Gravitational Waves From Spinning Neutron Stars: Development of a Directed Binary Search Technique and Spectral Characterization Tools

by

Ansel A. M. Neunzert

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2019

Doctoral Committee:

Professor Keith Riles, Chair
Associate Professor Hui Deng
Professor August Evrard
Associate Professor Eric Hetland
Professor Timothy A. McKay

Ansel A.M. Neunzert

neunzert@umich.edu

ORCID iD: 0000-0003-0323-0111

Ansel A. M. Neunzert, 2019

Except where otherwise noted,
this work is made available under a
Creative Commons Attribution-NonCommercial
4.0 license (international).
Further details are listed on page iv.

Acknowledgments

This thesis would not exist without the support of a great many people. Thanks are due to first to Keith Riles, my Ph.D. advisor, for years of reliable and insightful scientific guidance. His attention to detail, deep knowledge of gravitational wave studies, and appreciation for steady progress and robust infrastructure development have made this work possible.

Several former graduate students from Michigan laid the groundwork for much of the data analysis described in this thesis: Evan Goetz, who wrote TwoSpect; Grant Meadors, whose scripts and documentation for TwoSpect searches have been very useful; and Jax Sanders, whose Fermi-LAT source prioritization work guides the X-statistic search. I also count all three of these people as friends, and as mentors who have helped to orient me to the collaboration and the field.

On the subject of line and comb investigations, there are more people to thank. Robert Schofield taught me a great deal about the Hanford detector and about noise hunting in general. Evan Goetz (once again), Pep Covas, and Brynley Pearlstone have been regular collaborators on line investigations, and in doing so they helped refine FineTooth into its current form. Greg Mendell maintains the Fscan infrastructure on which FineTooth is built. Stephanie Zhao, an undergraduate at Michigan, has tested and improved the FineTooth comb finding algorithm, eventually writing her own version of it.

The Hanford site, where I spent six months in 2016, is a place of constant activity and steep learning curves. Fortunately, the people there went out of their way to welcome me, and to make sure I learned my way around. Among this wonderful crowd, Nutsinee Kijbunchoo and Daniel Sigg deserve special mention.

A key prerequisite to writing a thesis is making it through the years of graduate school leading up to it; for support in that endeavor, I cannot thank my fellow graduate students and friends here at Michigan enough. Their camaraderie, commiseration, and companionship got me through (as did their board game nights and camping trips). I have also been struck time and again by their tireless commitment to improving the world around them—never content to simply keep their heads down and study—as well as the wisdom, practicality, and resilience they have displayed in dealing challenges large and small. These people give me hope for the future. On that note, special thanks are due to the Graduate Employees Organization at Michigan, which is both a critical driving force for quality of life among graduate students, and a place where

I have met many friends.

My family has cheered me on through grad school, as they have in all the years before. My parents and grandparents have supported me despite the distance from here to Oregon, and have always taken an active, enthusiastic interest in my life and my research. My oldest friends, Ashley and Alex, are family in every way that matters: I turn to them for advice and support as I have for so many years. Lastly, I owe many thanks to my younger sister, Sonya. It is impossible to quantify her influence on my life, but here is an attempt: over the course of my graduate school career alone, our written correspondence has amassed a word count that dwarfs Tolstoy's *War and Peace*.

Preface

Details on the license governing this work may be found at:

`https://creativecommons.org/licenses/by-nc/4.0/`

The following figures are drawn from external sources, and are not subject to this license:

- Figure 2.4 on page 19
- Figure 3.4 on page 28
- Figure 3.5 on page 30
- Figure 3.6 on page 31
- Figure 3.7 on page 32
- Figure 3.8 on page 33
- Figure 3.9 on page 34
- Figure 3.10 on page 37

Information on the source of each figure is given in the figure caption.

Table of Contents

List of Tables

List of Figures

Abstract

Observational gravitational wave astronomy is a young field, begun in 2015 with the detection of the binary black hole merger event GW150914 by LIGO. Although all gravitational waves observed so far have emanated from the mergers of compact binary systems, many other classes of potential gravitational wave sources exist. One important class of predicted gravitational waves includes those from non-axisymmetric spinning neutron stars, which are expected to continuously emit near-monochromatic waves. With Advanced LIGO having recently entered its third observing run, prospects for the detection of such continuous waves are ever improving. This thesis will describe two complementary projects that support the eventual detection of continuous gravitational waves, in the areas of data analysis and detector characterization. Built on the existing TwoSpect analysis pipeline, a sparse-sampling algorithm called the X-Statistic method is introduced. This method brings templated searches for continuous waves from spinning neutron stars in unknown binary systems into the realm of computational feasibility for the first time. An ongoing pilot search using the X-Statistic method is described. Additionally, a suite of tools for LIGO spectral analysis, called FineTooth, is introduced. These tools have helped to track and investigate narrow spectral artifacts, particularly periodic structures known as combs, over the second and now third LIGO observing runs, making significant contributions to data quality for continuous wave searches.

Chapter 1

Introduction

The first gravitational wave detection by the Laser Interferometer Gravitational-wave observatory in 2015 was a landmark in modern astronomy. It confirmed a nearly century-old prediction of General Relativity: that masses moving through space should generate ripples in its fabric, and that these ripples travel across the cosmos at the speed of light, carrying with them information about the events that birthed them. The 2015 result also demonstrated that it is possible to *detect* these ripples using instruments on Earth: an engineering and data analysis feat of impressive proportions, given the tiny effect of a passing wave, and the many sources of terrestrial noise that threaten to confound its measurement. Perhaps most importantly, it yielded information on an astrophysical system never before observed: two massive black holes locked in a binary orbit, which spiralled together and coalesced into a single object, releasing in the process a burst of gravitational energy so powerful that its reverberations could be measured on Earth, more than a billion light years away.

The detection of that first merger was by no means an isolated event. In the years since, LIGO has continued to observe gravitational waves from the mergers of compact binary systems, and the rate of these observations continues to increase as the sensitivity of the detectors improves and the search range widens. Meanwhile, searches are underway using LIGO data for a variety of other types of signals.

This thesis will focus on the effort to detect continuous gravitational waves from spinning neutron stars in our galaxy, an as-yet undetected class of gravitational waves signals. Neutron stars are massive, dense objects which can spin hundreds of times per second, generating gravitational waves at a fixed frequency, like a single note played continuously. Unfortunately, this note is expected to be very quiet indeed: much quieter than the gravitational wave signals so far discovered. It will require both innovative data analysis techniques and a robust understanding of instrumental noise sources in order to confidently identify such a signal. Success in this endeavor would yield a wealth of scientific information which cannot be gleaned from electromagnetic observations alone: on the internal dynamics of neutron stars, their populations, and the mechanisms that cause them to spin as rapidly as they do.

Chapters 2, 3, and 4 of this work present background material on gravitational waves, LIGO, and neutron stars. Chapter 2 describes how gravitational waves arise from the structure of General Relativity, and reviews their basic properties and the effect they have on objects as they pass. Chapter 3 then introduces the tools necessary to detect gravitational waves, building from the fundamentals of interferometry to the complexity of the Advanced LIGO detectors as they currently operate. Chapter 4 returns to astrophysics, showing how neutron stars are expected to give rise to gravitational waves, and describing how gravitational wave observations can enrich neutron star astronomy.

In the second part of this thesis, chapters 5, 6, and 7 describe the author's contributions to the search for gravitational waves from spinning neutron stars. Chapter 5 documents a suite of tools developed for monitoring narrow spectral artifacts in LIGO data, and the ways in which these tools have contributed to noise identification and mitigation in the last three years. This work contributed to a publication in Physical Review D on narrow spectral artifacts [17]. Chapter 6 describes a new algorithm which builds on the existing data analysis tool TwoSpect, intended to search for signals from spinning neutron stars in binary systems. Chapter 7 then provides a snapshot of an ongoing pilot search using the new method, including source selection and preliminary results. Finally, chapter 8 closes with a summary of the work, and discusses the outlook for future measurements.

Chapter 2

Gravitational waves from general relativity

2.1 General relativity

Classical mechanics treats gravity the same way that it treats all other forces: as a three-dimensional vector that acts on an object, changing its motion according to $\vec{F} = m\vec{a}$. Motion takes place on a static, uniform background of space and time. Time is assumed to move forward in a consistent manner, like the frames of a movie, no matter who is measuring it. Choosing a set of spatial coordinates is as simple as picking an origin and orienting three perpendicular measurement axes. As long as everyone agrees on the convention, "moving in the x direction" will mean exactly the same thing regardless of whom you ask. Physical systems with spherical or cylindrical symmetry may be more easily described in coordinate systems adapted to their symmetries, but these coordinate systems are just new versions of an imaginary grid laid over the same physical reality. It is always possible to convert them back to familiar rectangular coordinates.

Albert Einstein's theory of relativity drastically deviates from the classical model. Rather than treating space and time as immutable and unambiguous, Einstein proceeded from a different set of axioms. The speed of light, recently understood to arise from the fundamental laws of electrodynamics, was taken to be constant for all observers, which required that space and time be consolidated into a single four-dimensional entity, and birthed special relativity (introduced in 1905). Next, gravity was described in *geometrical* terms. Instead of being treated as an entity separate from the static background of space and time, gravity was understood as a consequence of the (mutable, curved) structure of spacetime itself. This theory of gravitation is known as *general relativity*, or GR (introduced in 1915).

2.2 A conceptual overview of the math of general relativity

A detailed introduction to GR and the derivation of gravitational waves is well outside the scope of this work, and in any case, such calculations can be found in many reference texts. Instead, this section provides an overview of the key mathematical ingredients of GR, physical ways to think about them, and important

steps in the process of mathematically discovering gravitational waves within the framework of GR. It draws from a number of textbooks, [26, 18] notes, [14, 32, 27] and papers [54, 53, 30] which approach the material in different ways.

2.2.1 The spacetime metric

In relativity, the structure of spacetime is described by a two-index *metric tensor*, denoted in general as $g_{\mu\nu}$. For example, the Minkowski metric, which describes the flat spacetime of special relativity, is typically called η and written out in matrix form as follows:[1]

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.1}$$

The indices μ and ν for both rows and columns in this matrix are $\{t, x, y, z\}$, where t is a time coordinate, and x, y, and z are spatial coordinates, so that metric describes a 4-dimensional spacetime. In more precise terms: if two points are separated from each other in time by dt, and in space by dx, dy, and dz, the *spacetime interval* (called ds) between these two points is given by the expression:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \ , \tag{2.2}$$

where $dx^\mu, dx^\nu = \{dt, dx, dy, dz\}$ [2] This interval is analogous to a "distance" between points in spacetime, but it is not a spatial distance. Spatial distances are *not* invariant in relativity; different observers can measure different distances between two points, depending on their relative motion. The spacetime interval *is* invariant, and its form reflects the shape of spacetime as well as the chosen coordinate system.

Continuing with the example of Minkowski spacetime,

$$ds^2_{\text{Minkowski}} = -dt^2 + dx^2 + dy^2 + dz^2 \ . \tag{2.3}$$

All the off-diagonal terms (such as $dxdy$, $dxdt$, etc) are zero, and all the on-diagonal terms are either 1 or -1. Minkowski spacetime is the simplest possible *relativistic* metric; the only difference between Minkowski spacetime and familiar 3-dimensional rectangular coordinates (where $dr^2 = dx^2 + dy^2 + dz^2$) is that Minkowski

[1] There are different conventions for the sign of the metric; some authors flip the signs of each component. This choice is known as the signature; either signature can be used consistently to derive the same physics.

[2] Einstein summation notation is used to indicate a sum over all repeated indices, so there are 4x4=16 terms in this expression, correponding to the 16 entries in the matrix form. Indices are raised and lowered by contraction with the metric tensor, e.g $g_{\alpha\beta} x^\beta = x_\alpha$.

spacetime includes a time component. Importantly, this time component has a negative sign. This change encodes the phenomena of special relativity, such as the length contraction and time dilation observed by parties moving relative to each other.

2.2.2 The Einstein field equations

The other key element of general relativity is the Einstein field equation (or rather, equations):

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{2.4}$$

This statement is misleadingly compact, since both $G_{\mu\nu}$ and $T_{\mu\nu}$ are also 16-component tensors, which contain a great deal of information. However, the compact form of the equation illuminates the basic structure of the theory. The quantity on the right of the equation, $T_{\mu\nu}$, is the stress-energy tensor. It contains information on the distribution and motion of mass and energy in spacetime. The quantity on the left, $G_{\mu\nu}$, is the Einstein tensor, which contains information on the curvature of spacetime. $G_{\mu\nu}$ can be derived directly from the metric $g_{\mu\nu}$.

With the meanings of $G_{\mu\nu}$ and $T_{\mu\nu}$ in mind, the Einstein field equation tells a clear story: the form of spacetime (encoded in $G_{\mu\nu}$) depends on its matter and energy content (encoded in $T_{\mu\nu}$). The reverse is true as well: the distribution of matter and energy, and its motion, depends on the curvature of spacetime.

The stress-energy tensor can be further understood by breaking it down into four regions, labeled A, B, C, and D below:

$$T_{\mu\nu} = \begin{pmatrix} A & B & B & B \\ C & D & D & D \\ C & D & D & D \\ C & D & D & D \end{pmatrix} \ . \tag{2.5}$$

The cell labeled A corresponds to the energy density (a scalar). The cells labeled B form a vector which describes the energy flux (divided by the speed of light c), and the cells labeled C describe the momentum density (also divided by c). The rest of the cells, labeled D, describe a stress tensor, containing pressures (on the diagonal) and shear stresses (off diagonal). The physical meaning of the stress tensor will be discussed further in the context of gravitational waves.

2.2.3 From the metric to the Einstein tensor: a roadmap

Without delving into the specifics of the calculation, let us work backwards and unpack the Einstein tensor $G_{\mu\nu}$ into its components, to see how it connects with the metric tensor $g_{\mu\nu}$. This is purely to provide an overview of the steps; physical meaning will be discussed in the next section.

The quantity $G_{\mu\nu}$ is usually written in terms of the Ricci tensor $R_{\mu\nu}$, and the Ricci scalar R:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{2.6}$$

R is just a contraction of $R_{\mu\nu}$ with the metric tensor, so this expression can be expanded out as follows:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} R_{\alpha\beta} \tag{2.7}$$

The Ricci tensor is, in turn, a contraction of the Riemann tensor:

$$R_{\mu\nu} = R^{\gamma}_{\mu\gamma\nu} \tag{2.8}$$

The Riemann tensor is an unwieldy object with 4 indices, and thereby $4^4 = 256$ components, but fortunately it has a number of symmetries and only 20 degrees of freedom. It is defined from the Christoffel symbols, denoted Γ:

$$R^{\alpha}_{\beta\gamma\delta} = \frac{\partial \Gamma^{\alpha}_{\beta\delta}}{\partial x^{\gamma}} - \frac{\partial \Gamma^{\alpha}_{\beta\gamma}}{\partial x^{\delta}} + \Gamma^{\alpha}_{\gamma\epsilon}\Gamma^{\epsilon}_{\beta\delta} - \Gamma^{\alpha}_{\delta\epsilon}\Gamma^{\epsilon}_{\beta\gamma} \tag{2.9}$$

And the Christoffel symbols are in turn defined by derivatives of the metric tensor:

$$\Gamma^{\delta}_{\beta\gamma} = \frac{1}{2} g^{\alpha\beta} \left(\frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} + \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} + \frac{\partial g_{\beta\gamma}}{\partial x^{\alpha}} \right) . \tag{2.10}$$

From this quick overview, we can see that the $G_{\mu\nu}$ can be found directly from $g_{\mu\nu}$, and in particular that it contains *derivatives* of $g_{\mu\nu}$. This means that the Einstein field equations are a set of differential equations, exact solutions for which can only be found under specific circumstances. Analytical calculations in general relativity are usually concerned with the simplest possible cases, to make the problems tractable. Numerical approaches can deal with a broader range of problems, but often with great difficulty.

2.2.4 From the metric to the Einstein tensor: important stops along the way

The preceding section lays out the math, but does not discuss physical meaning. This section will briefly interpret the Christoffel symbols and the Riemann tensor, drawing largely from [41].

2.2.4.1 Christoffel symbols

The Christoffel symbols describe the way that coordinate unit vectors change between points. For example, one can calculate the Christoffel symbols for the familiar system of 3-dimensional spherical polar coordinates, and find that some of them are nonzero. This is true even though spherical polar coordinates are just another way to describe flat space; there is nothing inherently relativistic, or even specific to curved space, about nonzero Christoffel symbols.

Christoffel symbols are necessary to properly perform differentiation in non-rectangular coordinate systems. In such cases, the *covariant* derivative of a vector V is defined:[3]

$$\nabla_\beta V^\alpha = \partial_\beta V^\alpha + V^\gamma \Gamma^\alpha_{\gamma\beta} \, . \tag{2.11}$$

This is fundamentally an application of the product rule; the first term takes derivatives of the vector components, while the second takes derivatives of the coordinate system.

Anything which impacts derivatives in general will also impact expressions for the acceleration of a test particle. The Christoffel symbols appear in the equation of motion for a test particle in general relativity as follows:

$$\frac{d^2 x^\alpha}{d\tau^2} = -\Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} \tag{2.12}$$

2.2.4.2 Riemann tensor

Unlike the Christoffel symbols, which can be nonzero in flat space, the Riemann tensor encodes *physical* curvature. So, for example, the components of the Riemann tensor for 3-dimensional spherical polar coordinates will all be zero, indicating flatness, even though the Christoffel symbols are not.

The Riemann tensor can be considered as a commutator of covariant derivatives, a notion which is deeply tied to the concept of parallel transport. In order to take a difference between vectors (including an infinitesimal one, as in a derivative), it is necessary to transport them to the same point in space, without rotating them. In flat space, parallel transport of a vector is straightforward. It does not matter whether the vector is transported first in the dx direction or the dy direction or the dz direction; it will always end up pointing the same way. However, in curved space, this is not true. For example, on a globe, attempting to parallel transport a vector along a triangular path will return it to the starting point with a different direction. Moreover, the change is dependent on the order in which the transport is performed (see figure 2.1).

[3] Using the convention that $\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}$.

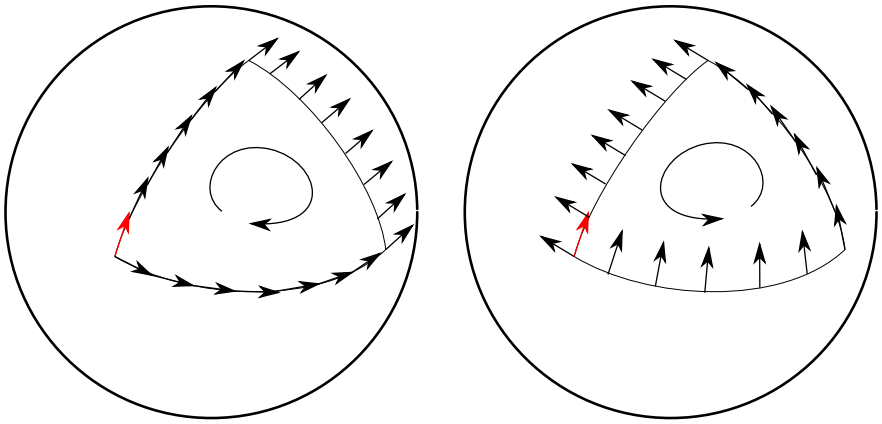


Figure 2.1: Parallel transport of a vector on a globe, along a triangular path with three right angles. This example shows that the direction in which the vector is transported along a given path matters. An infinitesimal version of this picture can be used to understand the Riemann tensor.

The Riemann tensor describes the difference between taking two covariant derivatives in a specific order, versus reversing that order. It allows one to find the deviation of a vector transported around a small parallelogram, which is a result of the curvature of the spacetime.

2.3 Gravitational waves in linearized gravity

The simplest way to see gravitational waves arise is to consider a nearly flat spacetime, introduce a small perturbation, and see if there is any form of the perturbation that will propagate like a wave. This is known as the *linearized gravity* approach.

First, we take the flat Minkowski metric $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ and add a very small perturbation called $h_{\mu\nu}$ on top of it ("very small" meaning that $|h_{\mu\nu}| \ll 1$ for all components). The metric is then written as:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \ . \tag{2.13}$$

From $g_{\mu\nu}$, the calculation of $G_{\mu\nu}$ proceeds as described earlier, so that $G_{\mu\nu}$ can be plugged into the Einstein equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{2.14}$$

Meanwhile, the right hand side of the equation, containing $T_{\mu\nu}$ ought to be zero (to first order) in vacuum.

The Einstein field equations in this case simplify to:

$$-\Box h_{\mu\nu} + \partial_\mu V_\nu + \partial_\nu V_\mu = 0 \ , \tag{2.15}$$

(following the notation convention of [26]), where

$$V_\mu \equiv \partial_\gamma h^\gamma_\mu - \frac{1}{2}\partial_\mu h^\gamma_\gamma \, . \tag{2.16}$$

Note that $\Box$ is the d'Alembertian operator:

$$\Box \equiv \eta^{\alpha\beta}\partial_\alpha\partial_b \tag{2.17}$$

or, in a more familiar form:

$$\Box \equiv -\frac{\partial^2}{\partial t^2} + \vec{\nabla}^2 \, . \tag{2.18}$$

Something like a homogeneous wave equation is beginning to emerge here, but there are two extra terms in the way: $\partial_\mu V_\nu$ and $\partial_\nu V_\mu$. Is it possible to simplify those terms away? The answer, fortunately, is yes; this is done by making use of a clever gauge choice.

The idea is that we can slightly alter the form of $h_{\mu\nu}$ by shifting coordinates, in a way that forces $\partial_\mu V_\nu$ and $\partial_\nu V_\mu$ to 0, but doesn't change the basic assumption that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Specifically, we need to shift the coordinates by some small amount:

$$x'^\alpha = x^\alpha + \xi^\alpha(x) \, , \tag{2.19}$$

requiring that ξ is as small as h (to preserve the assumption of near-flat spacetime), and is carefully chosen with the intent to get rid of the extra terms. It can be shown that it is always *possible* to find an appropriate ξ to achieve these goals. In fact, it is possible to go a step further, and force $V_\mu = 0$. [26] As it turns out, the condition which sets these terms to zero is very much analogous to the Lorenz gauge condition used in electromagnetism, and is called the Lorenz gauge in the context of GR as well.

With the assurance that we *can* get rid of the extra terms using *some* ξ, it is not necessary to actually find the appropriate form of ξ. After all, no form has been written for h yet, and it is acceptable to assume that we are already in the appropriate coordinate system, as long as one exists. So the terms containing Vs are simply dropped, and at last, we get a homogeneous wave equation:

$$\Box h_{\mu\nu} \equiv \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) h_{\mu\nu} = 0 \, . \tag{2.20}$$

Solving an equation like this is a standard and familiar problem in physics, and upon recognizing it, we can jump directly to the next step: look for a plane wave solution. (More technically, this is done using a

Green's function formalism, as it is in electrodynamics.) The solution should be of the form:

$$h_{\mu\nu}(x) = a_{\mu\nu} e^{i(k_\alpha x^\alpha)} \,, \tag{2.21}$$

where, as a reminder, k_α and x^α are four-vectors, so that $k_\alpha x^\alpha$ contains both the spatial *and* time dependence of the wave. Meanwhile, $a_{\mu\nu}$ will be a symmetrical 4x4 matrix of constants. Because it is symmetrical, there are only 10 independent components. The matrix has the form:

$$a_{\mu\nu} = \begin{pmatrix} A & B & C & D \\ B & E & F & G \\ C & F & H & I \\ D & G & I & J \end{pmatrix} . \tag{2.22}$$

It also turns out that it is possible to make *even more* infinitesimal changes to h, without disturbing any of the conditions imposed so far. Specifically, we are allowed to make 4 more "wave-like" changes to the coordinate system (wave-like in the sense that we can add another four-component ξ to our coordinates x, as long as ξ itself fulfills the wave equation $\Box\xi_\alpha = 0$). We can use these changes to cancel out up to 4 components of h.

It is not initially obvious which components of h are best set to zero, but a standard and mathematically-useful choice has been established, known as the transverse-traceless, or TT, gauge. Using the first three available degrees of freedom, we set $h_{ti} = 0$ (and thereby also $h_{it} = 0$), reducing a down to:

$$a_{\mu\nu} = \begin{pmatrix} A & 0 & 0 & 0 \\ 0 & E & F & G \\ 0 & F & H & I \\ 0 & G & I & J \end{pmatrix} . \tag{2.23}$$

Once we make this choice, several previously imposed conditions (the wave equation $\Box h_{\alpha\beta} = 0$, and the Lorenz gauge condition that set $V_\mu = V_\nu = 0$) combine to simplify a further. First, $a_{tt} = A = 0$. Next, we get the condition $k_\mu a^{\mu\nu} = 0$, which means that the distortion will be *perpendicular* to the direction of the wave vector. At this point, it is convenient to pick a direction for our wave so that we can write its form down explicitly. Let us examine a wave propagating in the z direction, so that the condition of perpendicularity sets $a_{zi} = a_{iz} = \{G, I, J\} = 0$. Now a is of the form:

$$a_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & E & F & 0 \\ 0 & F & H & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.24}$$

Recall that setting $h_{ti} = h_{it} = 0$ used only three of the four remaining degrees of freedom. For the fourth choice, we do not actually restrict a specific component to zero, but simply require that the trace of h (the sum of its diagonal components) be zero. This tells us that $h_{xx} = -h_{yy}$, or, in the above labeling, $E = -H$.

With the form of a now highly simplified, and a direction for the wave itself chosen, we can write out the solution. The nonzero components of h are relabeled from E and F to their standard notation, h_+ and $h_\times$, and k is set to $(\omega, 0, 0, k)$ as appropriate for a z-propagating wave with angular frequency ω:

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{i(kz-\omega t)} . \tag{2.25}$$

The amplitude has two independent components, h_+ and $h_\times$, which correspond to the amplitudes of the two possible wave polarizations. The whole tensor h is referred to as the *strain tensor*.

2.3.1 Physical interpretation of the strain tensor

The concept of a strain tensor is not specific to relativity; it has a much older incarnation in which it describes material deformations. A piece of material responds to *stresses* (forces per unit area) by deforming, and the fractional deformation is described in terms of *strain*. Stress and strain in three dimensions can both be written as as 3x3 tensors. In this application, the $\mu\nu$ component of the stress tensor describes the μ component of the force acting on the ν-normal face of an infinitesimally small cube. For small forces and deformations, the strain is proportional to the stress.

While the relativistic $h_{\mu\nu}$ is defined even when no matter is present, and represents the warping of spacetime rather than forces acting through the bulk of a material, it is not named "strain" without reason. To translate the relativistic strain tensor into physical meaning, it is useful to take its lower right 3x3 subset (the spatial part of the tensor) and imagine it as the result of a stress tensor acting on an infinitesimal cube of material to create the strain tensor.

When h_+ is nonzero, two things must happen to this hypothetical cube. First, the xx component being

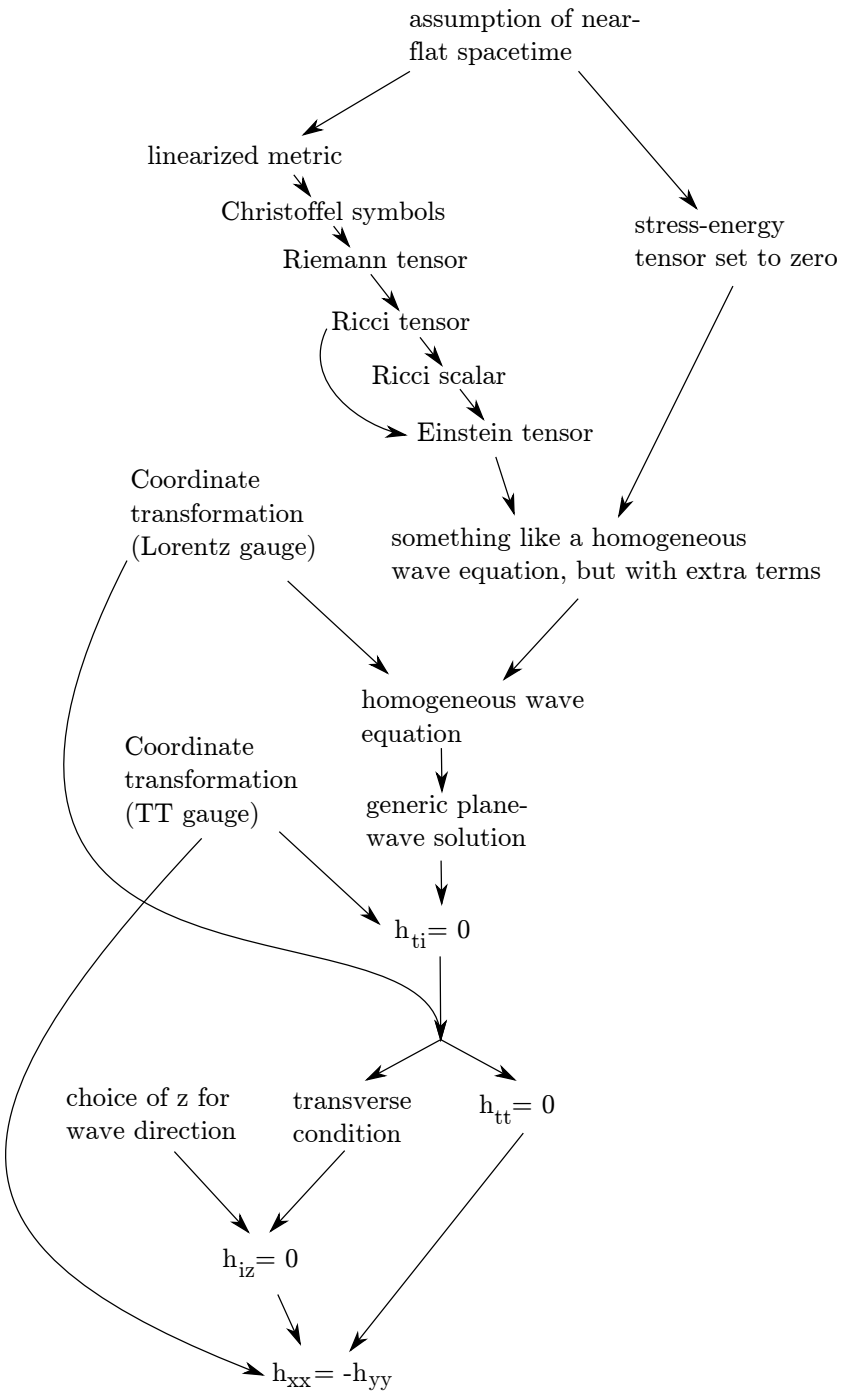


Figure 2.2: A flowchart showing how gravitational waves arise from linearized gravity.

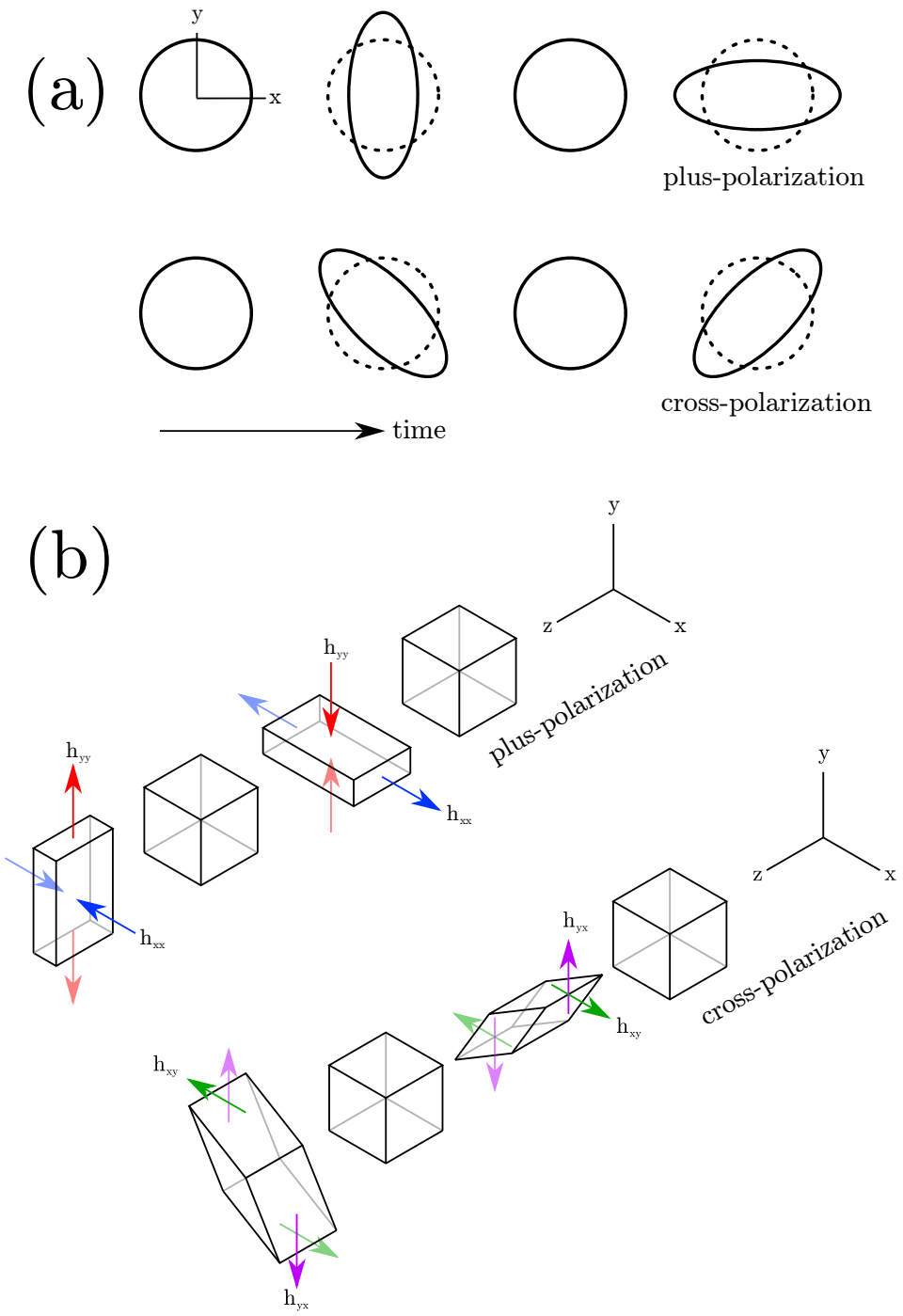


Figure 2.3: (a) Standard image of the effect of a gravitational wave on a ring of free test masses. The wave propagates normal to the page. (b) Equivalent visualization showing distortions of on an infinitesimal set of cubes. The gravitational wave here propagates along the z axis, and snapshots are frozen in time. The arrow labels indicate which component of the $h_{\mu\nu}$ is responsible for each component of the distortion. For clarity, translucent arrows are used to indicate forces acting on the hidden faces of the cube. Note that all the forces are transverse (there is no z-direction distortion), and that on-diagonal terms (h_{xx}, h_{yy}) describe pressures while off-diagonal terms (h_{xy}, h_{yx}) describe shears.

