational energy of the star with the power due to a rotating magnetic dipole- see Section 3.2.3.

Based on the derivation in 3.2.3, we can also see immediately that:

$$\dot{\Omega} = -\frac{B^2 R^6 \sin^2 \alpha}{6\pi c^3 I}\Omega^3 \tag{10–1}$$

A rotating magnetic dipole emission mechanism gives us a braking index of $n = 3$.

Additionally, if the remnant was a black hole, fall back accretion would occur and could in theory power the electromagnetic emission. Fall back accretion happens when neutrino emission causes ejecta matter to cool very quickly so that matter falls back into the remnant (the direction of matter is reversed by rarefaction wave), if the velocity of the matter is not much larger than the escape velocity (*84*). But this has been show to only explain emission at early times, where a spinning-down neutron star can explain the full spectra: the magnetic field required in $O(10^{12})$ G, but the ellipticity is high $O(10^{-3})$.

No single electromagnetic emission model can explain the observed spectra in the UV/Visible/ infrared bands (*30*). However, detecting a neutron star as a remnant of this merger could provide many insights, despite the ambiguous electromagnetic evidence. We know the distance just from GW observations: the mass ratio can be determined from the time/frequency evolution and the strain on the detector tells us about the strength of the signal. The location is know somewhat well with just LIGO/Virgo but extremely well with electromagnetic observations. But what we still don't know is: what is the EoS of the neutron star, the percentage of their energy that goes into GWs, how deformed they are, their moments of inertia, and if there is a difference between pre- and post-merger equations of states. We can only understand these things through GW observations: if we are able to measure the mass of the neutron star remnant, we can then constrain the product $I_z\epsilon$. We describe our search procedure for a remnant of GW170817 that was done jointly with three other pipelines (*30*).

10.3 Parameter Space Explored

Since we knew the location and the exact time at which to search, we were able to explore a relatively large range of the physical parameter space of a neutron star. We search across braking indices ranging from $n = [2.5, 7]$, spindowns from $[1/256, 1/4]$ Hz/s, initial frequencies $f_0 = [100, 2000]$ Hz, in three different configurations (three different $T_{FFT} = 2, 4, 8$ s starting on average 1, 2 and 5 hours after the merger). We cover different frequency and spindown ranges for each configuration, but we always cover the same range of braking indices.

We start the search 1 hour after the merger for a few reasons: (1) we expect that immediately after the merger, the physics will be very complicated and not well-described by a power law frequency/time evolution, (2) the spindown of the source should be smaller, which means that we can use longer T_{FFT} and observe for a longer time, and (3) we can probe signals with high initial spindowns and frequencies that may have moved into our frequency/spindown window within this one hour delay.

We construct which time/frequency bands to analyze for each value in our grid on the braking index on the basis of the spindown and the starting frequency of a hypothetical signal. Essentially, we assume that if there is a signal that starts at $f_0 = 500$ Hz with a maximum spindown of $1/4$ Hz/s, we determine the amount of time it takes for the signal's spindown to reach a value of $1/4^2 = 1/16$ Hz/s, which would be when we can change the T_{FFT} to be 4 seconds. We record the amount of time this takes, and the frequency band that would be covered by the signal. This is illustrated in Figure 10-2, as well as how different braking indices affect the bands that we analyze.

Each time/frequency band for each initial frequency we consider at each braking index is slightly different. This means that when we would like to analyze which portions of the parameter space we are sensitive to, we must account for the fact that we will have different T_{FFT} and durations for each class of signal we analyze.

Figure 10-3 shows which portions of the parameter space we are most sensitive to 1 hour after the merger, where "most sensitive" refers to a hypothetical signal starting with the

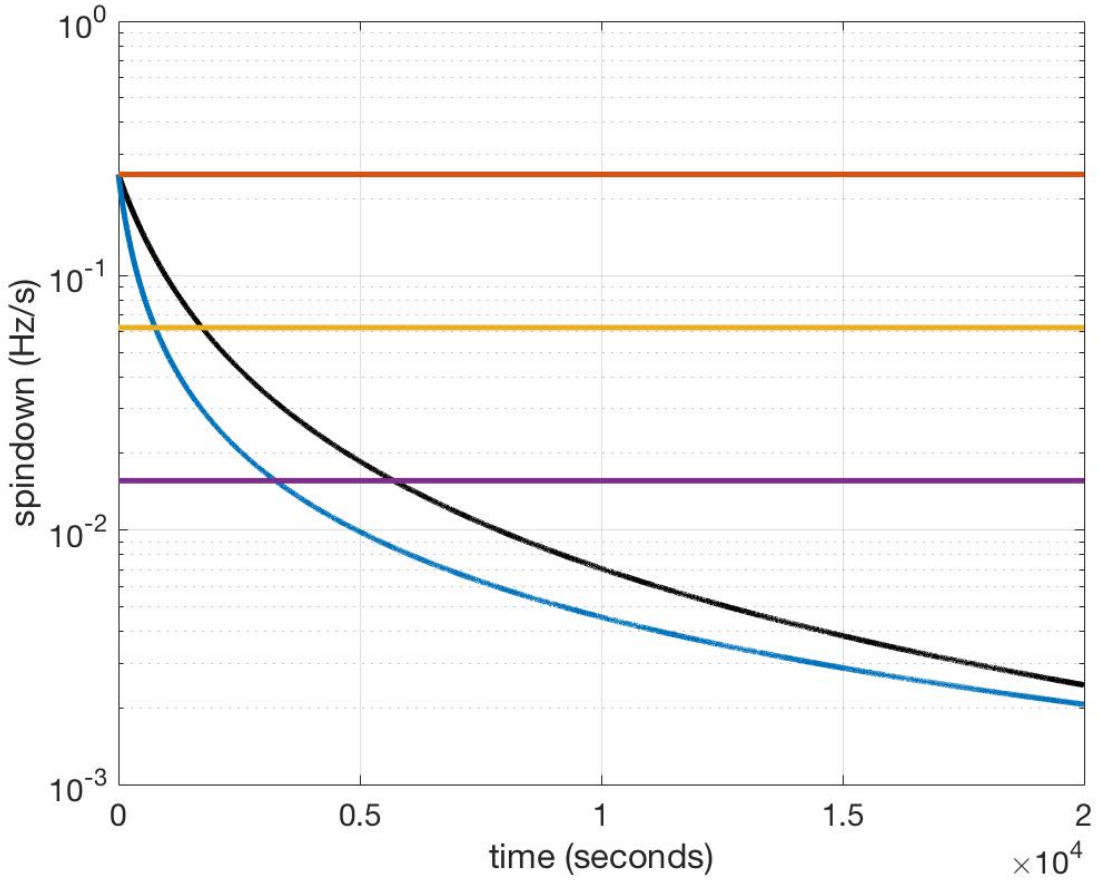


Figure 10-2: The spindown of two signals with $f_0 = 500$ Hz with different braking indices (black/top: $n = 2.5$; blue/bottom: $n = 7$) as a function of time. The horizontal lines correspond to spindowns $\dot{f} = 1/T_{FFT}^2$, $\dot{f} = 1/(2T_{FFT})^2$, and $\dot{f} = 1/(4T_{FFT})^2$. When a signal reaches a small enough spindown (intersects one of the horizontal lines), we can then analyze it for a longer duration with a higher T_{FFT}. Within each rectangle, the times and frequencies that will be analyzed with a given T_{FFT}, which will be different depending on n, f_0 and $\dot{f}_0$. We explore all signals whose paths lie between the curves corresponding to $n \geq 2.5$ and $n \leq 7$.

specified frequency/spindown as being fully contained in the frequency band we are analyzing. The white space ("the holes") that are visible in Figure 10-3 are indicative of signals that are for the most part contained in the frequency band we analyze, but spin out a little. We are still sensitive to this signals, but not as sensitive as if we extended the frequency band analyzed. But, we can still recover these signals, since they are contained within the map that we analyze.

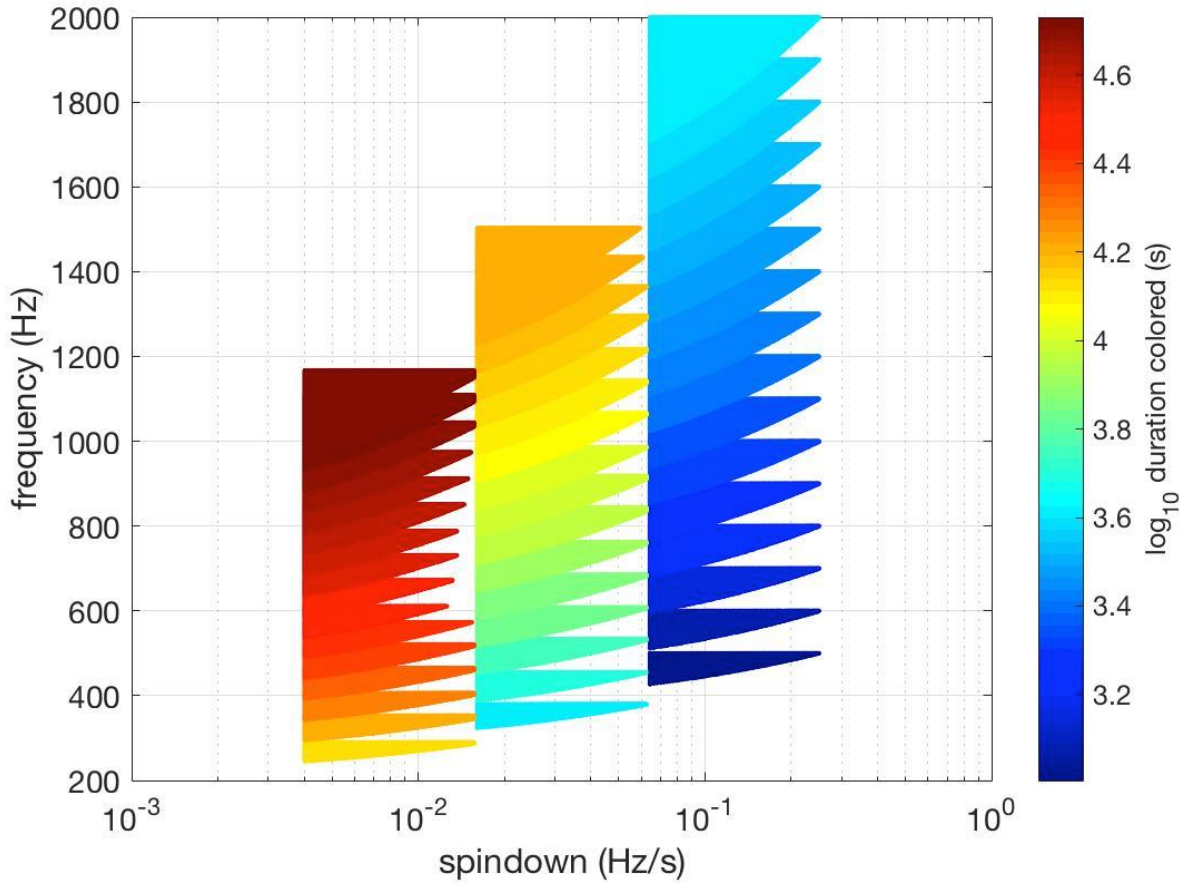


Figure 10-3: The frequency/spindown space that is explored by analyzing different portions of frequency and time with different T_{FFT} for $n = 5$. The durations T_{obs} of the analysis are colored, and the three separate blocks correspond to $T_{FFT} = 2, 4, 8$ s moving from right to left.

Figure 10-4 shows the corresponding source parameter space, where we have "back-tracked" the GW spindowns/frequencies in the Figure 10-4 to determine the true initial frequencies/spindowns. It is clear that we can probe very high frequencies and spindowns that are not normally accessible to us because of the noise present in the detector, and because of the amount of sensitivity we would lose to try to make $T_{FFT} < 1$ s. The so-called spindown time scale is colored in this figure, defined as:

$$\tau = \frac{1}{k f_0^{n-1}(n-1)} = \frac{f_0}{\dot{f}_0(n-1)} \tag{10–2}$$

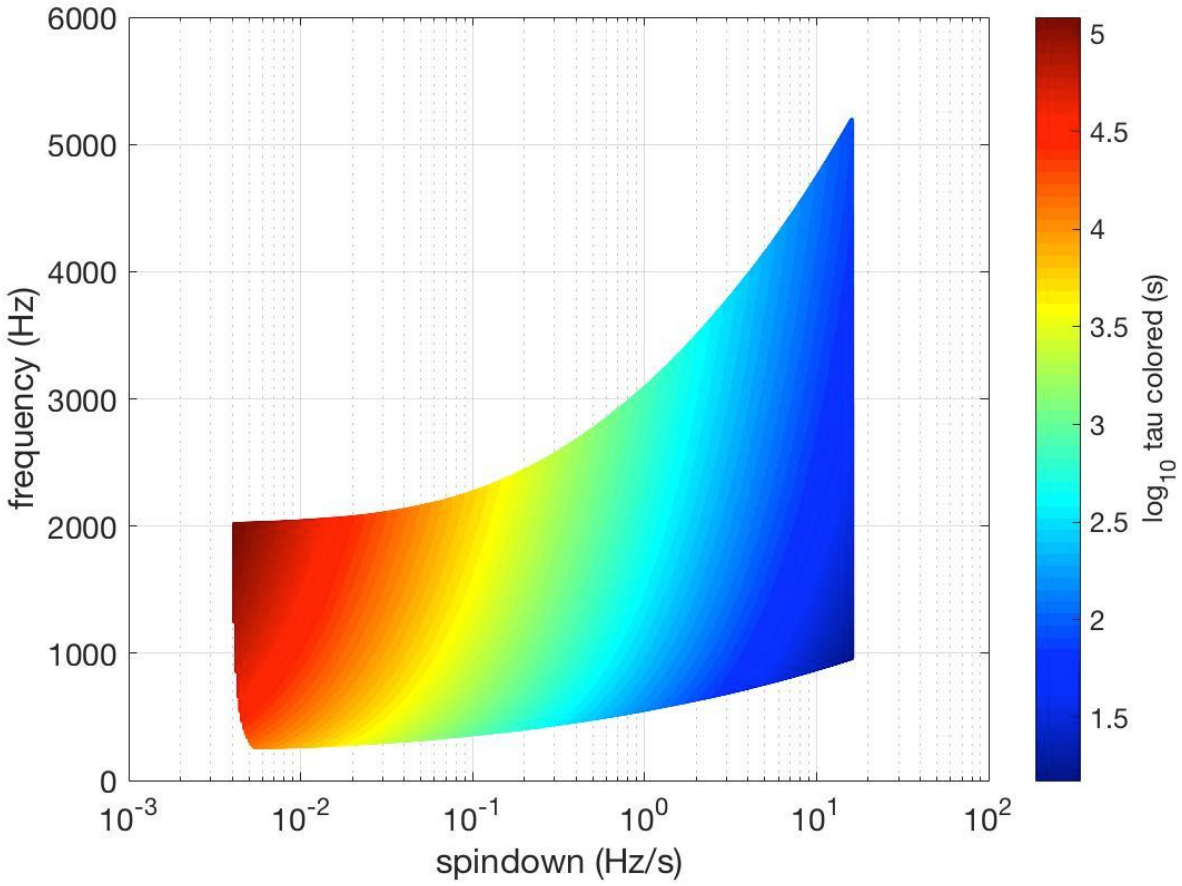


Figure 10-4: The source frequencies and spindowns that we are most sensitive to in our search scheme (Figure 10-2), with a choice to start analyzing the data one hour after a merger, for $n = 5$. The spindown timescale τ is colored. This plot includes signals that lie in the "holes" of Figure 10-3.

To determine these initial parameters, we try many different initial frequency/spindown combinations to determine which parameters will fall within one frequency bin of the frequencies given in Figure 10-3 after allowing this simulated signal to decay for a certain amount of time.

10.4 Candidate Selection

In each Hough map, we must select a certain number of candidates, chosen as a compromise between losing sensitivity, decreasing false alarm probability and decreasing computational cost of the follow-up. A candidate is defined as having an f_0 $\dot{f}_0$, n, longitude and latitude. In the Generalized Hough pipeline, we choose candidates with the highest number

counts/critical ratios in each box of the Hough map, where the size of each box is determined by the number of candidates to be chosen. To give an idea, if we only select a small, O(20) candidates, from the map, the boxes are large; if we select O(200) candidates, the boxes are smaller.

We employ this "box selection" method because of the non-uniformity of the noise. We found that if we simply select candidates in the standard way (in rectangular strips), we are predisposed to select only candidates in the highest frequency (low x value) space with low spindown (low k). With the box selection method, we can select candidates with more uniform f_0 and k, which are both physical parameters. Additionally, we select the second strongest candidate in the box if it is sufficiently far away in k. This is done so that we are not blinded by noise artifacts.

10.5 Follow-up Procedure

Since we constructed the search in three configurations, we would first like to compare candidates recovered in each one. Recall that each configuration is done at a different average time after the merger. What we can then do is do coincidences between candidates that we recover in each configuration. We simply take the recovered parameters in, for example, the $T_{FFT} = 2$ s configuration, allow the signal to spin down for a certain amount of time (depending on how long the analysis window was), and then compare the signal's parameters (x_0,k and n) to the candidates' parameters recovered in the second block ($T_{FFT} = 4$ s). The coincidence distance is

$$d = \sqrt{\left(\frac{x_{0,2} - x_{0,1}}{\delta x_0}\right)^2 + \left(\frac{k_2 - k_1}{\delta k}\right)^2 + \left(\frac{n_2 - n_1}{\delta n}\right)^2}, \tag{10–3}$$

which is the same as Equation 9–5, except that we now can consider a variation in the braking index n. We chose to impose the constraints that n and k should not vary, by definition. However, it is entirely possible to look for signals that do have variation in these parameters, and if we want to explore signals with varying braking indices, we would need to account for this.