$+1$ tells us that the right-facing ($+x$-normal) face will be pulled to the right ($+x$ direction). Second, the yy component being -1 tells us that the top-facing ($+y$-normal) face will be pushed downward ($-y$ direction). This action is symmetrically mirrored on the opposite sides of the cube. The cube is stretched in the x direction and contracted in the y direction. As the wave proceeds, the signs flip, and the cube is stretched in the y direction and contracted in the x direction.

When $h_\times$ is nonzero, the xy component tells us that the top-facing (y-normal) face is be pushed rightward ($+x$ direction), creating a shear. At the same time, the yx component tells us that the right-facing (x-normal) face will be sheared upward ($+y$ direction). This stretches the cube diagonally up and to the right, and down and to the left, contracting it along the perpendicular axis. As the wave proceeds, the signs flip, and the cube will be stretched along the perpendicular diagonal.

A more relativistic way to see this behavior is to write out the spacetime interval described by this metric tensor. When h_+ is positive, The $dxdx$ term is positive, indicating that distances will increase between points that are separated in the x direction. The $dydy$ term is negative, indicating that distances will decrease between points that are separated in the y direction. And the $dxdy$ terms are positive if dx and dy share the same sign: i.e., distances will increase along one diagonal direction and decrease along the other.

2.4 History of debate over the existence of gravitational waves

Probably the most subtle and non-intuitive part of the linearized-gravity derivation of gravitational waves is the use of clever gauge choices to simplify the mathematics. Two instances of gauge fixing are required in this process: first, the Lorenz gauge, and then the TT gauge. In the early part of the 20th century, it was not clear whether the resulting form of $h_{\mu\nu}$ indicated a physical wave, or if it was merely a trick of the coordinates used.

The physical reality of gravitational waves was a point of serious contention for decades after the presentation of general relativity (see [15], which in turn draws from [29], for a summary of the debate). Einstein himself wavered on the matter, claiming in 1936 that gravitational waves do not exist, before changing his mind shortly afterward. It was not until the 1950s that solid arguments were presented in favor of physical waves arising from GR. The Chapel Hill Conference in 1957 is usually cited as a turning point in the debate.

Of particular importance to the resolution of the problem was Felix Pirani's work on the physical significance of the Riemann tensor, specifically its connection to *geodesic deviation* (the general relativistic analog of the tidal force), which is physically measurable. As later noted by both Hermann Bondi and Richard Feynman, the geodesic deviation due to a passing wave could cause a bead sliding on a rod to move, and if there is friction between the bead and the rod, this will in turn generate heat. Since energy is conserved,

this indicates that gravitational waves must carry energy, and must be detectable, at least in principle.

2.5 Sources of gravitational waves

The derivation of gravitational waves in a weak-field, linearized regime actually tells us little about the *generation* of gravitational waves. It merely indicates that gravitational waves, from some unknown source, can propagate and carry energy.

Recall that gravitational waves were previously found by solving the homogeneous wave equation:

$$\Box h^{\mu\nu} = 0 \, , \tag{2.26}$$

on the basis of the assumption that $T_{\mu\nu} = 0$. That is, we were looking at the behavior of the wave far from the source, in otherwise empty space. Adding a source, which has nonzero $T_{\mu\nu}$ requires dealing with a more complicated situation:

$$\Box h^{\mu\nu} = -\frac{16\pi G}{c^4} T^{\mu\nu} \, . \tag{2.27}$$

Again, the corresponding problem from electrodynamics is well understood; we can jump immediately to the particular solution for this equation:

$$h^{\mu\nu} = \frac{4G}{c^4} \int \frac{T^{\mu\nu}(t - |\vec{x} - \vec{x}'|/c, \vec{x}')}{|\vec{x} - \vec{x}'|} d^3\vec{x}' \, . \tag{2.28}$$

This equation tells us to integrate the stress-energy tensor over some source distribution (primed coordinates) which is changing in time. The effect of the changes at the source will be felt at a distance $|\vec{x} - \vec{x}'|$ in the time that it takes for the wave to propagate, $|\vec{x} - \vec{x}'|/c$. If we are observing the source from very far away, $|\vec{x} - \vec{x}'|$ will be approximately the same for every point $\vec{x}'$ in the source; we can call this distance r, simplifying the equation to:

$$h^{\mu\nu} = \frac{4G}{c^4 r} \int T^{\mu\nu}(t - r/c, \vec{x}') d^3\vec{x}' \, . \tag{2.29}$$

We now have to deal with the actual integral. One way to do this is to recognize that if we work in the TT gauge as described above, we only care about spatial components of $h^{\mu\nu}$. These spatial components can be found in terms of T^{tt} by applying the conservation law $\partial T^{\mu\nu}_{\ \ ,\mu} = 0$ (this conservation law, being 4-dimensional, contains within it both energy and momentum conservation). After generating an expression for the spatial components T^{ij} in terms of T^{tt} and integrating by parts, we will obtain:

$$h^{ij} \approx \frac{2G}{c^4 r} \frac{\partial^2}{\partial t^2} \int x'^i x'^j \tau^{00}(t - r/c, \vec{x}') d^3\vec{x}' \tag{2.30}$$

(see e.g. [18] for details).

The physical meaning of this result can be understood by considering it in terms of a multipole expansion. By analogy with electrodynamics, we can write out the multipole moments of the source, with the charge density replaced by the energy density T^{tt}.

$$M(t - r/c) = \int T^{tt}(t - r/c, \vec{x}') d^3\vec{x}' \qquad \text{(monopole)} \tag{2.31}$$

$$D^i(t - r/c) = \int x'^i T^{tt}(t - r/c, \vec{x}') d^3\vec{x}' \qquad \text{(dipole)} \tag{2.32}$$

$$I^{ij}(t - r/c) = \int x'^i x'^j T^{tt}(t - r/c, \vec{x}') d^3\vec{x}' \qquad \text{(quadrupole)} \tag{2.33}$$

The quadrupole term is clearly what appears in the expression for h_{ij}, so in fact we have:

$$h^{ij} \approx \frac{2G}{c^4 r} \ddot{I}_{ij}(t - r/c) \tag{2.34}$$

There are two conceptual arguments that help explain why it is the *quadrupole* moment that matters, rather than a dipole moment. First, the wave distortion itself is quadrupolar; this suggests a quadrupolar source. Second, and perhaps more concretely, the monopole and dipole terms are disallowed by conservation laws in a straightforward way. An oscillating monopole moment would correspond to an oscillation in the total amount of mass-energy present; for an isolated system, this would violate conservation of energy. A changing dipole moment would correspond to moving a mass back and forth; for an isolated system, this would violate conservation of momentum. (Electromagnetic dipole radiation relies on the existence of opposite charges, which can exchange places and produce a changed dipole without violating conservation of momentum. There is no "opposite charge" for mass.) The first term which is physically plausible for an oscillation, then, is the quadrupole term.

It is not difficult to find examples of isolated systems with changing mass quadrupole moments. Any rotating system which is not symmetric about the axis of rotation will do the trick. It is much harder to find systems where the second time derivative of the mass quadrupole moment is *large* enough to produce gravitational waves that have any hope of being detected.

2.6 Detectable gravitational waves?

> "It is obvious that [the amplitude of gravitational radiation] has, in all imaginable cases, a practically vanishing value." - Albert Einstein [15]

In 2019, the scientific world is well aware that gravitational waves can indeed be detected, thanks to numerous such detections by the LIGO-Virgo Scientific Collaboration. But this was by no means an obvious conclusion in decades past.

Debate over the physical reality of GWs in the 1950s and earlier had hinged on the question of whether they were *hypothetically* observable; even with that resolved, the question of whether they were *practically* observable was another issue altogether. The second derivative of the mass quadrupole moment, $\ddot{I}_{ij}$, has to be enormous to compensate for the factor of $2G/c^4 = 10^{-44}\frac{\mathrm{s}^2}{\mathrm{kg\cdot m}}$ in order to obtain an observable strain h, even more so in the presence of a large observation distance r.

To get a sense of the amplitude of gravitational waves, consider the following. An estimate of the magnitude of h for two orbiting masses is given by Saulson [53] in terms of their Schwarzchild radii[4], r_{s1} and r_{s2}:

$$|h| \approx \frac{r_{s1} r_{s2}}{Rr}\,, \tag{2.35}$$

where R is the radius at which the masses orbit each other, and r is again the distance to the observer. From this expression, we can see that a large value of h (relatively speaking) will require high-mass objects with a low orbital separation, as close as possible to the observer. Continuing with Saulson's example: the radius of a 1.4 solar mass (1.4 $M_\odot$) neutron star is about 10 km, making it highly massive and highly compact. Assuming a separation between the stars of just 20 km, so that they are almost touching, and a distance to the observer of about 15 Mpc, this system yields a strain amplitude of about 10^{-21}. Even in this optimistic case, the physical effect is tiny: the passing gravitational wave would distort the diameter of Earth by an amount less than the radius of a proton.

The miniscule amplitude of a gravitational wave does not mean that it carries a miniscule amount of energy. On the contrary, gravitational waves are (proportionally) quite energetic. Spacetime is stiff, in the sense of a spring being stiff: it requires a great deal of energy to distort it even slightly. As a very rough estimate, it is possible to calculate an effective (albeit frequency-dependent) Young's modulus for spacetime, which indicates that it is about 10^{20} times stiffer than steel. [43] In the example above, the energy flux passing Earth would be around 1.6 mW/m^2 [50], on the same order of magnitude as the flux of

[4] The Schwarzschild radius $r_s = 2GM/c^2$ is the radius of the event horizon of a black hole mass M. There is a characteristic r_s for every mass, regardless of whether the object in question is actually a black hole.

electromagnetic radiation reflected to Earth from the full moon.

2.6.1 The first observational evidence for gravitational waves

The fact that gravitational waves carry a significant amount of energy makes their indirect detection much easier than direct detection. The rotation of the source is affected by its loss of energy through gravitational radiation, and this change can be observed electromagnetically. Observations of the Hulse-Taylor binary system, starting in 1974, [20] provide a classic and illustrative example.

The Hulse-Taylor binary is a system located about 21,000 light years away from Earth, composed of two neutron stars in orbit. One of those two stars is a pulsar: a neutron star which emits a beam of electromagnetic radiation, at an angle from its spin axis. As the pulsar spins, the beam sweeps past a distant observer, creating the effect of a "pulse" of observed light. In this case, the pulsar spins about 17 times per second.

Hulse and Taylor measured a periodic variation in the pulse arrival time, which shifted faster and slower with a period of 7.75 hours. This was interpreted as an effect of the Doppler shift from the binary's orbit, and hence as a measurement of the orbital period. Continued observation of the source indicated that the orbital period was very slightly changing, a trend which has continued since. The binary orbital period has in fact changed by less than a minute between 1974 and today, but even this small change is important: it indicates that the binary is losing energy through some invisible mechanism. In fact, it is losing energy at a rate which very precisely matches the predicted energy loss through gravitational waves.[62]

For more than 40 years, the Hulse-Taylor binary remained one of the strongest pieces of evidence (albeit indirect evidence) for the existence of gravitational waves, which would not be directly detected until 2015.

2.6.2 The first direct observation of gravitational waves

On September 14, 2015, the first direct observation of gravitational waves was made [60]. The signal registered first at the LIGO detector in Livingston, Louisiana; then, seven milliseconds later, at the LIGO detector in Hanford, Washington. It was a strong signal, with a peak strain of 10^{-21} and SNR 24, corresponding to a false alarm rate of less than 1 in 200,000 years.

The signal, named GW150914, had come from the merger of two black holes, approximately 39 and 26 solar masses, more than a billion light years away. As the two objects orbited one another, they radiated gravitational waves, slowly shrinking the orbit and increasing the orbital frequency until they coalesced into a single object. This coalescence event released a huge amount of energy in the form of gravitational radiation: more than two solar masses worth. The waveform detected by LIGO was a characteristic "chirp" pattern

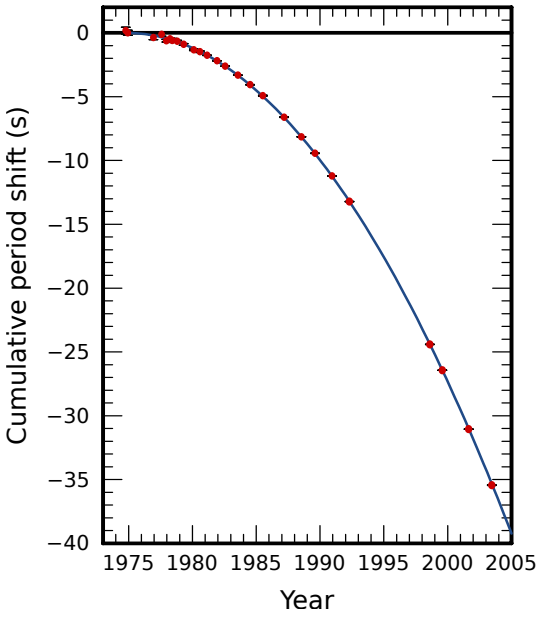


Figure 2.4: Cumulative orbital period shift of the Hulse-Taylor binary, 1974-2004. The red points are observational measurements; the blue line is the general relativistic prediction. Figure from [62], via WikiMedia Commons. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

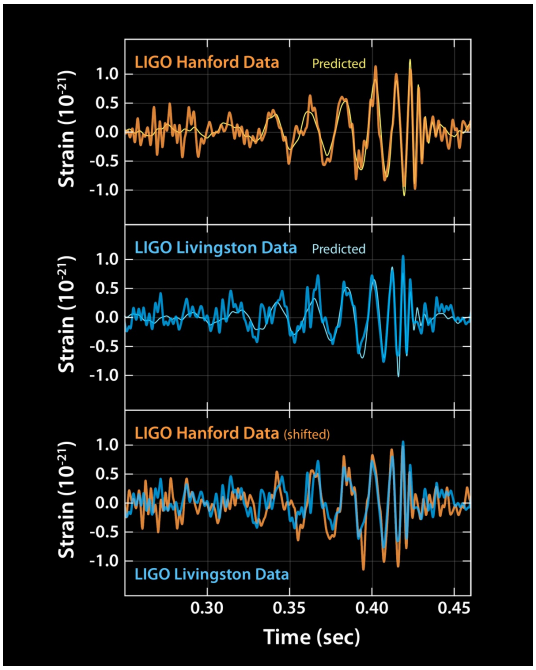


Figure 2.5: Signals observed at Hanford and Livingston for the binary black hole merger GW1501914. Figure from [6]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

with increasing frequency and amplitude, in excellent agreement with numerical relativity simulations (see figure 2.5).

The detection of GW150914 marked the beginning of gravitational wave astronomy, and was the first of many detections [59] made by the LIGO and Virgo scientific collaborations.

Chapter 3

Gravitational wave searches with LIGO

3.1 Principles of interferometry for gravitational wave detection

3.1.1 Michelson interferometer

The Michelson interferometer, which forms the basis for LIGO and other earth-based gravitational wave detectors, predates LIGO by about a century. Its first notable use was in Albert Michelson and Edward Morley's 1887 attempt to measure the speed of Earth's motion, relative to the stationary "luminiferous aether" which was then thought to pervade space. Despite the innovative experimental setup, Michelson and Morley failed to detect any motion whatsoever. Their null result would later be explained by special relativity, and would spell the beginning of the end of luminiferous aether theory.

Interferometry, then, played a role in testing relativity even *before* the theory was formulated, and long before interferometric gravitational wave detectors were conceptualized. It is not a simple historical coincidence that interferometry and relativity are so entwined. Interferometry uses light to measure space, and light speed is built into the fabric of relativity's description of space. Interferometry is also very sensitive, allowing for measurement precision at or below the wavelength of the light used. It therefore is a good candidate for measuring (or failing to measure) extremely tiny changes in length, on the order necessary to test relativistic effects.

The basic principle of a Michelson interferometer is as follows: a beam of light (generally a laser, at least in modern times) is split in two, and the beams travel along separate paths before recombining. The interference at the point of recombination, measured by a photodetector, indicates the difference in path length traveled by the two beams.[1] If both beams travel the same distance, they will interfere constructively, leading to a bright fringe. If the relative distance changes by half a wavelength, they will interfere destructively, leading to a dark fringe. In the region between dark and bright fringes, the power reaching the detector can be

[1] It actually indicates the difference in path length modulo the wavelength of light used, but interferometry is typically used to measure very tiny changes in relative displacement, which are smaller than the wavelength.

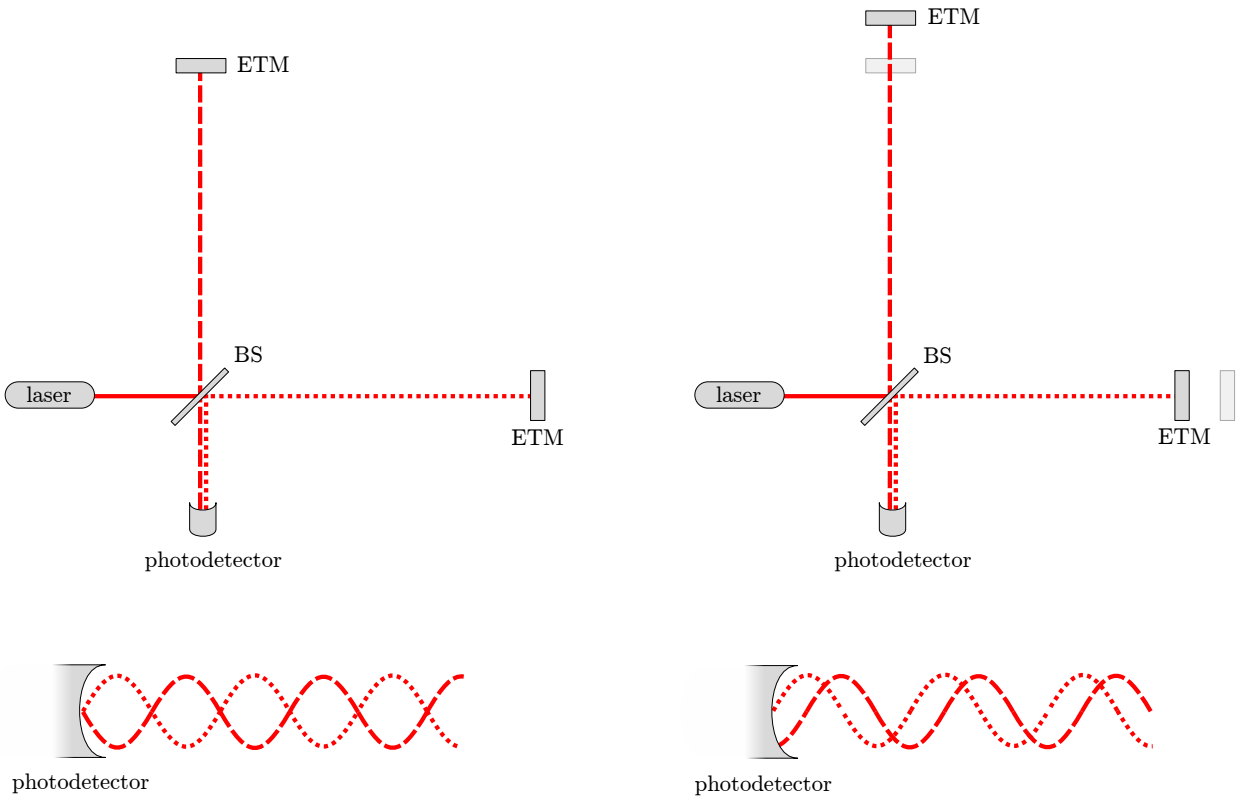


Figure 3.1: Diagram of a Michelson interferometer (left), showing the effect of a passing gravitational wave (right, highly exaggerated). A single laser beam (solid red line) is split into two beams (dotted lines of different styles), which reflect from mirrors at the end of the arms before returning to the beam splitter and recombining. The combination of the two beams is read by the photodetector. A passing gravitational wave, propagating perpendicular to the field of view and polarized in the + direction, causes a change in the relative path length of the two beams, changing the amount of destructive interference at the photodetector. This interferometer is configured to operate on the dark fringe (destructive interference) in the absence of a passing wave.

translated into a measurement of relative fractional length change h:

$$h = \frac{\Delta L}{L} \tag{3.1}$$

where L is the length of the detector arms and ΔL is the change in length. Figure 3.1 provides a diagram of this effect.

Interferometers are typically operated on the dark fringe (destructive interference at the output port). There are several reasons for this choice. First, it is technically easier to measure a small fluctuation in power when the total power is low. Second, in this configuration, any instrumental disturbances that are common to the two arms (such as laser noise) have a minimal effect, while differential oscillations are maximized. [12]

An interferometer is well suited to making direct measurements of gravitational waves, because it measures exactly what gravitational waves induce: a change in the relative length of two perpendicular paths. However, this does not mean it is straightforward to measure a strain of 10^{-21} or less, as is necessary for detecting gravitational waves. As an order of magnitude estimate, consider a basic Michelson interferometer with arms on the order of 1 km, and a laser wavelength of about 1 micron. The strain that will cause this detector to cycle through one fringe (a phase of 2π) is $\lambda_{\text{laser}}/L = 10^{-9}$. In other words, it remains 12 orders of magnitude away from the required sensitivity.

3.1.2 Photon counting and the limits of fringe splitting

Fortunately, it is possible to measure a change in the relative phase that is much less than 2π, with a sensitive photodetector. The main limitation is shot noise, or photon counting noise. Suppose we have a photodiode which counts the arriving number of photons, N_{photons}, over some small fixed time. There will be a natural fluctuation of $\sqrt{N_{\text{photons}}}$ in the number of photons observed, since this process is governed by Poisson statistics. In order to detect a change in the amount of light arriving at the photodiode, we will need a change in the path length on the order of:

$$\Delta L \approx \frac{\sqrt{N_{\text{photons}}}}{N_{\text{photons}}} \lambda_{\text{laser}} \ . \tag{3.2}$$

The number of photons is related to the laser power, and to the time over which we observe. If we are observing a gravitational wave, the observing time is limited by its frequency, f_{GW}. Since each photon has energy $h_{\text{Planck}}c/\lambda_{\text{laser}}$, we have approximately:

$$N_{\text{photons}} \approx \frac{P_{\text{laser}}}{h_{\text{Planck}}c/\lambda_{\text{laser}}} \frac{1}{f_{\text{GW}}} \ , \tag{3.3}$$

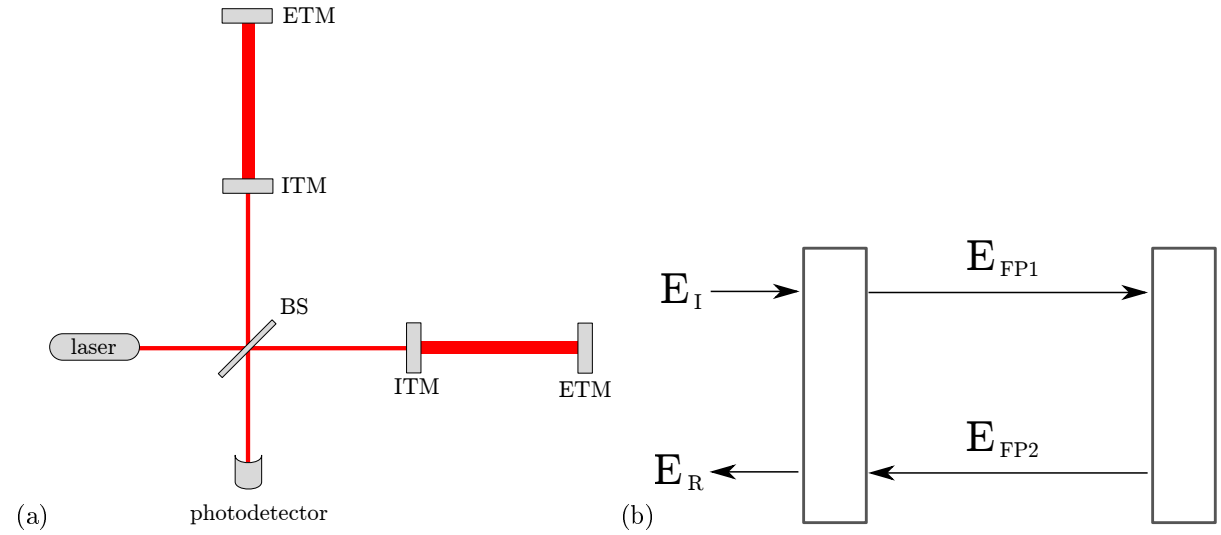


Figure 3.2: (a) An interferometer with a Fabry-Pérot cavity on each arm. The bold lines between the mirrors indicate much higher laser intensity in the cavity than the rest of the optics. (b) Diagram of cavity fields.

And the ΔL required to detect a change is then:

$$\Delta L \approx \sqrt{\frac{\lambda_{\text{laser}} h_{\text{Planck}} c f_{GW}}{P_{\text{laser}}}} \,. \tag{3.4}$$

We are, of course, interested in strain ($h = \Delta L/L$) rather than just ΔL. Rearranging to determine the required laser power in terms of strain and the other parameters, we get:

$$P = \frac{\lambda_{\text{laser}} h_{\text{Planck}} c f_{GW}}{(hL)^2} \,. \tag{3.5}$$

For an interferometer with a 1 micron laser, 1 km arms, observing a 100 Hz gravitational wave with strain amplitude 10^{-21}, this works out to about 20 MW of laser power, which is clearly technologically untenable.

The other parameter over which we have some control is the arm length of the interferometer, L. Increasing the length (or the effective length seen by a photon traveling through the interferometer) will help greatly. One way to do this, without making an unreasonably large detector, is to reflect the light beam back and forth between a pair of mirrors.

3.1.3 Increasing the effective arm length with Fabry-Pérot cavities

One way to increase the sensitivity of a Michelson interferometer is to add a Fabry-Pérot cavity on each arm, by inserting an additional mirror. The mirrors on the ends of the arms are called the end test masses (ETMs) while the mirrors nearest the beam splitter are called the input test masses (ITMs). The beam resonates in the cavity, increasing the light storage time, and thereby the cumulative phase shift of the light

in the presence of a passing gravitational wave. This is equivalent to increasing the effective length of the arm. A detailed treatment of Fabry-Pérot cavities as used in gravitational-wave detectors can be found in numerous texts, e.g. [18, 12, 54].

To simplify the situation, we first assume that the optical components are lossless, so that $|r|^2 + |t|^2 = 1$, and also that the ETM is perfectly reflective, so that $r_{\rm ETM} = 1$ and $t_{\rm ETM} = 0$. Then the circulating field inside the cavity will be (see figure 3.2b):

$$E_{\rm FP} = \frac{t_{\rm ITM}}{1 - r_{\rm ITM} e^{-i2kL_{\rm FP}}} E_I \tag{3.6}$$

[18, 52] where E_I is the incident light on the ITM and $L_{\rm FP}$ is the length of the Fabry-Pérot cavity. This quantity is maximized when:

$$e^{-i2kL_{\rm FP}} = 1 \, , \tag{3.7}$$

or when $L_{\rm FP} = N\pi/k = N\lambda/2$. In other words, the cavity will resonate when its length is an integer multiple of a half-wavelength. Meanwhile, the reflected light from the cavity is given by:

$$E_R = \frac{r_{\rm ITM} - e^{-2ikL_{\rm FP}}}{1 - r_{\rm ITM} e^{-2ikL_{\rm FP}}} E_I \, . \tag{3.8}$$

On resonance, this reduces to $E_R = -E_I$, or a phase shift of π between the incident and reflected light.

Suppose now that the cavity is pushed slightly off of resonance by the passing of a gravitational wave, so that

$$L_{\rm FP} = L_{\rm res} + \Delta L \, . \tag{3.9}$$

Then we have the following approximate expression for the reflectance:

$$r_{\rm FP}\left(L_{\rm res} + \Delta L\right) \approx r_{FP}(L_{\rm res}) + \frac{dr_{\rm FP}}{dL}|_{L_{\rm res}} \Delta L \, , \tag{3.10}$$

$$r_{\rm FP}\left(L_{\rm res} + \Delta L\right) \approx -\exp\left[-2ik\Delta L \left(\frac{1 + r_{\rm ITM}}{1 - r_{\rm ITM}}\right)\right] \, . \tag{3.11}$$

If there were no Fabry-Pérot cavity, and instead we were dealing with a simple Michelson interferometer, then the phase shift acquired by light traveling round-trip through an arm length change of ΔL would be just $-2ik\Delta L$. In the case of the Fabry-Pérot cavity, we get an additional factor of $\left(\frac{1+r_{\rm ITM}}{1-r_{\rm ITM}}\right)$. This factor is known as the arm *cavity gain*.

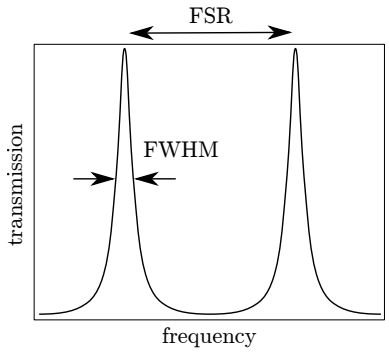


Figure 3.3: Schematic plot showing the free spectral range (FSR) and full width half max (FWHM) of a cavity on a plot of transmission versus frequency. Finesse is a ratio of these two values.

A more common figure of merit used in the literature is cavity *finesse*, which is a measure of the sharpness of the cavity's resonance peak:

$$\mathcal{F} \equiv \frac{f_{\mathrm{FSR}}}{f_{\mathrm{FWHM}}} \,, \tag{3.12}$$

where f_{FSR} is the free spectral range, and f_{FWHM} is the full width at half max of the peak. This is related to the optical cavity gain (approximately) as:

$$\frac{\pi}{2} G_{\mathrm{arm}} \approx \mathcal{F} \,. \tag{3.13}$$

(See [18] for details. The key point is that higher finesse yields higher optical gain.)

3.1.4 Power recycling

In the case of perfect destructive interference, all the laser light is redirected back toward the laser, and none toward the output unless the interferometer is disturbed. This light would effectively be wasted, were it not for the addition of a power recycling mirror. The power recycling mirror (PRM) forms a resonant cavity with the ITMs, and returns light toward the beam splitter. This increases the laser power in the interferometer.

3.1.5 Resonant sideband extraction

The earlier discussion of Fabry-Pérot cavities might lead one to believe that increasing the cavity gain indefinitely is a sound goal. Unfortunately, this comes with a tradeoff. Although longer light storage time in the arms builds up a cumulative phase shift, it also washes out high-frequency signals. A storage time of τ_s creates a lag in the response of the cavity, which then acts as a low-pass filter with corner frequency at the cavity pole:

$$f_{\text{pole}} = \frac{1}{4\pi\tau_s} \,. \tag{3.14}$$

The amplitude above this point falls off as $1/f$. [52] Furthermore, increasing the cavity gain comes with practical downsides: more light in the arms means more radiation pressure on the test masses, and more displacement noise as a result. An alternative to increasing the storage time would be to increase the power incident on the beam splitter (either by increasing the laser power or through use of a power recycling mirror), but this comes with its own set of downsides, including thermal lensing effects.

A clever solution to this problem is resonant sideband extraction. In order to understand this mechanism, it is necessary to think of the laser light in terms of a *carrier* with *sidebands*. The main laser frequency, f_{laser}, is the carrier. When a gravitational wave passes the detector, it modulates the phase of the carrier, introducing frequency sidebands on either side of the central frequency, $f_{\text{side}} = f_{\text{laser}} \pm f_{\text{GW}}$. On or near the dark fringe, the sidebands alone will reach the output, as the carrier is redirected back toward the laser.

By placing a mirror before the output, it is possible to create a new cavity in which the sidebands resonate. This new mirror forms a coupled cavity with the ITMs, and lowers the effective ITM reflectivity for the signal sidebands, but not for the carrier. The lowered reflectivity decreases the storage time for the signal sidebands, even when the Fabry-Pérot cavity gain is high and the carrier has a long storage time. [46]

3.2 Advanced LIGO

3.2.1 Detector network

The Advanced LIGO (aLIGO) interferometer [35] is the successor to Enhanced LIGO, which in turn succeeded Initial LIGO. The upgrade to aLIGO was completed in 2014. There are two LIGO detectors currently in operation. One is located in Hanford, Washington (called LHO, for LIGO Hanford Observatory), and the other in Livingston, Louisiana (called LLO, for LIGO Livingston Observatory). These detectors are identical in design, and as close to identical as possible in operation. A third detector is planned for construction in India. LIGO operates in a network with the Virgo detector in Italy,[2] as part of the LIGO-Virgo Scientific Collaboration (LVC).

A network of detectors is far more useful than any single instrument, for a number of reasons. A true signal will be coincident in multiple detectors, accounting for the speed-of-light travel time between them, which can be used to confirm its astrophysical origin. Multiple detectors also combine to improve the sky localization for incident signals. Gravitational wave detectors are not telescopes; they do not point. Instead,

[2]The Virgo detector will not be discussed extensively in this thesis, because its sensitivity is considerably worse than either of the LIGO detectors, and it only joined for a small section of the most recent observing run (see figure 3.7).