After doing coincidences between $T_{FFT} = 2$ s and $T_{FFT} = 4$ s, we were left with some candidates using a coincidence distance of $d = 3$ bins. Then, we did coincidences between the surviving candidates and $T_{FFT} = 8$ s, but no candidate survived. Therefore, we can conclude that there was not a detectable signal present in all three portions of the parameter space we analyzed.

Our search returned about 500 coincident candidates (too numerous to list) with an average critical ratio above 5. We chose a threshold of $CR = 5$ because it corresponds to a false alarm rate of $< 0.001\%$, and this is the value typically used in CW searches. We then looked at the critical ratio recovered in Hanford and Livingston, and rejected coincident candidates whose critical ratio was more than 4 times different in one detector over the other. We expect a true GW signal to have a difference in critical ratio that is proportional to (1) the orientation of the detector with respect to the source location and (2) the sensitivity difference between the two detectors. For this search, most of the candidates were vetoed because the critical ratio difference was too great. We only performed the follow-up on one candidate. The procedure for the follow-up is the following: we correct the raw time-series data for the phase evolution of the candidate we are following up as described in 9.4 in each detector. We then decide how much to increase the T_{FFT} by (we have chosen a factor of 2, and would have increased it more if the candidate survived this phase of the follow-up). Now, we create another peakmap for each detector and perform the original FH on it. This is because we have hopefully removed most of the frequency variation from the signal, therefore the signal should be almost monochromatic. After performing the FH, we select candidates in the new maps, and do coincidences. In the case of the search, the candidate did not survive; therefore, all candidates were vetoed.

10.6 Upper Limits Procedure

When a search doesn't find anything, we perform upper limits, that is, we ask the question, "what is the minimum amplitude of a signal we could have found?" Or, "how far away could we have seen?" Now, there is a key difference between upper limits and sensitivity:

the ULs are computed using the critical ratios of candidates in each frequency band that we received for a search. That is to say, we essentially perform the same procedure as with sensitivity, but we require that a given h_0, that the critical ratio of 90% of a detected signal has to be greater than the maximum critical ratio in the frequency band in which the search was performed. So we compute upper limits in the following way: we inject 100 signals every 50 Hz in each configuration ($T_{FFT} = 2, 4, 8$ s) at 1, 2 and 5 hours, to be reflective of the search that was actually done. We only inject signals that each configuration analyzed: for example, we did not use $T_{FFT} = 8$ s to analyze frequencies above 1200 Hz 5 hours after the merger, so we do not injects at $f = 1200$ Hz in that configuration. We include in Figure 10-5 our upper limits on the amplitude h at 1, 2, and 5 hours after the merger. The maximum possible spindown was used when injecting these signals, because our theoretical sensitivity estimates show that the sensitivity is worse for signals with higher spindowns, so our ULs are conservative. We also randomized over $\cos\iota$, and injected signals with frequencies f to $f + 50$ Hz randomly.

Additionally we plot the upper limits obtained by all four pipelines, given in (*30*), in Figure 10-6. Each method had its own way of doing injections (different signal durations, parameters over which the analysis was randomized, etc.); however, each performs similarly. It is evident that our sensitivity was not good enough to see a remnant of a neutron star merger: our reach is between 0.5-1 Mpc and GW170817 was at 40 Mpc. This analysis would have been more successful if Einstein Telescope (*68*) or Cosmic Explorer (*69*) were online.

10.7 Prospects

No signal coming from a remnant of GW170817 was detected by our pipeline or others. But we knew this would happen before doing the analysis: the point was to be ready for future BNS mergers, especially if one happens closer to us. Moreover, third generation detectors are predicted to have high enough sensitivity to detect signals at 40 Mpc away: in fact, if Einstein Telescope was already built and on during GW170817, we may have been able to detect a signal from the remnant, if it was a neutron star.

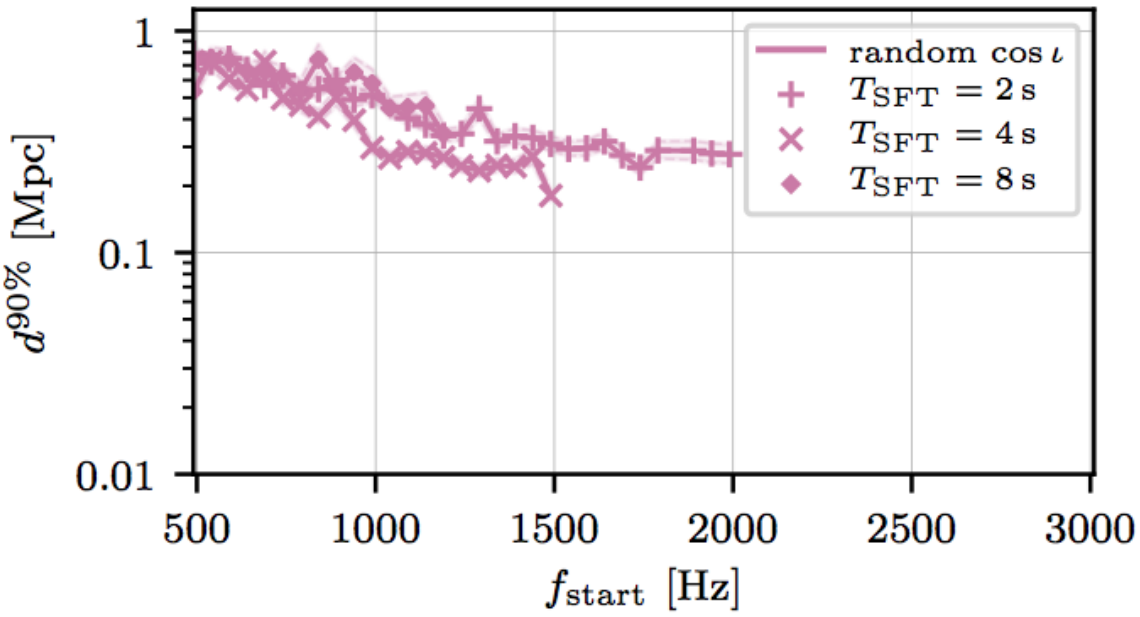


Figure 10-5: Upper limits placed in each configuration. Each point represents 100 injections done with the maximum possible spindown that an FFT can handle: $\dot{f} = 1/T_{FFT}^2$ between f and $f + 50$ Hz. The durations of signals analyzed are also different, ranging from about 1 hour to 18 hours. Higher frequencies tend to have longer signals, since the spindown decays quicker for lower frequencies. Note that the frequency on the x-axis is the frequency referenced to 1, 2, or 5 hours after the merger ($T_{SFT} = 2, 4, 8$ s respectively) The behavior of the upper limits is consistent with the sensitivity estimation in white noise, given in Figure 9-11.

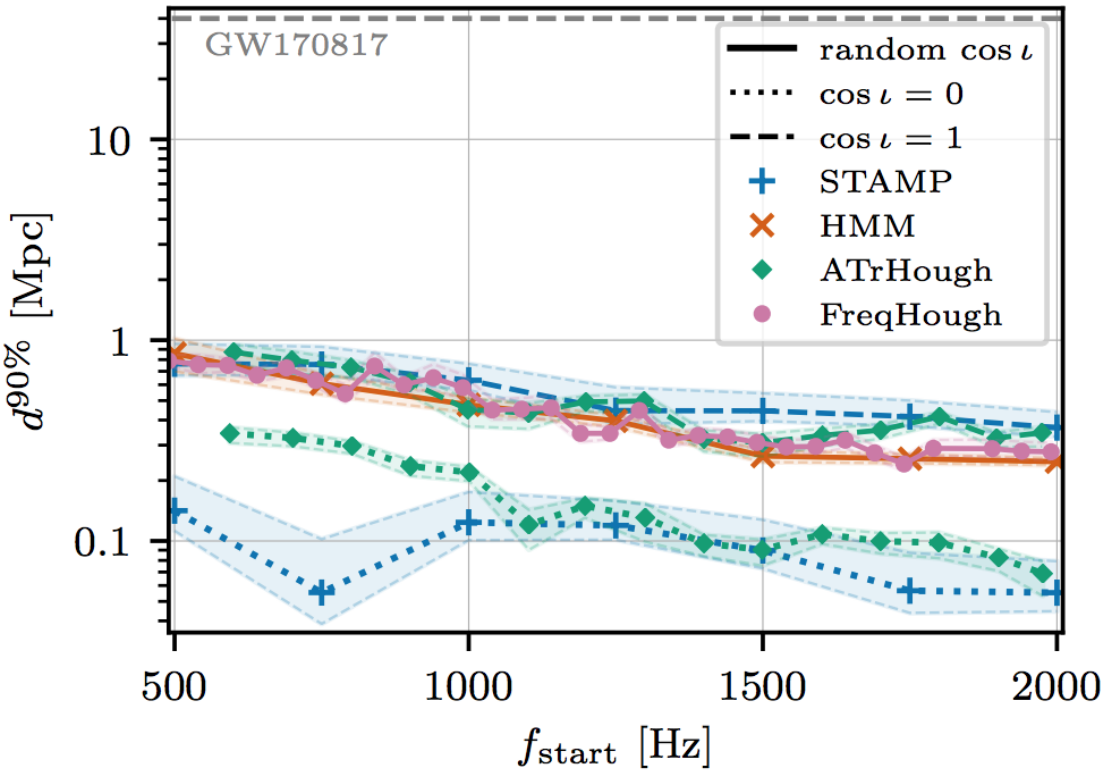


Figure 10-6: Upper limits placed by each pipeline that ran a search for a remnant of GW170817. Our method is labelled FreqHough. The methods all perform comparably; however, the merger happened at a distance that is beyond the reach of current analyses. If future detectors were online, we would have had a better chance of seeing a GW signal.

CHAPTER 11
TWO-DIMENSIONAL FAST FOURIER TRANSFORM LOW PASS FILTER

11.1 Introduction

If we think about the time-frequency map as an image, the r-mode waveform is a continuous, low frequency signal. The noise that surrounds the waveform is quite discontinuous, and can be expressed as a high frequency signal (many Fourier coefficients are needed to model the rapidly changing frequency of the signal). Hence, a low-pass filter would allow us to reduce the effect of the noise by filtering out the high frequency noise and enhancing the low-frequency r-mode signal. Our implementation of the low-pass filter is similar to matched filtering, which is a data analysis technique used to enhance the SNR of potential gravitational wave signals. However, matched filtering requires $O(10^6)$ templates of the signal because if the parameters of the filter are even slightly off, we miss the signal. Our filtering technique obtains an enhancement of the quality of the image even if the signal we are trying to find does not match the signal with which we are filtering. In Section 11.2, we describe the two-dimensional FFT image enhancement technique. Section 11.3 showcases some preliminary results we have obtained, showing that this technique can be used to enhance images. We then describe the future work that needs to be done in Section 11.4.

11.2 Method

To simulate the signals, we inputted amplitudes based on Equation 3–34 and frequencies following 3–33. We then integrated Equation 3–33 calculate the phase evolution of the signal. After this, we created a time/frequency spectrum by taking a one-dimensional FFT of the amplitudes- see the top panel of Figure 11-1.

The low-pass filter takes the form of a two-dimensional Fast Fourier Transform (2D FFT) of a simulated time/frequency behavior r-mode signal. The 2D FFT is defined as:

$$A(u,v) = F(a) = \Sigma\Sigma a(t,f)e^{-i(ut+vf)} \tag{11–1}$$

where $a(t, f)$ are the amplitudes of the signal in the time/frequency plane (the colors in the top panel of Figure 11-1), t is time, and f is frequency. u and v are the 'new frequencies' in Fourier space and don't have any physical meaning. The notation $F(a)$ denotes the two dimensional FFT of an image a, which is a matrix of amplitudes at each time/frequency.

When a low-frequency signal is transformed from the time-frequency plane to the Fourier space by a 2D FFT, it becomes concentrated at the corners of the new image, and the high-frequency signals in the original image are spread throughout the new image. To visualize this technique, we center the information that is actually present in the corners of the image, see the bottom panel of Figure 11-1).

To do the filtering, we use the following equation:

$$Y = F^{-1}(|F(X)| \times |F(M)|) \tag{11–2}$$

where the bars $|..|$ refers to taking the modulus of the image, Y is the new image, X is the original image that we want to enhance, and M is the filter (the ideal model of the signal).

We calculate the following factor before and after applying the filter:

$$A_n = \frac{\Sigma h^2}{M} \tag{11–3}$$

where Σh^2 represents summing the square of the amplitudes along the path of the signal in time-frequency space, and M is the median of the image. To calculate this factor after the enhancement, we normalize the image, apply a threshold of 0.9 to the image and then sum the square of the surviving time/frequency amplitudes. To calculate the enhancement, we then divide the two factors.

11.3 Preliminary Results

We have found that signals spinning up and spinning down in the time/frequency plane carve out an "ellipse" in the Fourier space, when looking at the centered image; hence, by combining multiple signals, we can cover much of time/frequency parameter space with only a few signals- see Figure 11-1. This enhancement resembles a "matched filter", in the sense that

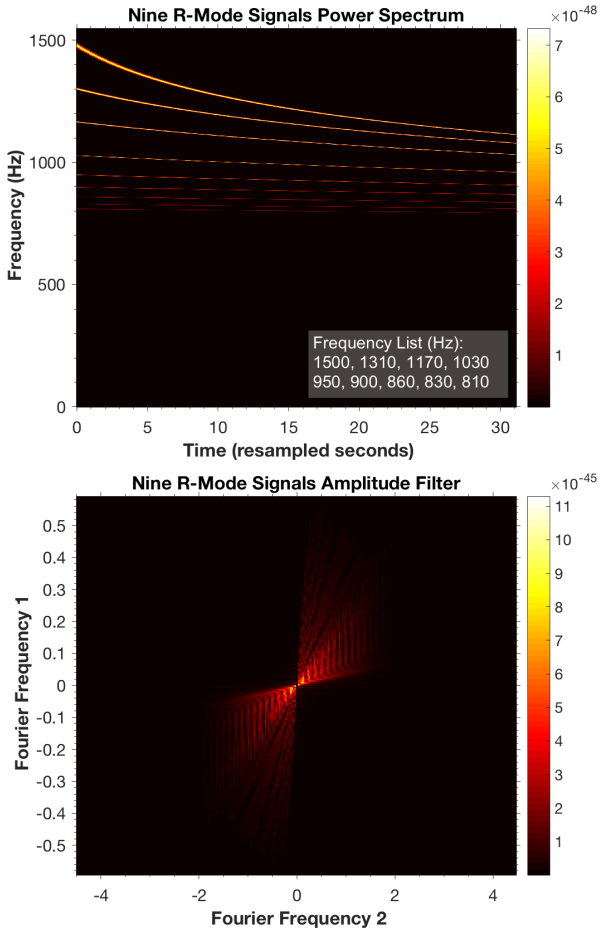


Figure 11-1: Top: time/frequency map of the ideal signals used as a filter. Bottom: the "ellipse" carved out in the new parameter space after performing a 2D FFT on the top image.

we are using our knowledge of the waveform of the signal; however, unlike matched filtering, we do not need to use the exact signal to obtain an enhancement (*85*).

Figure 11-2 shows an example of what the filtering technique does to noisy images. We see that the filter used (the right panel) does not match the signal we are searching for (the left panel). Yet after the filter, the signal looks better than it did before (middle panel), and is enhanced based on our calculation of the enhancement ratio. Additionally, we have quantified the level of enhancement that we receive using six filters that do not match the signal. See Figure 11-3.

We applied a filter containing nine different signals to the case where both an r-mode signal and a comparatively stronger glitch are present together. See Figure 11-4. We can see

clearly that the r-mode signal is enhanced relative to before, and the glitch remains at about the same strength.

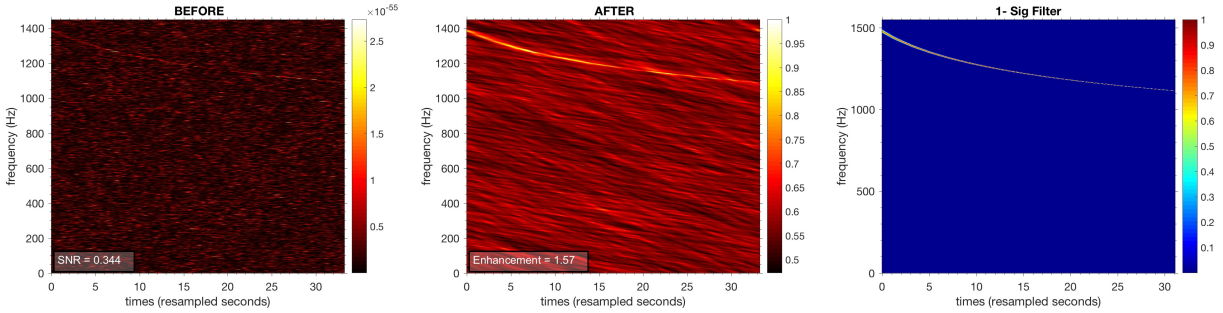


Figure 11-2: Left: r-mode signal before the filter is applied; middle: r-mode signal after the filter is applied; right: filter applied to transform left image to middle image. We see in the middle image that the r-mode signal is clearly enhanced, even though the filter we used did not match the signal we were looking for.

11.4 Outlook

The input SNR of the image determines how well the signal can be enhanced: if the signal is too strong, the signal is not enhanced so much, but the noise is, lowering the SNR. If the signal is too weak, then the filter cannot help us enhance the signal. This problem has been further explored and solved in (*86*) through a different calculation of enhancement. However, there is a range of input SNRs for which the signal can be enhanced by a factor of 2-8 based on how we calculate SNR_n. In the future, we would like to remove the threshold described in 11.2 and sum just along the line defined by only the signal to calculate the SNR_n. We have begun developing "masks" to perform this summation, whereby we are able to extract just the signal amplitudes corresponding to the actual signal frequencies from the image.

It is clear that a combination of filters enhances the image but the enhancement is low compared to filtering with one signal. Hence we anticipate using a combination of filters to obtain a coarse estimation of whether a signal is present in the data, then follow up this procedure with one-signal filters.

Additionally, filtering using Equation 11–2 is the simplest filter to use but not necessarily the best one, so we plan to experiment with other types of filters to see if they further enhance

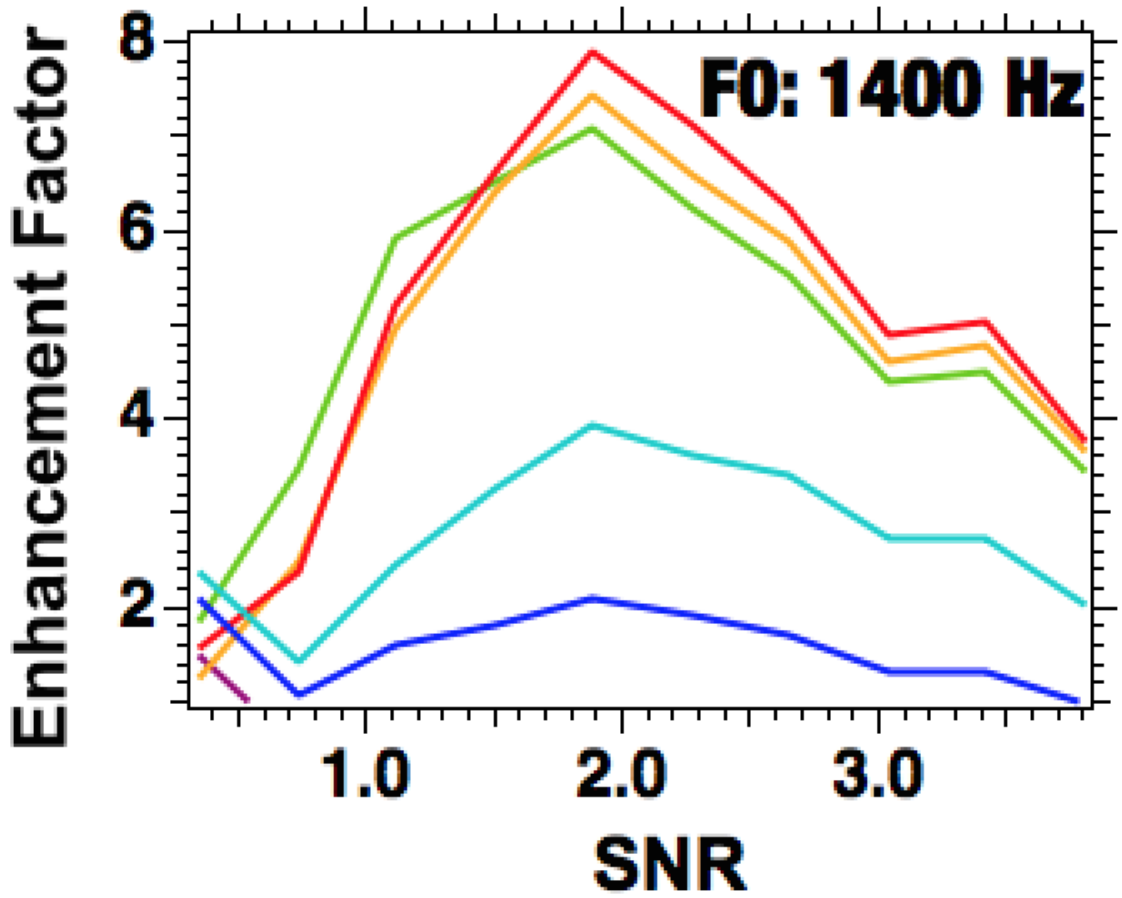


Figure 11-3: Enhancement of $f_0 = 1400$ Hz, $\alpha = 1$ signal using different filters. The colors correspond to the different filters used. Red: $f_0 = 1500$ Hz, one signal; orange: $f_0 = 1400$ Hz, 1 signal; green: $f_0 = 1300$ Hz, 1 signal; light blue: $f_0 = 515$-1500 Hz, 4 signals; dark blue: $f_0 = 500$-1500 Hz, 9 signals; brown: $f_0 = 1000$ Hz $\alpha = 0.1$, 1 signal .

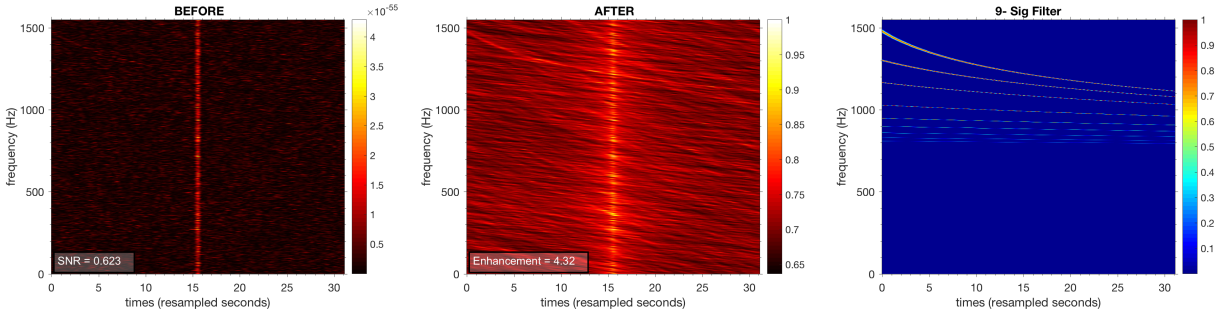


Figure 11-4: Left: a weak r-mode signal obscured by a glitch; middle: filtered image, showing r-mode signal has been enhanced; right: filter used to transform left image to middle image.

the images. For example, we see in Figure 11-3 that the 1000 Hz signal does not really enhance the 1400 Hz signal, mostly because the shapes of the signals in the time/frequency plane and the new Fourier plane after the 2D FFT are too different. However, some preliminary tests indicate that a more robust filter can be created such that signals that appear sufficiently different in time/frequency can still give enhancements. This work has been extended in (*86*).

CHAPTER 12
MACHINE LEARNING ALGORITHMS

12.1 Introduction

In the LIGO/Virgo collaborations, many methods have been developed to search for transient signals that last $O(s)$ and for continuous waves (CWs), quasi-infinite signals from spinning neutron stars (*87*–*89*). For intermediate duration signals lasting $O(hours - days)$, a cross-correlation-based method was developed some years ago (*90*, *91*). But now, two categories of these types of signals exist, requiring two different analysis procedures: (1) CW-like signals with still slowly varying frequencies (*92*, *93*) and (2) signals with large frequency variations. Only recently has the scientific community taken an interest in running a search for signals in the second category (*30*), which required the development of new methods, both modeled (*31*, *32*) and unmodeled (*33*). We expect that recently born neutron stars (i.e. after a BNS merger or supernova) will emit strong gravitational radiation on this timescale (*82*).

As the LIGO/Virgo detectors become more sensitive in the coming years, we should be able to not only detect more binary neutron star mergers, but also CWs from neutron stars. The all-sky and directed searches for CW signals are computationally expensive, since one must search the entire data set in every position in the sky. Even for long duration transient signals, signals that last for $O(hours - days)$, modeled searches such as those based on the Generalized Frequency Hough (GFH) (*31*) and the Adaptive Transient Hough (*32*) consume weeks of computing time, especially at intermediate spindowns, and require that the signal follow approximately a power law.

We therefore explore machine learning as a way to cut computation costs without losing sensitivity. The key idea of machine learning is that one gives it many examples of something that one wants to be characterized (hand-written digits, cats/dogs, gravitational-wave signals, etc.), and the algorithm learns how to distinguish between the different possibilities.

After sufficient training on different GW signals, MLAs can generalize and can perform well relative to existing methods. We describe the type of methods we have employed to

supplement the GFH: Artificial neutral networks in Section 12.2, convolution neural networks in Section 12.3 and random forests in Section 12.4.

12.2 Artificial Neural Networks

Artificial neural networks (ANNs) (*94*) are a biologically-inspired machine learning algorithm (MLA) that takes as input an image and outputs a statistic that determines if the image contains a pattern or not. It consists of an input layer, which is a collection of images as in Figure 12-1 that is flattened, one or more hidden layer with a certain number of nodes that allows nonlinear combinations of the pixels in the image, and an output layer. The hidden layer essentially assigns weights to each pixel in an image, and the weights are constantly updated throughout the training process as more examples are added. At this level, an activation function (hyperbolic tangent, softmax, or logistic) is used to introduce non linearity into the network. Typical activation functions include sigmoid (*95*):

$$\sigma(a_j) = \frac{1}{1 + e^{-a_j}}, \tag{12–1}$$

where a_j is the output of a layer in the network, softmax:

$$p(a_j) = \frac{e^{a_j}}{\sum e^{a_j}}, \tag{12–2}$$

and rectified linear unit:

$$f(a) = max(0, a) \tag{12–3}$$

The input data takes the form of a data matrix X, where each row corresponds to an unrolled image (an A x B-sized image becomes a $1\text{x}AB$ vector), and AB corresponds to the number of dimensions in the time/frequency map.

In standard neural network architectures, there is first the "forward pass" in the network (*96*), where a guess for whether an image has a signal or not is computed. We show an

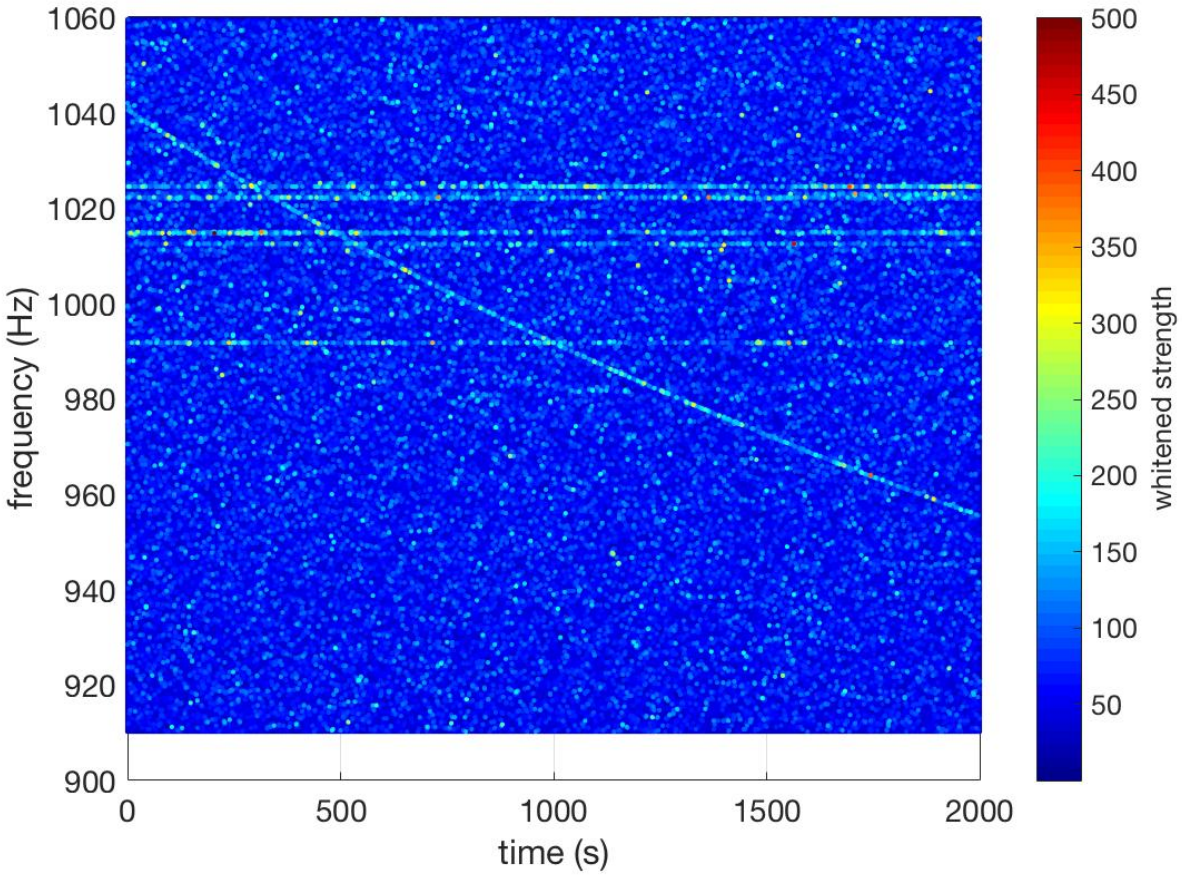


Figure 12-1: Injected signal with $f_0 = 1040$ Hz, $h_0 = 7.17 \times 10^{-23}$. This image (a matrix) is flattened into a row vector and given as input to the Artificial Neural Networks. We will see this image again in Figure 13-1.

example for one input image x_{1j}. To begin, the weight matrix $w_{kj}^{(1)}$ is multiplied by the unrolled image and a bias node $w_{k0}^{(1)}$ is added, and then this output is fed into Equation 12–1.

If we have N images with dimension $d = AB$, so an input matrix of N x $(d+1)$ dimensions, then the first weight matrix $w_{ij}^{(1)}$ has size $(n_{hidden} + 1)$ x $(d + 1)$. Matrix multiplication is performed between the input matrix and the first weight matrix, and then an activation function is applied to smooth the output. Each hidden layer in the network will have an associated weight matrix.

In the simplest case of one hidden layer, we would now have a matrix with size N x $n_{hidden} + 1$. The final step of the network requires another weight matrix $w_{ij}^{(2)}$ of size $n_{hidden} + 1$

x n_{class}, where $n_{class} = 2$ (signal or noise). Equation 12–4 is repeated again for this weight matrix, and the output of the network is a N x n_{class} matrix, where all reference to the hidden layer has been removed. Now, this matrix is fed into the softmax activation function 12–8.

Let us consider one training example. In equations, the forward propagation (*97*) of the networks looks like this:

From input to hidden layer 1:

$$x'_{1k} = \sigma\left(\sum_{j=1}^{d} w_{kj}^{(1)} x_{1j} + w_{k0}^{(1)}\right) \tag{12–4}$$

and from hidden layer 1 to output:

$$p_{1k} = p\left(\sum_{j=1}^{d} w_{kj}^{(2)} x'_{1j} + w_{k0}^{(2)}\right) \tag{12–5}$$

where p_{1k} is the outputted probability after the forward propagation. The error is then calculated between the returned and expected values for the image, and through back-propagation (*98*), the error is attributed in different amounts to each of the nodes in the hidden layer. The error per example δ_k is:

$$\delta^{(2)} = a_k - p_k \tag{12–6}$$

Then, this error is then converted into error on the weights $w_{ij}^{(1)}$:

$$\delta^{(1)} = w_{ij}^{(1)} \delta^{(2)} \frac{\partial}{\partial a_j} \sigma(z'_{1k}) \tag{12–7}$$

where z'_{1k} is given by the term inside the sigmoid function in Equation 12–4.

Changes to the weights are done according to which nodes were most or least responsible for the error at the output layer. The goal is to find the optimal weights such that the error on the validation set (a set of data used to avoid overfitting) is smaller than a certain value. A basic neural network can be seen in Figure 12-2.

The training data are typically divided into smaller sets, called batches, and each batch is pushed through the network. One can choose the number of times that the network should see the full dataset, which is called the number of epochs. If the network is trained too much, overfitting occurs, which means that the network learns the training data too well and can therefore not generalize to new situations. To prevent this, an early stopping criteria can be employed: if the sum of the mean squared error between the outputs of the forward propagation within each epoch continues to decrease (until a threshold on minimum error is reached), we continue to train the network (*47*).

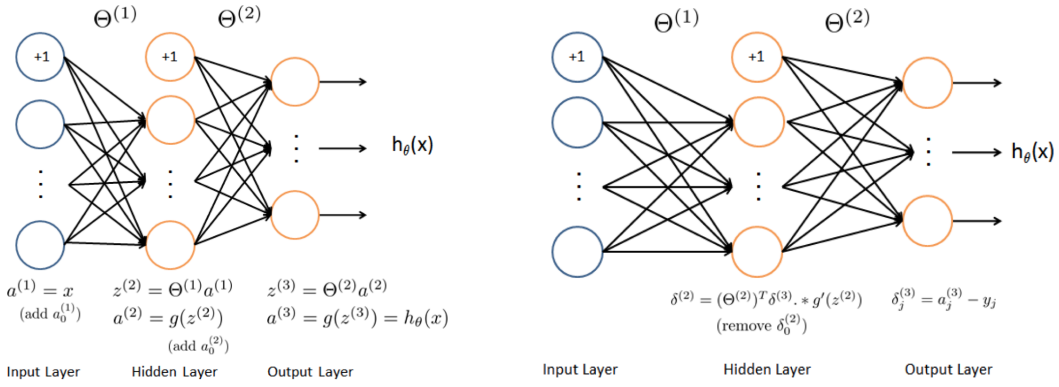


Figure 12-2: This diagram shows the layout of an artificial neural network. The left-hand panel shows forward propagation through the network, while the right panel shows back propagation. The $\Theta_{1,2}$ are matrices of weights that transform the input images into a guess as to which class an image belongs. g is the sigmoid function. The Δs represent the errors that are propagated back through the network after a guess is made in order to adjust the matrices of weights. This was taken from (*99*).

12.3 Convolution Neural Networks

Convolutional neural networks (CNNs) are essentially ANNs but instead of using the typical matrix dot product to pass between layers, the image is convolved with a weight matrix called a "kernel". The kernel moves over subsets of the image and returns a measure of overlap between the kernel and the image, before this output is fed into an activation function.

The principle of determining the ideal weights for ANNs at each layer remains the same for CNNs, except now one can use many convolutional layers, different activation functions, a random pruning of the network, etc. in order to create a better classifier.

The motivation to use CNNs in gravitational wave physics come from their quick and successful implementation in other fields. For example, AlexNet (*100*) was able to quickly train (on a Graphics Processing Unit (GPU)) to classify images into 1000 separate classes with a much smaller error rate than previous algorithms. This work revolutionized the field of artificial intelligence, and made deep learning possible on a reasonable timescale. CNNs and other deep learning algorithms can be found in almost every field now, from speech recognition (*101*) to medical imaging (*102*).