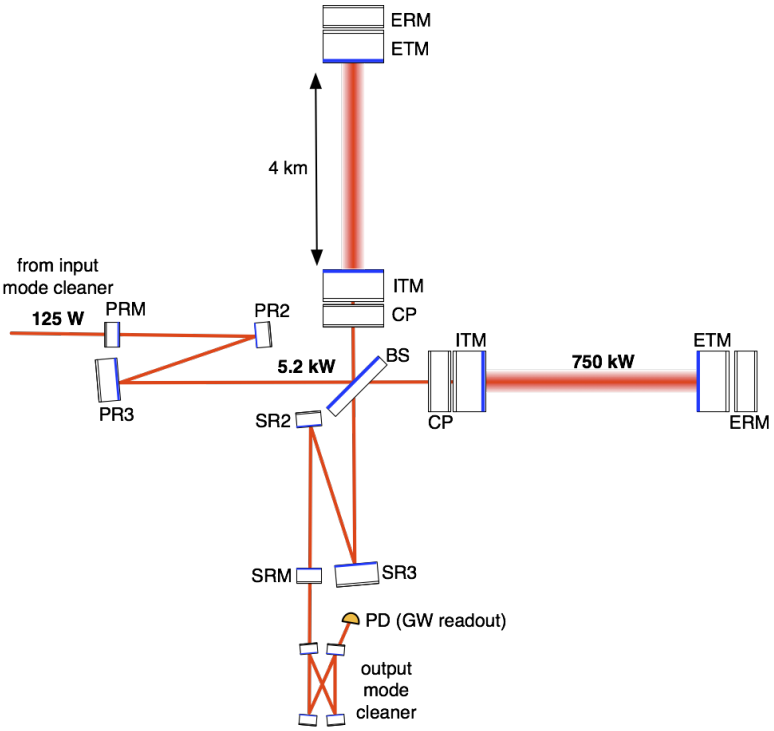


Figure 3.4: The basic design of Advanced LIGO. Figure from [13], adapted from [35]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

they act like antennae, receiving signals from many directions at once. Every detector added to the network improves the ability of the network to triangulate the source of the signal. Finally, detectors do not run constantly; they shut down for any number of reasons, from maintenance and repairs to local events like earthquakes or flooding. More detectors in the network increases the chance of having a coincident detection at multiple sites, and allows the wave polarization to be reconstructed. [3]

3.2.2 Detector schematic

The aLIGO detectors employ all three of the measures discussed earlier in this section to achieve their sensitivity: Fabry-Pérot cavities (which provide a cavity gain of two orders of magnitude), power recycling, and resonant sideband extraction. Somewhat confusingly, the signal extraction cavity and mirror are referred to as the "signal recycling mirror" (SRM) and "signal recycling cavity" (SRC), a term which has also been used to refer to a scheme for increasing the signal storage time rather than decreasing it [46]. In any case, the combination of this technique with power recycling is referred to as *dual recycling*.

The detectors have 4km arm lengths, and each arm contains a Fabry-Pérot cavity with a finesse of 450.

[3] There are some caveats to this statement. First, detection depends on the sensitivity of each detector. Also, detectors have antenna patterns; it is possible for a detector to miss a signal that lies in one of its blind spots, although this is not always a downside. In the case of the first BNS detection, the *lack* of detectable signal in the Virgo interferometer helped to rapidly narrow down the sky location of the source.

The entire beam tube and the interferometer optics are maintained in vacuum, to decrease both displacement noise of the test masses and scattering of the laser. The dominant residual gas component, hydrogen, persists at a pressure of just 4×10^{-7} Pa.

The main laser is an Nd:YAG with a wavelength of 1064 nm, capable of supplying up to 180 W, although 125 W is the maximum power to be used at design sensitivity. Along with measures to stabilize the laser frequency, beam direction, and intensity, it comprises the pre-stabilized laser (PSL) system. The beam size ($1/e^2$ radius) on the core optics ranges from 2.1 mm to 6.2 cm, with the widest beam spots on the ETMs in order to reduce thermal noise.

The test masses are 40 kg fused silica, and are designed for no more than 75 ppm round-trip optical loss in the arm cavities. All the test masses are suspended by pendulum systems, using fused silica fibers. Behind each ETM is the end reaction mass (ERM), and behind each ITM is a compensation plate (CP). The CP and the ERM each have a pattern of gold electrodes, which enables electrostatic actuation on the masses.

The seismic isolation system for aLIGO has two parts: active and passive. The passive seismic isolation takes the form of a quadruple suspension for the end test masses, which naturally supresses low-frequency seismic noise. Each stage is a pendulum, which suppresses seismic noise by a factor proportional to the square of the driving frequency (above the resonant frequency of the pendulum, which for aLIGO is about 0.76 Hz). Advanced LIGO also employs an active seismic isolation system, including both internal (in vacuum) and external stages. This system senses and actively compensates for ground motion.

3.2.3 Readout

Recall that a gravitational wave passing the detector creates signal sidebands, with frequencies $f_{\text{side}} = f_{\text{laser}} \pm f_{\text{GW}}$. This signal frequency is in the terahertz band; it cannot be straightforwardly detected by the photodetector. In order to decouple the audio-band gravitational wave components from the terahertz laser frequency, the signal must be compared with a *local oscillator*, and the beat frequency between them detected. This oscillator could be light from a separate beam (heterodyne detection) or, ideally, from the laser light itself (homodyne detection).

Homodyne detection is used in aLIGO. Initial LIGO was configured for near-perfect destructive interference at the output (dark fringe operation), but this prevented the use of homodyne detection, as the laser light did not reach the output to be used as a local oscillator. Advanced LIGO makes homodyne detection possible by holding the interferometer slightly off the dark fringe.

The photodiode readout provides a measure of the relative length between the detector arms; this data is known as DARM, for differential arm, and is the main channel of detector output. DARM data is acquired

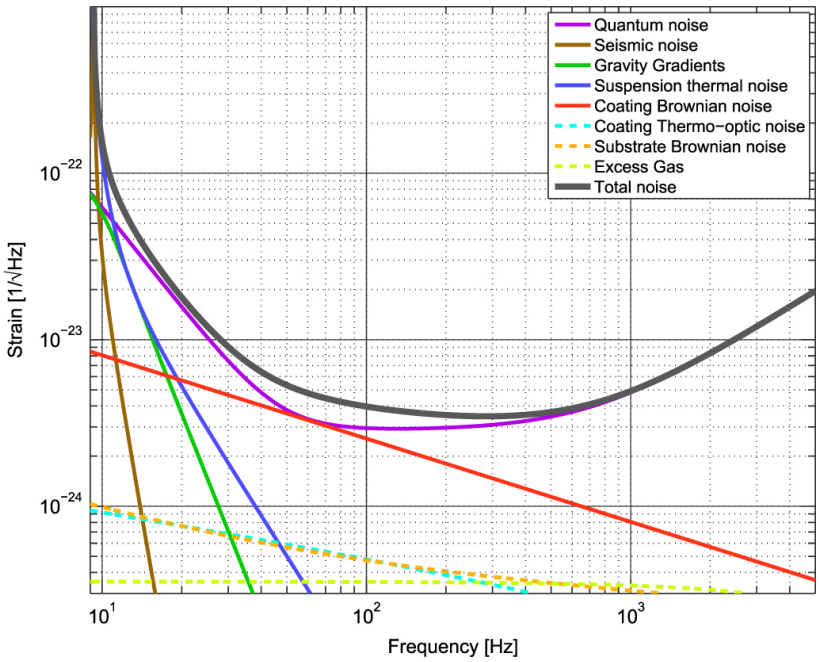


Figure 3.5: Contributions to the noise spectrum for aLIGO, as designed. Note that "quantum noise" encompasses not only fluctuations in photon counting at the photodiode, but also changes in radiation pressure on the optics due to photon number fluctuations. Figure from [35]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

at a rate of 16,384 samples per second, and is normalized by the length of the detector arms to recover the dimensionless strain, h.[4]

3.2.4 Sensitivity

3.2.4.1 The aLIGO design noise curve

The dominant sources of noise in aLIGO depend on the frequency range in question. At very low frequencies, seismic noise dominates. At very high frequencies (above 100 Hz), shot noise dominates, as photon counting statistics are lower for rapid oscillations. The intermediate zone, often called the "bucket", is the most sensitive band. The dominant noise sources here are less straightforward, but notably include thermal effects and degrees of freedom in the alignment of the optics.

3.2.4.2 The aLIGO noise curve in reality

The design sensitivity curves are misleadingly clean. In reality, the spectrum is complicated, filled with persistent noise artifacts that arise from a variety of sources (some well-understood, others less so). Figure 3.6 provides a more realistic picture of the sensitivity achieved in the O2 run. For example, there is an

[4] In fact there are several DARM channels per detector, with different whitening and calibration filters applied. Currently, there are three 16 kHz DARM channels acquired at Livingston, and 5 at Hanford (the additional two at Hanford are calibrated).

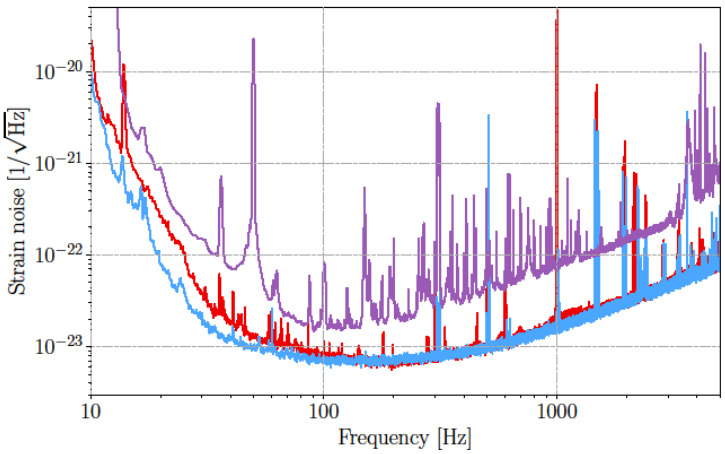


Figure 3.6: Noise spectra for Hanford (red), Livingston (blue) and Virgo (purple), representative of the best sensitivity achieved in each detector during the second observing run (O2). Figure from [58]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

unavoidable peak at 60 Hz due to the power mains (for Virgo, this is 50 Hz). There are also, as mentioned earlier, peaks corresponding to vibrational modes of the test mass suspensions. By design, the suspensions have a high mechanical Q, and their resonant frequencies are high (~500 Hz), well away from the bucket. This choice limits the amount of the spectrum that they contaminate, and keeps them out of the most sensitive region. Other spectral features, both narrow and broad-band, arise from a variety of instrumental and environmental sources, and change over timescales both long and short. Transient are not visible as lines in the spectra, but are very much present, and problematic for transient gravitational wave searches.

3.2.4.3 Binary neutron inspiral star range

A standard shorthand measure of strain sensitivity is the observation range at which the aLIGO detectors would be sensitive (SNR=8) to a binary neutron star inspiral in which both component masses are $1.4M_{\odot}$. This is known as the BNS inspiral range. BNS ranges are reported as an average over the orientation of the source, and the observation volume. The BNS range for aLIGO at design sensitivity is 190 Mpc. In practice, the detectors have not yet reached this range. Figure 3.7 shows the BNS range over the course of the most recent observing run.

3.2.5 Data channels and detector characterization

The differential arm length (DARM) is far from the only data collected and stored at the detector sites. A variety of detector systems and sensors also generate information, and this information is often critical for evaluating the quality of DARM data. Each stream of data is known as a *channel*; the LIGO Channel

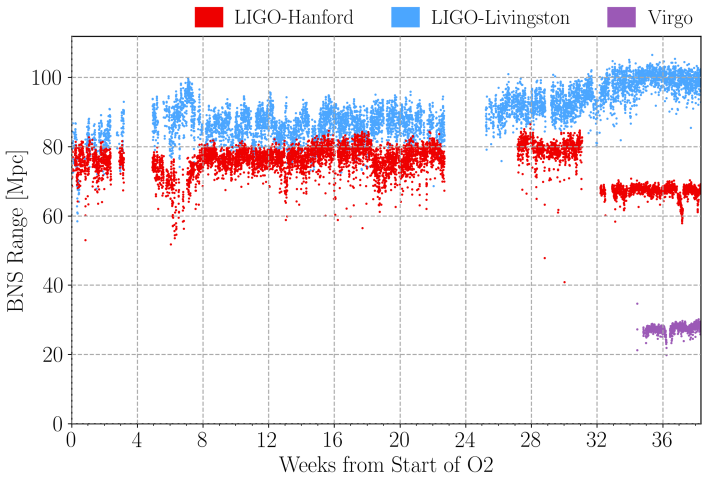


Figure 3.7: Binary neutron star inspiral range for Hanford (red), Livingston (blue) and Virgo (purple), over the course of the second observing run (O2). The Hanford drop in sensitivity at week 31 corresponds to an earthquake in Montana, which degraded the detector sensitivity. Virgo joined the observing run near the end. Figure from [58]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

Information System catalogs more than 400,000 entries between the two detectors.

An important subset of channels are associated with the physical and environmental monitoring (PEM) system. The PEM system includes accelerometers, magnetometers, temperature sensors, microphones, weather monitors, voltage monitors, and more. By matching noise artifacts in DARM with similar artifacts in PEM channels, it is often possible to narrow down the probable cause of the both transient and persistent noise sources. A map of the PEM system at Hanford is shown in figure 3.8.

In addition to automatic data collection, the status of each detector is logged continuously by the staff on site, including maintenance, upgrade, and troubleshooting measures. This information is collected in an online logbook, known (for aLIGO) as the *alog*. [1]

This wealth of data from the detectors provides the basis for extensive detector characterization and signal vetting efforts. More often than not, the main problem for detector characterization is an *excess* of information to filter through. This motivates automation. A number of automated data quality checks have been developed to identify transient noise sources, motivated by the need for low-latency measures of detector performance: it is important to know immediately which signals are real so that alerts can be issued to electromagnetic observers. Fewer automated checks currently exist for persistent noise.

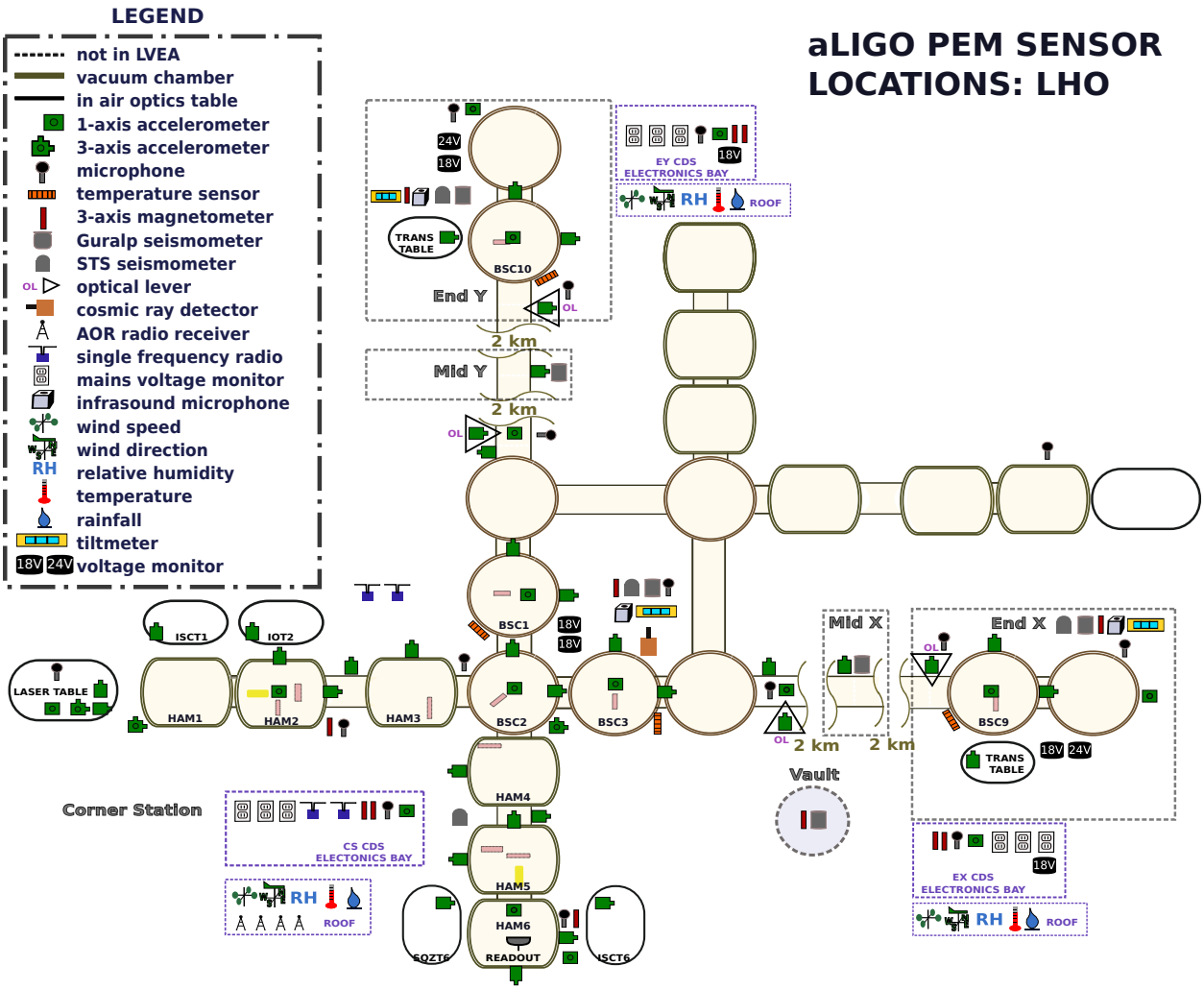


Figure 3.8: Map of the physical and environmental monitoring system at the LIGO Hanford Observatory. Figure via [10]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

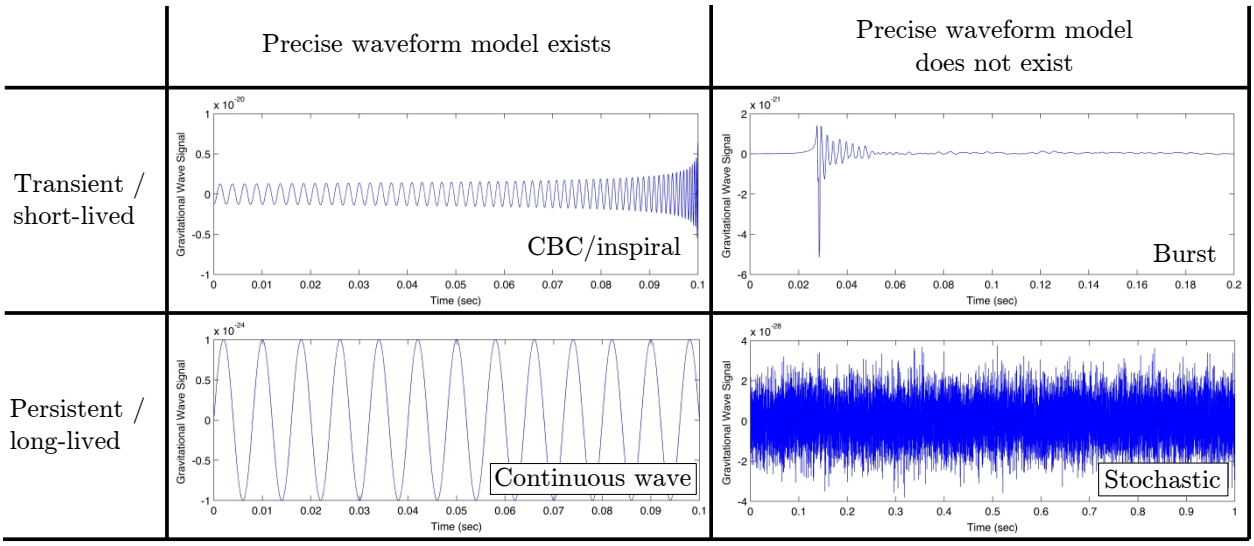


Figure 3.9: Example waveforms for CBC, burst, continuous, and stochastic signals, organized according to duration and the existence of well-modeled expected waveforms. This figure is adapted from [7], with credit to A. Stuver / LIGO. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

3.3 Overview of sources and searches

3.3.1 Organization of LIGO searches

Gravitational wave searches performed by the LIGO scientific collaboration are typically divided into four categories, along two axes: strong transient sources *versus* weak persistent signals, and well-modeled versus unmodeled signals.

The distinction between transient and persistent signals is critical; very different search methods are necessary for the two categories. Transient signals must be found and vetted quickly, so that alerts may be issued to electromagnetic observers outside the LVC. There is no opportunity to revisit these sources using data taken at a later or earlier time. Persistent signals, on the other hand, allow for data analyis to proceed at a slower pace, and the sources can be revisited. However, these signals are expected to be much weaker, and therefore require even more sensitive search techniques. There is also a difference in detector characterization needs between the two categories. Unsurprisingly, searches for transient signals are primarily affected by transient noise sources, whereas searches for persistent signals are primarily affected by persistent noise sources.

3.3.1.1 Compact binary coalescences

The inspiral and collision of two compact objects is known as a compact binary coalescence, or CBC. All LIGO detections so far have been CBC signals. Each of the objects may be a neutron star or a black hole; so far, only binary black hole (BBH) mergers and binary neutron star (BNS) mergers have been clearly observed. In the first observing run (O1), LIGO detected 3 BBH mergers. In the second observing run (O2), an additional 7 BBH mergers and one BNS merger (August 2017) were detected [58]. At the time of writing,[5] there have been 10 merger candidates so far in O3. Of the ten, seven were likely BBH mergers; each of the remaining three likely contain at least one neutron star [5].

CBC signals are divided into three stages: the inspiral, in which the two objects approach each other; the merger, in which they coalesce; and (in the case of black hole) the ringdown, in which the resulting black hole sheds excess angular momentum in the form of gravitational waves. The inspiral phase is generally modeled using a post-Newtonian approximation, while the merger requires a full numerical relativity simulation.

The most sensitive form of CBC data analysis is *matched filtering*, in which predicted waveform templates are directly compared to strain data. Matched filtering is optimally sensitive, but it represents a computational challenge, because the unknown signal must be compared with a large number of predicted waveforms in order to find a match. A large number of templates must be placed across the parameter space of interest, such that no incoming signal is too far from *some* template (where "too far" is defined in terms of acceptable loss of signal).[6]

3.3.1.2 Bursts

A generic transient excess of power in the strain channel is known as a burst. Searching for bursts in general will also catch strong CBC signals, so burst-focused analyses (which are typically faster than CBC-specific analyses) often provide valuable low-latency CBC alerts. In fact, GW150914 was first registered by a burst analysis pipeline.

In general, however, burst analysis is concerned with those signals that are *not* CBCs. A standard example of a burst source is a supernova core collapse.[7] A core collapse is an extremely complicated event, requiring sophisticated relativistic magneto-hydrodynamics to model. The best simulations do not approach the precision and reliability that would be needed for a templated search method. Supernova core collapses are also not alone in the category of unmodeled or poorly-modeled transient signals. Other possible burst

[5] May 17, 2019

[6] This is a general problem with *all* templated searches, and will later be discussed in detail in the context of continuous-wave analysis with TwoSpect.

[7] Recall that a *spherically symmetrical* core collapse will not produce gravitational waves, because there is no changing quadrupole moment. Fortunately, there is evidence that such collapses are asymmetrical, and there are many proposed mechanisms for the asymmetry.

sources include the sudden release of energy from a magnetar, or even more exotic sources like cosmic strings [50].

3.3.1.3 Continuous waves

The canonical source for a continuous wave is not a compact binary, but rather a neutron star with some amount of asymmetry about the spin axis. Such a system is expected to emit gravitational waves continuously at a fixed frequency, but even in the most optimistic scenario, the signals will be several orders of magnitude weaker than the characteristic amplitude for a CBC. Continuous wave (CW) analysis relies on summing data over a long period of time (months to years) in order to recover small signals from a noisy background. Continuous waves will be discussed in detail in chapter 4.

3.3.1.4 Stochastic background

Stochastic background searches do not attempt to distinguish individual signals, but rather to measure the spectrum associated with a superposition of many incoherent signals. The stochastic gravitational background will simply appear as noise in a single detector, so stochastic searches rely on correlations between detectors for their analyses. One source of the stochastic background is the combination of inspirals, bursts, or continuous waves originating throughout the universe. There may be anisotropy in the stochastic background from such sources, depending on their density in a given region of space.

Another component of the stochastic background, and one with great scientific relevance, is the gravitational radiation from the early universe. The gravitational wave background is analogous to the cosmic microwave background, being the redshifted signals from a period early in the history of the universe when gravitational waves decoupled from matter (around the Planck time, 10^{-43} s). Measurements of its spectrum would similarly probe the inflationary era of cosmic expansion.

3.3.2 LIGO in the context of future experiments

LIGO can observe in only a small segment of the gravitational wave spectrum. In particular, it is limited at low frequencies by seismic noise, which is an unavoidable problem for interferometers on Earth. Fortunately, a number of projects are in development which promise to open up new observational bands. Space-based interferometers will obviously not be limited by seismic noise, enabling them to make observations of low-frequency sources such as massive binaries and extreme mass ratio inspirals.

Pulsar timing arrays are another opportunity for low-frequency gravitational wave searches. As a gravitational wave passes Earth, the stretching of space and time should delay the pulses in a characteristic

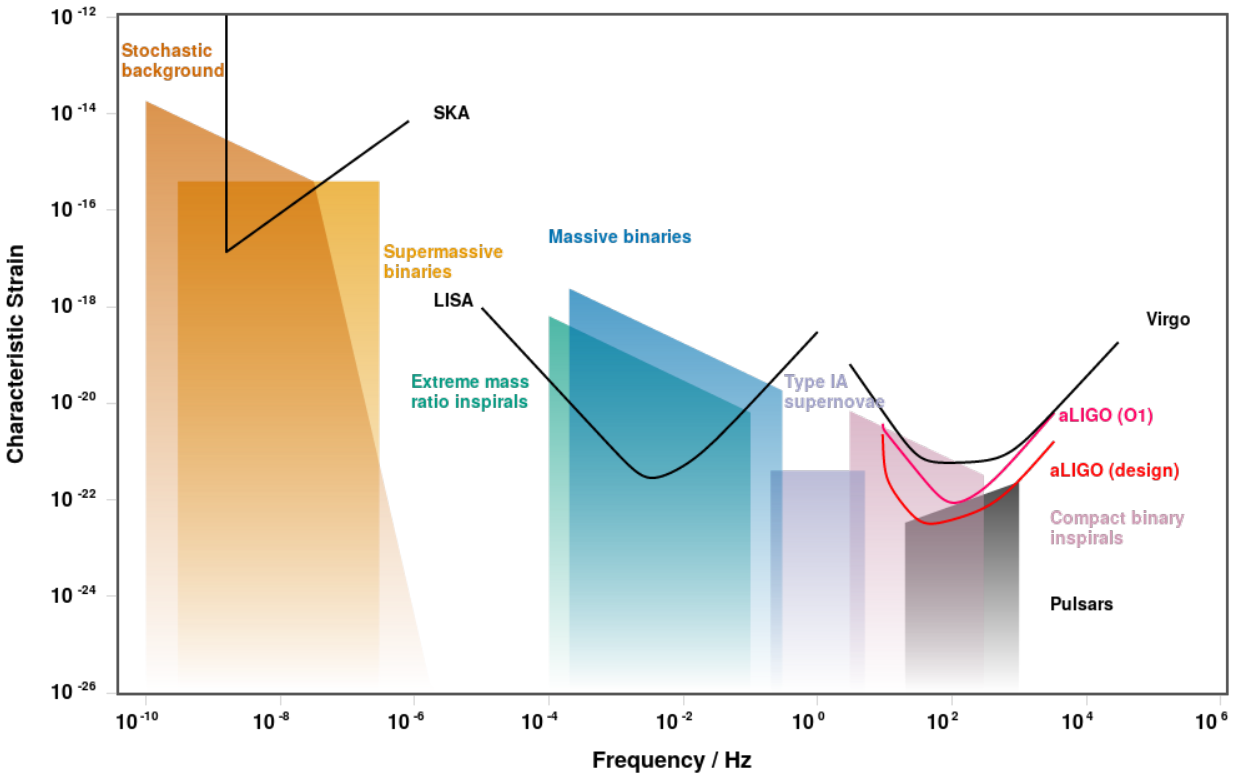


Figure 3.10: A comparison of detector sensitivities for various types of gravitational-wave experiments, compared to the expected characteristics strain ranges for various sources. LISA is the Laser Interferometer Space Antenna, and SKA is the Square Kilometer Array (a radio telescope project which will be used, among other things, for pulsar timing). Both LISA and the SKA are future projects; their sensitivity curves are predicted, as is the design sensitivity curve for aLIGO. In contrast, the O1 aLIGO curve reflects the actual sensitivity of the interferometers during the first observing run (2015). Figure via [47]. *This figure is not subject to the Creative Commons license that has been applied to the rest of this work.*

quadrupolar pattern across the sky. By carefully measuring the timing residuals of a large number of pulsars over the course of years, this pattern could be detected. Pulsar timing is appropriate for observing even lower gravitational wave frequencies than possible with space-based interferometers, potentially allowing the technique to probe supermassive binaries and the stochastic gravitational-wave background. Pulsar timing is already an active area of research, although so far no detections have been reported.

Figure 3.10 shows the expected sensitivities of some future gravitational wave experiments, alongside aLIGO and Virgo. One key insight from this comparison is that increased strain sensitivity is not the only way to reach scientifically interesting results. Many low-frequency sources have higher characteristic strains, enabling less-sensitive instruments to detect them.

Chapter 4

Continuous waves from spinning neutron stars

Neutron stars are the remnants of massive stars that collapsed in violent supernovae, creating such high-pressure conditions that protons and electrons in their cores were forced to combine into neutrons, with emission of neutrinos. These objects are not only massive and compact, with around 1.4 times the mass of the sun contained in a ~ 10 km radius, but also rapidly spinning: up to hundreds of times per second. With even a tiny non-axisymmetry, spinning neutron stars become good candidates for detectable GW generation.

Neutron stars are well studied, but not well understood. Despite an intense scientific focus on these objects since their first observations in the 1960s, and the identification of more than 2,500 neutron stars, there remain many open questions about their physical properties, population, and evolution.

In particular, the equation of state (EOS) of a neutron star is unknown [51]. The EOS governs the exact relationship among pressure, volume, temperature, and other thermodynamic quantities within the star. These properties in turn influence the large-scale and potentially observable properties of the star: the relationship between its mass and radius, and the amount of ellipticity that the neutron star can maintain (see e.g. [34] for a review of this subject). The latter quantity, ellipticity, is key for gravitational wave generation by spinning neutron stars.

Gravitational wave observations have already begun to probe the neutron star EOS. Gravitational waves from BNS mergers carry information about the outer regions of neutron stars, which are in the process of being torn apart by gravitational forces. The first BNS merger observation yielded constraints on the radii of the component stars, and favored "soft" EOSs over "stiff" ones [57].

Observations of gravitational waves from spinning neutron stars are expected to provide independent constraints on the NS EOS. An observation of this type could yield information about an entire neutron star in equilibrium, unlike the BNS merger measurements.

The vast majority of known neutron stars are detected as pulsars. Pulsars are spinning neutron stars, that emit a beam of electromagnetic radiation at some angle from their rotation axis. As the beams sweeps past a distant observer, the observer records a pulse of light (typically radio, although x-ray pulsations and

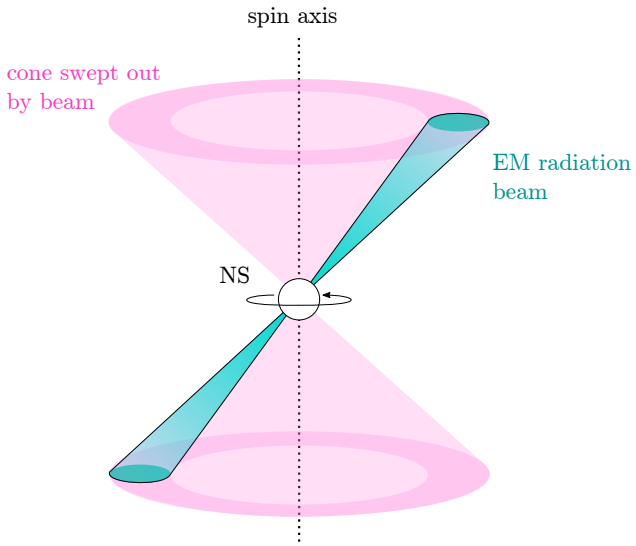


Figure 4.1: Diagram of the "lighthouse model" of a pulsar. An observer located in the path of the beam will observe pulses of light as the pulsar spins.

gamma-ray pulsations have also been observed in recent years). Figure 4.1 shows a diagram of a pulsar.

Unfortunately, the beam subtends a fairly small angle; the beaming fraction of a normal pulsar is about 20% [40], although it may be larger for millisecond pulsars [31]. There must be many neutron stars for which the beam simply does not point toward Earth, rendering them invisible to radio searches.

Gravitational wave searches can be guided by existing electromagnetic observations of neutron stars, and are simultaneously poised to complement those observations with valuable new data. Because gravitational waves are emitted in all directions, they could be used to detect neutron stars that are invisible to electromagnetic observers. And because gravitational waves are generated by non-axisymmetries in the star itself, observing them would yield insight into the kinds of deformations that neutron stars can maintain, and thereby into their underlying nuclear physics.

4.1 Gravitational waves from a rotating triaxial ellipsoid

A neutron star can be modeled as a rotating triaxial ellipsoid (see [18]). While this is a simplistic model, it captures the key physics: amplitude estimates, gravitational wave frequency, and polarization properties.

There is substantial similarity between the gravitational waves from compact binary coalescences (CBCs) and those from a rotating ellipsoid. In both cases, the gravitational wave frequency will be twice that of the system's rotational frequency, and in both cases, circular polarization is observed on-axis and linear polarization is observed perpendicular to the spin axis. The main differences lie in the amplitude and the time evolution of the signal. While CBC signals undergo rapid changes immediately before and during

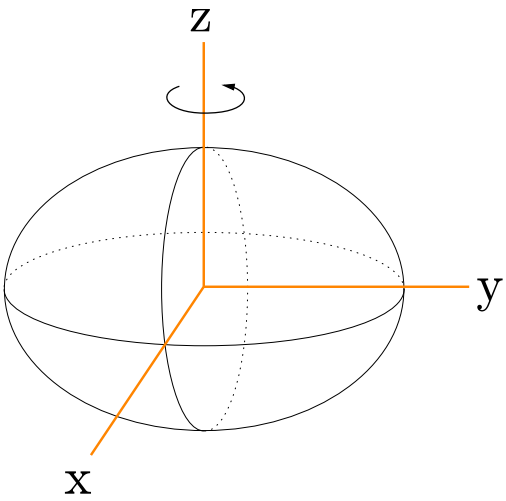


Figure 4.2: Diagram of a rotating triaxial ellipsoid, with its major axes aligned to a rectangular coordinate system, spinning about the z axis.

coalescence, and reach a high peak amplitude, CW signals are generally modeled as unchanging (or very slightly changing) and have a much lower but sustained amplitude.[1]

4.1.1 Gravitational waves on the spin axis

Consider an ellipsoid which starts out with its three axes aligned along x, y, and z axes, (see figure 4.2) so that its quadrupole tensor is:

$$I_0 = \begin{pmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix}. \tag{4.1}$$

We then set the ellipsoid spinning about the z axis with an angular frequency ω. A rotation by ωt can be performed by applying the appropriate rotation matrix:

$$R_z(t) = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) & 0 \\ -\sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

to rotate I:

$$I(t) = R_z(t) I_0 R_z^{-1}(t) \tag{4.3}$$

[1] Note that while work focuses on neutron stars as a source for CW signals, it is not true that CW sources need be neutron stars, nor is it true that binary systems are only detectable as CBCs. For example, binaries which are a long way from coalescence will effectively be continuous wave sources for LISA.

which yields

$$I(t) = \begin{pmatrix} \frac{1}{2}\left(I_{xx}+I_{yy}\right) + \frac{1}{2}\left(I_{xx}-I_{yy}\right)\cos 2\omega t & -\frac{1}{2}\left(I_{xx}-I_{yy}\right)\sin 2\omega t & 0 \\ -\frac{1}{2}\left(I_{xx}-I_{yy}\right)\sin 2\omega t & \frac{1}{2}I\left(I_{xx}+I_{yy}\right) - \frac{1}{2}\left(I_{xx}-I_{yy}\right)\cos 2\omega t & 0 \\ 0 & 0 & I_{zz} \end{pmatrix}. \quad (4.4)$$

Taking two time derivatives yields:

$$\ddot{I}(t) = 2\omega^2\left(I_{xx}-I_{yy}\right)\begin{pmatrix} -\cos(2\omega t) & \sin(2\omega t) & 0 \\ \sin(2\omega t) & \cos(2\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.5)$$

which is often rewritten in terms of the ellipticity $\varepsilon = (I_{xx} - I_{yy})/I_{zz}$:

$$\ddot{I}(t) = 2\omega^2 I_{zz}\varepsilon\begin{pmatrix} -\cos(2\omega t) & \sin(2\omega t) & 0 \\ \sin(2\omega t) & \cos(2\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.6)$$

Noting that $h_{ij}^{TT} = \frac{2G}{c^4 r}\ddot{I}_{ij}$, we can read off the two polarizations in the TT gauge, which are :

$$h_+(t) = -\frac{4G\varepsilon I_{zz}\omega^2}{c^4 r}\cos(2\omega t) \quad (4.7)$$

$$h_\times(t) = \frac{4G\varepsilon I_{zz}\omega^2}{c^4 r}\sin(2\omega t) \quad (4.8)$$

One key message here is that the gravitational waves will radiate at twice the spin frequency of the neutron star. When the direction of observation is aligned with the spin axis, the wave will be circularly polarized: a combination of + and × waves with equal amplitude, offset in phase by a quarter cycle.

4.1.2 Gravitational waves off the spin axis

To see what happens for an observer located off the spin axis, it is necessary to apply another rotation, and then to project the result into the transverse-traceless (TT) gauge.

Rotating the system by an inclination angle ι about the x axis requires another rotation matrix:

$$R_x(\iota) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\iota & \sin\iota \\ 0 & -\sin\iota & \cos\iota \end{pmatrix} \tag{4.9}$$

$$\ddot{I}(\iota, t) = R_x(\iota)\ddot{I}(\iota = 0, t)R_x^{-1}(\iota) \ . \tag{4.10}$$

After computing this, the transverse traceless gauge projection is obtained by setting $I_{iz} = I_{zi} = 0$ ("transverse") and then subtracting off half the trace ("traceless") from the diagonals. (To recall why this is allowed, see section 2.3.) This yields:

$$h_+(t) = -\frac{4G\varepsilon I_{zz}\omega^2}{c^4 r}\left(\frac{1+\cos^2\iota}{2}\right)\cos 2\omega t \tag{4.11}$$

$$h_\times(t) = \frac{4G\varepsilon I_{zz}\omega^2}{c^4 r}(\cos\iota)\sin 2\omega t \ . \tag{4.12}$$

This result retains circular polarization on the z-axis ($\iota = 0$), when the rotation is viewed face-on. When viewed from the side ($\iota = \pi/2$), however, only + polarized waves remain. This makes physical sense. The system is symmetrical along the z axis, so there is no clear way to pick out a preferred alignment for the polarization when viewed along that axis. Viewed from the side, the "spinning" motion is projected onto the plane of the observer's view; the only transverse component is aligned with a + polarized wave.

4.1.3 Strain from ellipticity

Consider a neutron star with $I_{zz} = 10^{38}\,\mathrm{kg}\cdot\mathrm{m}^2$ and ellipticity 10^{-6}, spinning at a frequency of 100 Hz, at a distance of 1 kpc. Based on the calculation above, this would yield a strain amplitude of $\sim 4 \times 10^{-26}$.

Unfortunately, the actual ellipticity of a typical neutron star is unknown. Even the theoretical maximum value for the ellipticity of a neutron star is unknown; this is part of the motivation for GW searches targeting neutron stars in the first place.

The value of 10^{-6} may well be overly optimistic. The maximum deformation supported by crust rigidity is approximately [49]

$$\epsilon_{\max} \approx 5 \times 10^{-7}\left(\frac{\sigma}{10^{-2}}\right) \ , \tag{4.13}$$

where σ is the breaking strain. The breaking strain for the best terrestrial alloys is around 10^{-2}, but it is not at all clear that this is a viable value to use for a neutron star. The breaking strain for a neutron star

could be as high as 0.1, able to support an ellipticity of $\sim 4 \times 10^{-6}$ [28]. Nor is it clear in the first place that this model for crustal deformation in terms of the breaking strain is accurate. Finally, the question of maximum sustainable ellipticity is very different from the question of typical ellipticity.

4.1.4 Mechanisms for non-axisymmetry

A variety of mechanisms [49] have been proposed by which a neutron star could gain some non-axisymmetry large enough to produce detectable gravitational waves. It is possible that stresses could build up in the crust of the neutron star as it spins down, eventually breaking the crust. This model has also been proposed as an explanation for the occasional sudden jumps observed in pulsar frequencies, known as pulsar glitches. Such breaks could leave the neutron star with a non-axisymmetric mass distribution for some time after the event.

Alternatively, the strong magnetic field of a neutron star could induce non-axisymmetric distortions of the star directly. Magnetic field distortion is particularly likely for young neutron stars, which are likely to be born with large internal fields and correspondingly large ellipticities. [33]

Internal oscillations modes could also create significant neutron star ellipticities. The unstable r-mode is one of the likeliest options. This type of oscillation is unusual because gravitational radiation feeds back into it, increasing the amplitude of the oscillation rather than decreasing it. The details of r-mode oscillations, and the window of neutron star spins in which they are likely to be significant, depend sensitively on the neutron star equation of state and the dynamics of the crust.

In cases where a neutron star has an orbital companion, it may accrete matter from the companion. The path of the infalling matter would be guided by the neutron star's magnetic field, if strong enough, landing near the magnetic poles and creating "hills" on the surface. The magnetic field of the star could also help maintain the height of such hills. [49]. This mechanism provides more motivation to search for gravitational waves from accreting neutron stars in binary systems.

4.2 Neutron star and pulsar astronomy

Our ability to estimate the characteristic strain from a spinning neutron star is limited by model uncertainty. The various proposed mechanisms for non-axisymmetry are not well enough understood to make confident predictions about observable gravitational waves. A more productive option is to turn to electromagnetic observations of neutron stars, and to consider the prospects for gravitational wave detection based on simple energy conservation arguments. This is also a useful course of action because electromagnetic observations of particular sources can be used to guide gravitational wave searches. Unlike transient gravitational wave

searches (CBCs and bursts), continuous gravitational wave searches benefit greatly from focusing on a specific point on the sky, and from leveraging existing measurements of the source properties to constrain the search parameter space. Finally, an understanding of the astrophysics of neutron stars, and the many questions that still exist about neutron star populations and formation, motivate the search for gravitational wave signals from such sources.

4.2.1 Pulsar populations

The majority of observations of neutron stars have been from pulsars. There are two main populations of pulsars: "normal" pulsars, and millisecond pulsars (MSPs). Normal pulsars have a relatively long spin period ($\sim$ 1 Hz) and high magnetic fields ($\sim 10^{12}$ G). Over time they slow down, with the spin period $P_{\rm NS}$ decreasing on the order of 10^{-15} s/s. Normal pulsars have a characteristic age of about 10^7 years. By contrast, MSPs spin much faster ($P_{\rm NS}$ measured in milliseconds, hence the name), but have weaker magnetic fields ($\sim 10^8$ G). Their spindown is also slower, on the order of $\dot{P}_{\rm NS} \lesssim 10^{-19}$ s/s, and they are older, on the order of 10^9 years. [40]

The vast majority of normal pulsars are isolated stars, with very few found in binary systems. A much larger fraction of MSPs (about 80% of all those observed) are found in binary systems. This fact is explained by the formation mechanism of MSPs versus normal pulsars, discussed in section 4.2.2.

4.2.2 Pulsar formation

This section overviews the current understanding of pulsar formation in binary systems, which seeks to explain the observed populations of normal and millisecond pulsars. The material is drawn primarily from the review in [40].

The current understanding of pulsar formation begins with a binary system, in which the more massive star goes supernova, forming a neutron star. The binary is very likely to be disrupted in this process, separating the neutron star from its orbital companion. In this case, the neutron star may be observable as a normal radio pulsar, which spins down to a period of several seconds over the next 10^{7-8} years, after which point its radio emission becomes weak enough that it is not observable

On the other hand, if the binary does not disrupt, and the newly-formed neutron star's companion is sufficiently massive and nearby, then eventually the companion will evolve into a giant and overflow its Roche lobe. The neutron star will begin accreting matter and gaining angular momentum (transferred from the orbital angular momentum of the binary system). This "spins up" the neutron star, increasing its rotational frequency and moving it into the population of millisecond pulsars. During accretion, frictional heating of

infalling matter will lead to X-ray emission. During this phase, the system is likely to be an observable X-ray binary.

X-ray binaries are distinguished into two categories, based on the mass of the companion [40]. A high-mass system will eventually see the companion go supernova, possibly producing a second NS, but probably also disrupting the system. On the off chance that the system is not disrupted, the result can be a neutron star binary. However, the newer NS has no way to be spun up by accretion, so it is unlikely to be visible as a radio pulsar for as long as its companion.

Low-mass systems, on the other hand, provide a long time for matter to transfer onto the NS, so the star can be spun up to periods as short as a few milliseconds. Eventually, the companion sheds its outer layers and becomes a white dwarf.

Pulsars that have been given a new lease on their (observable) lifespans by accretion are referred to as "recycled" pulsars. In this model, the population of millisecond pulsars is made up of recycled pulsars, which explains why they are far more likely to be found in binary systems than normal pulsars are. The process of accretion is also thought to attenuate the strong magnetic fields with which pulsars are born, explaining the weaker fields associated with millisecond pulsars.

Gravitational waves could theoretically be observed from a spinning neutron star at any point in its life span, so long as the wave frequency is in one's detection band and the star's ellipticity is high enough. Whether the neutron star emits in radio, x-ray, or neither; whether it is in a binary system or not, any non-axisymmetry should cause it to emit in gravitational waves. However, electromagnetic observations often guide GW data analysis, motivating searches for specific promising sources, and constraining the parameter space over which those searches must take place. They also indicate what kinds of systems ought to be searched, for in general. In particular, the large fraction of MSPs in binary systems is a key motivating factor for the techniques discussed later in this thesis.

4.2.3 Strain estimates from electromagnetic observations and conservation laws

In the absence of solid theoretical predictions for neutron star ellipticity, several different "best-case scenario" estimates are used. The spindown limit is appropriate for isolated neutron stars with a measured spin frequency derivative (e.g. radio pulsars). The torque balance limit is appropriate for accreting systems which may not exhibit pulsations, but which have a measured X-ray luminosity (e.g. LMXBs). Both limits are based solely on energy balance, rather than a detailed understanding of the mechanism of gravitational radiation.

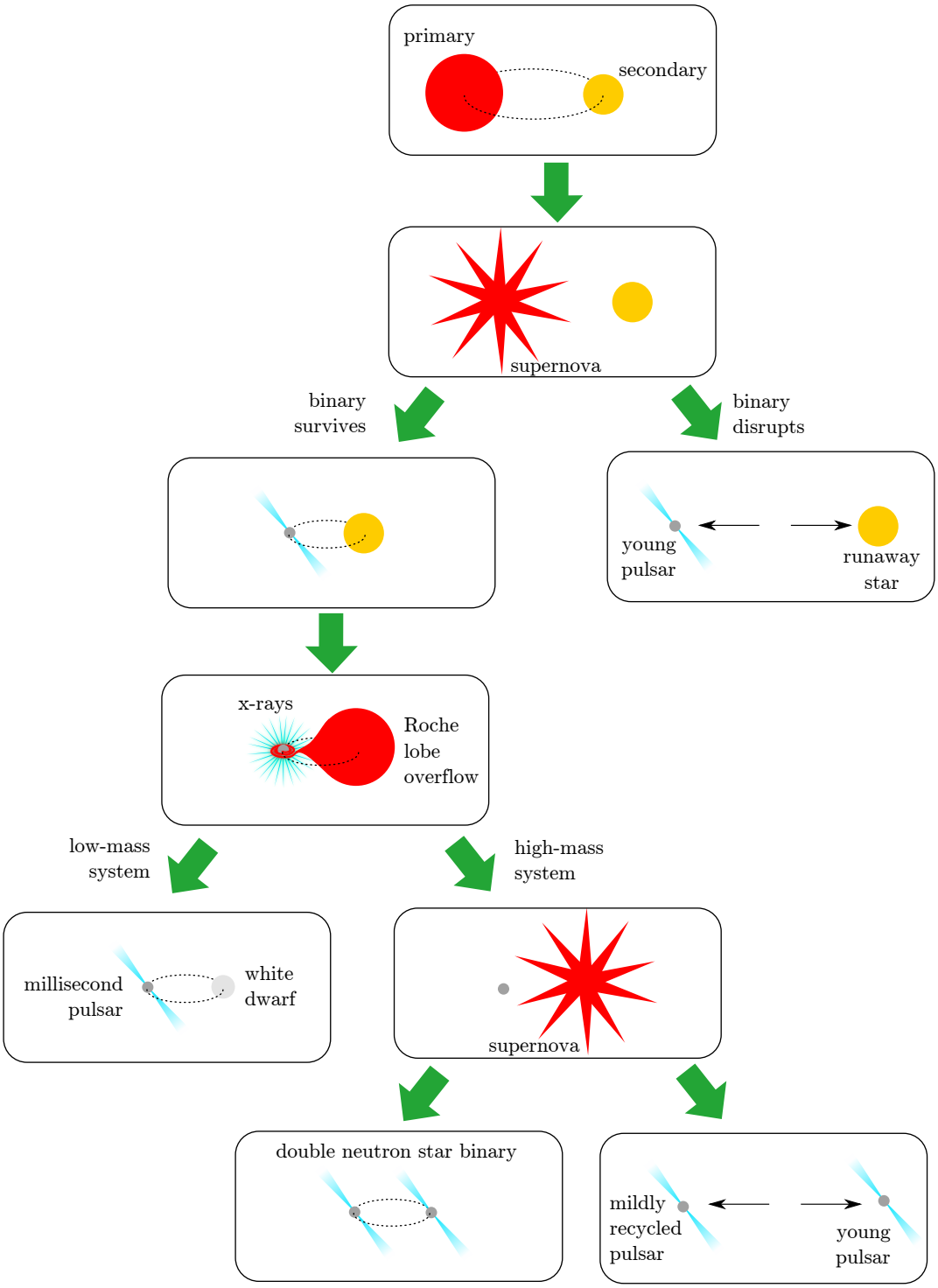


Figure 4.3: A model of pulsar evolution in binary systems. Figure adapted from [40].

4.2.3.1 Spindown limit

The spindown limit assumes that the measured slowing of a neutron star's spin is entirely due to gravitational radiation. This is not realistic; electromagnetic braking is expected to play a large role in slowing neutron stars. The spindown limit sets an upper bound, however, on the possible detectable strain, and also marks a point at which strain upper limits from gravitational wave searches become scientifically interesting. The spindown limit is given by:

$$h_{\mathrm{spindown}} = \frac{1}{r}\sqrt{\frac{5G}{4c^3} I_{zz} \left(\frac{-\dot{f}_{\mathrm{GW}}}{f_{\mathrm{GW}}}\right)}\,, \tag{4.14}$$

where r is the distance to the source and f_{GW} is the gravitational wave frequency (twice the spin frequency).

The best upper limits from continuous-wave searches now beat the spindown limit for a handful of sources.

4.2.3.2 Torque balance limit

The torque-balance assumes that all the angular momentum added to a neutron star by accretion is radiated away in the form of gravitational waves. Once again, this is unlikely to be a realistic estimate, but it sets an upper bound on the detectable strain from such systems. The torque-balance limit also rests on the understanding that the X-ray luminosity of the source is a measure of the mass accumulation on the surface. It is given by:

$$h_{\mathrm{torque}} = \left(5 \times 10^{-27}\right) \sqrt{\left(\frac{600\mathrm{Hz}}{f_{\mathrm{GW}}}\right)\left(\frac{\mathcal{F}_{\mathrm{X}}}{10^{-8}\mathrm{erg/cm^2/s}}\right)}\,, \tag{4.15}$$

where $\mathcal{F}_{\mathrm{X}}$ is the energy flux of X-rays from the source, as observed from Earth. (The X-ray flux contains a dependence on the distance to the source, which is why r does not appear in the torque-balance equation.)

While it is not actually reasonable to assume that *all* the angular momentum from accretion is lost to gravitational waves, there are observational hints that that gravitational radiation does play an important role in the energy balance of accreting neutron stars. The fastest known pulsar, PSR J1748-2446ad, spins at 716 Hz. It is possible that faster-spinning pulsars exist, but none has so far been observed (see figure 4.4). This presents something of a puzzle, since the predicted breakup speed for a neutron star is considerably higher: around 1 kHz for most theorized equations of state. Something else must be limiting the spin frequency of MSPs, but it is not apparent what that mechanism could be. One possibility is that gravitational radiation provides a braking mechanism for the fastest-spinning neutron stars.

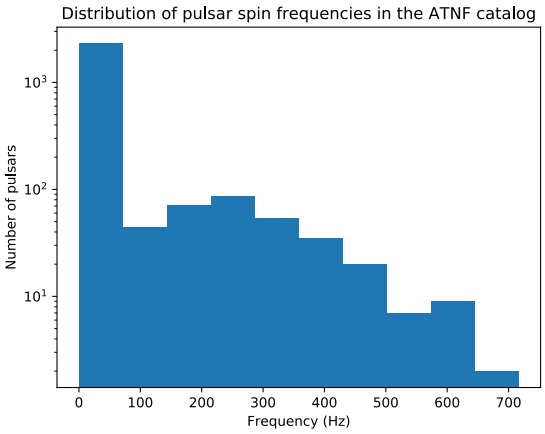


Figure 4.4: Distribution of known pulsar spin frequencies. Data is taken from the ATNF pulsar catalog ([11], originally published in [42])

4.3 Searches for continuous gravitational waves

Continuous wave searches are fundamentally looking for a very simple signal: a near-constant single-frequency wave emanating from some point on the sky. These signals may be very weak, but they persist over a long period of time, making it possible to sum up data over the course of months or years. Unfortunately, there are a variety of effects that complicate this simple picture.

This section draws extensively from [51], which provides an overview of CW searches and the associated computational challenges.

4.3.1 Changes in source frequency

The first computational challenge for CW searches is that neutron star spin frequencies are not strictly constant; as noted from pulsar observations, neutron stars tend to slow down over time (hopefully, driven in part by gravitational radiation).[2] Although the spindown is very small, it accumulates over time, so that a source with a given frequency at the beginning of an observing run may have an appreciably different frequency by the end.

This change is often expressed in terms of phase evolution rather than frequency evolution (phase being the more fundamental quantity). The gravitational wave phase $\phi(t)$ can be expanded out in a Taylor series:

$$\phi(t) \approx \phi_o + 2\pi \left[f_{gw}(T - T_0) + \frac{1}{2}\dot{f}(T - T_0)^2 + \frac{1}{6}\ddot{f}(T - T_0)^3 \right] , \tag{4.16}$$

[2]Other factors can also affect the source frequency, such as variable rates of accretion (e.g. in LMXBs) and neutron star glitches. Most searches necessarily operate on the assumption of "typical" behavior, and do not consider glitches or spin wandering. Characterizing search robustness against spin wandering and glitching is an area of active inquiry, as is developing new analysis pipelines that can better handle these effects.

where $\phi(T_0) = \phi_0$ is the phase at some reference time T_0, f is the source frequency, and T is the arrival time of the signal at the solar system barycenter. Consider a search that matches the signal to a predicted model, keeping the phase agreement within some fixed acceptable error margin (known as a *fully coherent* search). It is necessary to search over a number of steps in f_{gw} proportional to the total observation time T_{obs}, and to search over a number of $\dot{f}_{gw}$ steps proportional to T_{obs}^2, and a number of $\ddot{f}_{gw}$ steps proportional to T_{obs}^3. In total, then, the parameter space volume will be proportional to T_{obs}^6. This type of search is sensitive, but rapidly becomes computationally intractable for large T_{obs}.

An alternative approach is to use a semi-coherent method, which splits the data into many short segments of coherence time $T_{\rm coh}$, and tracks the phase evolution only within segments (not between them). This sacrifices some information and hence some sensitivity, but allows for a longer total observation time to be used. The improved statistics gained through longer observation can compensate to an extent for the loss of phase information.

4.3.2 Signal modulation due to the motion of Earth

A second issue is Doppler modulation due to the motion of Earth around the Sun. The observed frequency of the source will shift depending on the relative motion between it and the Earth-based LIGO observatories. Earth's motion imposes a daily relative frequency modulation on the order $v_{\rm rot}/c \approx 10^{-6}$, and an annual frequency modulation on the order of $v_{\rm orb}/c \approx 10^{-4}$. To search for a signal from a particular sky location, it is necessary to demodulate the signal, correcting for the Doppler effect over the course of the observing time. In practice, this means that a separate calculation must be performed for each of thousands to millions of points on the sky, multiplying the computational cost accordingly. The angular resolution required is approximately:

$$\delta\theta \approx \frac{0.5c\delta f_{gw}}{f_{gw}[v\sin\theta]_{\rm max}} \,, \tag{4.17}$$

where θ is the angle between the detector's velocity and the sky location. Approximating that the relative frequency shift $[v\sin\theta]_{\rm max}/c$ is about 10^{-4}, and noting that the frequency bin width δf is the inverse of the coherence time $T_{\rm coh}$,

$$\delta\theta \approx 9\times10^{-3}{\rm rad}\left(\frac{30\,{\rm minutes}}{T_{\rm coh}}\right)\left(\frac{300\,{\rm Hz}}{f_{gw}}\right) \,. \tag{4.18}$$

4.3.3 Signal modulation due to orbital motion of the source

A third issue arises when the neutron star is located in a binary system. Another Doppler modulation is introduced by the motion of the source as it orbits a companion. The orbital period of the system, and the radius of the orbit, are generally unknown. This adds at least two more dimensions to the parameter space,

even for a circular orbit; additionally, orbital phase and eccentricity may be searched over. A method for handling signals from neutron stars in binary systems will be discussed in further detail in chapter 6.

4.3.4 All-sky, directed, and targeted searches

Continuous wave searches are roughly divided into three categories, based on the amount of information available about the source from electromagnetic observations. The more information is available, the more the search can be narrowed, often allowing for better sensitivity at fixed computational cost.

All-sky searches proceed from no specific information; as implied by the name, they search the entire sky for unknown sources. All-sky searches in O1 data [37] reported strain amplitude upper limits as low as $\sim 1.5 \times 10^{-25}$ for circularly polarized waves (best case) and $\sim 4 \times 10^{-25}$ for linearly polarized waves (worst case).[3]

Directed searches cover only a single point on the sky, where a promising source is expected to be located based on electromagnetic observations. A directed search must still search, however, over the other unknown parameters of the source. This notably includes the spin frequency and its time derivative, and in the case of a binary system, the orbital parameters. Scorpius X-1 is an example of a promising source for directed searches. Its high X-ray flux yields a high torque-balance limit, [51]

$$h \sim \left(3.5 \times 10^{-26}\right) \sqrt{\frac{600\,\mathrm{Hz}}{f_{gw}}}\,, \tag{4.19}$$

and the orbital period of the binary is well measured, greatly decreasing the parameter space of the necessary search. Its actual rotation frequency remains unknown, though. Directed searches in O1 data [39] reported strain amplitude upper limits as low as $\sim 8 \times 10^{-26}$ for circular polarized waves (best case), and $\sim 2.3 \times 10^{-25}$ marginalized over source inclination/polarization. These results do not yet reach the torque-balance limit, but are closing in on it, to within a factor of a few, at low frequencies.

Targeted searches for known sources are generally the most sensitive type of CW search, because they can leverage existing measurements of the source parameters. Known, relatively nearby pulsars with measured spins and spindowns are regularly targeted by CW searches, although so far these searches (like all CW searches) have yielded only upper limits on the gravitational wave strain amplitudes. Targeted searches in O1 data [38] surpassed the spindown limit for eight of the 200 pulsars searched; O2 searches improved on this record, surpassing the spindown limit for 20 of 222 pulsars searched. [16]

[3] Note that these upper limits correspond to searches for *isolated* neutron stars, not those in binary systems. There are few existing data analysis pipelines for all-sky binary searches. The TwoSpect algorithm is a notable exception, which will be discussed in detail in section 6.2.

Chapter 5

Continuous wave data quality and FineTooth

5.1 Effect of narrow spectral artifacts on CW searches

Continuous wave searches must contend with a variety of terrestrial noise sources and detector artifacts. The most troublesome forms of noise are those which appear as narrow "lines" in the spectrum of DARM. These are long-lived oscillations near a single frequency, much like an expected signal from a spinning neutron star. Unlike an astrophysical signal, there is no Doppler modulation due to Earth's motion. Nonetheless, search algorithms can yield a false positive for noise artifacts of this type, reading the artifact as a low-modulation signal. Each false positive represents a significant time investment for the person running the search, and in some cases significant computational cost as well, because the outlier must be investigated further.[1] It is important to understand the DARM spectrum in detail, and to make this information broadly available to those running searches, so that outliers associated with terrestrials sources can be quickly identified and rejected.

Instrumental noise couples to DARM in a variety of ways [17]. Periodic power draws from one electronic component can couple to another through a shared power supply; direct coupling through magnetic and electrostatic fields are another common culprit. Mechanical coupling, such as resonances in the suspension wires for the detector optics, and data acquisition artifacts, such as aliasing, can produce lines and combs in DARM.

5.2 Key detector characterization goals

5.2.1 Vetted line lists

Continuous wave analyses are typically performed after an observing run ends, so that the analysis can take advantage of all the data from the run to increase its sensitivity. Line lists are therefore also generated

[1] Chapters 6 and 7 will provide more detailed information on handling outliers, with some specific tests discussed in 7.5.2.4.

after each observing run. Line lists attempt to catalog all known narrow lines in the run-averaged DARM spectrum, including frequency, spectral width, and physical origin (when known). This is no small task, given the number of lines that appear. Vetted line lists for the first and second observing runs can be found on the LIGO-Virgo Open Science Center. [8, 9]

Line lists may be divided into "vetted" and "unvetted" categories. Vetted lines are those which are confirmed beyond reasonable doubt to be terrestrial in origin. Unvetted lines are catalogued for reference, and are very likely to be terrestrial (as they are typically much stronger than an expected pulsar signal, and lack any apparent modulation from Earth's motion), but their causes remain unknown.

5.2.2 Line mitigation

Line mitigation happens before or during an observing run. This often requires time and effort on the part of off-site data analysts, as well as on-site scientists and engineers. The goal is straightforward: rapidly identify sources of noise as they arise, prioritize those that will have the largest effect on search sensitivity, determine their causes, and take action to remove them.

5.3 Combs

A single-low-frequency noise source can contaminate a large number of spectral bins, as its higher harmonics create an evenly-spaced pattern of lines in the spectrum at integer multiples of the fundamental frequency. This structure is known as a "comb", and the individual lines are known as "teeth." Combs are a significant problem for CW searches, since narrowly-spaced combs can contaminate tens or hundreds of bins, often in the most sensitive part of the LIGO detection band. Identifying and mitigating a comb is therefore a detector characterization achievement with a large impact on searches.

Determining the origin of a comb and working out how to mitigate it before an observing run begins (or shortly after it appears during the run) is, of course, the ideal outcome. This minimizes the amount of data contaminated by the artifact. Identifying combs after a run is completed, once the opportunity to fix the problem has passed, is not ideal but can still be very useful, because membership in a comb is evidence that a line is terrestrial in origin. Many unknown lines can be grouped and moved immediately from "unvetted" to "vetted" status, once it is established that they are just higher harmonics of a low-frequency noise source. This is true even if the physical cause of the noise is unknown; the point is that (for example) a line appearing at 300 Hz is not reliable evidence of a 300 Hz continuous gravitational wave if it lies on the 300th tooth of a 1-Hz comb.

A comb can be described in the following manner, which is convenient for computational purposes. The tooth frequencies are given by:

$$f_n = nf_s + f_o \tag{5.1}$$

where n is an integer, f_s is the comb spacing, and f_o is an offset ($f_o < f_s$). This is an empirical description which does not always provide the best physical insight; for example, a comb with $f_s = 1.0$ Hz and $f_o = 0.5$ Hz is most likely comprised of the odd harmonics of an oscillation with fundamental frequency 0.5 Hz. However, tempting as it is to describe combs in terms of their fundamental frequencies and the indices of visible teeth, this quickly becomes difficult as a standard practice. There are some combs for which f_s is not a multiple of f_o; it is often unclear how these structures arise, or what the fundamental frequency of the oscillation actually is. In order to specify combs in a universal format, the f_s and f_o comb parameterization will be used exclusively in this work.

5.4 Overview of FineTooth tools

5.4.1 Visualization

The majority of line and comb investigations are not automated at this time. Spectra must often be inspected by eye, a process which is difficult for high-resolution, wide-bandwidth spectra with important features at many scales. Furthermore, the LIGO Data Grid, where DARM data is stored and most LIGO data analysis takes place, does not have a graphical user interface with acceptable response time. The most common way to access plots generated on the LDG is to place them in a user's public html folder, which is mirrored in a web-accessible directory (requires LIGO log-in credentials). This tends to limit spectral plots to static image files at a fixed scale. Identifying combs by eye can be particularly difficult, because the teeth are often hidden amidst stronger noise artifacts. In the case of low-strength, high-spacing combs, individual teeth may not be resolvable at the scale of a plot necessary to show multiple teeth.

To address the practical problems described above, FineTooth makes use of a Python data visualization library called Bokeh. Bokeh generates HTML/javascript webpages, which provide interactivity without any server-side computation. This is well-suited for publishing plots with fixed data, which nonetheless provide features like zooming, panning, and hover text for the user's convenience. A screenshot of one of these plots is shown in figure 5.1.

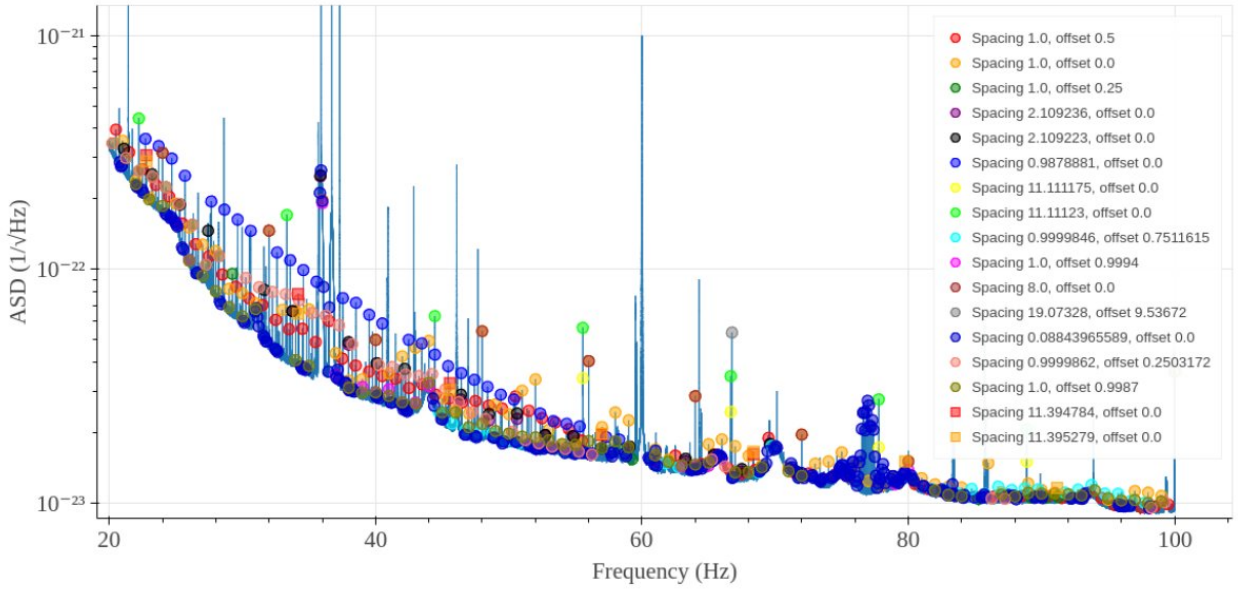


Figure 5.1: Screenshot of a FineTooth plot showing low-frequency combs in the O2 run-averaged spectrum. This plot was previously published in [17].

5.4.2 Sources and storage of data

FineTooth stores spectra in HDF5 files, using a simple custom format (two arrays, one of which holds the frequency values, and the other the spectral data). It is straightforward to convert data from almost any source to this format through Python, using h5py. FineTooth also provides the ability to automatically load data from two standard sources, and save it in the proper format.

First, FineTooth can load data from a monitoring tool called Fscan. During engineering and observing runs, Fscan generates daily spectra and plots for DARM and a variety of auxiliary channels. FineTooth accepts a list of channels to investigate and a date range, and navigates the Fscan-generated file tree to select the appropriate spectra. Fscan does not run continuously; there are stretches of time, particularly between observing runs, when spectra are not generated. This mode of operation saves storage space and computational time, and time for the maintainer of the Fscan code, but means that FineTooth must often contend with data gaps.

Second, FineTooth can use GWpy, a fairly comprehensive Python package for gravitational-wave astrophysics, to load data from any detector channel and generate a spectrum. This is often useful when investigating a specific noise artifact, which may require a higher time resolution (hours or minutes rather than days, to pinpoint the appearance of a noise artifact), or data from a channel or time span not included in the Fscan spectra.