We implement a CNN shown to work for both supernovae (*103*) and CW signals. The key aspects of this CNN are 5 convolution blocks, described in detail below, each containing zero-padding, convolutions, ReLU and max pooling steps, and 2 fully-connected and dropout layers followed by a softmax layer at the end.

1. Input (N x 62 x 37 x 1)
2. 5 convolution blocks
    - ZeroPadding 2D
    - Convolution 2D (3x3x16)
    - ReLU
    - MaxPool(2x2)
3. Flatten
4. 2 blocks of:
    - Fully-Connected (16 nodes each)
    - Hyperbolic Tangent
    - Dropout
5. Fully-Connected (2 nodes)
6. Softmax

7. Output

The zero-padding layer ensures that the size of the image stays constant throughout the convolution block by surrounding the images with zeros. This is necessary because the maximum pooling layer takes the maximum value of every non-overlapping 2x2 block of the image, effectively cutting the size of the image in half each time.

The flatten layer unrolls an H x W x N kernel into a vector of length $H \cdot W \cdot N$.

In the fully connected layers, each output of the previous layer is used as in input to each of the 16 nodes in this layer. These layers are essentially what each hidden layer in an ANN is. The last fully-connected layer in our CNN outputs a vector whose size is just the number of classes in which we are trying to classify: 2 in this paper (signals or noise). The values correspond to the CNNs' confidence in classification of the image.

The dropout layer is used to prevent overfitting of the network. It essentially zeroes the activation of randomly chosen hidden nodes of the model at each iteration of the learning phase, akin to pruning a decision tree.

To convert the "confidence" of the CNN into a probability distribution P_i, the final layer is a softmax activation function:

$$P_i = \frac{e^{x_i}}{\sum_{n=1}^{N} e^{x_n}}, \tag{12–8}$$

where N is the number of classes (signal or noise) and x_i are the outputs of the fully-connected layer (essentially the result of convolutions of the input image with the kernel).

12.4 Random Forests

In machine learning, decision trees are a predictive method for classifying data. They are created to make a series of decisions based on the parameters of the input data. The trees grow in size and complexity in order to fully separate data points that belong to one class or another. However, individual decision trees are likely to overfit the data. Therefore, we employ a "forest" of trees, where each tree is as independently generated as possible in order to obtain

a monotonic increase in classification accuracy as more trees are added to the forest (*104*, *105*). We describe in Subsection 12.4.1 how we construct our random forest.

We can see in Figure 12-3, taken from (*106*), an example of a random forest composed of 5 decision trees. The two classes (signal and noise) are defined by the two colors light blue and black, and at each numbered node, a decision is made to split the training set, given in the trunk of the tree, into different leaf nodes. For our purposes, each of these nodes will be further split, until each leaf contains only blue dots or black dots.

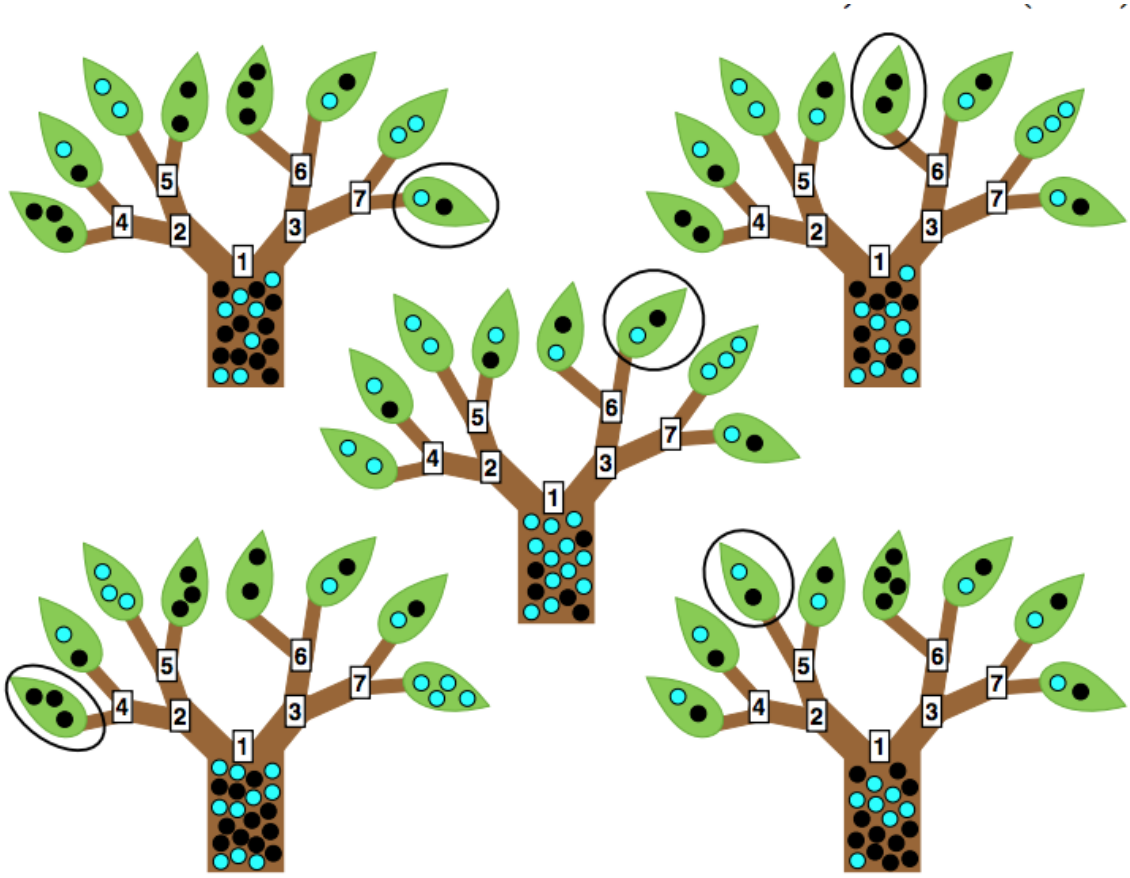


Figure 12-3: A random forest composed of 5 trees (taken from (*106*)). The training set, in the bark of the tree, is composed of two classes: light blue dots and black dots, representing noise and signal. At each numbered node, a decision is made, and the data are split. In our random forests, we will actually continue this figure until all leaf nodes contain only instances of one class, so each leaf will have exclusively blue dots or black dots.

12.4.1 Constructing a Decision Tree

To make a decision tree, input data and labels are required, as in the case of ANNs/CNNs. But one additional piece is required: a splitting function. The splitting function determines in what way to separate the training data into the two classes we would like to identify, and this happens at each node of the tree.

To begin, we take the entire (down-projected) data matrix X and give it to the "root node" of the tree. The splitting function is then applied, and X is split into two groups X_1 and X_2 and y is split into y_1 and y_2 based on some criterion (described in a subsequent section). Then the two individual groups are again given to the splitting function, and more nodes are created until the classes are fully split, that is, until each leaf node contains elements from only one class.

12.4.2 Constructing a Forest with the Random Subspaces Method

The random subspaces method (RSM) was proposed in (*104*, *105*) to generate independent decision trees that will have 100% accuracy on their corresponding training sets. Previous implementations had relied on randomly selecting entire training examples and giving them to different trees; however, it is not guaranteed that each tree will classify the entire training set correctly. In the random subspaces method, each feature vector is down-projected, meaning a certain number of features are randomly selected and the rest of the features are set to 0 in the unselected dimensions. For example, if we want to create a subspace of 8 dimensions, and we have a 16 dimensional feature vector:

$$f = \begin{bmatrix} 1 & 2 & 3 & 4 & \dots & 16 \end{bmatrix}$$

one tree will randomly select

$$f = \begin{bmatrix} 5 & 2 & 6 & 8 & 3 & 1 & 16 & 15 \end{bmatrix}$$

and another tree in the forest will select

$$f = \begin{bmatrix} 4 & 3 & 1 & 5 & 11 & 9 & 15 & 14 \end{bmatrix}$$

All training examples (all feature vectors) are down-projected into the desired subspace dimension before being sent into the tree. Each tree will be the same number of dimensions, but will live in randomly different subspaces. It was shown in (*105*) that these trees will have different generalization errors and be accurate on the entire (down-projected) training set.

12.4.3 The Splitting Function: Central Axis Projection

Central axis projection (CAP) (*104*) is a fast and feature-independent algorithm that has been used to make splits in decision trees. Many conventional splitting techniques randomly select a single feature to split at each node, which does not allow us to take into account any relationships that exist between the individual peaks. To begin, all training examples are simultaneously given to the root node. The classes whose means are farthest apart by Euclidean distance are found, and then a line is drawn between the coordinates of those means. All the data are then projected onto this line, called the central axis, following the standard equation (projecting $\vec{a}$ onto $\vec{b}$):

$$P_b^a = \frac{\vec{a} \cdot \vec{b}}{|b|^2}\vec{b} \tag{12–9}$$

We start at the beginning of the line, and take discrete steps, placing a hyperplane at each step that is perpendicular to the line. We then count the number of points on each side of the hyperplane, and choose the hyperplane that minimizes the number of points that are not on the side of the hyperplane where the majority of that class's points fall. This choice of hyperplane causes the split in the dataset, separating some points correctly and some not. Each daughter node of the tree will then perform the same selection of hyperplane, and the process will continue until the tree is fully split. For example, if the feature space

is two-dimensional, a line is drawn perpendicular to the central axis at each step; in a three dimensional space, a plane is drawn- see Figure 12-4 for an example.

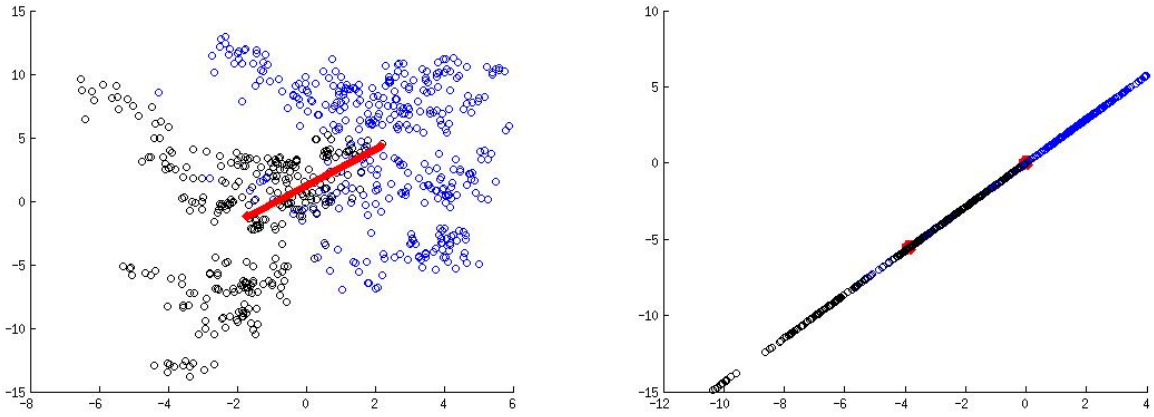


Figure 12-4: Left: original data in a 2D subspace with central axis drawn between means of the two classes. Right: all data in this 2D subspace have been projected onto the central axis- its endpoints are shown in red. In a finite number of steps, we will move from the first endpoint to the second one, determining first which side the majority of each class's points fall on. Then we minimize the number of points on the wrong side of the hyperplane.

12.4.4 Voting and Predicting

Once many trees have been created in a forest, we need a way to combine their estimates on new testing data, that is, their classification if the new data have a gravitational wave signal or not. Each testing feature vector is down-projected into the corresponding subspaces of each tree in the forest, and since the trees are fully split, each feature vector will end up at a leaf node corresponding to whether it is a signal or not. After the feature vector is sent down each tree, we count the number of votes for signal and the number for noise. In general, the class with the greatest number of votes wins, but in this case, it is simply that the class with the majority of votes wins.

CHAPTER 13
CHARACTERIZATION OF CONVOLUTIONAL NEURAL NETWORKS AND RANDOM FORESTS

13.1 Introduction

MLAs have been shown to achieve extremely good sensitivities in both white and real noise, comparable with matched filtering and some unmodeled methods (*103*, *107*–*111*). However, what is lacking from these studies is a quantitative analysis of (1) how to train, (2) how much to train, (3) robustness towards signals on which the networks were not trained, (4) differences in architectures' effects on detection efficiency/false alarm probability, and (5) applying the networks to a real search. Moreover, these algorithms have only focused on BBH signals or supernovae, and have not yet been expanded to search long duration transient signals from isolated neutron stars. We address these points for ANNs, developed in (*47*) and CNNs, developed in (*103*). We characterize the performance of the networks under a variety of conditions, and propose a possible design for a search using CNNs to provide triggers that would be followed up by another method. We also compare the ANNs and CNNs to an established method to search for long duration transient signals, the GFH. The work in this chapter is based heavily on (*112*).

In Section 13.2, we define what are the inputs to the algorithms. Then we quantify how much we should train as a function of source parameters in Section 13.3. We experiment with using a forest of ANNs in Section 13.4. Next, the robustness of the methods towards different signals on which they were trained is quantified in Section 13.5. The detection statistic we use to determine success of the algorithms or not, as well as the effects of non-stationary noise on the algorithms' performance, is described in Section 13.6. The error associated with efficiency and false alarm probability is empirically determined in Section 13.7 for a set of $O(1000)$ models. A comparison is made between ANNs/CNNs and the GFH in Section 13.8. Then we design a possible hierarchical search using CNNs as a way to quickly generate triggers to be followed up by the GFH in Section 14.2. Finally we do a quick analysis of decision trees' ability to detect GW signals in Section 13.9.

13.2 Construction of Input Images: Reduced Resolution Time/Frequency Maps

The pattern on which we train the networks is in the time/frequency domain, so we first construct a time/frequency map by fast Fourier transforming each chunk of data (interlaced by half and windowed). Then, we estimate the noise power spectral density by using an auto-regressive spectrum to whiten the data (*77*). Afterwards we plot the result in the time/frequency plane, as shown in the left panel of Figure 13-1. Ideally, the more information that we can give to the networks, the better they can learn to identify and distinguish signals from noise. However, there are too many pixels on which to attempt to train the ANNs/CNNs, so we reduce the resolution of each time/frequency map in an analogous way as done in (*47*). We split the time/frequency map into different square blocks and take the maximum value that appears in each 16 x 16 block. The size of this box was chosen based on an optimal resolution reduction factor determined in (*47*). Additionally, we apply a threshold to each local maximum we select: if it is less than a threshold of 2.5 (this choice of threshold is described in (*78*)), we put a 0 at this point in the reduced time/frequency map. The value 2.5 is equivalent to selecting pixels that are at least 2.5 standard deviations away from the mean level of noise in the map. Selecting a local max and applying a threshold gives us the so-called peakmap, constructed in a similar way as described in (*78*), but our method of resolution reduction ensures that the peakmaps have the same size, which is necessary for the MLAs to work well. In Figure 13-1 we show the before and after resolution peakmaps.

13.3 Training Efficiency

To run a real search, we need to train the CNNs to recognize both signals and noise. However, when we train on noise maps, we have instructed the networks to see these time/frequency maps as not having signals. We would therefore have to throw out this data, so we wish to minimize how much data we need to use for training. Thus we quantify at what point additional training of CNNs no longer adds appreciably to the detection efficiency. Figure 13-2 shows a CNN trained with different numbers of injections and noise maps on real data. We apply a threshold on the probability p_{thr} outputted by Equation 12–8: $p_{thr} = 0.9$, chosen

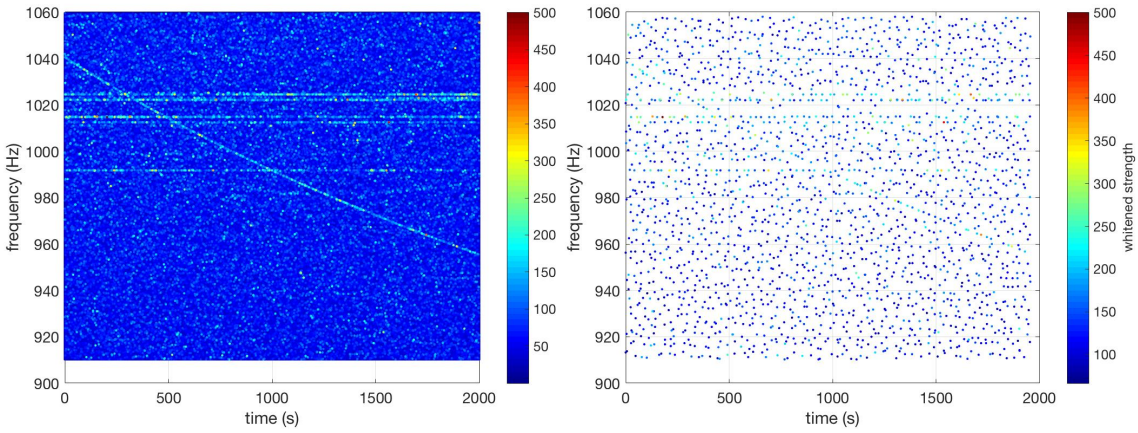


Figure 13-1: Left: injected signal with $f_0 = 1040$ Hz, $h_0 = 7.17 \times 10^{-23}$. Right: reduced resolution peakmap by a factor of 16 on each axis. This reduced peakmap is the input to the machine learning algorithms. When we reduce the resolution of the image, we pick the pixel in each 16x16 block that has the highest value, but if this value is less than a threshold (2.5, essentially not higher than 2.5 standard deviations from the mean), we zero this pixel.

based on the false alarm probability analysis done in Section 13.6 . Note that $p_{thr} = 0.9$ does not imply a detection efficiency of 90%; rather, p_{thr} is our detection statistic. Even for training on ~ 20000 injections, we see that the detection efficiencies are comparable to training on ~ 100000 injections. This appears to be independent of signal strength. However, when training on fewer injections, we are in fact training on fewer noise maps, which means the false alarm probability is higher. These results depend on which p_{thr} we use, but they offer a guideline for training a single network on a variety of amplitudes.

13.4 Using a Forest of Neural Networks

A single ANN was able to achieve high detection efficiency at the expense of a high false alarm probability (*47*). It is therefore necessary to consider ways to lower this false alarm probability. One way is to train on more noise maps, but there are practical limitations to this since the length of an observing run is finite, and we do not want to lose too much data. The solution we propose is to use a forest of ANNs that vote on whether each map contains a signal or just noise. But naively training the entire data set on each ANN will not decrease the false alarm probability, so we must consider that each ANN should see only an independent portion of the image to train on. Down-projecting the image into a subspace allows us to

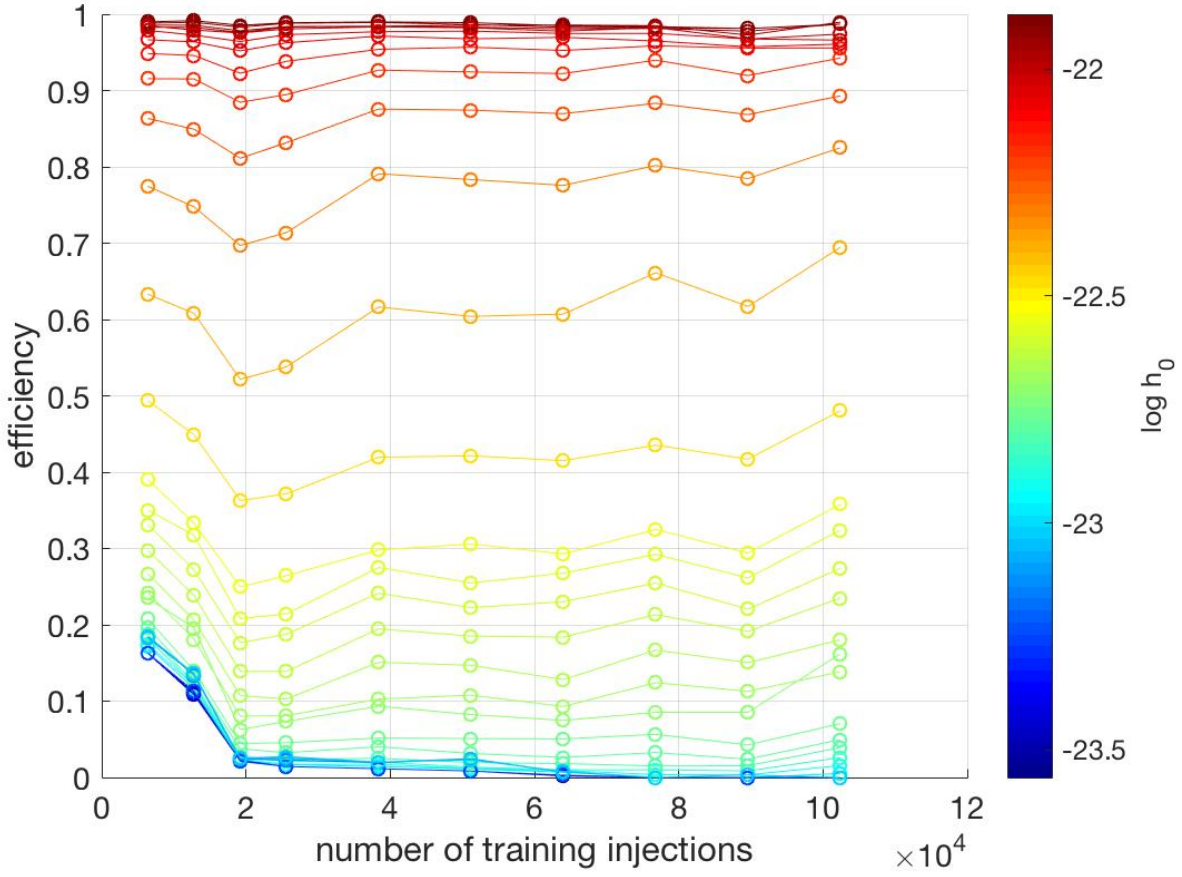


Figure 13-2: A CNN was trained on a variety of amplitudes in real noise with a different number of signal injections. It appears that the detection efficiency does not change much as the number of training injections is increased. Therefore, it is not necessary to do many injections to achieve high detection efficiencies. For this plot, $p_{thr} = 0.9$. Each of 32 amplitudes is represented by a line in the above figure.

preserve accuracy on the training set and to generate independent ANNs. While the detection efficiency does not change that much, the false alarm probability drastically improves- see Figure 13-3.

13.5 Robustness Towards Signals Not Trained on

Machine learning methods have become useful because they can be trained to identify certain patterns, and can also generalize to find similar ones, but we must quantify if and how the detection efficiencies will change. The signals we have described thus far have had constant braking indices, which means that the GFH was a strong method to search for them.

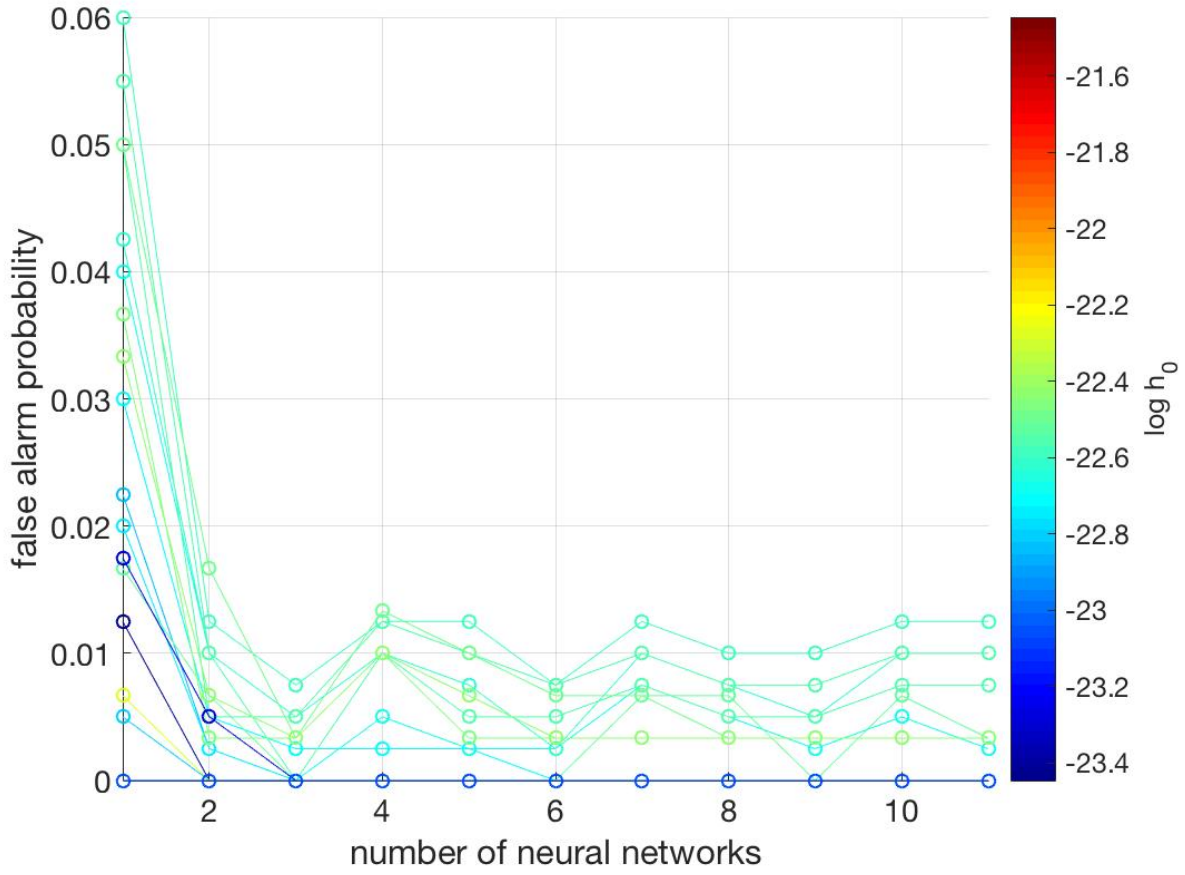


Figure 13-3: False alarm probability as a function of the number of neural networks we use in our forest for different amplitudes in white noise. The false alarm probabilities were calculated with a probability threshold of 0.7. A forest of neural networks was trained for each signal amplitude. It is clear that even with the addition of 1 or 2 independently trained neural networks, the false alarm rate greatly decreases. At this probability threshold, a single neural network trained on the full images has a false alarm rate $\sim \quad 10$ times greater than a forest of neural networks each trained on a randomly chosen portion of the image.

However, if the braking index varies too quickly in time, the GFH cannot recover the signal, no matter how strong the signal is, because the signal deviates too much from the power law model in Equation 3–33. Figure 13-4 shows two sets of sensitivity curves for a CNN trained on signals with a braking index of $n = 7$. In the left panel, the curves are the result of testing on injections with fixed braking indices ($n = 5$ or $n = 7$). The CNNs performed worse at higher frequencies, which is consistent with the sensitivity curve of the LIGO detector. In the right panel, the CNNs were tested on injections with varying braking indices, with variation $\delta n/\delta t$

between $[-10^{-4}, 10^{-4}]$ /s. In total, this variation corresponds to a change $\Delta n = T_{obs}\frac{\delta n}{\delta t} = 0.1$: the GFH cannot detect signals with such large Δn. Comparing both plots in Figure 13-4, we see that despite being trained only $n = 7$ in the $[600, 750]$ Hz band, the CNNs can detect signals with or without large $\delta n/\delta t$ and have comparable efficiencies across all frequency bands.

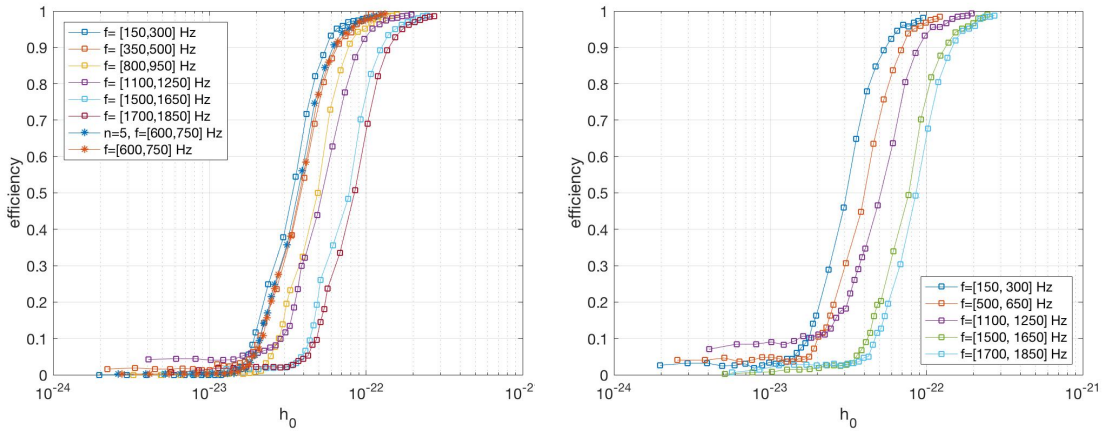


Figure 13-4: A CNN was trained on injections with $n = 7$ and tested on 16000 (500 injections/amplitude) with fixed braking indices (left) and varying braking indices (right) in different frequency bands. The random variation in braking index is between $\delta n/\delta t = [-10^{-4}, 10^{-4}]$Hz/s. This variation over the observation time is too large for the GFH to recover a signal. A threshold on probability $p_{thr} = 0.9$ was used in both plots. The sensitivity curves for the different frequency bands are consistent with the detector's sensitivity curve (we are less sensitive at higher frequencies). The CNNs can obtain comparable sensitivities in both the fixed and varying braking index cases. However, even after cleaning both known and unknown noise lines, the threshold on probability must be high for some frequency bands to achieve a false alarm probability of 1%- see Figure 13-5.

13.6 Impact of Threshold Choice on False Alarm Probability

The output of the CNN can be interpreted as a probability that a time/frequency map contains a signal or not. If we wish to decrease the false alarm probability, we can use a threshold that is higher than $p_{thr} = 0.5$. To determine an optimal threshold, we test our network on 500 noise maps generated at random times. As shown in Figure 13-5, to guarantee a false alarm probability of 1%, we should have $p_{thr} = [0.7, 0.95]$ depending on the frequency band. We also varied the number of noise maps used for training, and found that $\sim$ 23 days of data (2000 noise maps) was enough to obtain a false alarm probability of $\sim 1\%$ at $p_{thr} = 0.9$.

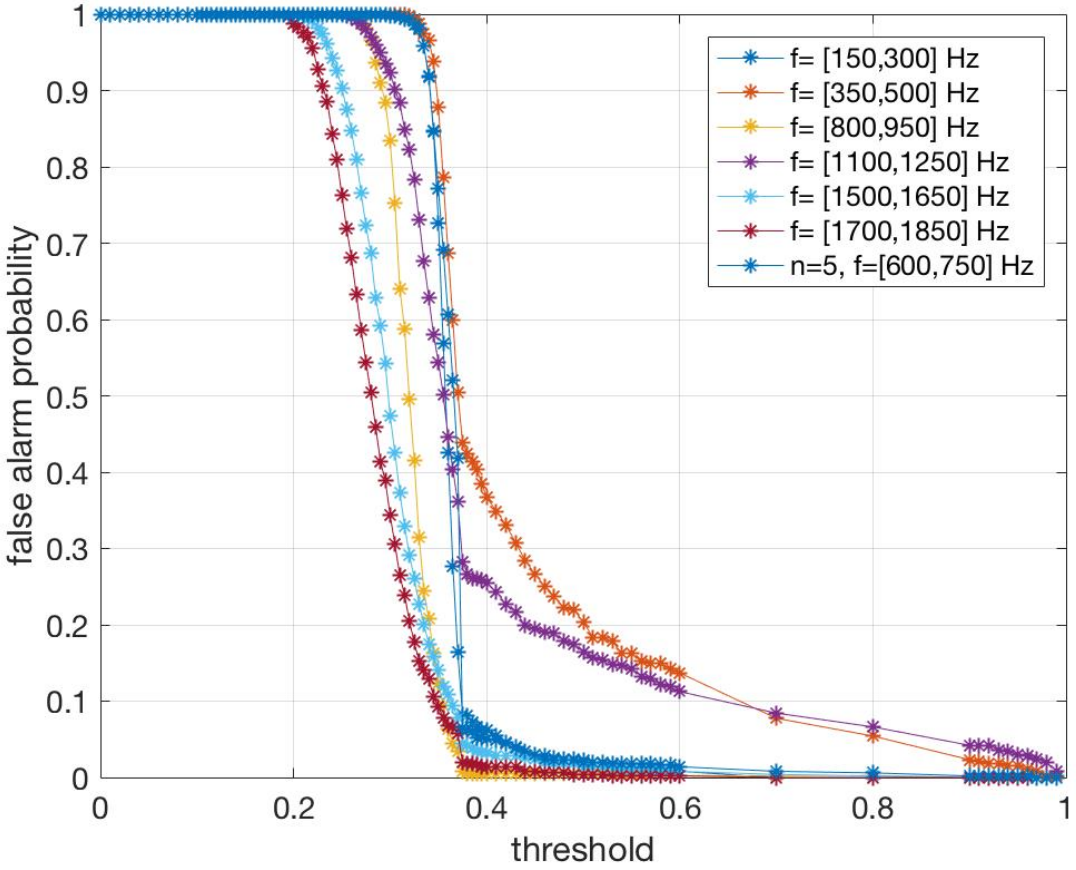


Figure 13-5: We plot the false alarm probability as a function of the threshold on the output of a CNN (Equation 12–8) for 500 noise maps per frequency band, which correspond to $\sim$ 11.5 days of new data. The CNN was trained on 2000 noise maps (23 days of data). Different frequency bands, which have different levels of noise, require different p_{thr} to achieve the same false alarm probabilities.

13.7 Error Analysis

The CNNs operate as a black box: we control the input, the structure of the network and the output. Because of this, there is a good degree of randomness in the possible outputs of the network, meaning that each measurement we make has an associated error. This randomness is caused by the dropout layer, which throws away half of the nodes in the network, and the fact that the network sees randomly shuffled data in different batches in each training session. We therefore train a network 1255 times with the same input data in order to quantify the range of measurements of both detection efficiency and false alarm probability.

We discovered that the errors do not obey a binomial distribution for $p_{thr} = 0.9$. For the [600, 750] Hz frequency band, we train on 38400 injections and 1200 noise maps. The sensitivity and false alarm probability curves are shown in Figure 13-6.

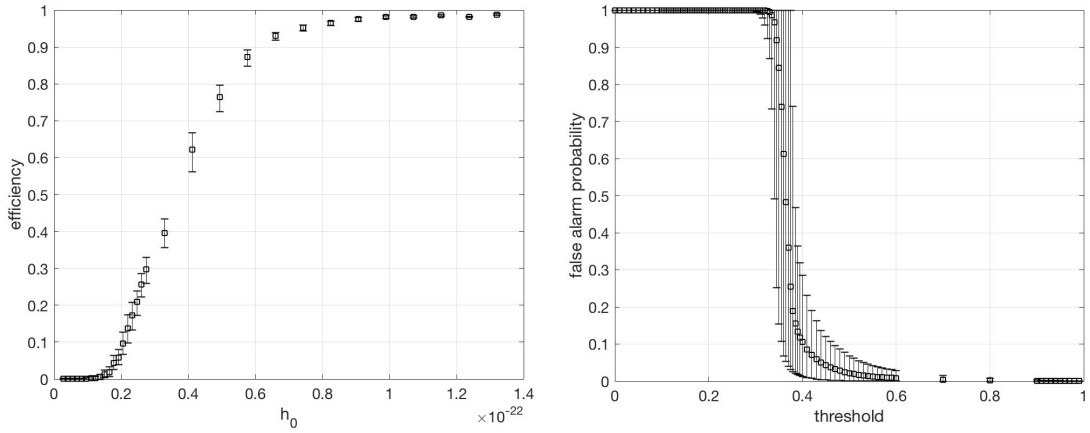


Figure 13-6: 1255 CNNs were trained on 38400 injections with $n = 7$ and 1200 noise maps and tested on 44800 (1400 injections/amplitude) and 1400 noise maps. The error bars represent the 90% confidence level. Left: efficiency is plotted as a function of amplitude. Intermediate strength signals seem to have the greatest spread in error. Right: false alarm probability is plotted as a function of threshold. The errors are largest around 0.4, and show that higher thresholds are necessary to truly guarantee low ($\sim 1\%$) false alarm probabilities.