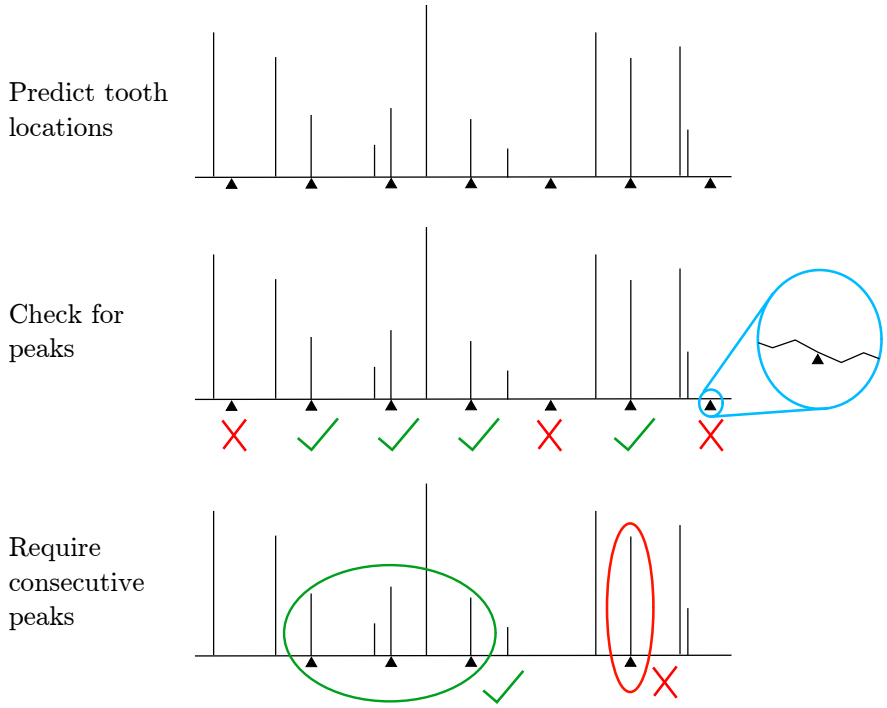


Figure 5.2: Diagram of three stages of comb tooth marking.

5.4.3 Comb tooth marking

5.4.3.1 Motivation

One of the largest practical problems for comb studies is the simple fact that it is difficult to pick out which lines, in a spectrum filled with lines, are actually the teeth of a specific comb. This is true even when the spacing of a comb is already known. Not only can it be visually difficult to discern a comb, but comb teeth may overlap or be confused with other lines. Suppose, for example, that a comb is visible in the spectrum at 1 Hz intervals between 80 and 100 Hz. Is a line at 105 Hz part of the comb, or is it an unassociated noise source? Is a series of lines spanning 65 to 70 Hz part of the comb, despite the gap? There are no perfectly reliable answers to these questions, but heuristics can be developed to address them, based on practical experience with line investigations.

5.4.3.2 Criteria

FineTooth uses two criteria for "marking" comb teeth, i.e. for identifying the bins in a spectrum that are likely to contain real teeth of a given comb, as opposed to noise or unassociated line artifacts. The two criteria are: (a) the bin must contain a peak, and (b) the peak must be part of a consecutive set of peaks. A diagram of the algorithm is shown in figure 5.2.

For a given spacing f_s and offset f_o, over a frequency range $f_{\min}$ to $f_{\max}$, the predicted position of each

tooth $f_{\min} \leq f_n \leq f_{\max}$ is generated using $f_n = f_o + nf_s$. In a real spectrum, of course, frequency bins are discrete. Each f_n of the comb must be associated with the nearest spectral bin, $f_{n,\text{near}}$. This process generates a list of "predicted" frequency bins where teeth of the comb may exist.

Next, FineTooth determines whether a peak exists in each of the predicted bins. The criterion for a peak is that the power in the predicted bin must be larger than the power in the adjacent bin(s) on either side. This criterion is appropriate for high-resolution spectra, where narrow peaks are less likely to be immediate neighbors. For some user-specified integer m, FineTooth checks each bin $f_{(n,\text{near})\pm i}$ where $1 \leq i \leq m$ and compares it to the bin $f_{n,\text{near}}$. If any neigboring bin contains more power than the predicted bin, the nth tooth is excluded from membership in the comb, on the grounds that no peak is apparent.

In practice, m is usually set to 1, although in cases where it is hard to distinguish the comb from noise, a more strict requirement of $m = 2$ or 3 can be imposed. In cases where the user wishes to see all predicted teeth of the comb, m can be set to 0.

Finally, FineTooth requires each identified peak to be part of a group of p consecutive teeth. Isolated peaks are unlikely to be real members of the comb. The indices n that survive the peak test are consecutively ordered, and anywhere that gaps exist, the list is split into sub-lists. If the nth predicted peak is found in a sub-list less than p long, it excluded from membership in the comb. In practice, p is usually set to ~ 3, although exploratory studies often loosen this requirement, sometimes to the point of removing it altogether ($p = 0$).

Comb teeth which pass both checks can be "marked" as real members of the comb, according to the user-specified criteria.

5.4.3.3 Optional threshold criterion

To avoid marking very low-strength peaks, a threshold criterion may also be used for peak identification. This requires calculating a background value at each point (as was done to normalize the spectrum), and a standard deviation for the background. The user can then supply a number of standard deviations by which a point must exceed the background, in order to be considered a peak.

5.4.3.4 Use of comb marking

Comb marking has several uses. First, it can be used as a visual aid, where a user supplies FineTooth with f_s and f_o and receives a plot highlighting the likely extent of the comb (even if the comb covers a small section of the spectrum, or is interrupted in the middle by other noise sources). This is helpful organizationally, particularly when dealing with a large number of combs at the same time.

It is also helpful for fitting more precise values for f_s. For example, a quick check that the optimal f_s has been identified is to plot three combs at once: with f_s, and with some slightly changed spacings $f_s \pm \delta f$. In some cases, previously unnoticed higher teeth of the comb will be marked with $f_s \pm \delta f$, refining the spacing estimate. This is especially likely to occur when f_s is estimated using a relatively low-resolution spectrum, and subsequent investigations turn to higher resolutions.

Second, comb marking can be used as a step in an automated comb finding algorithm. This will be discussed further in the section on comb finding.

5.4.4 Tracking combs in daily spectra

Examining the history of a comb over weeks, months, or even years can yield important insights into its origin. The strength and pervasiveness of combs shift over time, and obvious change points can sometimes be linked to changes in the detector configuration. The point in time at which a comb appears is often a key clue as to its cause.

As noted earlier, daily spectra are available for a variety of channels through the Fscan monitoring tool, and it is straightforward to convert these spectra to Bokeh plots using FineTooth, with any combs of interest highlighted. Howevever, sorting through hundreds of daily plots is not an efficient approach. Instead, it is useful to construct measures of comb strength that can be plotted over time.

Comb strength measurements are complicated by several factors. First, the teeth of the comb must be identified, as discussed in 5.4.3. If comb strength is measured by (for example) taking the mean or median of all predicted bins, it is likely that the result will be skewed by large parts of the spectrum where the comb is not visible. Second, the over all level of noise and the amount of available data can vary from day to day. As a result, a comb of constant strength may appear to fluctuate, as it is buried and revealed by different levels of background noise.

FineTooth generates comb tracking plots with appropriate data to help the user understand changes in comb strength. Figure 5.3 shows a screenshot of one such plot. The yellow dots are the main figure of interest. Their height indicates the strength of the comb, defined as the average power in all marked bins, and their size indicates the pervasiveness of the comb, defined as the number of marked bins. The blue dots are a "sanity check," showing the average power in all predicted bins. For combs which do not span the entire bandwidth of interest, this value is expected to be considerably lower than the average power in all marked bins. The red and black traces are a simple check of the overall noise; they show the 1st and 99th percentile of the power in the entire bandwidth of interest. Sudden changes may mean that there is very little data available for a particular day, or that there is some other error with the data. In short, they warn the user

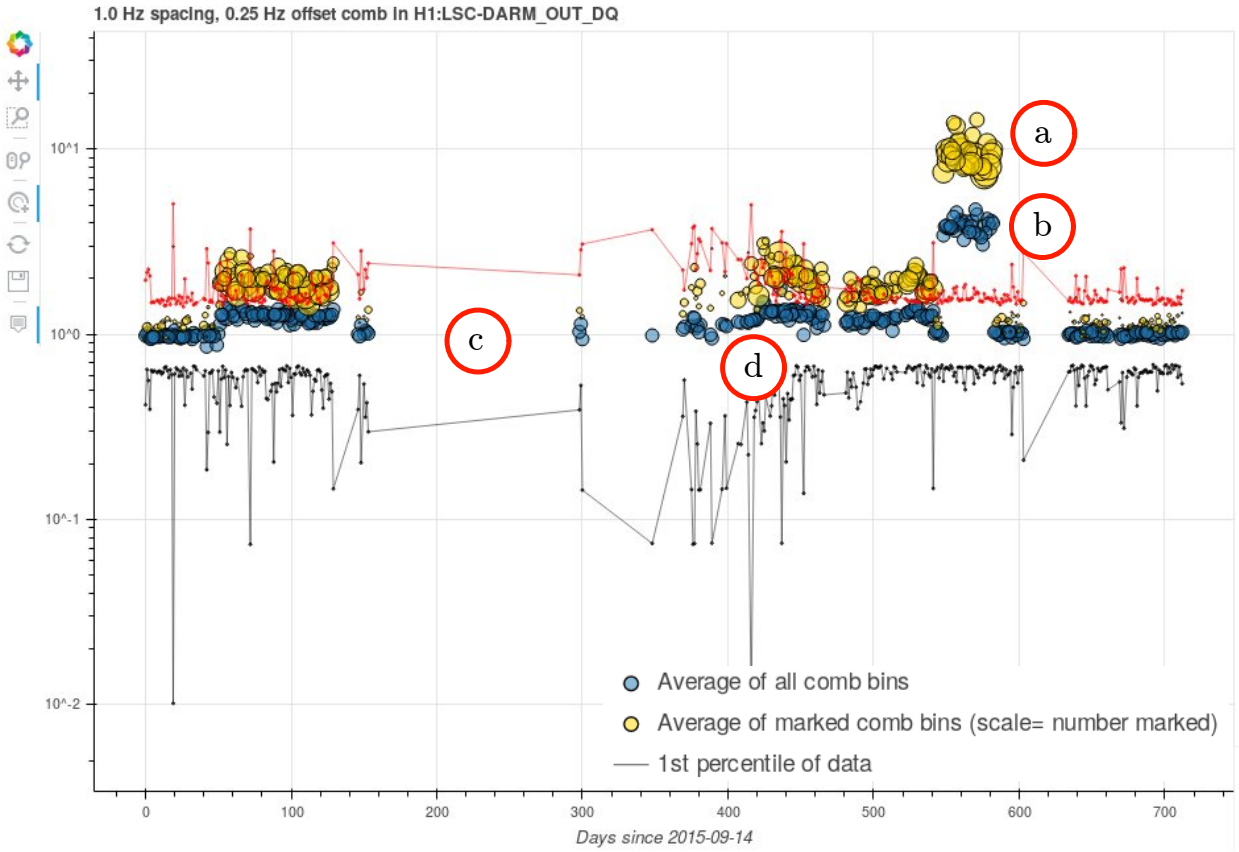


Figure 5.3: Annotated screenshot of a FineTooth comb tracking page for a comb with $f_s = 1.0$ Hz and $f_o = 0.25$ Hz, in DARM daily spectra from the beginning of O1 to the end of O2. This plot is a representative example of DARM comb tracking. Interactive controls are shown on the left side of the screen. Annotations: (a) yellow points show the average strength of the marked comb teeth, and are scaled according to the number of marked comb teeth; (b) blue points show the average strength of all of the comb based on spacing and offset; (c) a data gap corresponding to the end of O1 and the subsequent period of commissioning; (d) a change point in the behavior of the comb which would be difficult to identify without specifically considering the *marked* comb teeth.

not to take the comb measurement for that day at face value.

Like the individual spectrum plots, the tracking plots are interactive, and can be zoomed and panned as necessary. Furthermore, they can be linked to the individual spectrum plots. This allows someone to click on a data point and immediately see the full spectrum for the day in question.

5.4.5 Tracking combs in cumulative spectra

In some cases, combs are so weak that they do not appear in daily spectra at all. It may require days or even weeks of averaging the daily spectra before they begin to rise slightly above the rest of the noise. This is cause for concern; if only daily spectra are used for tracking, then it is possible to miss an important comb for a long period of time.

Therefore, FineTooth also incorporates the ability to average multiple days of spectral data together, generating a cumulative plot of all data since some reference start time. This functionality is dependent on a program called spec_avg_long [61] which can handle the noise-weighted averaging of large spectra in an efficient manner.

Cumulative spectra can be just as easily used for tracking as daily spectra. The difference, of course, is that a comb that is actually quite constant will appear to increase in strength in the cumulative spectra, as it rises out of the other noise.

5.5 Comb tracking pages, post-O1 to pre-O3

Starting between O1 and O2, through the current pre-O3 and O3 era, FineTooth has been used to generate summary pages for low-frequency comb monitoring. These summary pages are updated each day, matching the cadence of Fscan updates. The features of the comb tracking pages are described in this section. Section 5.7 will discuss case studies in comb tracking, which in a number of cases have used the summary pages.

The main summary page contains a table, with a subset of magnetometer channels[2] on one axis and a list of combs to monitor on the other.

The comb list is not automatically generated. Instead, the comb list must be manually updated when a new comb is identified. This is done for several reasons. First, even the best existing algorithms to identify combs are not reliable enough to be used as a basis for the tracking list without human vetting, in part due to the large variations in observed comb morphologies (heights, widths) and the presence of other instrumental lines. Second, it is often prudent to add combs that were previously (but are not currently) observed to the list, so that they are immediately highlighted in the event that they reappear. Third, when adding combs to the list, it is possible to add notes (in HTML format) with supplementary information about each comb, which will appear when the user mouses over the column header for that comb. For example, this space can be used to link to alog entries relevant to the comb, to describe its known or suspected origin, or any other relevant information.

The magnetometer channels include entries from the corner station and from the end station. Each of the three measurement axes of each magnetometer has its own channel of data; channels from the same instrument are grouped together. At a higher level, magnetometers in the same region of the detector are grouped. Each channel name is hyperlinked to a plot of the most recent daily spectrum for that channel, with all monitored combs highlighted in that data (likely comb teeth, selected by the method described in 5.4.3, are marked with colored dots).

[2] Any channel from the Fscan monitoring list could be added to the summary page with minimal effort. The magnetometer channels, however, have historically been the most helpful for comb investigations, so they are prioritized for the tracking pages.

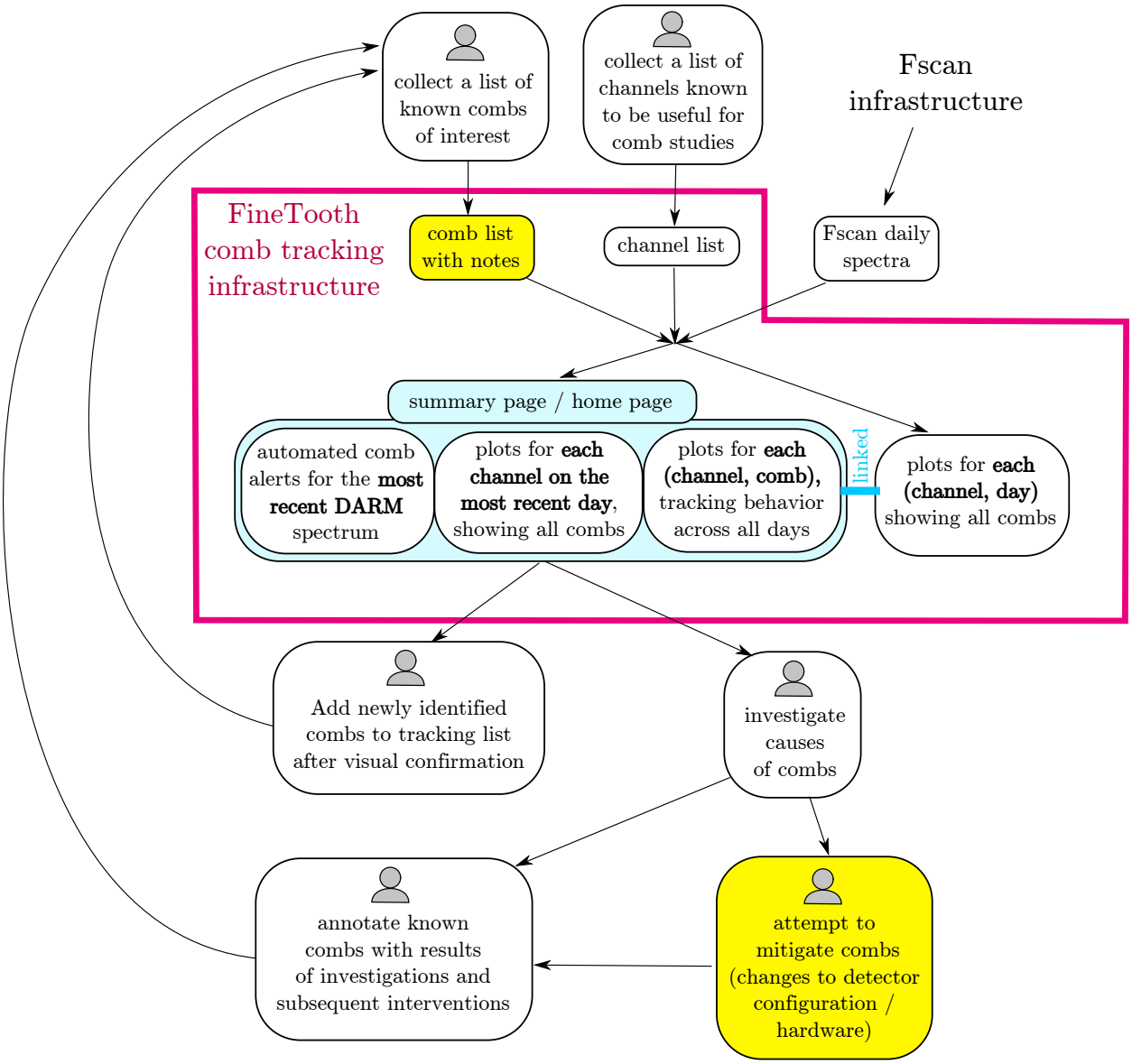


Figure 5.4: A flow chart showing how the FineTooth comb tracking pages fit within the cyclical workflow of comb analysis. Boxes marked with a gray "person" icon involve judgment calls that must be made by a human. The two highlighted boxes represent key goals of comb analysis: first, to improve the detector data quality by fixing the underlying causes of combs; and second, to construct and maintain a list of combs over the course of an observing run, so that these combs may be added to a vetted lines list. Note that some of the flow chart nodes contain a large amount of non-trivial analysis work, which typically draws on a variety of different tools; FineTooth is just one of them. The most important example is the node labeled "investigate causes of combs." Almost all the components of the FineTooth comb tracking infrastructure (e.g. plotting spectra, or creating individual tracking pages for channels and combs) can be used as standalone tools to aid in such investigations.

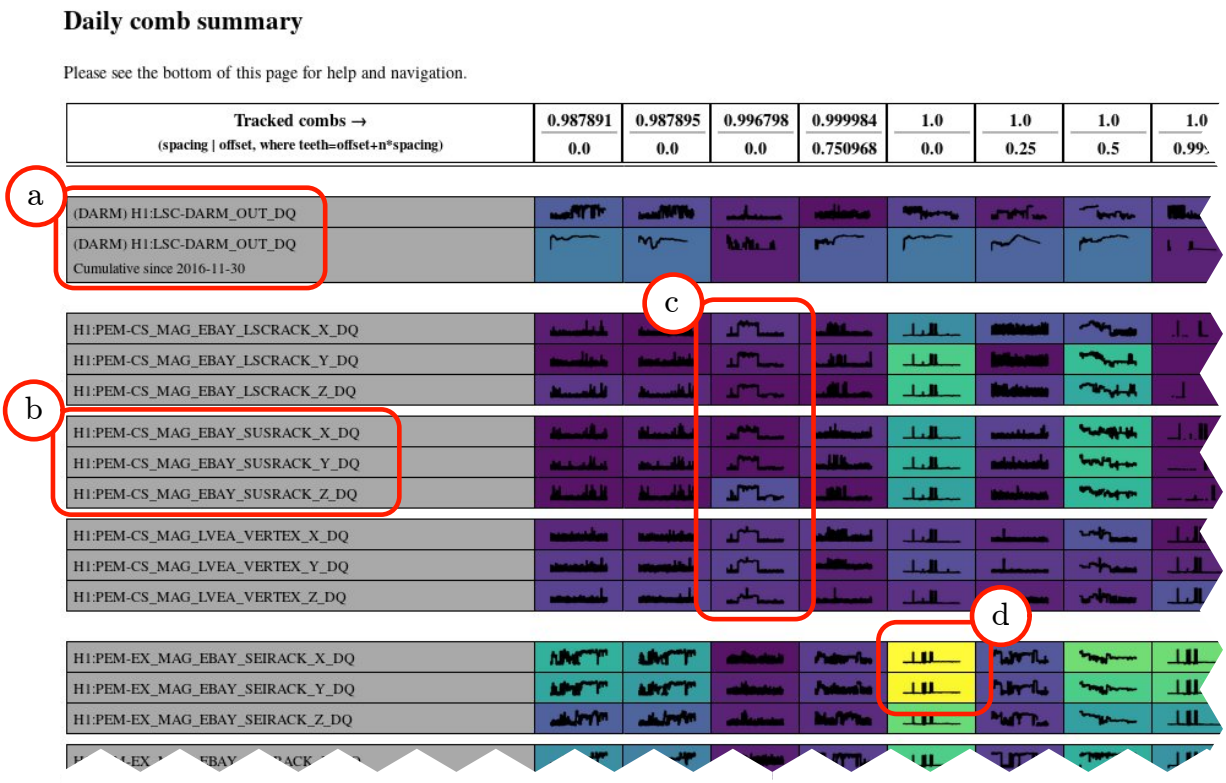


Figure 5.5: Annotated screenshot of the FineTooth comb tracking page for the Hanford detector at the end of O2. Annotations: (a) daily and cumulative DARM data; (b) daily data from three axes of the same magnetometer, in this case located in the corner station ("CS") electronics bay ("EBAY") near an electronics rack housing equipment associated with the test mass suspensions ("SUSRACK"); (c) preview traces for a comb that was once strong in the corner station magnetometers, before being successfully mitigated in October 2016; (d) bright yellow background indicating that the $f_s = 1.0$ Hz, $f_o = 0.0$ Hz comb is currently strong in a particular magnetometer, in this case located at the X end station ("EX") electronics bay ("EBAY") near a rack housing equipment associated with the seismic system ("SEIRACK"). The list of channels and the list of combs are both truncated.

By default, all tracking uses daily spectra, but for DARM data specifically there are also one or more "cumulative channels." These are not true LIGO channels, but are generated from the standard DARM channel by averaging over multiple days (section 5.4.5). During an observing run, the start date for cumulative averaging is typically set to the beginning of the run. There is also a "cumulative channel" using data from the previous ten days, for which the start date updates daily.

Each cell in the table is linked to a tracking plot, (section 5.4.4) which shows the evolution of the corresponding comb in the corresponding channel. A small preview image in each cell shows the general trend of the comb strength over time, while the background is colored according to the most recent measurement of the comb's strength. This allows a user to quickly scan all the channels for a given comb, identifying at a glance where the comb is currently strong, and the general historical trends for its behavior.

Finally, the comb tracking pages contain a somewhat experimental feature which will be described in detail in the next section: automated comb alerts.

5.6 Automated comb identification method

Automatically finding combs in the DARM spectrum is deceptively difficult. Algorithms that work well on simple simulated data, e.g. a comb on a background of Gaussian noise, often fail when confronted with real data, or become computationally unfeasible for all but the lowest-resolution spectra.

The requirements for a useful comb-finding algorithm are as follows:

- Identifies combs that can be found by eye (no worse than visual investigation).
- Identifies combs that were *not* previously found by eye, but which can be confirmed by eye.
- Identifies very weak but pervasive combs, which may be dwarfed by more powerful spectral artifacts.
- Guards against the common issue of contamination by strong peaks, while also allowing for some tooth height variation. (In short: if there exists a very strong peak in the spectrum, it will vastly inflate the apparent amount of power in any comb which overlaps with it. On the other hand, combs are not constant throughout the spectrum, and the heights of real comb teeth can vary considerably).
- Runs quickly, so that it can be used for daily testing, or interactively to explore any spectrum of interest. This is a difficult requirement, given the high resolution of the spectra typically required for comb analysis. Combs often require five or six significant figures in the spacing f_s and offset f_o in order to fit all visible teeth. For example, suppose that we wish to test each possible comb from some small minimum $f_s \lesssim 0.01$ up to 10 Hz, spacing our trial f_s values 10^{-5} Hz apart. Then we have

approximately $(10^6)^2/2 = 5 \times 10^{11}$ combinations to consider, noting that $f_o < f_s$. Depending on the complexity of each test, this may easily exceed the reasonable limitations of real time or computational time, or both.

The FineTooth method for comb identification was constructed with these goals in mind, and applied to real data (the O1 run-averaged spectrum) from its earliest stages of development. It is therefore a very empirically-motivated algorithm, which prioritizes rapidly producing useful, if imperfect, results. It relies on expert human judgment to vet those results. Despite its limitations, this method has been used to identify a number of artifacts which were missed by previous thorough spectral studies. For example, in O1, a total of about 30 combs were identified across both detectors prior to trying automated comb finding; 5 more were then identified using the comb finder.

5.6.1 Comb finding algorithm

This section describes, step by step, the process of the FineTooth comb finder. Capital letters (A, B, C...) stand in for tunable parameters, which can be set by the user, although default values are supplied with the code which are appropriate for a run-averaged DARM spectrum.

- Calculate a running median A bins wide across the spectrum. This is a measure of background noise, not due to lines.
- Remove any lines of known origin (list supplied by the user) from the spectrum, by setting the values for the corresponding bins to the background value. This list may include combs, in which case all marked peaks of the comb will be removed.
- Divide each value by the background value. This normalizes the spectrum and flattens any broad variations in the noise level.
- Identify all peaks in the spectrum, using the simple criterion that peaks must be taller than their B nearest neighbors in frequency. This is the same criterion described in section 5.4.3.
- Order the peaks by height and select the top C entries.
- Compare each peak to its D nearest neighbors in *height* (not frequency). This takes advantage of a common property of combs: consecutive teeth of a real comb are close to each other in height. [3] Each comparison generates a spacing $f_s = |f_i - f_j|$, where f_i and f_j are the frequencies of the peaks being

[3] Generally speaking, strong combs exhibit more absolute peak height variation, so it may be necessary to broaden the allowable peak height variation for strong combs. Fortunately, in a typical spectrum, there are fewer bins containing very high power. Therefore, for a fixed number of nearest neighbors in height D, high peaks will naturally be compared with a more varied array of peak heights than low peaks.

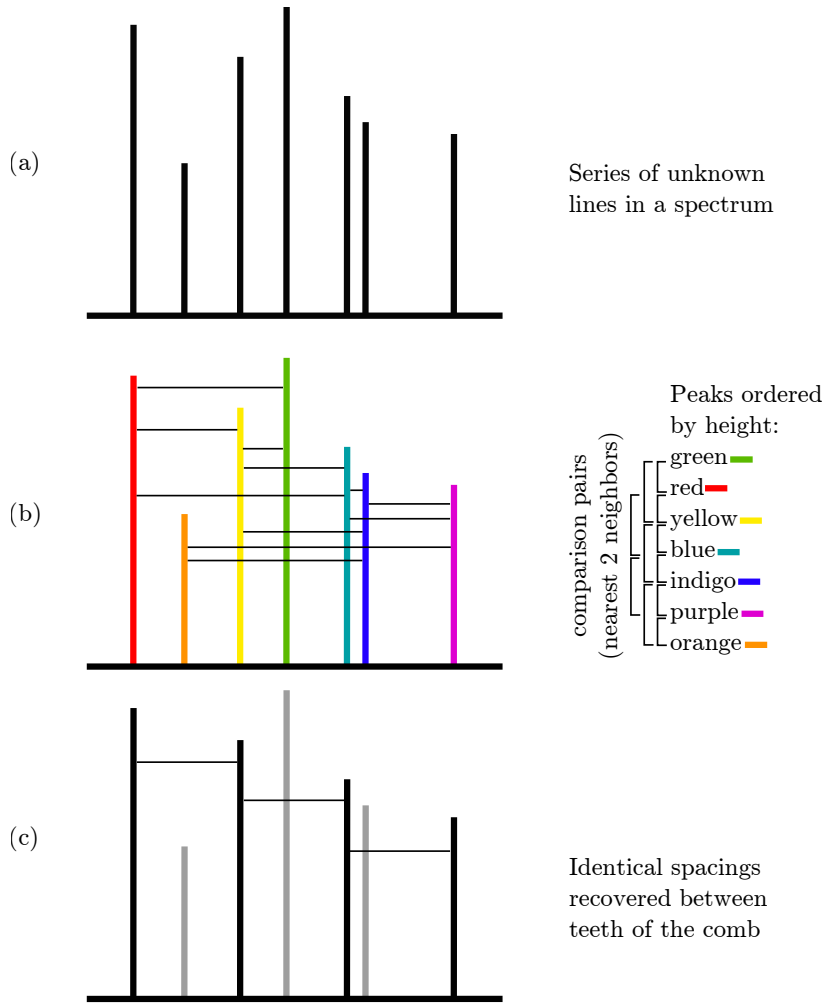


Figure 5.6: Diagram of comb candidate generation based on the spacings of D nearest-neighbor peaks in height. In this case, $D = 2$. Part (a) shows an unknown series of lines of a given spectrum. In (b) the lines are given arbitrary color coding for visual clarity, and ranked by their heights. The nearest $D = 2$ neighbors in the height ranking generate frequency spacings, shown as black horizontal bars in the diagram. Duplicate spacings are identified; in this case there is one set of 3 identical spacings, identifying a single comb candidate. This candidate is highlighted in (c), shown (along with its spacing markers) in black, while all other lines in the spectrum are shown in gray.

compared, and an offset $f_o = f_i \mod f_s = f_j \mod f_s$. This f_s, f_o pair becomes a candidate comb, to be tested further in the next stage of the algorithm. See figure 5.6 for a diagram of this process.

- Throw away any candidate combs which are outside the maximum and minimum comb spacing search range (supplied by the user).

- Count the number of appearances of each (f_s, f_o) pair. For a real comb, each pair of consecutive teeth should yield the same (f_s, f_o) values. Select the E most commonly-recovered combs, and throw away the rest.

- Generate additional comb candidates to account for the (likely) possibility that some of the current candidates are slightly mis-identified, due to the discrete nature of the spectrum.

  - For comb candidates with a small nonzero offset value, a similar comb candidate with zero offset is generated. This reflects an assumption that combs with zero offset are more likely to appear than those with nonzero offset, an assumption which is both physically and observationally grounded. Alternate comb candidate generation is done in the following manner. First, the teeth of the comb are marked.The frequency corresponding to the highest marked harmonic is divided by the comb spacing and rounded to the nearest integer, providing an measure of the highest marked harmonic number. Next, the frequency corresponding to the highest harmonic is divided by the approximate harmonic number to generate a new spacing. The new spacing, with zero offset, is added to the comb candidate list.

  - For comb candidates with a zero offset value (including those generated in the previous step), more comb candidates are generated to account for the possibility that the current candidate is actually a subset of a real comb. For example, it is possible that the current candidate identifies only every other tooth of a real comb, or every third tooth, etc. Comb candidates with spacing $f_s/2$, $f_s/3$, $f_s/4$, and $f_s/5$ are generated and added to the comb candidate list.

- For each comb candidate, generate a statistic describing how well it conforms to the expected behavior of a real comb.

$$L_{\text{comb}} \equiv \frac{\text{avg(marked tooth heights)}}{\text{std(marked tooth heights)}} \frac{(\text{number of marked teeth})^2}{\text{number of predicted teeth}} \tag{5.2}$$

  This statistic depends heavily on the process of comb marking. This statistic selects for the following qualities of a comb:

  - strong (average of marked comb teeth is high)

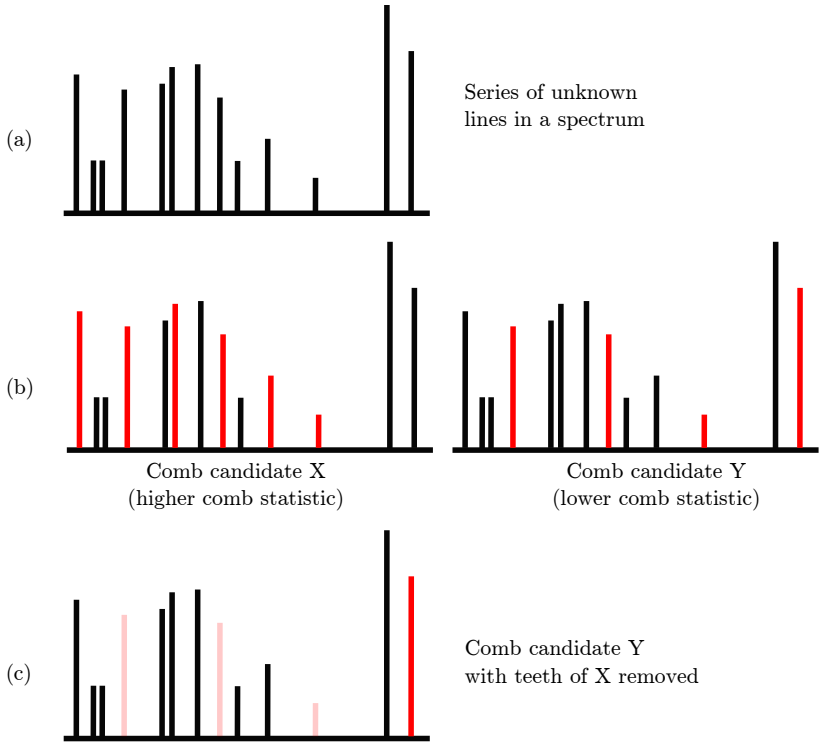


Figure 5.7: Diagram of iterative zeroing routine. In this case, from a series of unknown lines shown in (a), two possible comb candidates were generated, shown in (b). X has a better comb statistic than Y. But is Y an independent structure from X, which just happens to overlap slightly, or is it a spurious candidate? The answer can be seen easily in (c) by removing the teeth of Y that are also teeth of X. (Removed lines are shown as translucent.) Y is likely not an independent structure; it is probably due to the overlap of an isolated strong line with an odd harmonic of X.