13.8 Comparison to Generalized Frequency Hough

We have compared the detection efficiencies of our CNNs to those of the GFH (*31*), to AlexNet CNNs and to ANNs, developed in (*47*). In Figure 13-7, we plot the sensitivity curves for each of these four methods, where we have used $p_{thr} = 0.9$ for the ANNs and both CNNs. The ANN performs the worst, not even reaching 90% efficiency, while both architectures of CNNs and the GFH achieve similar efficiencies. The CNNs are even slightly more efficient than the GFH at some amplitudes. However, the CNNs cannot perform parameter estimation and the false alarm probabilities are about 100 times greater for both CNNs than for the GFH, see Figure 13-5. A "detection" for the GFH is defined as when the parameters $x_0 = 1/f_0^{n-1}$ and k of the loudest candidate in the map are within a distance of 3 bins of the injected signal. For the CNNs, a detection occurs when the returned probability is greater than $p_{thr} = 0.9$.

We also show in Figure 13-7 the ideal efficiency of a CNN (green curve). This is the best-case scenario because we train and test the CNN at a fixed GPS time, meaning that the CNN learns perfectly the noise structure and hence the false alarm probability is 0. The green curve represents essentially an upper limit on the CNN's sensitivity. In contrast, the red curve corresponds to the realistic case, where the CNN is affected by the non-stationary noise of the LIGO/Virgo detectors. This result shows that we must be aware of the changing noise in the detector when determining how efficient a CNN will be in detecting power-law signals.

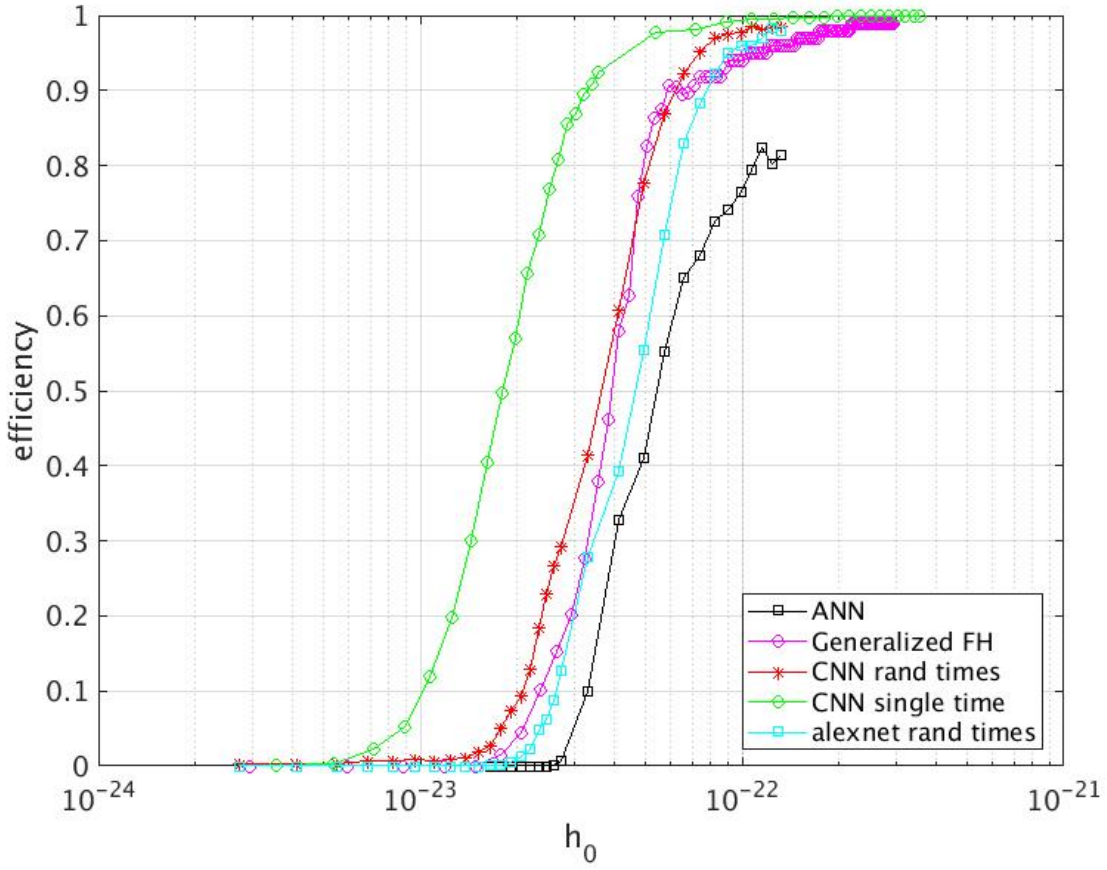


Figure 13-7: We show sensitivity curves for an ANN, the GFH, a CNN and AlexNet CNN trained on 2000 injections/amplitude and tested on 1000 injections/amplitude in real noise. $p_{thr} = 0.9$. It appears that different networks obtain similar efficiencies to the GFH, though the false alarm probabilities vary: $\sim 1\%$ for the ANN/CNNs, and $\sim 0.01\%$ for the GFH. The green curve, corresponding to a CNN trained and tested at a fixed GPS time, represents what the efficiency would be if the noise was the same throughout the run.

13.9 Preliminary Characterization of Random Forests

The previous work on MLAs (*47*) resulted in high false-alarm probabilities using neural networks, support vector machine (SVM), and constrained subspace classifiers. Random forests are advantageous here because the trees separate features from each training example until no further separation is possible, growing in complexity as the data grow in complexity. It was shown that a random forest generation package could be used to search for GWs from GRBs, and that the forest could generalize well when signals of different morphologies are used in the testing set, with good false alarm rates (*113*). Moreover, random forests, and specifically the random subspace method (explained later), allow us to use all the data in the creation of each tree- no resolution reduction of the initial time/frequency maps is needed. The number of subspaces available to down-project into grows with the dimensionality of the data (2^d, where d is the dimensionality of the data). As we increase the number of trees in the forest, we are exploring all parts of the feature space in a random, non-redundant way. Additionally, the random forest can be generalized to use other types of splitting functions, for example, a full SVM at each node. The splitting function that minimizes computation time to create, store (and predict with) each tree, and that maximizes detection efficiency, could be investigated. Random forests provide a convenient way to explore different types of MLAs as splitting functions while preserving generalization accuracy.

We have trained and tested our random forest algorithms on injected signals in white noise with the following parameters: durations about ~ 1700 s, $f_0 = 777$-945 Hz, $\alpha = 10^{-2}$-10^{-1}, $d = 10^{-2}$-10^{-1} Mpc, $SNR \sim$ 10-260, $n = 5$, size of training set$= 1800$, and size of testing set$= 200$. Each tree was created in a different 10 dimensional subspace of the feature space, whose resolution was 585 dimensions. Originally, the resolution of the image was much higher, O(10,000) dimensions; however, the random forest has a hard time working with such large data streams, since the data are all given to the root node of the tree at the same time to be split. In Figure 13.9, we show the results of our analysis. It is clear that as the number of trees in the forest increases,the detection efficiency improves. But note that each

tree only analyzed such a small (random) portion of the data, and still a detection efficiency of $\sim 90\%$ was achieved. This is primarily due to the fact that a portion of the testing sample was composed of signals that were too weak to be seen by any algorithm.

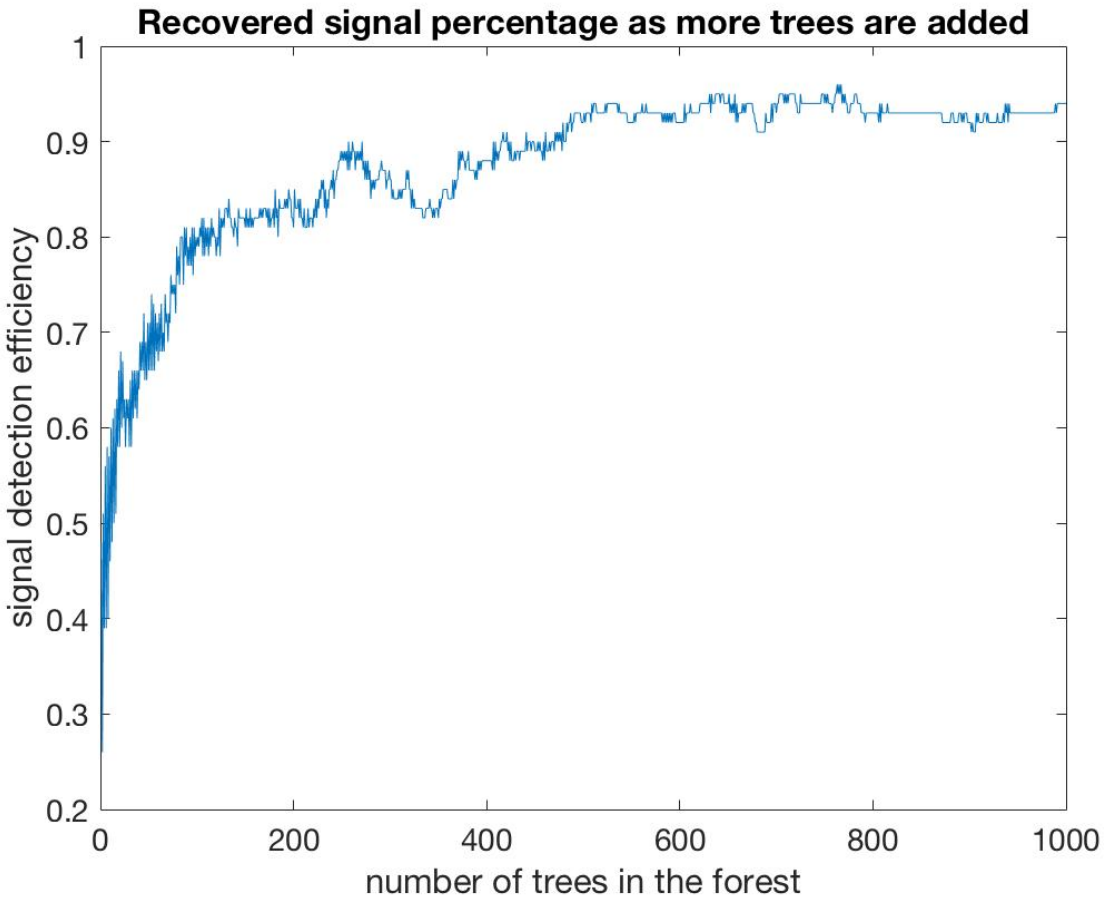


Figure 13-8: Detection efficiency of the forest as a function of the number of trees present. The leveling off at 90% occurs because there were some signals analyzed by the random forests that were too weak to be seen by any algorithm.

CHAPTER 14
MACHINE LEARNING-BASED SEARCH FOR A REMNANT OF GW170817

14.1 Introduction

To test our procedure, we ran a search on one week of O2 Livingston and Hanford data after GW170817, spliced in 2000 s time/frequency maps, following the search design outlined in the previous section, to look for a remnant of this binary neutron star merger. Though we did not expect to find a signal, we wanted to ensure that we are prepared for future binary neutron star mergers.

Performing this search took about 30 minutes. The CNNs took about 2 minutes to analyze one week of data from [100,1900] Hz and to generate triggers. Then, the follow-up of those 50 triggers by the GFH took about 15 minutes, when running this follow-up on 14 cores. Finally, the follow-up for the candidate remaining after the GFH took less than a minute.

In Section 14.2, we describe the design of our machine learning/ GFH search for a remnant of GW170817. Afterwards in Section 14.3, we display the portion of the parameter space that our search covered. We discuss the results of the search in Section 14.4, and given that no candidate signals were found, we place upper limits on the possible GW emission from a remnant for both fixed and time-varying braking indices in Section 14.5. These upper limits are consistent with those found in (*30*).

14.2 Search Design

We plan to use the CNNs in a real search in order not only to speed up the analysis, but also to probe a portion of the parameter space to which the GFH is not sensitive. It seems that the CNNs are of comparable efficiency to the GFH, so they can be used to quickly analyze time/frequency maps to determine if a signal is present or not. CNNs can analyze a time/frequency map in a microsecond, while the FH and GFH could take ~ 1 minute (*31*, *78*). If the map contains a signal, we can then apply the GFH to estimate the parameters of the signal. Moreover, the CNNs are robust towards varying braking indices and different frequency

bands with a limited amount of training data. This means that the CNNs can be trained on a small frequency band/time range and then tested across many different frequencies/times.

Once a CNN is trained on ~ 20000 injections (at different amplitudes) across 150 Hz and 2000 s maps, for about 22 days of data for a false alarm probability of 1% (which is about 2000 noise maps), we can apply the following procedure for each detector:

1. Load short fast Fourier transform Databases (*77*).
2. Choose appropriate FFT length depending on maximum $\dot{f}_0$, create complete time/frequency map and peakmap (*78*).
3. Clean known noise lines and persistent lines.
4. Reduce resolution by a factor of 16 on each axis.
5. Run CNNs on each reduced time/frequency map.
6. Repeat for each detector's data. Do coincidences between time/frequency maps that produce triggers.
7. If a trigger remains after the coincidence step, run GFH on the peakmap and obtain parameters of the signal.
8. Perform coincidences between candidates returned by the GFH, follow-up, etc. using methodology described in (*31*).
9. Repeat for each 150 Hz/2000 s map.

We must emphasize that if a coincidence occurs between time/frequency maps at the level of the CNNs, the GFH can estimate the parameters of a signal only if it follows a power-law behavior. Otherwise, we can only constrain that the variation in braking index is no more than a certain δn.

As stated in Section 13.3, we cannot test a CNN on the same noise maps on which we train, since the CNN has learned that those maps do not contain a signal. But, if we create different CNNs that are trained and tested on different portions of the data, we can avoid this problem. The logic is the following: suppose we have data for 90 days, and that we need to train a CNN on only 2000 noise maps (23 days of data) to guarantee reasonable false alarm

probability/detection efficiency comparable to the GFH. At a minimum, we can train on the first 23 days, and then run an analysis on the last 67 days. Afterwards, we could train on days 24-47, then analyze days 1-23 and 48-90. This would then require the same number of injections, just a doubling of the training time, which takes ~ 30 minutes. The more we repeat this "shifting", the more times that a different CNN will see the same data.

We have shown that the CNNs are robust towards signals in different frequency bands in Figure 13.5; hence we have not used injections in all bands. Instead, we have just trained on signal and noise maps in one band.

14.3 Parameter Space Explored

We display the parameter space to which we are most sensitive in our search, assuming a power-law signal. Using the criteria that a signal must not spin out of the frequency band we are analyzing with the CNNs and the GFH, we determine which signal parameters we have the best change of observing. Figure 14-1 plots the initial frequencies, spindowns and spindown timescales to which we are most sensitive for our search configuration, fixing $n = 7$. The spindown timescale is colored.

14.4 Description of Search Results

Coincidences between the outputs of Livingston and Hanford resulted in 50 time/frequency maps having a value of $p > p_{thr} = 0.9$. A grid over n was constructed for each of these maps from $n = [2.5, 7]$, in each 150 Hz band between $f = [100, 1900]$ Hz. We then applied the GFH individually to the peakmaps generated at these times (not the maps that the network saw, but the maps constructed by the procedure in (*78*)). The GFH returned 431 coincident candidates, where a coincidence occurs if the Euclidean distance between candidate parameters x_0/k in each detector is less than 3 bins. Most candidates had a very low detection statistic, called the critical ratio (CR), defined as:

$$CR = \frac{y - \bar{y}}{\sigma} \tag{14–1}$$

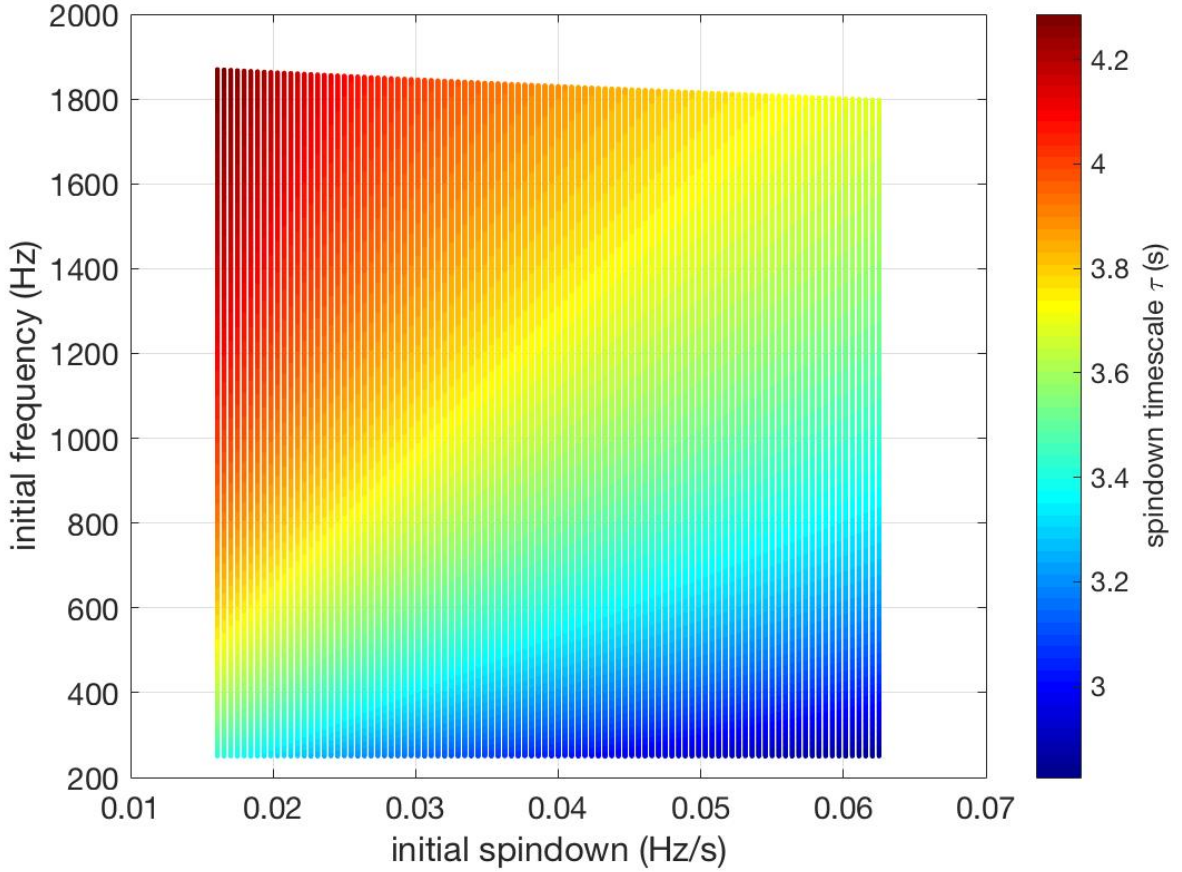


Figure 14-1: Parameter space explored in our search for a remnant of GW170817, for $n = 7$, invoking the criteria that a signal must be fully contained within a frequency band analyzed to be detectable.

where y is the number count returned by the GFH, $\bar{y}$ and σ are the mean and standard deviation of the full Hough map, respectively.

We required $CR > 3.5$, corresponding to a false alarm probability of 0.02%. After applying this threshold, 1 candidate remained (for $T_{FFT} = 4$ s). We then use the full Band-Sampled-Data (BSD) (*81*) follow-up procedure described in (*31*). The procedure corrects the peakmaps for the time/frequency evolution of the signal, resulting (theoretically) in a monochromatic signal. T_{FFT} is then increased to improve the signal-to-noise ratio and the original Frequency Hough is applied. If the critical ratio of a candidate is less now than before the follow-up, the candidate is vetoed. In each iteration, T_{FFT} is doubled. The candidate

remained for a few iterations until we used a $T_{FFT} = 64$ s, when it was finally vetoed. We therefore conclude that there was no detectable gravitational wave emission from a remnant of GW170817.

14.5 Upper Limits

We determine the weakest signals we can see by performing 250 injections, each lasting 2000 s, at different amplitudes with parameters uniformly distributed between $n = [2.5, 7]$, $f_0 = [250, 1900]$ Hz, $\dot{f} = [1/64, 1/16]$ Hz/s, $\cos\iota = [-1, 1]$, polarization angle $\psi = [-90, 90]$ degrees, throughout the entire week after the merger. Another set of injections was also done with $\delta n/\delta t = [-10^{-4}, 10^{-4}]$ /s, analyzed only by the CNNs. A detection is obtained if a candidate returned by the GFH is within 3 bins of the injections and if its $CR > 3.5$, the requirement used in the follow-up in the search. Figure 14-2 reports upper limits on distance on a possible power-law signal coming from a remnant of GW170817 at the 50% confidence level. We have plotted five curves: one using the CNNs followed up by the GFH to estimate the parameters of the signal (magenta), one curve using just the CNNs to detect time-varying braking index signals (blue), and three other curves for comparison purposes. The red one is from the original search (*30*), adjusted to be at the 50% confidence level, and the other two curves were obtained by just using the CNNs (green) and GFH (black) to calculate the upper limits. It seems that the previous upper limits are better because the signals were simulated for 2000-16000 s consistent with the search length at $n = 5$ (*30*). This should imply an increase of at most 40% in distance reach at higher frequencies, and $n = 5$ corresponds to the best sensitivity by about 20% relative to other braking indices at lower frequencies for the GFH, determined by using Equation 9–22 and looking at Figure 9-10. Additionally the grid in k constructed during the GFH is not optimal for signals of such a short duration, especially at high frequencies, so we are currently investigating ways to improve this grid.

For each curve, we determined the distance reach by assuming a moment of inertia of $I_{zz} = 100 M_{sun}^3 G^2/c^4 \approx 4.34 \times 10^{38}$ kg$\cdot$ m^2, and calculating the maximum allowed ellipticity ϵ by energy conservation, as was done in (*30*).

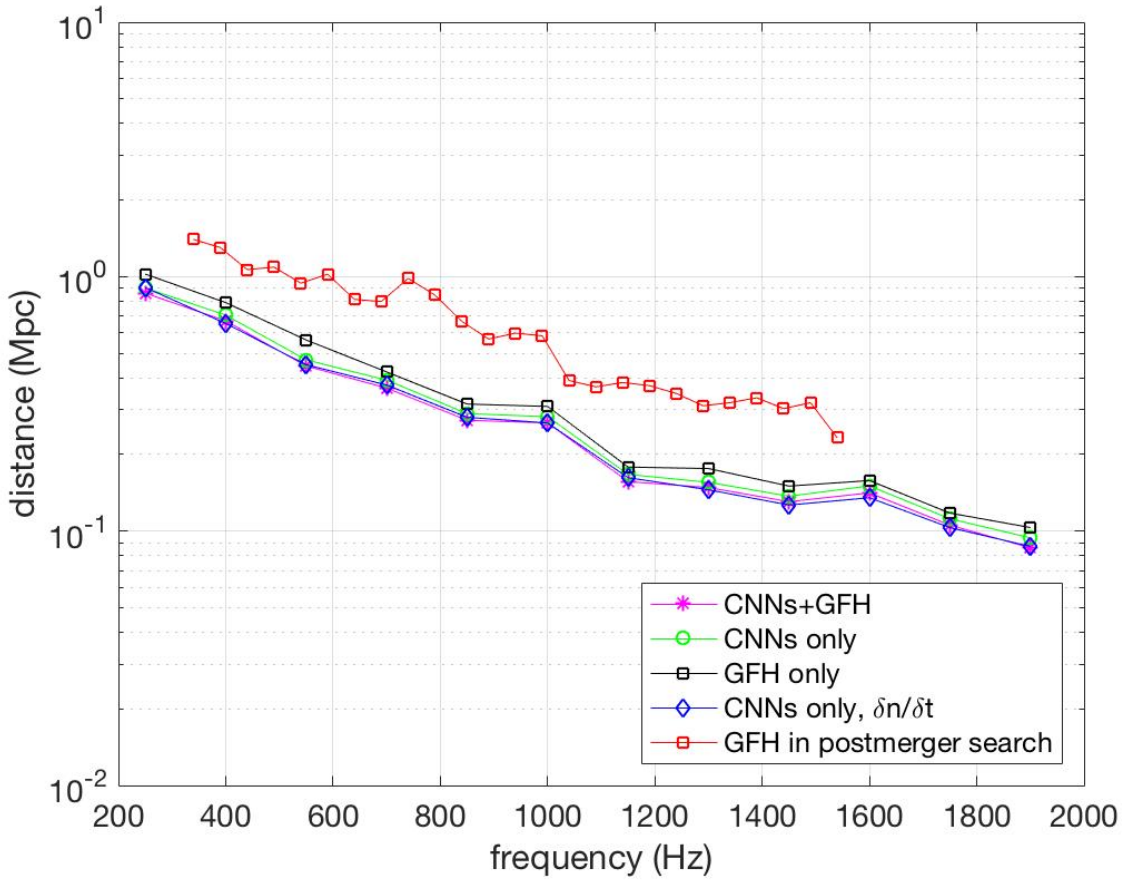


Figure 14-2: Upper limits on distance reach at 50% confidence for the search design presented in Section 14.2. CNNs followed up by the GFH (magenta curve) and only CNNs trying to detect time-varying braking index signals (blue). For comparison curves obtained by using just the CNNs (green), the GFH in this search (black) and the previous post-merger search (red) (*30*) are shown. The CNNs are sensitive to time-varying braking indices as well, even though these upper limits are done on fixed braking indices. Reasons why the GFH performed better in the first post-merger search include the use of a fixed $n = 5$, boosting sensitivity at low frequencies, and longer duration signals at high frequencies, improving sensitivity there (*31*). Additionally the GFH in this search performs worse at higher frequency because the grid in k is much finer and for signal durations of 2000 s, the signal is split among neighboring k bins. Alternative ways to optimally construct the grid in k are currently being investigated.

CHAPTER 15
ALL-SKY SEARCH FOR LONG GRAVITATIONAL WAVE TRANSIENTS IN O3 DATA

15.1 Motivation

As stated in Section 6.2, many GRBs are observed above and below threshold each day by Fermi-GBM and Swift-BAT (*57*, *58*). We do not know how far away these sources of GRBs are, nor do we know what caused them. Due to the fact that GW170817 generated an almost sub-threshold GRB event, it is possible that mergers of binary systems are producing these GRBs. If this is the case, long transient signals similar to those we searched for after GW170817 may be occurring frequently.

Because we expect that these signals last for less than $O(days)$ and have large spindowns, we do not need to be concerned about the Doppler modulation. This means that our searches are by default all-sky searches, or at least require so few grid points in the sky that any signal not "properly corrected for" would simply be shifted uniformly in time/frequency space, retaining the characteristics of the power-law signal for a particular braking index. Moreover, our frequency bins are quite large compared to CW searches, so any kind of Doppler modulation is confined to one frequency bin.

To illustrate that our search is essentially independent of position in the sky , we begin with the Doppler shift equation:

$$f(t) = f_0 \left(1 + \frac{\vec{v}_{orb} \cdot \hat{n}}{c}\right) \tag{15–1}$$

where v_{orb} is the velocity of the earth around the sun and $\hat{n}$ is a unit vector pointing from the earth to the source of GWs. The maximum Doppler shift occurs when the earth is moving directly towards or away from the source; hence:

$$\Delta f_{dopp,max} = f_0 \frac{v_{orb}}{c} \approx 10^{-4} f_0 \tag{15–2}$$

We impose that this shift must be less than a frequency bin df_{max}:

$$\Delta f_{dopp,max} <= df_{max} = \frac{1}{T_{FFT,max}} \tag{15–3}$$

Hence, for a signal of $f_0 = 2000$ Hz, $T_{FFT,max} <= 5$ s. As long as we restrict our search to frequencies at 2000 Hz or below and use $T_{FFT} <= T_{FFT,max}$, we can perform an all-sky search without having to worry about Doppler modulations shifting potential GW signals by more than one frequency bin.

15.2 Search Setup and Parameter Space Explored

Though the CNNs can perform an inherently unmodeled search, they are clearly likely to pick up signals similar to those on which they were trained. So, even though we only train on a subset of the parameter space we are searching for, the CNNs can generalize to find new ones, as we showed in 13.5. Therefore we are sensitive to all power-law and power-law-like signals, regardless of f_0, $\dot{f}_0$ and n. However, our sensitivity is actually limited by how much time we can observe and which T_{FFT} we use.

We search 150 Hz x 4000 s time/frequency maps created using $T_{FFT} = 4$ s for signals lasting at most 4000 s, whose initial frequencies could be between 100 Hz and 2000 Hz and whose spindowns could be no greater than $\dot{f}_{0,max} = 1/16$ Hz/s.

One CNN was trained on the first calibration of data, C00, which had not been cleaned, for each detector (Livingston and Hanford). For Livingston, we used 55756 signal maps and noise maps from 09-Apr-2019 00:38:17 to 07-May-2019 09:35:19.For Hanford, we used 54172 signal maps and noise maps from 08-Apr-2019 20:56:39 to 06-May-2019 07:11:05. The networks were trained only on science data from each detector.

The analysis of coincident science time data began on data taken after 07-May-2019 09:35:19 and ended on July 1, which spans almost 2 months of data.

15.3 Search Progress and Preliminary Results

Time/frequency maps covering approximately 2 months of Livingston/Hanford data were analyzed by the CNNs. Thirty-six of these maps were in coincidence. On these 36 maps, we perform the GFH follow-up as described in 14.2, covering a range of braking indices from 2.5

to 7. This results in 180 Hough transforms to do, which is in sharp contrast to the $O(10^5)$ Hough transforms needed in the three configurations covering only a day of data in which the first GW170817 post-merger search was run (see Chapter 10 for more details). After performing coincidences among the candidates returned by the GFH, imposing that the average critical ratio of candidates must be greater than 5, removing candidates near known noise lines, and vetoing candidates whose critical ratio in Hanford and Livingston differs by greater than a factor of 4 (all these procedures were employed in (*30*)), we are left with 53 coincident candidates to follow up. Considering a small window around the parameters of each coincident candidate, and all combinations of these parameters, we will follow up 837 candidates. The parameters of these candidates are shown in Figure 15-1. These potential signals occur at the following times: 19-May-2019 05:38:31, 20-May-2019 00:09:01, 03-Jun-2019 10:54:09, 06-Jun-2019 05:51:13, 07-Jun-2019 08:36:57, 16-Jun-2019 18:24:57, and 23-Jun-2019 07:03:17.

A coincidence distance equal to 2 bins will be used now, and a greater critical ratio threshold will be employed, since the signals should be stronger and be closer in the parameter space in both detectors at this stage. Unfortunately the data structures needed to perform the full follow-up procedure are not ready at the time of writing; in the meantime, preparing for the likely scenario of no detection, we place stringent upper limits on the expected all-sky GW emission.

15.4 Upper Limits

Given that we do not expect candidates to survive our follow-up procedure, we place upper limits on our distance reach, at 50% confidence. We perform 100 injections across many different amplitudes in each frequency band analyzed, whose parameters were uniformly chosen in reference time t_0, $n = [3, 7]$, $f_0 = [f_{max} - bw/4, f_{max}]$, where $bw = 150$ Hz is the bandwidth in the time/frequency map, $\dot{f}_0 = [1/(2T_{FFT})^2, 1/T_{FFT}^2]$. The injections were first passed to the CNNs, and then the coincident maps were given to the GFH for further analysis, as done in the real search. The upper limits, seen in Figure 15-2, were calculated in the same way as done

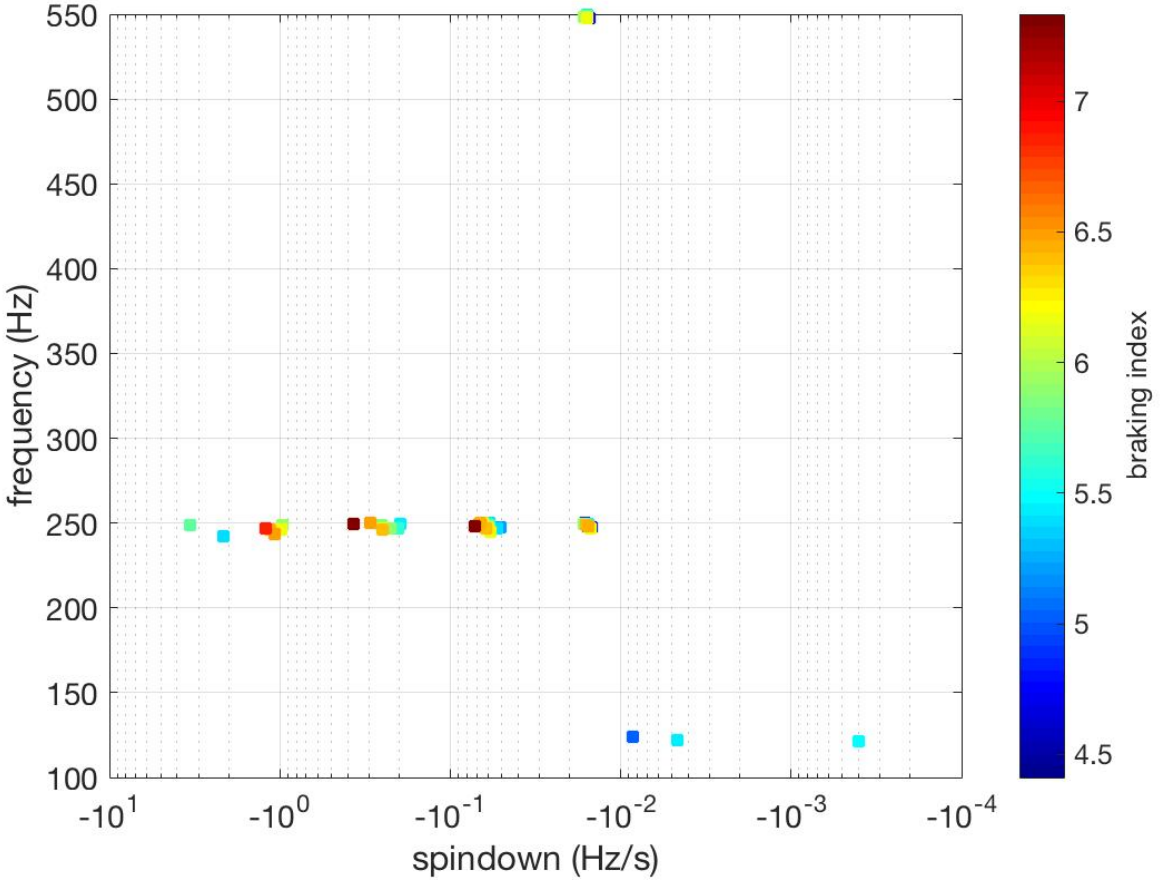


Figure 15-1: Candidates returned by the CNNs and GFH in the O3 all-sky search analysis. Around 250 Hz, the candidates are actually spread from 240 to 250 Hz. We believe that many candidates survived the coincident step around this frequency because the LIGO interferometers are most sensitive to potential signals the [100,250] Hz band.

in the O2 machine learning post-merger remnant search (see Section 14.5 for more details). These are more stringent upper limits than obtained after the O2 post-merger searches.