  - consistent (standard deviation of marked comb teeth is low)
  - long (numerous teeth)
  - high fraction of teeth marked [4]

- Rank candidates by L_{comb}.
- Check whether each candidate is a subset of a candidate with a higher L_{comb}. This check proceeds as follows. First, for each candidate with a higher L_{comb} than the one under consideration, set all marked bins of the spectrum to zero. Next, recalculate L_{comb} for the candidate under consideration. See figure 5.7 for a diagram of this process.
- Re-rank candidates by the updated L_{comb}, and report those at the top of the list (as many as the user

[4] The square on the number of marked teeth in the expression for L is motivated as follows. It is not sufficient merely to measure the number of marked teeth, because that would excessively prioritize narrowly-spaced combs for a fixed bandwidth of interest. It is also not sufficient to merely measure the fraction of predicted teeth which were marked, because that would give equal priority to a comb X with spacing f_s, and a comb Y with spacing Nf_s for which all marked teeth are *also* marked teeth of X. In other words, it would be unable to distinguish between a given comb candidate and a comb candidate comprised of its every Nth harmonic.

specifies).

5.7 Comb case studies

Line studies and mitigation efforts are carried out by a variety of groups and individuals in the LIGO collaboration, using a variety of different tools, typically on an as-needed basis. The following are case studies, presented chronologically, which demonstrate the wide variety of comb-related investigations at the LIGO detector sites.[5] Many of these case studies also demonstrate the motivation for, and utility of, the comb tracking and investigation tools described in the preceding section.

Further information about narrow spectral artifacts and the various tools used to identify them (including FineTooth) can be found in [17].

5.7.1 Blinking LED comb

In the spring of 2016, investigations began at both Hanford and Livingston sites into a DARM comb with $f_s = 1.0$ Hz, and offset $f_o = 0.5$ Hz. This structure was presumed to be the odd harmonics of an underlying 0.5 Hz oscillation, an assumption corroborated by studies of folded data.[6] This comb contaminated a large number of bins across the spectrum, and most problematically, appeared at both sites, and was synced with the GPS timing system, making it extremely likely to interfere with continuous-wave and stochastic search analyses. Researchers at the Hanford site discovered coherence between the DARM comb and several magnetometers located around the detector. Follow-up studies were done using portable magnetometers, which showed that the comb was particularly strong near equipment associated with the timing system. A test at the Livingston site around the same time showed that placing two timing chassis on independent power supplies reduced the comb in amplitude.

Based on these studies, a hypothesis was formed regarding the origin of the comb. Many components of the timing system have LED indicators, which are intended to provide visual confirmation that the system is receiving the timing signal and is properly synced with it. The default behavior of these indicators was to blink on and off, drawing power in a 2-s period square wave, and thereby minutely altering the voltage supplied by the power supply. Other sensitive electronics sharing the same power supply could be affected by the tiny variations. The large number of indicator LEDs (tens of chassis across each detector, often

[5] Most of the case studies are from Hanford, because I worked there for six months in 2016 (June-December), as part of the LSC Fellows program. The FineTooth infrastructure was developed during that time to aid in ongoing comb investigations, and its earliest applications were to Hanford noise artifacts.

[6] Folding studies involve breaking data into a small segments (in this case, 8 seconds), and averaging the segments, so that weak periodic structures which repeat on that time scale will be highlighted. For example, an 8-second folding period will also show structures with a period of 4 seconds, 2 seconds, 1 second, etc.

with several indicators) and the consistent, precisely-synced nature of the oscillation combined to make the contamination visible in DARM over integration times on the order of a day.

Rather than replacing the power supply for the master system, an update to the default LED behavior was used to mitigate the comb. The firmware on timing chassis around both sites was updated so that the LEDs remain constantly on when the system is properly synced. This intervention succeed in mitigating the comb by a factor of ~ 10, [17] but did not completely remove it. Various hypotheses exist as to why the comb was not fully eliminated, but one likely possibility is that some LEDs were missed during the reprogramming. For example, some weeks after the visible LEDs were reprogrammed, unrelated maintenance turned up another set of blinking LEDs on timing cards hidden inside the chassis.

An early application of the FineTooth comb plotting and tracking tools was to monitor the blinking LED comb before and after the intervention, particularly at Hanford, and to evaluate the effect of reprogramming the blinking LEDs.

5.7.2 CPS timing fanout comb

In May of 2016, and again in July of that year, a comb was observed in DARM with $f_s = 1.999951$ Hz, with teeth on approximately odd-integer frequencies. Examinations of DARM data indicated that the comb had appeared some time prior to April, but it was impossible to determine the exact start date because the detector was offline for a long period in the spring.

Fortunately, the comb was also clearly present in magnetometer data, specifically in a magnetometer channel at the end X station. Using data from that channel, it was possible to track down the appearance of the comb to March 14, 2016, and to cross-reference that date with the alogs. Notably, work had been done that day on the capacitive position sensor (CPS) timing system at both end stations. The CPS interface chassis shares an electronics rack and a power supply with the electrostatic drive (ESD), providing a plausible coupling mechanism to DARM. (Recall that the electrostatic drive is used to move the test masses themselves.)

As part of an ongoing (unrelated) investigation into transient glitches, a magnetometer had recently been placed near the ESD. This magnetometer was not on the usual list of channels to monitor, because it was intended to be temporary and for a specific purpose. After the CPS timing system was identified as a probable cause of noise, however, it was natural to examine the temporary magnetometer data. The comb in question was extremely strong in the temporary magnetometer data, providing solid evidence that the CPS timing hypothesis was correct. [3] The CPS timing fanout was subsequently reprogrammed, and placed on an alternate power supply separate from the ESD. This very clearly mitigated the comb in DARM.

5.7.3 Hartmann wavefront sensor comb

In July 2016, a strong comb was noted in the Hanford DARM spectrum, with spacing $f_s = 0.997698$ Hz and offset $f_o = 0$ Hz. It was easily the strongest comb in the Hanford spectrum at the time, and contaminated almost 200 bins at low frequencies, in the most sensitive part of the LIGO observation band. Unfortunately, there was not enough available DARM data to determine when the comb had appeared, and the comb coupled only intermittently into DARM, even when data was available. It was present in the data from ER9 (the ninth engineering run), and would be observed several more times for the next few months, but little progress was made on identifying its cause.

In September 2016, the newly-developed FineTooth tracking tools were used to search for the comb across magnetometer channels. This revealed that the comb was reflected in corner-station magnetometers. Because data from magnetometers are continuously acquired, unlike DARM data, it was possible to identify when the comb first appeared by tracking its presence in magnetometers. A very distinctive change point (see figure 5.8) was found at a much earlier date: between February 3rd and 4th of 2016. Cross-referencing with alog entries for that date showed that work had been performed on the Hartmann wavefront sensors (HWS) at the X and Y ITMs, specifically the installation of a temperature sensor. The alog yielded an approximate time of day for this work which corresponded with the appearance of the comb; a more detailed check on one of the channels associated with the sensor confirmed the estimate. [2]

Working from the hypothesis that the HWS shutter was responsible for the noise, and that the coupling most likely took place through a shared power supply with other electronics, an intervention was devised. The HWS was placed on an alternate, isolated power supply. This caused the comb to promptly disappear from the corner station magnetometer spectra, and it did not re-couple into DARM in the months following the intervention.

5.7.4 Hanford photon calibrator ethernet adapter

In March of 2017, a variety of new narrow lines were identified in the Hanford DARM spectrum, all close to a spacing of 1 Hz, with various offsets. These lines appeared after the normal Tuesday maintenance period on March 14. It was initially suspected that the Hartmann wavefront sensor (HWS) cameras were at fault, since work had been logged on the HWS system on the 14th, but no major changes to their configurations were made. Another point against the HWS hypothesis was the fact that the combs appeared in magnetometers at the end stations only; previous HWS-associated noise had come from the corner station, and was presumably associated with the ITM HWSs. Researchers at the Hanford site eventually came to suspect that similarly-timed work with the photon calibration cameras (located at the end stations) might be at fault. After the

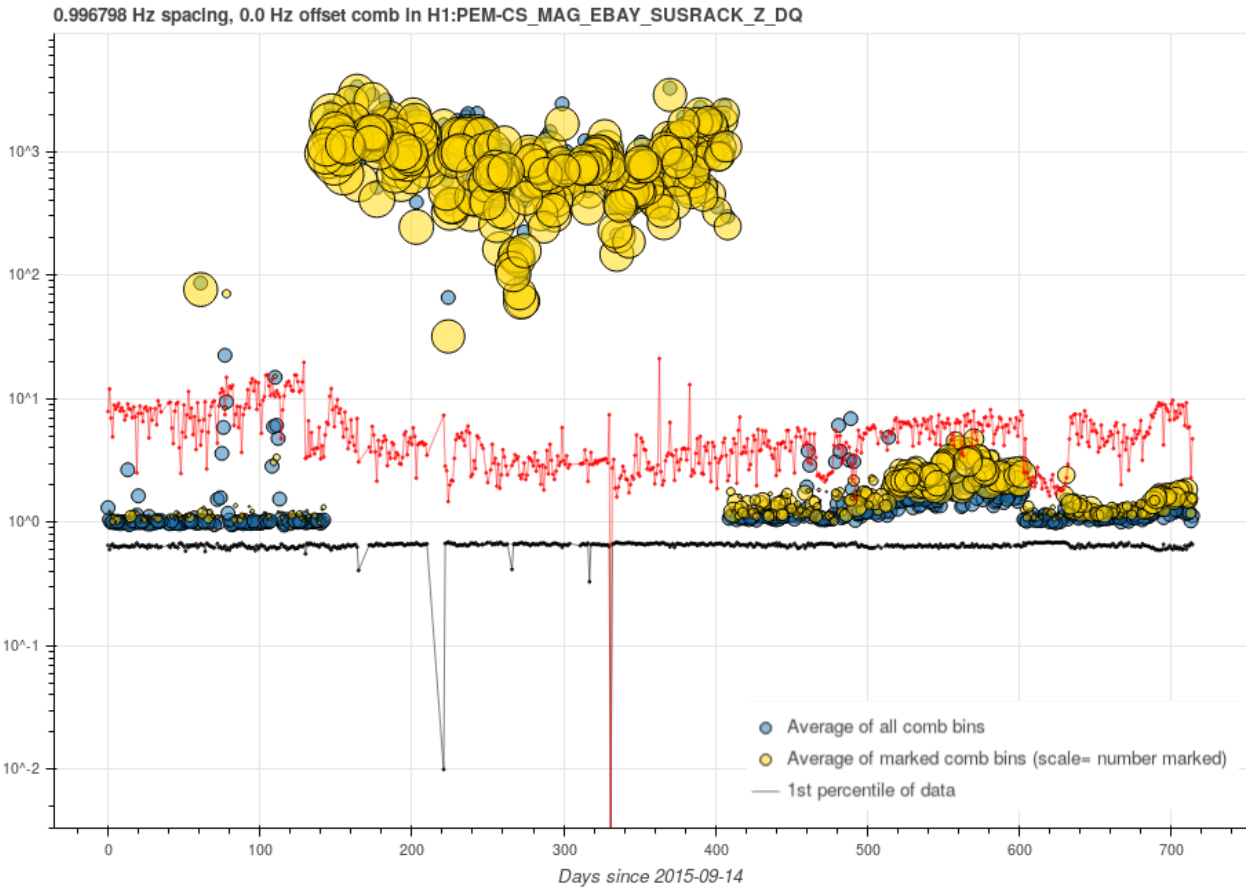


Figure 5.8: Screenshot of a tracking plot of a near 1 Hz comb, associated with the Hartmann wavefront sensor (HWS) shutter, as it appeared in a corner station magnetometer channel. The initial appearance of the comb helped to identify its cause; the sudden drop in strength corresponds to an intervention which placed the HWS on an isolated power supply.

camera ethernet adapters were turned off on April 18th, the combs in question were strongly mitigated. Comb tracking data allowed this change to be easily monitored and confirmed in the days following the intervention. [4]

Chapter 6

Searching for signals from neutron stars in binary systems with TwoSpect and the X-statistic

6.1 TwoSpect and X-statistic

6.1.1 Overview of the TwoSpect method

TwoSpect was originally developed in the Michigan Gravitational Wave Group [24, 22]. It is now one of the core analysis pipelines maintained by the continuous wave group, and has been used to set upper limits on detectable strain amplitude in several searches. [36, 45]. TwoSpect was developed to search for sources in binary systems. It operates by performing two Fourier transforms (hence the name), the first of which extracts information on the source frequency over short timescales, and the second of which extracts information on the orbital modulation frequency over long timescales. Descriptions of the TwoSpect method in this chapter draw from [24].

6.1.2 Overview of the X-statistic method

The X-statistic analysis technique, described in this thesis, builds on the framework of TwoSpect. It leverages a feature of TwoSpect's templated search mode in order to obtain partial information on signals by sparse sampling points in parameter space. The X-statistic method, while not fully mature, is currently being applied to a pilot search for promising unassociated sources from the Fermi-LAT gamma ray catalog, which will be described in detail in the next chapter.

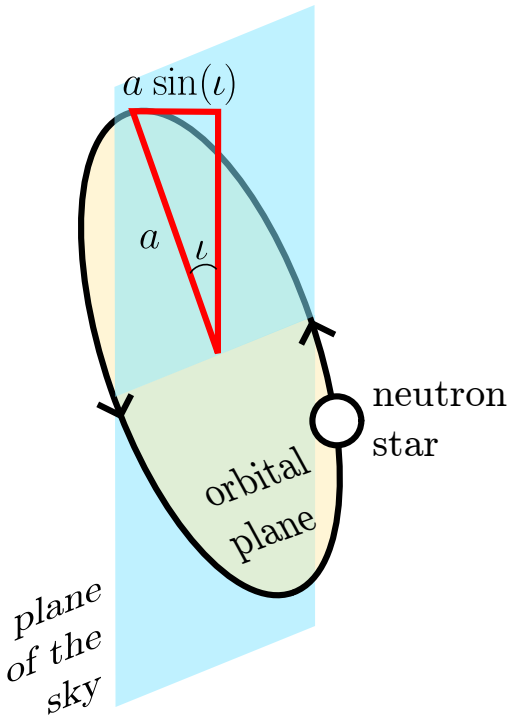


Figure 6.1: Diagram showing a binary orbit with semi-major axis a, inclined at some angle ι , such that the projected semi-major axis is $a \sin \iota$.

6.2 TwoSpect analysis

6.2.1 Parameter space

TwoSpect searches over five astrophysical parameters: two angles describing sky location (right ascension and declination), and three parameters describing the source itself. It does not take phase information into account, and it assumes that the orbit of the binary system is circular (although the method is robust against low orbital eccentricities).

The three source parameters are the frequency of emitted gravitational waves f_{gw}, the binary orbital period P_{orb}, and the projected semi-major axis of the binary, $a \sin \iota$ (see figure 6.1). These source parameters are typically rephrased for computational purposes in terms of three signal parameters: f_{gw}, P_{orb} (identical to the source parameters) and the modulation depth Δf_{gw}. The modulation is the maximum amount by which the observed frequency shifts from f_{gw} due to the binary's orbit,

$$\Delta f_{\mathrm{gw}} = \frac{2\pi a \sin \iota}{c P_{\mathrm{orb}}} f_{\mathrm{gw}} \tag{6.1}$$

This modulation is expected to be small; 0.1 Hz is considered a large value.

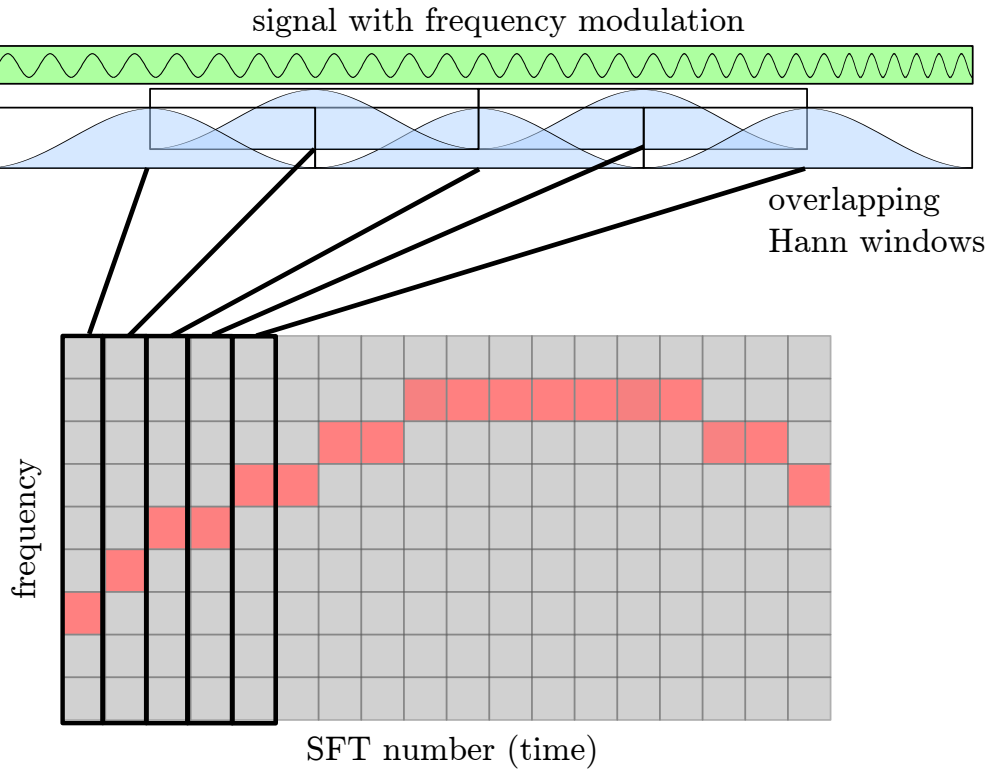


Figure 6.2: Diagram showing how SFTs are used to construct a time-frequency map. The modulated signal is "tiled" with overlapping Hann-windowed SFTs, each of which results in a 1-dimensional spectrum. The frequency of the signal is modulated, so the bin containing the most power changes between SFTs. Since the frequency modulation comes from the binary orbit, the bin containing the most power traces out a sinusoidal path. The amplitude of this sinusoid is the modulation depth, Δf_{gw}.

6.2.2 Short Fourier Transforms (SFTs)

As a first step in data processing, the strain data from each detector is split into a series of equal-length segments. The segment length is the coherence time for the search, T_{coh}. In the original TwoSpect, the segments from different detectors were combined incoherently after being independently coherently analyzed. Recently, a coherent SFT addition method has been developed, allowing for increased sensitivity when using combined data from multiple detectors [23].

Each segment of strain data is Fourier transformed, creating a spectral "snapshot" of the signal, known as a Short Fourier Transform (SFT). The frequency of the signal will change between consecutive SFTs, but T_{coh} is chosen so that *within* a given SFT, the signal frequency is approximately constant, and the signal power will be mostly contained within a single spectral bin. This sets a maximum bound for T_{coh}:

$$T_{\mathrm{coh}} \leq \sqrt{\frac{P}{2\Delta f_{\mathrm{gw}}}} \tag{6.2}$$

Outside of the constraint based on modulation depth, T_{coh} should be as long as possible to achieve maximum

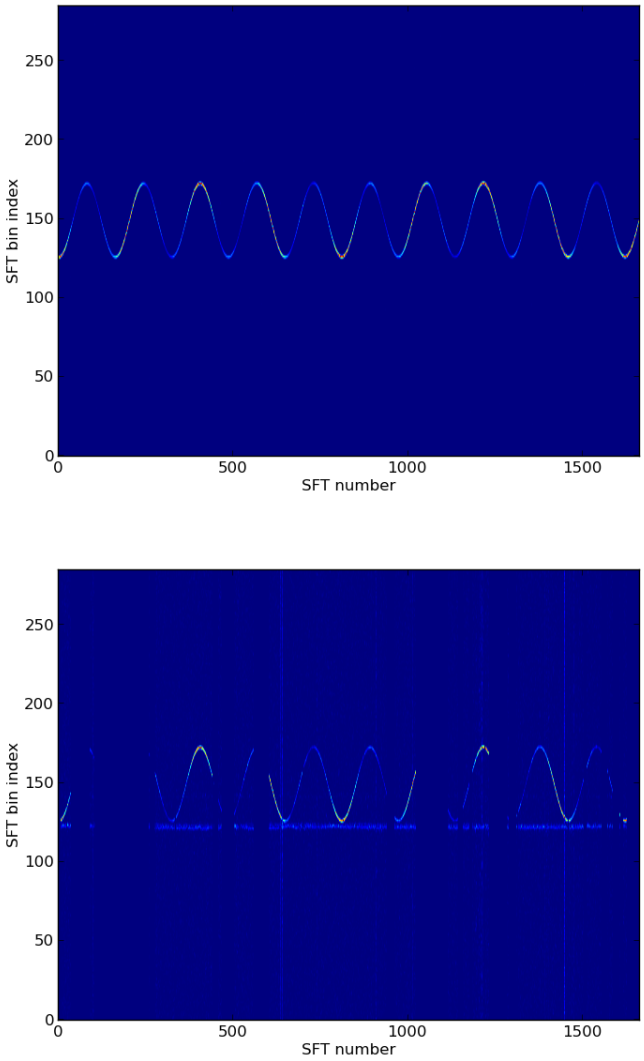


Figure 6.3: The time-frequency plane for a simulated signal in Gaussian noise (left) and for a simulated signal injected into O1 data (right). The horizontal axis is effectively a time axis, labeled in units of SFT number. The orbital modulation period P_{orb} is much shorter than the total observation time T_{obs}. Note the presence of an instrumental line in the real data, near the frequency minima of the signal. If this signal were real, its detection could be complicated by the proximity of the line.

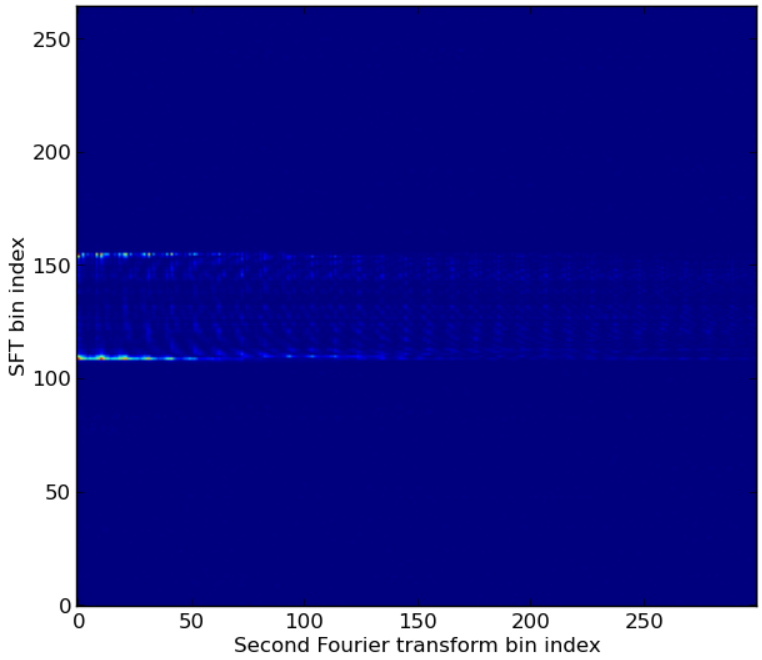


Figure 6.4: The frequency-frequency plane for a simulated signal. Note the symmetric structure about the central frequency.

sensitivity. Each SFT is Hann windowed, and they overlap by $T_{\mathrm{coh}}/2$ (see figure 6.2).

Next, the data is barycentered: frequency bins are adjusted to account for the Doppler shift due to the detector's motion relative to the source location, at the time the data was taken. This calculation must be done independently for each sky location considered in the search. For an all-sky search, this results in a large number of sky locations [24]:

$$N_{\mathrm{sky}} \approx 2 \times 10^4 \left(\frac{f_{\mathrm{gw}}}{100\,\mathrm{Hz}} \right)^2 \left(\frac{T_{\mathrm{coh}}}{1800\,\mathrm{s}} \right)^2 , \tag{6.3}$$

a factor which multiplies the total computational cost. For a targeted or directed search, by contrast, a single point is generally sufficient.

6.2.3 The second Fourier transform

The presence of a periodic signal in the first Fourier transform plane (i.e. the time-frequency plane; see figure 6.2) invites a second Fourier transform. In this step, each fixed-frequency row of the first Fourier transform plane is treated as a time series, and the second Fourier transform extracts periodicity in it, related to the orbital motion. A signal from a neutron star in a binary system will produce a characteristic pattern in the

second Fourier transform (frequency-frequency) plane. The "top" and "bottom" rows of the pattern contain the most power, corresponding to the minimum and maximum of the orbital sinusoid, where the signal spends most of its time. These rows are centered around the un-modulated signal frequency, f_{gw}. In the second Fourier transform direction, the fundamental and higher harmonics of the orbital modulation can be seen, with the higher harmonics trailing off in strength.

6.2.4 Incoherent harmonic summing and templated searches

After the second Fourier transform is computed, two options are available within the TwoSpect framework. The first is a method called incoherent harmonic summing (IHS), which searches for an excess of power in the frequency-frequency plane, after summing the first five orbital modulation harmonics. This technique is relatively computationally cheap, but it is not the most sensitive option. As a result, it is used mainly as a hierarchical stage for all-sky searches. Outliers found in in the IHS stage will be followed up by a small templated search, described below.

The second option for handling the frequency-frequency data is to go directly to a templated search. In this case, a large number of templates (predicted signals) are computed, organized in a grid over the parameter space of interest. As discussed earlier in section 6.2.1, this is a three-dimensional parameter space. The computational cost of this search is dominated by the number of templates computed. For an all-sky search, the number of required templates is multiplied by the number of different sky locations to search, rendering a fully-templated all-sky search computationally infeasible with the available resources. Even a fully-templated directed search (fixed sky location) is not feasible unless more information about the source is known. The main application of TwoSpect's fully-templated search has so far been to Scorpius-X1, a potential source for which the orbital period is well measured, and the projected semi-major axis is constrained. [45]

6.2.5 The R-statistic for templated searches

A templated search involves comparing a series of templates (predicted signals, defined by frequency $f_{\mathrm{gw}}^{\mathrm{temp}}$, a modulation depth $\Delta f_{\mathrm{gw}}^{\mathrm{temp}}$, and an orbital period P_{orb}^{temp}) to the actual data (comprised of n SFTs with k frequency bins each.) For each template, a detection statistic R is calculated:

$$R = \frac{\sum_{i=0}^{M-1} w_i \left[Z_i - \lambda_i\right]}{\sum_{i=0}^{M-1} \left[w_i\right]^2} \tag{6.4}$$

where i indexes each pixel, M is a maximum pixel index less than the total number of pixels $(n)(k)$, Z_i is the spectral power of the ith pixel, and λ_i is the expected noise power in the ith pixel. The template weights

w_i encode the signal prediction; they are calculated for each combination of $f_{\rm gw}^{\rm temp}$, $\Delta f_{\rm gw}^{\rm temp}$, and $P_{\rm orb}^{\rm temp}$. When heavily-weighted pixels overlap with high-power pixels, R is large, indicating that a signal is likely to be present.

6.2.6 Partial signal recovery from specific non-optimal templates

The R statistic straightforwardly compares a predicted pattern for the frequency-frequency plane (w_i) with the actual pattern ($Z_i - \lambda_i$). The best R statistic is recovered when the template precisely overlaps with a signal (or at least, more precisely than any neighboring template).

In certain cases, non-optimal templates can also recover part of the signal, yielding an R statistic which is not as high as the optimal template, but which is higher than the surrounding non-optimal templates. The reason for this can be seen by considering the top and bottom rows of the characteristic signal pattern described in section 6.2.3. These rows are located at $f_{\rm max} = f_{\rm gw} + \Delta f_{\rm gw}$ and $f_{\rm min} = f_{\rm gw} - \Delta f_{\rm gw}$. If a template is offset by some amount f_{off} in both frequency and modulation depth, then one of the following situations occurs:

$$
\begin{aligned}
f_{\rm max}^{\rm temp} &= \left(f_{\rm gw}^{\rm signal} - f_{\rm off}\right) + \left(\Delta f_{\rm gw}^{\rm signal} + f_{\rm off}\right) = f_{\rm max}^{\rm signal} \\
f_{\rm max}^{\rm temp} &= \left(f_{\rm gw}^{\rm signal} + f_{\rm off}\right) + \left(\Delta f_{\rm gw}^{\rm signal} - f_{\rm off}\right) = f_{\rm max}^{\rm signal} \\
f_{\rm min}^{\rm temp} &= \left(f_{\rm gw}^{\rm signal} + f_{\rm off}\right) - \left(\Delta f_{\rm gw}^{\rm signal} + f_{\rm off}\right) = f_{\rm min}^{\rm signal} \\
f_{\rm min}^{\rm temp} &= \left(f_{\rm gw}^{\rm signal} - f_{\rm off}\right) - \left(\Delta f_{\rm gw}^{\rm signal} - f_{\rm off}\right) = f_{\rm min}^{\rm signal}
\end{aligned}
$$

Or, in summary,

$$
f_{\rm gw}^{\rm temp} \pm \Delta f_{\rm gw}^{\rm temp} = f_{\rm gw}^{\rm sig} \pm \Delta f_{\rm gw}^{\rm sig} \tag{6.5}
$$

In any of the four scenarios above, one of the most heavily weighted rows of the template overlaps with one of the highest power rows of the signal, yielding a elevated R statistic compared to other non-optimal templates.

If a templated search is performed over an array of $f_{\rm gw}^{\rm temp}$, $\Delta f_{\rm gw}^{\rm temp}$ points about the actual signal, non-optimal templates will form a distinctive X-shaped pattern (with arms of slope 1) in the resulting map of R.

6.3 X-statistic algorithm

6.3.1 Motivation

The behavior of non-optimal templates described in section 6.2.6 means that information about the signal is spread across the frequency-modulation depth plane in a fully templated search. The X-statistic method is a novel attempt to leverage this information for a computationally efficient search.

6.3.2 Method

A normal fully-templated search would involve calculating a grid of templates in the $(f_{\mathrm{gw}}, \Delta f_{\mathrm{gw}})$ plane at each value of $P_{orb}^{\mathrm{template}}$. In the X-statistic method, only a small number of $\Delta f_{\mathrm{gw}}^{\mathrm{template}}$ points are sampled. There is no such restriction on $f_{\mathrm{gw}}^{\mathrm{template}}$, which spans the entire frequency range of interest. Therefore, each $\Delta f_{gw}^{\mathrm{template}}$ value defines a sampled transect across the $(f_{\mathrm{gw}}, \Delta f_{\mathrm{gw}})$ plane, which will intersect any X pattern in that plane at two points. [1]

The information gleaned from sampling fixed-Δf_{gw} transects can be used to reconstruct the position of the X. This is accomplished by testing a grid of proposed signal locations $\left(f_{\mathrm{gw}}^{\mathrm{test}}, \Delta f_{\mathrm{gw}}^{\mathrm{test}}\right)$ across the parameter space of interest. For each $\left(f_{\mathrm{gw}}^{\mathrm{test}}, \Delta f_{\mathrm{gw}}^{\mathrm{test}}\right)$ and each sampled $\Delta f_{\mathrm{gw}}^{\mathrm{template}}$, the two points of intersection are calculated. The nearest sampled template to the intersection point has a detection statistic $R_{\mathrm{intersect}}$. All the values of $R_{\mathrm{intersect}}$for a particular $\left(f_{\mathrm{gw}}^{\mathrm{test}}, \Delta f_{\mathrm{gw}}^{\mathrm{test}}\right)$ location are averaged to obtain the X-statistic for that point. Note that this calculation, a simple average of four numbers, is extremely computationally cheap, unlike calculating the R statistic for a template.

The test point with the best X statistic is then followed up by a small templated search around that point. The best R statistic among the follow-up points is considered to be the detection statistic for the search.

6.3.3 Template placement and computational cost

6.3.3.1 Requirements for template placement

Because the placement of templates for an X-statistic search is a modified version of the placement for a fully-templated TwoSpect search, this section largely follows [24] and some errata [21].

The placement of templates is based on the chosen acceptable mismatch μ for the search. Mismatch quantifies the maximum allowable loss of signal. If there is a signal in the data, then a template perfectly matched with that signal would have some value R_{ideal}. In reality, it is unlikely that any of the tested

[1] Or one point, if $\Delta f_{\mathrm{gw}}^{\mathrm{temp}}$ is coincidentally equal to $\Delta f_{\mathrm{gw}}^{\mathrm{signal}}$.

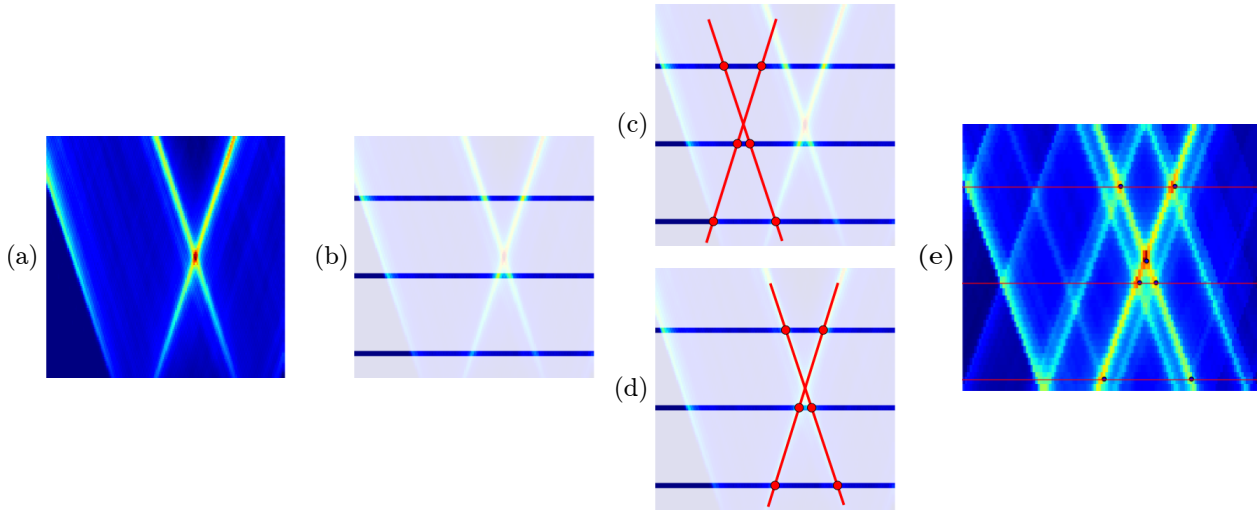


Figure 6.5: A visual explanation of the X-statistic method. Sub-figure (a) shows a plot of the R statistic for a fully templated search spanning part of the $(f_{gw}, \Delta f_{gw})$ plane. This map corresponds to a simulated signal of high strength, located at the center of the X pattern. In (b), a diagram shows the selection of three transects across the plane. In (c), a diagram shows a non-optimal location of the X, and its intersections with the sampled transects. In (d), a similar diagram is shown for the optimal location of the X pattern. If tests like those in (c) and (d) are repeated across the full plane, a map of the resulting X statistic can be generated; this data is shown in (e). Unlike (b, c, and d), (e) is not a diagram, but rather a quantitative graph showing the actual map of X-statistic values generated using this technique.

templates will perfectly match the signal; instead, the best detection statistic will be some value $R_{\rm best}$, associated with the nearest template. Templates must be placed closely enough that $R_{\rm best} \geq (1-\mu)R_{\rm ideal}$. The standard acceptable mismatch for TwoSpect is $\mu = 0.2$.