15.5 Prospects

This search was done for only one choice of T_{FFT}, but this is certainly not optimal if the GW signal has a smaller spindown than $1/T_{FFT}^2$ or lasts for longer than the time we analyzed. We therefore plan to employ different configurations, each containing time/frequency maps with different $T_{FFT} = 2, 4, 8, 16$ s and durations to cover this parameter space with better sensitivity.

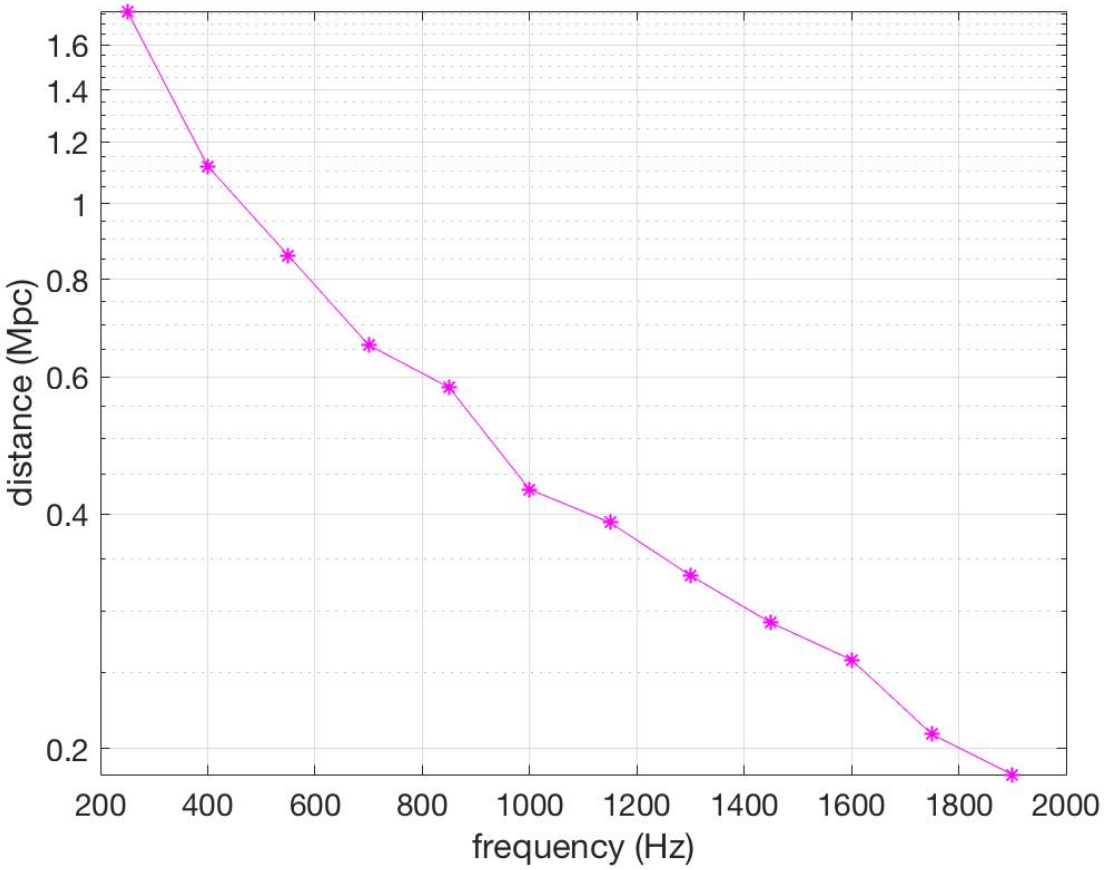


Figure 15-2: Upper limits calculated for our all-sky machine learning-based search for long transient GW signals. Distance reach is at 50% confidence. We injected 100 signals per amplitude, across a range of braking indices, initial frequencies, spindowns, and GPS start times in line with the search parameter space.

CHAPTER 16
CONCLUSIONS

We have described in detail the background of GW physics, beginning from the basics of general relativity, and neutron star astrophysics, demonstrating the problem of determining the equation of state. R-modes, and the GWs they emit from being excited in a neutron star, were also discussed. Recent results from real GW detections by the LIGO/Virgo collaborations were presented, and physical conclusions concerning the speed of the graviton and the equation of state were drawn. We also explained how the LIGO/Virgo detectors work at an introductory level, and described various data analysis strategies used both in our procedures and within the LIGO/Virgo collaborations. Moreover, implications of a detection of GWs from an isolated neutron star, and extrapolations of our sensitivity in future detectors, were discussed. The major outputs of this thesis are two new, extensively developed methods to search for GWs from isolated neutrons lasting $O(hours - days)$ and two searches of O2 data for a remnant of GW170817, and one almost-finished all-sky search of O3 data for long transient signals. It is clear from our results that machine learning can be useful in GW physics but needs to be used carefully, and that the GFH is a robust parameter estimator for power-law signals. However, both methods can be improved and applied on future data when a supernova occurs or two neutron stars merge closer than 40 Mpc.

Bibliography

1. B. P. Abbott *et al.*, *Physical Review Letters* **119**, 161101 (16 2017).
2. B. P. Abbott *et al.*, *Physical Review Letters* **116**, 061102 (2016).
3. M. P. Hobson, G. P. Efstathiou, A. N. Lasenby, *General relativity: an Introduction for Physicists* (Cambridge University Press, 2006).
4. W. Li *et al.*, *Monthly Notices of the Royal Astronomical Society* **412**, 1441–1472 (2011).
5. S. P. Reynolds *et al.*, *The Astrophysical Journal Letters* **680**, L41 (2008).
6. E. Müller, H.-T. Janka, *Astronomy and Astrophysics* **317**, 140–163 (1997).
7. K. N. Yakunin *et al.*, *Physical Review D* **92**, 084040 (2015).
8. S. Klimenko, I Yakushin, A Mercer, G. Mitselmakher, *Classical and Quantum Gravity* **25**, 114029 (2008).
9. P. J. Sutton *et al.*, *New Journal of Physics* **12**, 053034 (2010).
10. S. A. Colgate, M. Herant, W. Benz, *Physics Reports* **227**, 157–174 (1993).
11. N. Christensen, *Reports on Progress in Physics* **82**, 016903 (2018).
12. A. L. Watts, B. Krishnan, L. Bildsten, B. F. Schutz, *Monthly Notices of the Royal Astronomical Society* **389**, 839–868 (2008).
13. G. B. Cook, S. L. Shapiro, S. A. Teukolsky, *The Astrophysical Journal* **423**, L117 (1994).
14. R. Wijnands, M. van der Klis, *Nature* **394**, 344 (1998).
15. L. Bildsten, *The Astrophysical Journal Letters* **501**, L89 (1998).
16. G. Ushomirsky, C. Cutler, L. Bildsten, *Monthly Notices of the Royal Astronomical Society* **319**, 902–932 (2000).
17. C Messenger *et al.*, *Physical Review D* **92**, 023006 (2015).
18. N. K. Glendenning, *Compact stars: Nuclear Physics, Particle Physics and General Relativity* (Springer Science & Business Media, 2012).
19. M. Camenzind, *Compact Objects in Astrophysics* (Springer, New York, 2007).

20. R. S. Foster, D. Backer, *The Astrophysical Journal* **361**, 300–308 (1990).

21. J. W. Hessels *et al.*, *Science* **311**, 1901–1904 (2006).

22. S. L. Shapiro, S. A. Teukolsky, *Black holes, white dwarfs, and neutron stars: The physics of Compact Objects* (Wiley, New York, 1983).

23. P. Arras, presented at the Binary Radio Pulsars, vol. 328, p. 317.

24. G. B. Cook, S. L. Shapiro, S. A. Teukolsky, *The Astrophysical Journal* **424**, 823–845 (1994).

25. F. Gittins, N. Andersson, *arXiv preprint arXiv:1811.00550* (2018).

26. J. D. Jackson, *Classical electrodynamics*, 1999.

27. B. P. Abbott *et al.*, arXiv: 1707.02669 (gr-qc) (July 2017).

28. B Knispel *et al.*, *Science* **329**, 1305–1305 (2010).

29. P. Jaranowski, A. Krolak, B. F. Schutz, *Physical Review D* **58**, 063001 (1998).

30. B. P. Abbott *et al.*, *The Astrophysical Journal* **875**, 160 (2019).

31. A. Miller *et al.*, *Phys. Rev. D* **98**, 102004 (10 2018).

32. M. Oliver, D. Keitel, A. M. Sintes, *Physical Review D* **99**, 104067 (2019).

33. L. Sun, A. Melatos, *Physical Review D* **99**, 123003 (2019).

34. N. Sarin, P. D. Lasky, L. Sammut, G. Ashton, *arXiv preprint arXiv:1805.01481* (2018).

35. B. J. Owen *et al.*, *Physical Review D* **58**, 084020 (1998).

36. M. G. Alford, K. Schwenzer, *Monthly Notices of the Royal Astronomical Society* **446**, 3631–3641 (2014).

37. N. Andersson, K. D. Kokkotas, *International Journal of Modern Physics D* **10**, 381–441 (2001).

38. B. J. Owen, *Physical Review D* **82**, 104002 (2010).

39. K. S. Thorne, *Reviews of Modern Physics* **52**, 299 (1980).

40. J Provost, G Berthomieu, A Rocca, *Astronomy and Astrophysics* **94**, 126 (1981).

41. L. Lindblom, B. J. Owen, S. M. Morsink, *Physical Review Letters* **80**, 4843 (1998).

42. T. G. Cowling, *Monthly Notices of the Royal Astronomical Society* **101**, 367 (1941).

43. L. Lindblom, R. J. Splinter, *The Astrophysical Journal* **348**, 198–202 (1990).

44. S. Yoshida, Y. Kojima, *Monthly Notices of the Royal Astronomical Society* **289**, 117–122 (1997).

45. S. Yoshida, U. Lee, *The Astrophysical Journal* **567**, 1112 (2002).

46. J. R. Oppenheimer, H. Snyder, *Physical Review* **56**, 455 (1939).

47. A. Mytidis, A. A. Panagopoulos, O. P. Panagopoulos, A. Miller, B. Whiting, *Physical Review D* **99**, 024024 (2019).

48. B. Abbott *et al.*, *Physical Review X* **9**, 031040 (2019).

49. B. P. Abbott *et al.*, *Astrophys. J. Lett* **848**, L12 (2017).

50. P. Pani, C. F. Macedo, L. C. Crispino, V. Cardoso, *Physical Review D* **84**, 087501 (2011).

51. L. S. Finn, P. J. Sutton, *Physical Review D* **65**, 044022 (2002).

52. A. G. A. von Kienlin, C. Meegan, the Fermi GBM, *GCN* **21520** (2017).

53. B. P. Abbott *et al.*, *The Astrophysical Journal Letters* **848**, L13 (2017).

54. G. Gonzales, *Moving pendulum*, http://www.phys.lsu.edu/faculty/gonzalez/Teaching/Phys7221/PendulumWithMovingSupport.pdf.

55. M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments* (Oxford University Press, 2008), vol. 1.

56. P. Shawhan, LSC/Virgo, "Fermi-GBM and Swift GCN Notice Statistics", LIGO Document T1900297-v4, 2019.

57. C. Meegan *et al.*, *The Astrophysical Journal* **702**, 791 (2009).

58. H. A. Krimm *et al.*, *The Astrophysical Journal Supplement Series* **209**, 14 (2013).

59. B. P. Abbott *et al.*, *The Astrophysical Journal Letters* **832**, L21 (2016).

60. V. Tiwari, *Classical and Quantum Gravity* **35**, 145009 (2018).

61. V. Tiwari, S. Fairhurst, M. Hannam, *The Astrophysical Journal* **868**, 140 (2018).

62. J. Abadie *et al.*, *Classical and Quantum Gravity* **27**, 173001 (2010).

63. E. S. Phinney, *The Astrophysical Journal* **380**, L17–L21 (1991).

64. R. Narayan, *The Astrophysical Journal* **319**, 162–179 (1987).

65. C. Kim, V. Kalogera, D. Lorimer, *The Astrophysical Journal* **584**, 985 (2003).

66. V. Kalogera, R Narayan, D. N. Spergel, J. Taylor, *The Astrophysical Journal* **556**, 340 (2001).

67. R. Bondarescu, S. A. Teukolsky, I. Wasserman, *Physical Review D* **79**, 104003 (2009).

68. M Punturo *et al.*, *Classical and Quantum Gravity* **27**, 194002 (2010).

69. B. P. Abbott *et al.*, *Classical and Quantum Gravity* **34**, 044001 (2017).

70. P. J. Sutton, *arXiv preprint arXiv:1304.0210* (2013).

71. H.-Y. Chen *et al.*, *arXiv preprint arXiv:1709.08079* (2017).

72. L. S. Finn, D. F. Chernoff, *Physical Review D* **47**, 2198 (1993).

73. M. Evans *et al.*, "Unofficial sensitivity curves (ASD) for aLIGO, Kagra, Virgo, Voyager, Cosmic Explorer and ET", https://dcc.ligo.org/LIGO-T1500293/public, 2019.

74. W. van Straten *et al.*, *Nature* **412**, 158 (2001).

75. I. I. Shapiro, *Physical Review Letters* **13**, 789 (1964).

76. J. H. Taylor, J. M. Weisberg, *The Astrophysical Journal* **345**, 434–450 (1989).

77. P Astone, S Frasca, C Palomba, *Classical and Quantum Gravity* **22**, S1197 (2005).

78. P. Astone, A. Colla, S. D'Antonio, S. Frasca, C. Palomba, *Physical Review D* **90**, 042002 (2014).

79. P. Astone, S. D'Antonio, S. Frasca, C. Palomba, *Classical and Quantum Gravity* **27**, 194016 (2010).

80. S. Caride, R. Inta, B. J. Owen, B. Rajbhandari, *Phys. Rev. D* **100**, 064013 (6 2019).

81. O. Piccinni *et al.*, *Classical and Quantum Gravity* **36**, 015008 (2018).

82. L. Baiotti, L. Rezzolla, *Rept. Prog. Phys.* **80**, 096901 (2017).

83. Y.-W. Yu, L.-D. Liu, Z.-G. Dai, *The Astrophysical Journal* **861**, 114 (2018).

84. S. A. Colgate, *The Astrophysical Journal* **163**, 221 (1971).

85. A. Vajpeyi, A. Miller, P. Astone, S. Frasca, *Bulletin of the American Physical Society* **62** (2017).

86. L. Pierini, *Ph.D. Thesis, Sapienza* (2019).

87. B. P. Abbott *et al.*, *Physical Review D* **96**, 062002 (2017).

88. B. Abbott *et al.*, *Physical Review D* **100**, 024004 (2019).

89. K. Riles, *Modern Physics Letters A* **32**, 1730035 (2017).

90. E. Thrane *et al.*, *Physical Review D* **83**, 083004 (2011).

91. E. Thrane, V. Mandic, N. Christensen, *Physical Review D* **91**, 104021 (2015).

92. R. Prix, S. Giampanis, C. Messenger, *Physical Review D* **84**, 023007 (2011).

93. D. Keitel *et al.*, *arXiv preprint arXiv:1907.04717* (2019).

94. T. Hill, L. Marquez, M. O'Connor, W. Remus, *International journal of forecasting* **10**, 5–15 (1994).

95. M. I. Jordan, T. M. Mitchell, *Science* **349**, 255–260 (2015).

96. Y. LeCun, Y. Bengio, G. Hinton, *Nature* **521**, 436 (2015).

97. H. Adeli, *Kybernetes* **28**, 317–318 (1999).

98. R. Hecht-Nielsen, in *Neural networks for perception* (Elsevier, 1992), pp. 65–93.

99. A. Ng, in *Machine learning*.

100. A. Krizhevsky, I. Sutskever, G. E. Hinton, presented at the Advances in neural information processing systems, pp. 1097–1105.

101. G. Hinton *et al.*, *IEEE Signal processing magazine* **29** (2012).

102. H. Greenspan, B. Van Ginneken, R. M. Summers, *IEEE Transactions on Medical Imaging* **35**, 1153–1159 (2016).

103. P Astone *et al.*, *Physical Review D* **98**, 122002 (2018).

104. T. K. Ho, presented at the Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1, pp. 278–282.

105. T. K. Ho, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**, 832–844, ISSN: 0162-8828 (Aug. 1998).

106. P. T. Baker *et al.*, *Physical Review D* **91**, 062004 (2015).

107. D. George, E. Huerta, *Physical Review D* **97**, 044039 (2018).

108. D. George, E. Huerta, *Physics Letters B* **778**, 64–70 (2018).

109. H. Gabbard, M. Williams, F. Hayes, C. Messenger, *Physical Review Letters* **120**, 141103 (2018).

110. C. Dreissigacker, R. Sharma, C. Messenger, R. Zhao, R. Prix, *arXiv preprint arXiv:1904.13291* (2019).

111. T. D. Gebhard, N. Kilbertus, I. Harry, B. Schölkopf, *arXiv preprint arXiv:1904.08693* (2019).

112. A. L. Miller *et al.*, *Physical Review D* **100**, 062005 (2019).

113. T. S. Adams *et al.*, *Physical Review D* **88**, 062006 (2013).

BIOGRAPHICAL SKETCH

Andrew Miller earned his Bachelor of Science in physics at The College of New Jersey in 2015. He decided to immediately pursue his Ph.D. in physics at the University of Florida and University of Rome, La Sapienza. His interests are in detecting gravitational waves from isolated neutron stars, machine learning algorithms, pattern recognition techniques, future generation gravitational wave detectors, and axion-like dark matter particle searches. He earned his Ph.D. in physics from the University of Florida and University of Rome, La Sapienza jointly in December 2019.