Template placement for the initial stage of an X-statistic searches mimics that for fully templated TwoSpect searches for two of the three source parameters ($f_{\rm gw}$ and $P_{\rm orb}$). The mismatch requirement is broken for $\Delta f_{\rm gw}$ in order to decrease computational cost (see section 6.3.2). The requirement *is*, however, applied for the second stage, where a small templated follow-up is performed about the point with the best X statistic.

6.3.3.2 Estimating the number of template points in frequency and modulation depth

Templates must be spaced by $(2T_{\rm coh})^{-1}$ in $f_{\rm gw}$, which straightforwardly yields a number of $f_{\rm gw}^{\rm temp}$ points:

$$n_f^{\rm temp} = (f_{\rm gw}^{\rm search\ max} - f_{\rm gw}^{\rm search\ min})\,(2T_{\rm coh}) \tag{6.6}$$

The number of grid points in $\Delta f_{\rm gw}$ would be treated similarly in a fully templated search,

$$n_{\Delta f}^{\rm temp} = (\Delta f_{\rm gw}^{\rm search\ max} - \Delta f_{\rm gw}^{\rm search\ min})\,(2T_{\rm coh}) \tag{6.7}$$

but for an X-statistic search, only a small number of $\Delta f_{\rm gw}^{\rm temp}$ points are used. In this case,

$$n_{\Delta f}^{\rm temp} = \text{floor}\left(\frac{\Delta f_{gw}^{\rm search\,max} - \Delta f_{\rm gw}^{\rm search\,min}}{\Delta f_{\rm gw}^{\rm transect\,spacing}}\right). \tag{6.8}$$

Since searches typically run to very low $\Delta f_{\rm gw}$, this can be approximated as

$$n_{\Delta f}^{\rm temp} \approx \text{floor}\left(\frac{\Delta f_{\rm gw}^{\rm search\,max}}{\Delta f_{\rm gw}^{\rm transect\,spacing}}\right) \tag{6.9}$$

The spacing of templates in orbital period is more complicated. Unlike $f_{\rm gw}$ and $\Delta f_{\rm gw}$, templates are not placed uniformly in $P_{\rm orb}$. Low orbital periods require templates to be much more closely spaced.

6.3.3.3 Estimating the number template points in orbital period

A mismatch between the template and the signal in $P_{\rm orb}$ means that the signal is effectively shifted some fraction of a second Fourier transform bin relative to the template, over the course of the observation time $T_{\rm obs}$. The fractional bin shift corresponding to a mismatch $\mu = 0.2$ is an empirically derived parameter, known as α. The next template should be placed so as to properly recover the shifted signal. In other words,

$$\frac{1}{P_{\rm orb}^{\rm temp} + \Delta P_{\rm orb}^{\rm temp}} = \frac{1}{P_{\rm orb}^{\rm temp}} + \frac{1}{\alpha T_{\rm obs}} \tag{6.10}$$

To first order in $P_{\rm orb}^{\rm temp}/(\alpha T_{\rm obs})$, this works out to:

$$\Delta P_{\rm orb}^{\rm temp} \approx \frac{\left(P_{\rm orb}^{\rm temp}\right)^2}{\alpha T_{\rm obs}} \tag{6.11}$$

This expression, however, is missing a necessary dependence on $\Delta f_{\rm gw}$. For a fixed $P_{\rm orb}$, an increase in modulation depth means a decrease in the signal at the turning points of the sinusoid in second Fourier transform data. Specifically, the signal scales with the square root of $\Delta f_{\rm gw}$. That, in turn, requires more closely spaced templates when $\Delta f_{\rm gw}$ is high. Absorbing the constant of proportionality (also empirical, and with appropriate units) into α:

$$\Delta P_{\rm orb}^{\rm temp} \approx \frac{\left(P_{orb}^{\rm temp}\right)^2}{\alpha T_{\rm obs}\sqrt{\Delta f_{\rm gw}}} \tag{6.12}$$

In actual practice, the search will be carried out over a series of fixed $P_{\rm orb}^{\rm temp}$ points, each of which will involve a search up to the maximum modulation depth $\Delta f_{\rm gw}^{\rm search\,max}$. Therefore, to be conservative, we should use this value for $\Delta f_{\rm gw}$ when calculating orbital period spacing.

Noting that templates are placed in orbital period with a density of $1/\Delta P_{\rm orb}^{\rm temp}$, the total number of template points can be approximated as:

$$n_P^{\rm temp} \approx \int_{P_{\rm orb}^{\rm search\ min}}^{P_{\rm orb}^{\rm search\ max}} \frac{1}{\Delta P_{\rm orb}^{\rm temp}} dP_{\rm orb}^{\rm temp} \tag{6.13}$$

$$n_P^{\rm temp} \approx \alpha T_{\rm obs} \sqrt{\Delta f_{\rm gw}^{\rm search\ max}} \left(\frac{1}{P_{\rm orb}^{\rm search\ min}} - \frac{1}{P_{\rm orb}^{\rm search\ max}} \right) \tag{6.14}$$

$$n_P^{\rm temp} \approx \frac{\alpha T_{obs} \sqrt{\Delta f_{\rm gw}^{\rm search\ max}}}{P_{\rm orb}^{\rm search\ min}} \tag{6.15}$$

6.3.3.4 Estimating the total number of templates and computational cost

Combining the estimates from the previous section,

$$n_{tot}^{\rm temp} = n_f^{\rm temp} n_{\Delta f}^{\rm temp} n_P^{\rm temp} \tag{6.16}$$

$$n_{\rm tot}^{\rm temp} = (f_{\rm gw}^{\rm search\ max} - f_{\rm gw}^{\rm search\ min}) (2T_{\rm coh}) \left(\frac{\Delta f_{\rm gw}^{\rm search\ max}}{\Delta f_{\rm gw}^{\rm transect\ spacing}} \right) \frac{\alpha T_{\rm obs} \sqrt{\Delta f_{\rm gw}^{\rm search\ max}}}{P_{\rm orb}^{\rm search\ min}} \tag{6.17}$$

$$n_{\rm tot}^{\rm temp} = \frac{2\alpha T_{\rm coh} T_{\rm obs} \left(\Delta f_{\rm gw}^{\rm search\ max}\right)^{3/2}}{P_{\rm orb}^{\rm search\ min} \Delta f_{\rm gw}^{\rm transect\ spacing}} (f_{\rm gw}^{\rm search\ max} - f_{\rm gw}^{\rm search\ min}) \, . \tag{6.18}$$

Note that this is a factor of $\left(\Delta f_{\rm gw}^{\rm transect\ spacing}\right) (2T_{\rm coh})$ smaller than the number of templates that would be required for a fully templated search. Depending on the transect spacing chosen, the computational cost of an X-statistic search is generally a few percent of the cost of a fully templated search.

6.4 Evaluating the potential of the X-statistic method

At their core, gravitational wave detection pipelines are binary classifiers: the goal is to achieve either a detection or a non-detection, as reliably as possible. This is done by generating a detection statistic, and comparing that statistic to a chosen threshold value. Detection statistics that exceed the threshold are considered evidence that a signal exists.

The details of the detection statistic, its statistical distribution, and the threshold value may be different for each pipeline. When approaching a new method, like the X-statistic, it is also not initially clear where the threshold ought to be set.

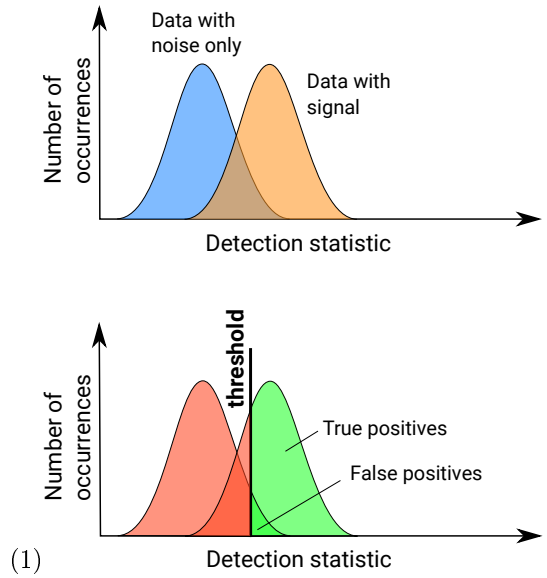


Figure 6.6: Diagram of using a threshold value to distinguish between two detection statistic distributions.

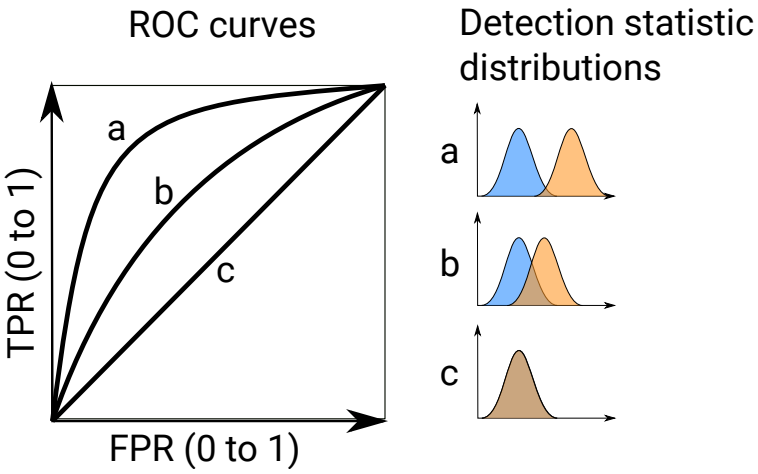


Figure 6.7: Diagram of ROC curves for different underlying distributions of detection statistic.

6.4.1 Characterizing search methods without a fixed threshold value

A Receiver Operating Characteristic (ROC) curve is a method-independent way to characterize a binary classifier, and therefore to compare performance between different pipelines. An ROC curve is generated by comparing two distributions of the detection statistic for a pipeline: one distribution corresponding to pure noise (zero signal), and one corresponding to a signal in noise. If the classifier works well, these distributions will be well separated; if it works poorly, they will be nearly indistinguishable. For a given threshold, a false positive and true positive rate can be calculated (see figure 6.6). The false positive rate (FPR) is the percentage of zero-signal trials which exceed the threshold, while the true positive rate (TPR) is the percentage of signal-containing trials which exceed the threshold. By varying the threshold value, a curve

comparing the FPR and TPR can be generated. In general, for real searches, the threshold will be set to achieve low FPR, due to the difficulty of following up outliers. (See chapter 7 for details.)

6.4.2 Characterizing search methods in terms of upper limits and sensitivity depths

The key scientific goal for a search is to either make a detection, or to set upper limits on the amplitude of gravitational waves emanating from a given source (or class of sources). Given the current sensitivity of the LIGO detectors and the extremely small strain amplitude expected for continuous-wave sources, non-detection is a plausible outcome for most real searches in the near future.

6.4.2.1 Upper limit calculations

The detection efficiency of a search can be measured using simulations: for a given number of simulated signals, the efficiency is the percentage of those signals that are successfully detected. This measurement requires that a specific detection statistic threshold be set, in accordance with the acceptable false positive rate. The detection efficiency is also a function of the signal strength, of course; strong signals will be detected nearly all the time, while weak signals will be detected only some of the time. The detection efficiency as a function of strain amplitude is reasonably well modeled by the logistic function:

$$C(h_0) = \frac{1}{1 + \exp\left(-c_1\left(h_0 - c_0\right)\right)} \tag{6.19}$$

In this context, the upper limit is the strain amplitude for which the efficiency meets some given requirement. For example, the 95% upper limit is the value of h_0 for which $C(h_0) = 0.95$. A signal with this amplitude, or a greater amplitude, would be detected at least 95% of the time, so a non-detection implies 95% confidence that no such signal above the upper limit exists.

Upper limits can be calculated using simulated signals, injected into the noise. For each injection, a binary determination (detection labeled as 1, or non-detection labeled as 0) is made using the search method in question, mimicking as closely as possible the process of a real search. This data is fit to a logistic curve, which is then used to calculate the upper limit.

6.4.2.2 Sensitivity depth calculations

While upper limits are the key astrophysical figure of merit, they are not an inherent property of the search method. Instead, upper limits scale with the sensitivity of the data. Identical search procedures using data from different observing runs will produce different results. A better figure of merit for characterizing the

search method itself is the sensitivity depth, which compares the upper limits with the overall noise level of the input data:

$$D^{95\%} = \frac{\sqrt{S_h}}{h_0^{95\%}} \tag{6.20}$$

where S_h is the noise amplitude spectral density. All of these quantities are generally functions of frequency; $h_0^{95\%}$ and $D^{95\%}$ are typically also functions of other search parameters.

6.4.3 The X-statistic method in comparison with the IHS hierarchical stage

A key goal for the X-statistic method is to outperform the incoherent harmonic sum (IHS) stage in terms of sensitivity. The potential to achieve this goal was evident from the earliest tests of the method.

Early tests of the X-statistic were performed on narrow frequency bands, first in simulated Gaussian noise, and then in real data from LIGO's first observing run (O1). In both cases, simulated signals (typically called injections) were added to the noise.

Although the strain amplitude h_0 was fixed for each trial, the orbital inclination angle ι was randomized. Specifically, injections were distributed uniformly in $\cos\iota$, as is geometrically appropriate for a real population. The effective strain seen by the detector is dependent on $\cos\iota$,

$$h_{0,\mathrm{eff}} = h_0 \sqrt{\frac{1}{2}\left(\frac{1}{4}\left(1+\cos^2\iota\right)^2 + \cos^2\iota\right)}\,, \tag{6.21}$$

so in actuality, a range of signal strengths were tested in this manner. The polarization angle ψ was also randomized.

Figure 6.8 shows the underlying distribution of detection statistics at varying injected strain amplitudes, for both methods. Some key features are apparent from these plots. First, we see a confirmation of the information in the ROC curves: X-statistic outperforms the IHS method. Second, the X-statistic detection statistic is more closely correlated with the effective strain amplitude $h_{0,\mathrm{eff}}$. The X-statistic method, however, has a longer tail on the upper end of the distribution of its detection statistic. This fact is not wholly surprising; an X-statistic search selects the best R statistic from a number of values, corresponding to the small templated follow-up. It therefore has many opportunities to pick up spurious noise.

The broadened distribution of detection statistics for the noise-only X-statistic search requires either setting the threshold higher to preserve a low false positive rate (sacrificing sensitivity), or setting the threshold lower and being confronted with a non-trivial number of outliers. This issue will be the largest challenge for the X-statistic method.

Figure 6.9 shows ROC curves for several tests of the X-statistic method, and equivalent tests for the IHS

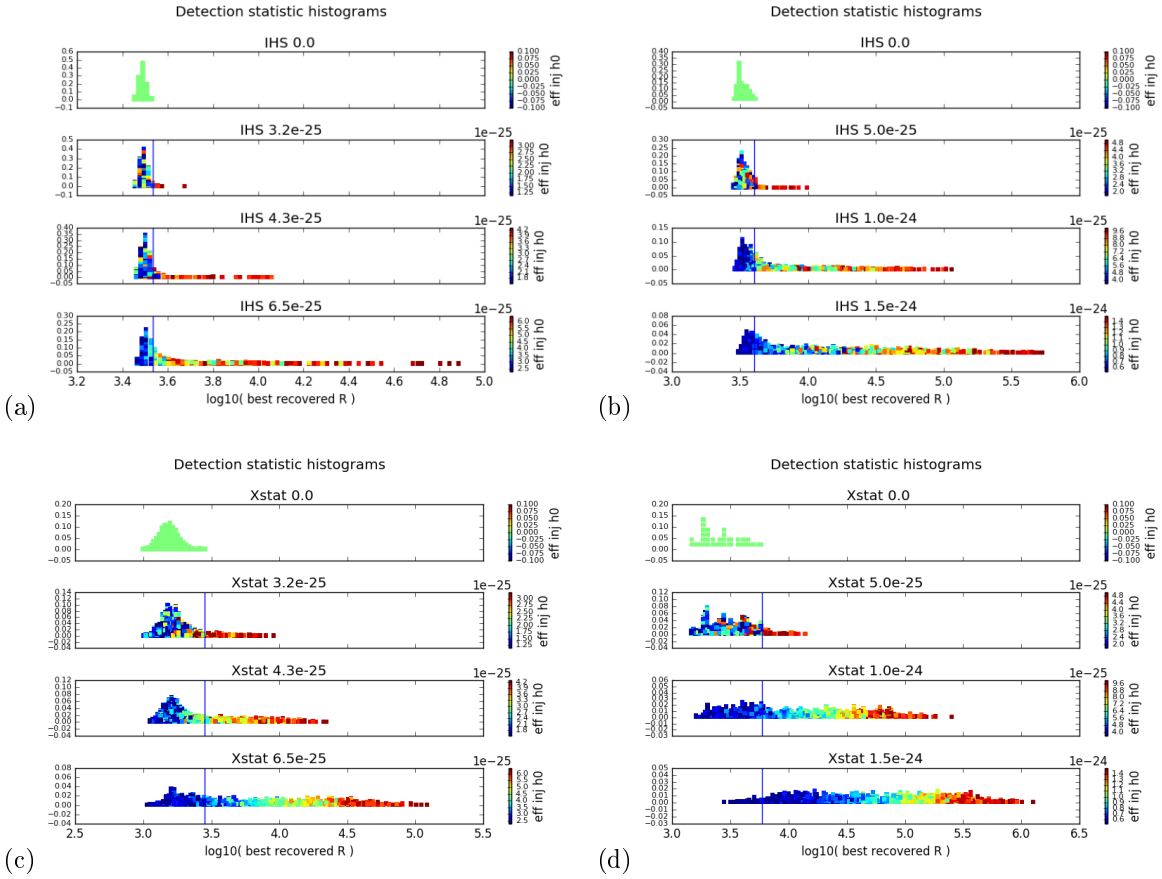


Figure 6.8: Histograms of the detection statistic for varying injected strain amplitude h_0. Sub-figure (a) shows distributions for the IHS method in Gaussian noise; (b) IHS method in O1 data; (c) X-statistic method in Gaussian noise; (d) X-statistic in O1 data. Injected signals are randomly distributed over the cosine of the orbital inclination angle, $\cos\iota$; each trial is color-coded by the resulting effective strain, $h_{0,\mathrm{eff}}$. For each plot, a threshold is plotted at the maximum value of the zero-signal distribution. Injections were randomly distributed over a 5-Hz band ranging from 102 to 107 Hz, with each 0.1 Hz sub-band analyzed separately, with modulation depths ranging up to 0.1 Hz. The length of the data used was $T_{obs} = 5 \times 10^6$, split into SFTs of length $T_{coh} = 840$s.

method. The left side of each plot, corresponding to low false positive rate, is the most relevant information, since a real search will aim to minimize false positives. Sample upper limits on $h_{0,\mathrm{eff}}$ at 1% false alarm rate may be calculated for this test data as well, although they should not be considered realistic (since thresholds must be set based on the O2 noise levels for an O2 search), and are primarily informative for comparisons between the IHS and X-statistic method. These tests show that the X-statistic performs 40-50% better than the IHS stage in terms of upper limits. As discussed earlier, however, the X-statistic method at this stage is limited by the long tail of high detection statistic outliers, especially in O1 data tests, which prevents lowering the threshold value. Outlier rejection methods, introduced in chapter 7, are expected to improve the X-statistic search sensitivity much more than they would the IHS search sensitivity.

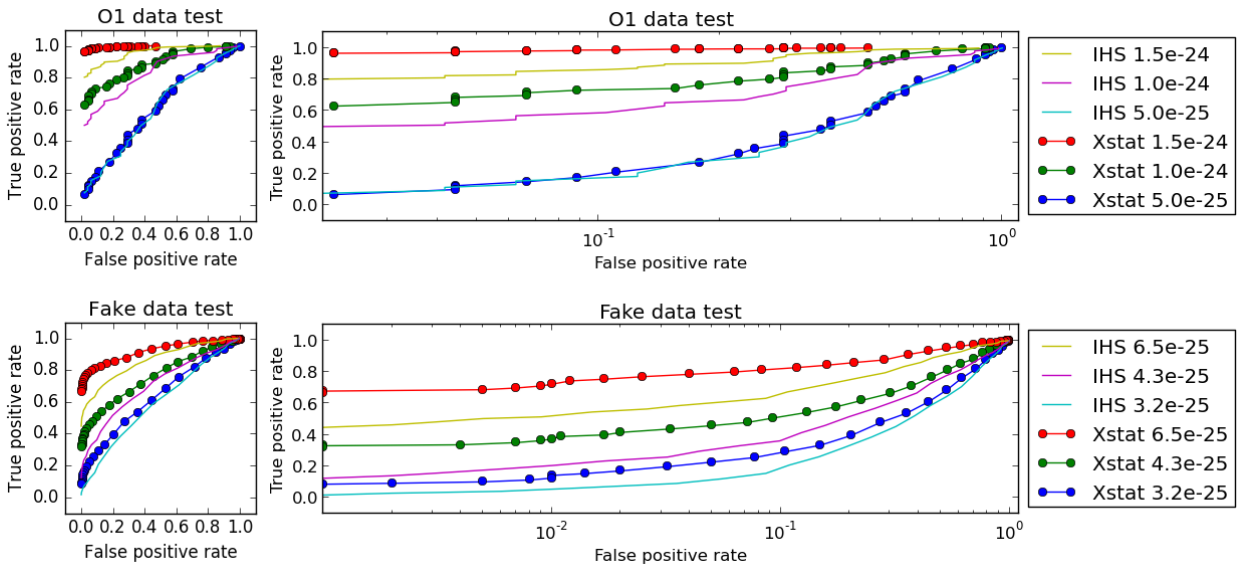


Figure 6.9: ROC curves for the X-statistic and the IHS method, linear scale (left), and log scale (right) to emphasize low false positive side of the plot. The first plot is based on signals injected into data from the first observing run of LIGO (O1), while the second plot is based on signals injected into pure Gaussian noise. The curves are labeled with the dimensionless injected strain amplitude h_0. From these data, it is apparent that the X-statistic method outperforms the IHS method, particularly at low false positive rate.

6.4.4 Signal parameter recovery with the X-statistic method

Since the X-statistic method relies on estimating the best-fit parameters of a signal, it is useful to characterize how well the algorithm recovers the injected signal parameters. In its simplest form, the error in parameter matching in the $(f_{gw}, \Delta f_{gw})$ plane can be measured as:

$$d\left(f_{\rm gw}, \Delta f_{\rm gw}\right) = \sqrt{\left(f_{\rm gw}^{\rm inj} - f_{\rm gw}^{\rm rec}\right)^2 + \left(\Delta f_{\rm gw}^{\rm inj} - \Delta f_{\rm gw}^{\rm rec}\right)^2} \tag{6.22}$$

Figure 6.11 shows an illustrative example. This test was done with a narrow (0.1 Hz) frequency band in O1 data, and at fixed $P_{\rm orb}^{\rm temp} = P_{\rm orb}^{\rm signal}$. The fixed $P_{\rm orb}$ means that this scenario measures only parameter matching for $f_{\rm gw}$ and $\Delta f_{\rm gw}$. The results, however, demonstrate some important features of the X-statistic algorithm.

First, it is clear that the algorithm functions roughly as intended. At low injected signal strengths, parameter matching is poor; no X pattern is reliably identified. As the signal strength increases, parameter matching improves. The most marked transition is around the same strength as the threshold value (derived from a comparable study of zero-signal trials). Above this signal strength, injections are reliably recovered with a high detection statistic and a low error in the signal parameters.

In figure 6.11(a), a few trials can be seen far in the upper right quadrant of the plot. These correspond to

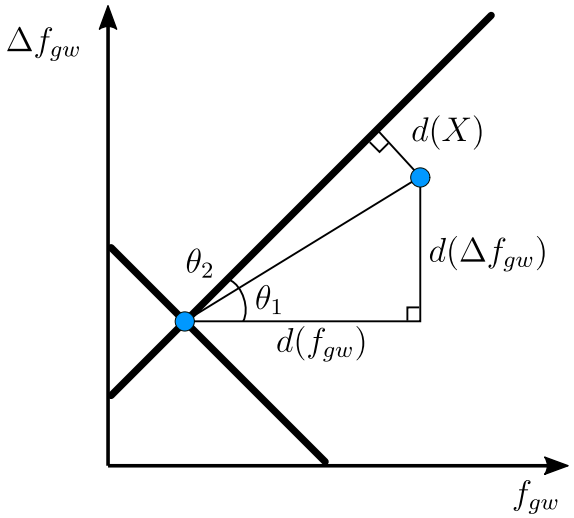


Figure 6.10: Relevant geometry for comparing a recovery point to the X-pattern created by the corresponding injected signal.

trials in which the detection statistic easily exceeds the threshold value, yet the parameter matching is poor. The cause of this behavior is illuminated by figure 6.11(b). Here, instead of measuring the parameter-space distance between recovery and injection, we measure the parameter-space distance between recovery and the X-pattern created by the injection. The relevant geometry is shown in figure 6.10. After some trigonometry, we see that

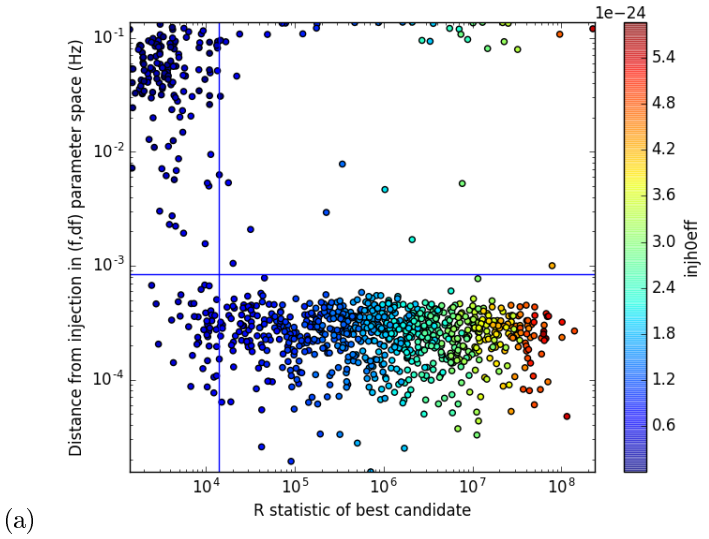
$$d(X) = \left| d\left(f_{\mathrm{gw}}\right) - d\left(\Delta f_{\mathrm{gw}}\right) \right| / \sqrt{2}\,. \tag{6.23}$$

In other words, the difference between the f_{gw} error and the Δf_{gw} error is a measure of how far from the X the recovery point is located.

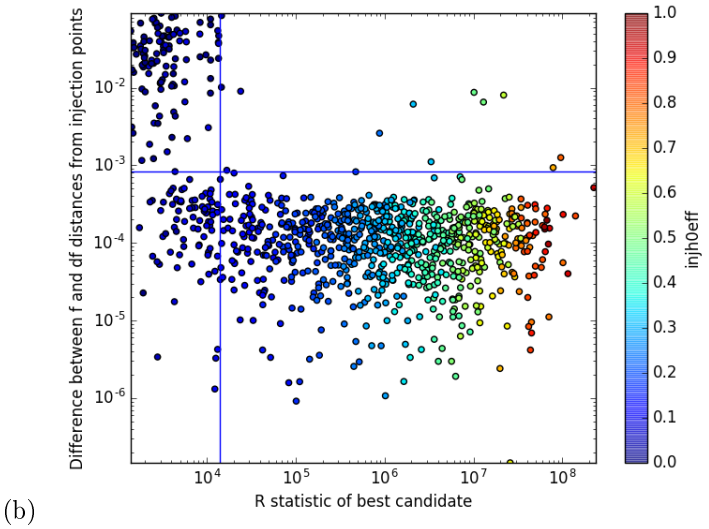
In figure 6.11(b), compared with (a), there are few points in the upper right quandrant. This indicates that the points with high detection statistic but poor parameter matching are actually located on the "arm" of the X pattern. This is generally due to one arm of the X interacting with a noise fluctuation that mimics an intersecting arm. (In the frequency-frequency plane, from which templates are calculated, this would be an interaction between the top or bottom row of the signal pattern, and noise fluctuation which creates high-power pixels in another row.)

6.4.5 Prospects for directed searches

From the initial tests of the X-statistic method described here, it is clear that the method has the potential to achieve better sensitivity than the IHS method, at a fraction of the computational cost required for a fully templated search. It is therefore well-suited for use in a directed search. More detailed information



(a)



(b)

Figure 6.11: Two plots demonstrating the agreement between injected parameters and recovered signal parameters. The horizontal axis shows the best recovered R statistic after a small templated follow-up search, which functions as the detection statistic for an X-statistic search. The vertical axis shows (a) the distance between the injected and recovered parameters in the $(f_{\rm gw}, \Delta f_{\rm gw})$ plane, and (b) the difference in the $f_{\rm gw}$ error and the $\Delta f_{\rm gw}$ error. The latter is a measurement of the distance between the recovered parameters and the X-pattern created by the injected signal. On both plots, a threshold is plotted for R corresponding to the 99th percentile (1% false alarm rate) of a comparison zero-signal distribution, and a threshold is plotted for the error which corresponds to two template spacing units.

regarding a pilot search using the X-statistic method will be presented in the next chapter.

Chapter 7

A pilot search with the X-statistic method

7.1 Overview and status

The X-statistic method was developed for application to directed searches. That is, it is intended to target points on the sky where a spinning neutron star is suspected to exist, based on electromagnetic observations, but no further details on the source are available.

The Fermi-LAT Point Source Catalog, currently on its fourth iteration (4FGL) [56], provides a list of high-energy gamma-ray sources detected by the Large Area Telescope. This list provides a promising starting point for directed continuous-wave searches. Many of the point sources in the catalog are unassociated, and could be previously unidentified spinning neutron stars.[1] A key challenge is to prioritize these sources in order of scientific interest, given the information available on them.

Previous work in the Michigan gravitational wave group [52] prioritized sources from the third iteration of the Fermi-LAT Point Source Catalog (3FGL) [55], for the purpose of a directed searches for isolated neutron stars. At the time, no computationally feasible method existed for a templated search over the relevant parameter binary parameter space. Source prioritization for an X-statistic search borrowed from this prior work, and a pilot search was started for a high-priority candidate, 3FGL J1258+2123.

After a small portion of the relevant data had been analyzed, however, the 4FGL catalog was released. In this new version the targeted source had been dropped,[2] rendering the initial data analysis obsolete, but new information was also available. Work on the pilot search was halted in favor of selecting a new target from the 4FGL catalog.

Therefore, this chapter will present the following in-progress snapshot of the pilot search. First, an

[1] The emission mechanisms for gamma-ray pulsars are not fully understood. The curved field lines of a neutron star's strong magnetic field will accelerate charged particles, leading to curvature radiation at gamma-ray energies. However, the specific locations of particle acceleration in the vicinity of a neutron star, as well as many additional mechanisms such as synchrotron radiation from electron-positron pair cascades, remain the subject of considerable study. A review of gamma ray pulsar physics can be found in [25].

[2] In addition to the larger data set used for the 4FGL catalog, the catalog construction procedure had a number of differences relative to the 3FGL procedure, as described in [56] section 3. Presumably, the source failed to meet the 4FGL criteria for catalog inclusion, although the specific reason is unclear.

overview of the original source prioritization method; second, an overview of the search technique and work flow; third, lessons from the first 10-Hz band analysis targeting the now-obsolete source; and fourth, an update to the prioritization method applied to the 4FGL catalog.

7.2 Source selection

7.2.1 Source selection from the 3FGL catalog

The following cuts were applied to select promising sources from the 3FGL (more detail on this process and its motivation can be found in [52]):

- The source must be unassociated.
- The source must *not* have the spectral type "Log Parabola". This spectral categorization is characteristic of active galactic nuclei, and not of pulsars. Pulsar spectra are also expected to be curved, but should have a power law with exponential cutoff form.
- The source must *not* have any data quality flags.
- The variability index of the source must be under 72.44. Variability is measured over long timescales; therefore, a rapidly-spinning pulsar in a steady state will actually have a low variability index. This specific threshold comes from the 3FGL catalog; sources with a variability index larger than 72.44 have less than a 1% chance of being a steady source. The variability index is calculated by comparing the likelihood under the null hypothesis (the source is constant over the observation time) and under the alternate hypothesis that the flux in each bin is optimized. See equation (4) in [48] for details.
- The spectral curvature significance σ_{curve} of the source must be 2 or greater. Note that the 3FGL threshold for switching spectral classifications is $\sigma_{\text{curve}} > 4$, so this cut effectively selects $2 < \sigma_{\text{curve}} < 4$.
- The absolute value of the source's galactic latitude must be greater than 15 degrees. This is one of the most important requirements, because the galactic latitude is used as a proxy for distance. The Milky Way is thin compared to its width; assuming that the source is located in the galaxy, a high latitude source is much more likely to be located nearby. Nearby sources are more promising for gravitational wave searches, since the observable strain amplitude h_0 would be larger. The use of galactic latitude as a proxy for distance is discussed in detail in [52].

Applying these cuts to the 3FGL catalog yields a list of 73 sources. Of these, [52] selected two for a pilot search, the first based on its small sky location uncertainty region,[3] and the second based on its particularly high galactic latitude.

For the X-statistic search, different criteria were used to prioritize among the 73 promising sources. The issue of sky location uncertainty region is less important for the X-statistc search, since TwoSpect allows a wider spacing between templates in sky location; only a single pointing would be necessary for most of the sources. Instead, the sources were ranked by gamma ray flux,[4] in the hopes of identifying a strong source of gravitational waves. Sources with high flux may be nearby, or may be highly energetic, or both. The top-priority candidate determined through this method was 3FGL J1258.4+2123 (galactic latitude: 84 degrees).

7.3 Data period

The second LIGO observing run (O2) suffered from severe spectral contamination in certain epochs, due to instrumental lines. The O2 lines list team identified a relatively "clean" epoch for each detector; the vetted line lists apply only to this epoch, and we do not recommend CW analyses outside of these periods due to the contamination. For the X-statistic pilot search, a span of 5,000,000 seconds (roughly two months) was selected within the clean epoch.

7.4 Parameter space selection

Table 7.1 shows the parameter space ranges selected for the X-statistic pilot search.

The choice of frequency minimum is driven by the LIGO sensitivity curve; seismic noise increases rapidly below this point. The choice of maximum frequency is driven by observations of pulsar spin frequencies from electromagnetic observations (keeping in mind that $f_{\mathrm{gw}} = 2f_{\mathrm{spin}}$ for the triaxial ellipsoid model of a neutron star); few pulsars spin faster than 500 Hz, and the computational cost of the search scales linearly with the selected frequency range.

The modulation depth minimum is the size of a single bin in the second Fourier transform, $1/2T_{\mathrm{coh}}$. This search targets binary systems, so a nonzero modulation depth is expected, but it can be very small. The modulation depth maximum is chosen based, again, on electromagnetic observations of pulsars. Most

[3] A small sky location uncertainty region was necessary to constrain computational cost for the technique described in [52]. The TwoSpect method is able to sample more widely in sky location, so sky location uncertainty was not an important factor in source prioritization for TwoSpect/X-statistic.

[4] 4FGL sources associated with pulsars have a wide spread in observed gamma-ray flux: about four orders of magnitude separate the brightest and dimmest sources. Observed gamma ray pulsar luminosities are also spread across four orders of magnitude [25].

Search parameter	Minimum	Maximum	ATNF binary pulsar coverage
Frequency f_{gw}	30 Hz	1000 Hz	85%
Modulation depth Δf_{gw}	0.0006 Hz	0.1 Hz	90%
Orbital period P_{orb}	0.8 days	8 days	30%

Table 7.1: Parameter space selection for the X-statistic pilot search. The last column compares the given parameter ranges with binary pulsars from the ATNF catalog. [42]

pulsars in binary systems have a modulation depth smaller than 0.1 Hz; again, computational cost scales linearly with this choice.

The orbital period minimum is driven primarily by computational cost. As noted in section 6.3.3.3, the computational cost scales inversely with $P_{\mathrm{orb}}^{\mathrm{search\,min}}$. The orbital period maximum has little effect on computational cost, but is limited by the total observation time T_{obs}; it is important that several orbits be completed within the observation time so that there is evidence of periodicity in the time-frequency plane.

The first 10-Hz test band of the pilot search, for which results are given in subsequent sections, spanned from 95 to 105 Hz in f_{gw}.

7.5 Search workflow

7.5.1 Overview

There are two parallel processes necessary for an X-statistic search, which are diagrammed in figure 7.1. The first is the search itself, which produces a list of outliers. This outlier list should be as short as possible, because of the human and computational time needed to examine an outlier in depth and determine whether it is actually scientifically interesting. The second process is a series of injection studies, which produces upper limits for the search. The injection studies must mimic the search as closely as possible, inheriting its parameter space ranges and the search threshold. Any automated outlier rejection tests must also be applied to the injections just as they are to the search outliers, so that the erroneous rejection rate is taken into account when calculating upper limits.

7.5.2 Search details

7.5.2.1 Search setup

The initial X-statistic search is performed in a series of frequency sub-bands, each 0.12 Hz wide, overlapping by 0.01 Hz. Each sub-band is handled separately. For each sub-band, a series of orbital period search points are determined iteratively following the procedure in section 6.3.3.3. At each orbital period, an X-statistic search is performed, fully covering the frequency sub-band and sampling from fixed transects over

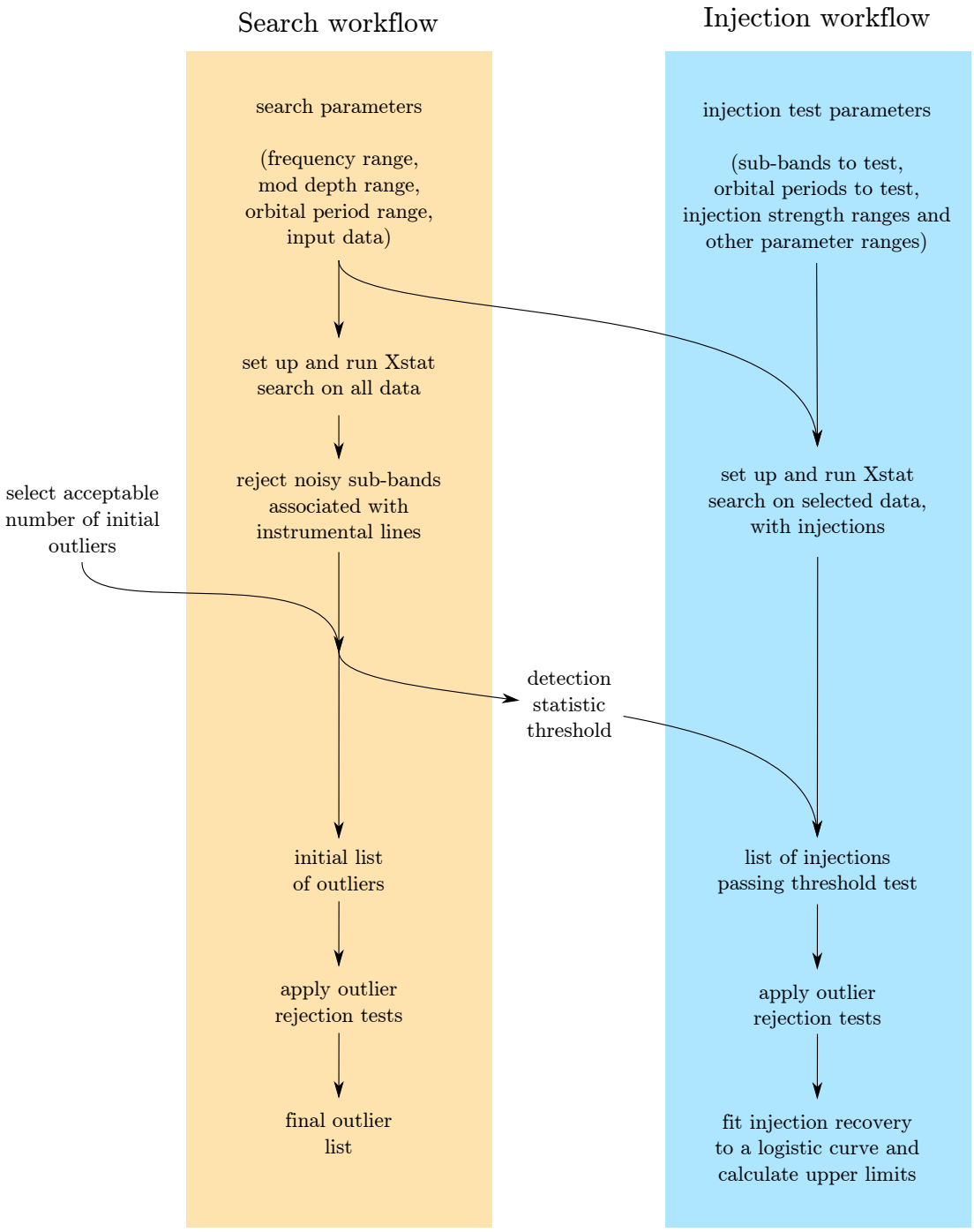


Figure 7.1: Schematic diagram of the X-statistic workflow.

the modulation depth range. More than 85,000 fixed-$P_{\rm orb}$ X-statistic searches were necessary to cover the 10-Hz test band. For each search, the best recovered R-statistic is recorded.

7.5.2.2 Search sub-band rejection

Scatter plots of the detection statistic versus frequency quickly reveal sub-bands with unusually high detection statistics. Contaminated bands are noted and excluded from all subsequent analysis; the upper limits calculated for clean bands will not apply here. For the 10-Hz test band, 5 of 100 sub-bands were rejected in this manner, spanning 99.8-100.1 Hz and 104.8-105.0 Hz. Both regions contain strong instrumental lines which are confirmed to be terrestrial in origin.[5]

7.5.2.3 Outlier clustering

A single spectral feature (whether signal or noise) is likely to result in many templates with high R statistics, both at the best-fit template point and also at neighboring points. It would be redundant and inefficient to treat all of these templates separately. To simplify the analysis, the parameter space locations for all the best recovered templates are compared, and those which are no more than four frequency bins ($2/T_{\rm coh}$) apart are grouped together. For each group, the best template (highest R) is treated as a representative of the entire group. Note that templates with very different $P_{\rm orb}$ and $\Delta f_{\rm gw}$ may be grouped together in this manner. This method is motivated by the qualitative observation of the search behavior around instrumental lines, which tend to yield many templates with high R, all grouped closely in $f_{\rm gw}$.

7.5.2.4 Rejection of outliers

After outliers have been identified and clustered, more detailed outlier-rejection tests can begin. This process aims to narrow down the list of outliers as much as possible.

7.5.2.4.1 Zero modulation depth rejection This type of rejection is implemented in the X-statistic search code, and results are reported at the end of this chapter.

In noise-only tests, the X-statistic search has a tendency to recover the best detection statistic at zero modulation depth. This fact provides a useful veto. A real signal from a binary system should have nonzero modulation depth, so outliers for which a template at $\Delta f_{\rm gw} = 0$ is better than all other sampled templates

[5] A number of comb peaks contaminate these bands. Both are contaminated by peaks of a precisely 1-Hz comb in Hanford data. The lower-frequency band also contains peaks of two combs near 11 Hz spacing (specifically 11.111175 and 11.11123 Hz), while the higher-frequency band contains a peak of a comb near 19 Hz spacing ($f_s = 19.07328$, $f_o = 9.53672$). In the Livingston data, no known combs contaminate the lower-frequency band, although there is a cluster of unidentified lines which were also seen in a number of physical and environmental monitoring channels during the run, and are likely to be terrestrial. The higher-frequency band contains a peak of a comb with spacing 0.99678913 Hz. None of these combs have a well-understood origin.

should be rejected. (Although the search parameter space is limited to include only nonzero modulation depths, the small templated follow-up from the X-statistic step is allowed to reach zero modulation depth, in part to allow this rejection method to be used.)

7.5.2.4.2 Single-detector rejection This type of rejection is largely implemented in the X-statistic code and looks promising (it may reject around 50% of outliers), but rigorous testing has not yet finished, and results are not reported here.

A real signal should be present in both detectors; both detectors should contribute to the combined SNR of the signal in the coherently summed data. That is, we expect that $R_{\mathrm{H1}} < R_{\mathrm{H1,L1}}$ and also that $R_{\mathrm{L1}} < R_{\mathrm{H1,L1}}$. If one of these is not true, then the outlier is most likely due to an artifact present in a single detector, and should be rejected. This rejection method requires re-running single-detector searches at the parameter-space point of interest, but adds only a small amount to the over all computational cost of the search, because a single template (or a very small templated search) may be used.

7.5.2.4.3 Sky location offset rejection This type of rejection is experimental; further work is needed to characterize its behavior.

A real signal should emanate from the sky location being studied; therefore, a template for that precise sky location should yield a better detection statistic than a template for a different sky location. As with the single-detector rejection method, this requires re-running a single template or small templated seach at the parameter-space point of interest, this time with an offset sky location.

Noise artifacts, however, may also result in templates with high detection statistics off the sky location of interest, which complicates the use of this rejection method. Future work will need to use injection studies in real data, to determine the optimal amount of sky location offset and the impact of this method on sensitivity depth.

7.5.3 Injection study details

7.5.3.1 Injection parameters

Injection studies are X-statistic searches that span selected sub-bands, and at fixed orbital period, with fake signals injected into the data to test recovery. Injected signals are randomly distributed across the same parameter space range as the search, albeit limited to the appropriate sub-band in frequency, and to an appropriate range around the tested orbital period, $P_{\mathrm{orb}}^{\mathrm{temp}}$. For the 10-Hz test band, two clean sub-bands were selected, spanning 103.3-103.4 Hz, and 97.2-97.3 Hz. Since in a real search templates are spaced in

Decreasing detection statistic threshold ↓	$D^{95\%}$ for f_{gw}= 103.3-103.4 Hz, $P_{orb} = 0.8$ days	$D^{95\%}$ for f_{gw}= 97.2-97.3 Hz, $P_{orb} = 0.8$ days	$D^{95\%}$ for f_{gw}= 103.3-103.4 Hz, $P_{orb} = 8.0$ days	$D^{95\%}$ for f_{gw}= 97.2-97.3 Hz, $P_{orb} = 8.0$ days	Number of outliers (% of total)	Number of outliers after clustering (% of total)	Number of outliers after rejecting $\Delta f_{gw} = 0$ (% of total)
	1.7	1.7	2.6	2.8	87 (0.1)	61 (<0.1)	8 (<0.1)
	2.6	2.6	4.3	4.2	87 (0.1)	60 (<0.1)	8 (<0.1)
	3.0	3.0	5.1	4.9	88 (0.1)	61 (<0.1)	8 (<0.1)
	4.9	4.8	7.6	8.4	405 (0.5)	73 (<0.1)	20 (<0.1)
	6.0	5.4	(no data)	10.2	5662 (6.7)	241 (0.3)	183 (0.22)

Table 7.2: Sensitivity depths and outliers for a 10-Hz test band, in order of decreasing threshold value.

orbital period by some ΔP_{orb} (P_{orb}), injected signals are distributed over the range $P_{orb}^{temp} - \Delta P_{orb}^{temp}/2$ to $P_{orb}^{temp} + P_{orb}^{temp}/2$.

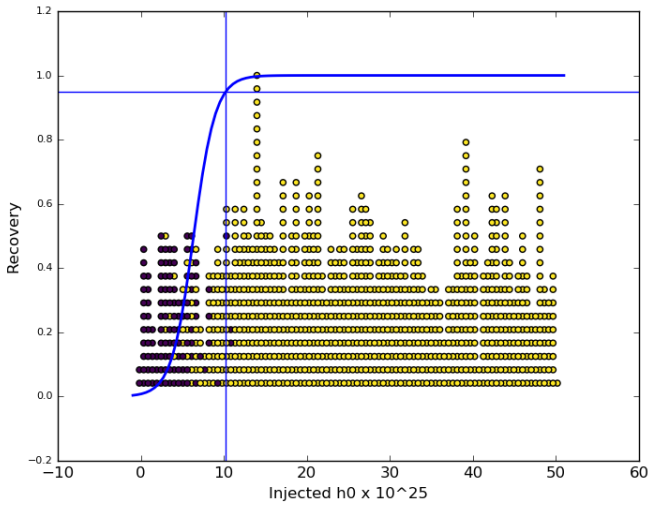
7.5.3.2 Rejections

The same rejection criteria are applied to the injections as are applied to the clustered outliers in the real search. Injections that fail to pass the criteria in section 7.5.2.4 are considered "not recovered" even if they pass the detection statistic threshold. This requirement ensures that the outlier rejection is not overzealous; if the rejection criteria also reject fake signals, the upper limits for the search will be impacted.

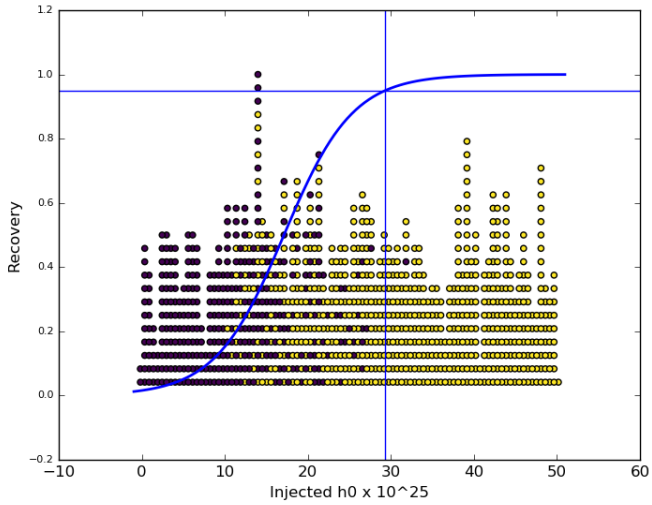
7.6 Lessons from a 10-Hz band in O2 data

The 10-Hz band study described in this chapter functions as a test of the X-statistic method, for use in a subsequent search targeting a sky location of greater scientific interest. One of its key functions is to demonstrate how the sensitivity depth of the search will scale with the number of remaining outliers after clustering and automated rejections. Outliers that survive the automated rejection process will have to be followed up manually (unless additional automated tests are developed), so they represent a significant burden in terms of time investment. On the other hand, allowing for more outliers will also provide a more sensitive search, since the threshold value can be lowered. Table 7.2 demonstrates this scaling, while figure 7.2 shows an example of how upper limits are calculated from injection studies.

Other key points from table 7.2 include the scaling of the sensitivity depth as a function of frequency and orbital period. Sensitivity depths for the two tested sub-bands are comparable; this meets the expectation that sensitivity depth calculations in one clean sub-band can be reasonably applied to other nearby clean-sub-bands, and will allow a full search to minimize the number of injection studies, thereby minimizing computational cost. Second, the sensitivity depths at the maximum of the orbital period range ($P_{orb} = 8$



(a)



(b)

Figure 7.2: Example plots showing how the recovery of injected signals varies with changing threshold. Both plots show data from the same set of injections, with f_{gw} from 103.3-103.4 Hz, searched at $P_{\mathrm{orb}} = 8$ days. The blue and yellow plot shows a histogram of the injected strain amplitude h_0, color-coded by whether each signal was recovered (yellow) or not (blue). This is overlaid with the fitted logistic curve, a horizontal line showing 95% recovery, and a vertical line showing the 95% upper limit.

days), are about 1.6 times better than those at the minimum of the orbital period range ($P_{\rm orb} = 0.8$ days). The best sensitivity depths will likely be achieved for long orbital periods in a full search.

Finally, it should be noted that these sensitivity depths and remaining outliers correspond to just one type of implemented outlier rejection method. Future work will likely improve these results with further automated outlier rejection, which will allow lower detection statistic thresholds to be set.

7.7 Source selection revisited

7.7.1 3FGL cuts in light of the 4FGL catalog

The cuts applied to the 3FGL catalog were undertaken with the intent to identify likely nearby neutron stars among the high-latitude unassociated sources. Information from the 4FGL catalog can be used to revisit the 3FGL cuts and understand what worked well, and what should be amended for future prioritization.

Between the 3FGL and 4FGL releases, a number of sources were newly associated with pulsars. Ideally, the cuts applied to 3FGL should predict which of the sources where likey to be neutron stars. In practice, this was not immediately seen to be the case, and some investigation was required to determine why.

The 3FGL catalog uses a cutoff on $\sigma_{\rm curve}$ to classify spectral type. All sources with $\sigma_{\rm curve} > 4$ were classified as log parabola, which, as noted earlier, is characteristic of AGNs rather than pulsars. However, many of the sources that were later identified as pulsars had $\sigma_{\rm curve} > 4$. This development indicates that the spectral classification based on $\sigma_{\rm curve}$ may not be a reliable way of identifying probable pulsars. Further support is given to this hypothesis by new information available in the 4FGL catalog. Rather than a generic $\sigma_{\rm curve}$, the 4FGL catalog provides two quantities: $\sigma_{\rm curve}^{\rm PLEC}$ and $\sigma_{\rm curve}^{\rm LP}$. The former describes the significance of curvature for a power law with exponential cutoff (PLEC) type spectrum, while the later describes the significance of curvature for a log parabola (LP) type spectrum. In the 4FGL catalog, it is apparent that many highly-curved spectra have *both* a high $\sigma_{\rm curve}^{\rm PLEC}$ and a high $\sigma_{\rm curve}^{\rm LP}$. In fact, because spectral classification is still done using a $\sigma_{\rm curve}$ cutoff in 4FGL (now using $\sigma_{\rm curve}^{\rm LP}$), there are a number of sources classified as "log parabola" for which $\sigma_{\rm curve}^{\rm PLEC} > \sigma_{\rm curve}^{\rm LP}$. It is likely that in 3FGL, many sources that were classified as having log parabola type spectra also had a high significance of PLEC-type curvature.

In light of this new information, it is worthwhile to see what would have happened if the spectral type requirement had been dropped from the original cuts. Table 7.3 shows a comparison between the general population of high-latitude (galactic latitude above 15 degrees or below -15 degrees) unassociated 3FGL sources, and the sub-population of those sources which passes the original and modified cuts.

This analysis suggests that the original spectral cuts worked well, with the exception of the cut on spectral

Population	Total number of sources	Number of sources associated between 3FGL and 4FGL	Number of sources associated with pulsars between 3FGL and 4FGL	Ratio of pulsar associations to total associations	Ratio of pulsar associations to total population
All unassociated high-latitude 3FGL sources	480	142	11	8%	2%
Unassociated high-latitude 3FGL sources passing original cuts	73	22	1	5%	1%
Unassociated high-latitude 3FGL sources passing modified cuts (spectral type requirement dropped)	96	31	8	26%	8%

Table 7.3: This table shows the rate of high-latitude unassociated 3FGL sources being associated with pulsars in 4FGL, for various populations.

classification. It motivates dropping the spectral classification requirement in the future, but otherwise repeating the same technique for 4FGL source prioritization.

7.7.2 New source selection from the 4FGL catalog

A number of differences exist between the 3FGL and 4FGL catalog. The following two differences are most notable for this application.

First, as discussed in the previous section, the 4FGL provides more detailed information about spectral curvature. Where the 3FGL catalog reported a generic significance of curvature $\sigma_{\rm curve}$, the 4FGL is more specific, reporting a significance for log parabola spectral shape, $\sigma_{\rm curve}^{\rm LP}$, and for a power law with exponential cutoff, $\sigma_{\rm curve}^{\rm PLEC}$. Therefore the $\sigma_{\rm curve}$ cut should logically be replaced by a $\sigma_{\rm curve}^{\rm PLEC}$ cut. Figure 7.3 shows the effect of this cut on the 4FGL catalog.

Second, the 4FGL catalog is currently a preliminary version, and does not yet contain the variability index for any of the sources. This information will be included in the final release of the catalog in a few months. In the meantime, the variability indices from the 3FGL must be used for 4FGL sources. This, unfortunately, limits the current 4FGL analysis to considering sources that are also associated with a 3FGL entry.

In light of these constraints, the following cuts are applied to the 4FGL catalog:

- The source must be unassociated.
- The source must *not* have any data quality flags.

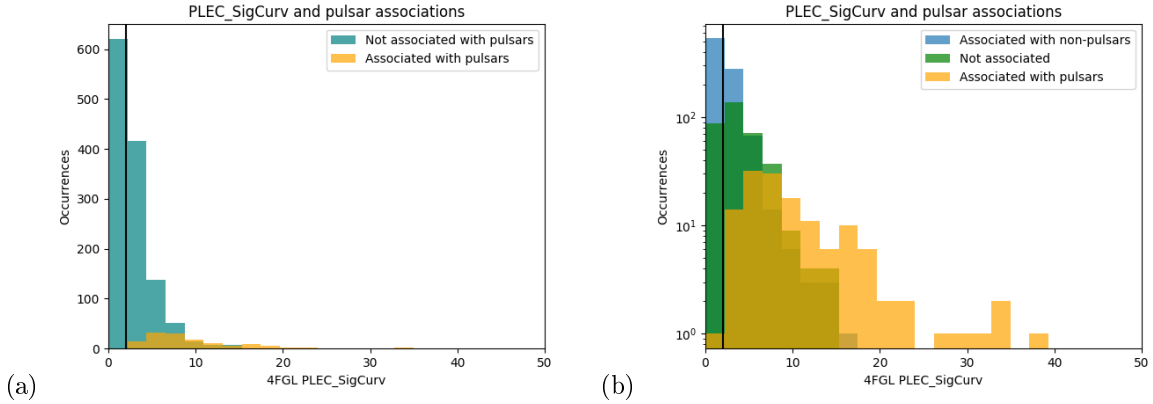


Figure 7.3: Figures showing the distribution of $\sigma_{\text{curve}}^{\text{PLEC}}$ in the 4FGL catalog (all sources, not limited to high latitude). Both plots are made from the same data; (b) splits the non-pulsar population into those which are unassociated, and those which have been associated with something other than a pulsar. In both plots, a vertical line at $\sigma_{\text{curve}}^{\text{PLEC}} = 2$ shows the cutoff used for source prioritization. (A very small number of sources have $\sigma_{\text{curve}}^{\text{PLEC}} > 50$; they are excluded from the plot for readability.)

- The source must be associated with an entry in the 3FGL, for which the variability index is under 72.44.
- The significance of PLEC-type spectral curvature, $\sigma_{\text{curve}}^{\text{PLEC}}$, for the source must be 2 or greater.
- The significance of PLEC-type spectral curvature, $\sigma_{\text{curve}}^{\text{PLEC}}$, for the source must be greater than the significance of LP-type spectral curvature, $\sigma_{\text{curve}}^{\text{LP}}$.
- The absolute value of the source's galactic latitude must be greater than 15 degrees.

After applying these cuts, 61 unassociated 4FGL sources survive. We are interested in prioritizing sources at very high galactic latitude (under the assumption that they are more likely to be nearby), and those with a high gamma ray flux, which may be promising gravitational wave sources. Figure 7.4 shows the available options.

A source of particular interest, 4FGL J1225.9+2951, is annotated in figure 7.4. The very high galactic latitude, and relatively high flux, are suggestive of an energetic, nearby system. This object is planned to replace 3FGL J1258.4+2123 as the highest-priority target for the X-statistic pilot search.

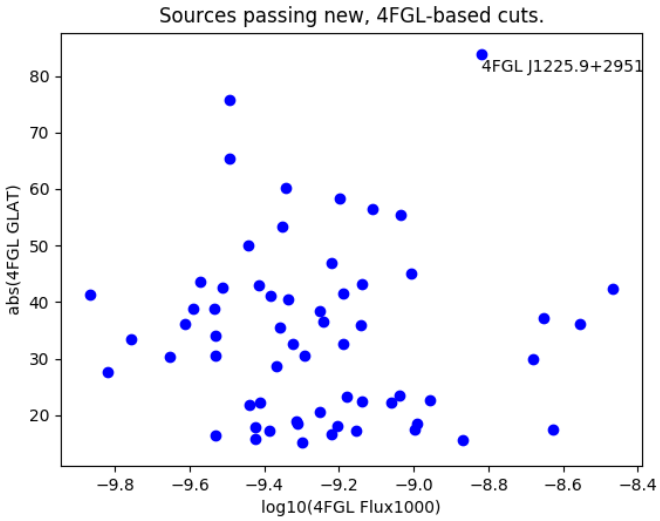


Figure 7.4: High-latitude unassociated sources which survive the modified cuts on the 4FGL catalog. The annotation marks a particularly promising source.

Chapter 8

Conclusions and future outlook

This work has described the development and status of two programs for continuous wave data analysis and spectral monitoring: the X-statistic method and the FineTooth tools.

LIGO has recently begun its third observing run (O3), with both the Hanford and Livingston detectors operating at unprecedented sensitivity. In addition to the cornucopia of compact binary coalescence signals which are expected to be observed, O3 represents the best opportunity yet for the detection of gravitational waves from spinning neutron stars. Meanwhile, electromagnetic observations, which are used to guide continuous-wave searches, are also improving.

8.1 FineTooth in O3 and beyond

In order to prevent contamination of O3's high-quality data by narrow spectral artifacts, continual line and comb monitoring is necessary. As noted in chapter 7, O2 saw an unfortunate amount of line contamination, which rendered long periods of the data taken unfit for continuous wave searches. This outcome must be avoided in O3. To this end, the FineTooth tools are already running daily, with results posted on internal collaboration web pages and made available to experts on and off the detector sites. Links to the comb tracking summary pages have recently been posted on the main detector summary pages, so they can be easily found by all members of the collaboration. These tools will soon be incorporated into regular data quality monitoring tasks; documentation on their use has already been written and provided to data quality shift coordinators. The author continues to act as the sole maintainer and developer of FineTooth, and as a contact person for its use in comb investigations.

A number of additions to the comb summary pages have been requested by line experts in the collaboration, including improved high-frequency spectral monitoring and the ability to incorporate data from other monitoring tools; these changes are currently being implemented. FineTooth is tracking a number of combs which will eventually be included in the O3 lines list, and its plotting tools have been used to identify and alert staff on site to several noise artifacts in need of mitigation.

8.2 X-Statistic pilot search continuation

The X-statistic pilot search has dealt with a number of challenges, some of them inherent to the method (the trade off between outliers and sensitivity) and others the result of external circumstance (the Fermi-LAT catalog update). Important information has been gained from all these challenges. The next stage of the X-statistic pilot search will involve re-running the 10 Hz test band using the newly identified top-priority source from the fourth Fermi-LAT catalog, before moving on to the rest of the search frequency range. Several methods of automated outlier rejection remain to rigorously investigated, but early results are promising. New tools for outlier investigation are also being developed in the Michigan group, such as a strain histogramming tool [63] which will soon be extended to work for binary systems as well. These developments may speed up the X-Statistic outlier analysis, and allow for lower thresholds to be set, thereby achieving better sensitivity. Meanwhile, when the new Fermi-LAT catalog is updated to include information on source variability, the current source prioritization scheme can be expanded to those sources which are new in 4FGL. This may allow the identification of further promising targets for a directed binary search.

In the further future, improvements in data quality between O2 and O3 will open the door to more senstive X-Statistic searches, partially due to lower noise, and partially due to the availability of longer periods of clean data. It is also notable that TwoSpect is not alone in having a characteristic X-like pattern in the space of templates surrounding a signal; for example, the CrossCorr pipeline [19] has a similar feature [44]. Continuing X-statistic investigations with TwoSpect may eventually pave the way for sparse-sampling methods built on other data analysis techniques.

8.3 Summary and scientific impact

The first detection of a continuous gravitational wave remains elusive. When it happens, it will considerably enrich the fields of neutron star astronomy and gravitational waves. Gravitational waves are poised to complement existing electromagnetic observations of neutron stars, providing new information about the internal dynamics, population statistics, and life cycles of these stars. However, the road ahead requires a great deal of work: data quality must be carefully monitored to catch and prevent needless spectral contamination, while search algorithms must continue to improve and explore new techniques for accessing as much of the source parameter space as possible, under the constraint of computational time available. Continued attention to electromagnetic observations is also necessary to prioritize search targets and parameter space ranges. Both FineTooth and the X-Statistic pilot search represent steps along this road.

As for the length of the journey to a first continuous wave detection, and what exactly lies at its end:

only nature and time will tell.

Bibliography

[1] aLIGO LHO Logbook. https://alog.ligo-wa.caltech.edu/aLOG/.

[2] aLIGO LHO Logbook: Comb updates from recent data. https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=30047.

[3] aLIGO LHO Logbook: End station CPS timing fan outs running on separate power, BSCs re-isolated. https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=28915.

[4] aLIGO LHO Logbook: Turned off Pcal camera ethernet adapter – maybe mitigates a 1 Hz comb? https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=35640.

[5] Gracedb: Gravitational wave candidate event database: Latest. https://gracedb.ligo.org/latest/, accessed May 17, 2019.

[6] GW150914: The first direct detection of gravitational waves. https://www.ligo.org/detections/GW150914.php.

[7] Introduction to LIGO and gravitational waves: sources of gravitational waves. https://www.ligo.org/science/GW-Sources.php.

[8] O1 Instrumental Lines. https://www.gw-openscience.org/o1speclines/.

[9] O2 Instrumental Lines. https://www.gw-openscience.org/o2speclines/.

[10] PEM Central. http://pem.ligo.org. Accessed: 2019-02-04.

[11] Updated ATNF catalog. http://www.atnf.csiro.au/research/pulsar/psrcat, accessed February 11, 2019.

[12] C. Bond, D. Brown, A. Freise, and K. A. Strain. Interferometer techniques for gravitational-wave detection. *Living Reviews in Relativity*, 19(1), dec 2016.

[13] A. F. Brooks, B. Abbott, M. A. Arain, G. Ciani, A. Cole, G. Grabeel, E. Gustafson, C. Guido, M. Heintze, A. Heptonstall, M. Jacobson, W. Kim, E. King, A. Lynch, S. O'Connor, D. Ottaway, K. Mailand, G. Mueller, J. Munch, V. Sannibale, Z. Shao, M. Smith, P. Veitch, T. Vo, C. Vorvick, and P. Willems. Overview of advanced LIGO adaptive optics. *Applied Optics*, 55(29):8256, oct 2016.

[14] S. M. Carroll. Lecture notes on general relativity, 1997.

[15] C.-M. Chen, J. M. Nester, and W.-T. Ni. A brief history of gravitational wave research. *Chinese Journal of Physics*, 55(1):142 – 169, 2017.

[16] T. L. S. Collaboration, the Virgo Collaboration, and additional authors. Searches for gravitational waves from known pulsars at two harmonics in 2015-2017 ligo data, 2019.

[17] P. Covas et al. Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of advanced LIGO. *Physical Review D*, 97(8), apr 2018.

[18] J. D. E. Creighton and W. G. Anderson. *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis*. Wiley-VCH, 2011.

[19] S. Dhurandhar, B. Krishnan, H. Mukhopadhyay, and J. T. Whelan. Cross-correlation search for periodic gravitational waves. *Phys. Rev. D*, 77:082001, Apr 2008.

[20] L. W. Esposito and E. R. Harrison. Properties of the hulse-taylor binary pulsar system. *The Astrophysical Journal*, 196:L1, feb 1975.

[21] E. Goetz. Template placement. `https://wiki.ligo.org/pub/CW/TwoSpect/templatePlacement.pdf`, 2014.

[22] E. Goetz and K. Riles. An all-sky search algorithm for continuous gravitational waves from spinning neutron stars in binary systems. *Classical and Quantum Gravity*, 28(21):215006, Sept. 2011.

[23] E. Goetz and K. Riles. Coherently combining data between detectors for all-sky semi-coherent continuous gravitational wave searches. *Classical and Quantum Gravity*, 33(8):085007, Mar. 2016.

[24] E. A. Goetz. *Gravitational Wave Studies: Detector Calibration and an All-Sky Search for Spinning Neutron Stars in Binary Systems*. PhD thesis, University of Michigan, 2010. `http://hdl.handle.net/2027.42/78838`.

[25] I. A. Grenier and A. K. Harding. Gamma-ray pulsars: A gold mine. *Comptes Rendus Physique*, 16(6-7):641–660, Aug. 2015.

[26] J. B. Hartle. *Gravity: An Introduction to Einstein's General Relativity.* Pearson Education, 2003.

[27] M. Hendry. An introduction to general relativity, gravitational waves, and detection principles. 2007.

[28] C. J. Horowitz and K. Kadau. Breaking strain of neutron star crust and gravitational waves. *Physical Review Letters*, 102(19), May 2009.

[29] D. Kennefick. *Traveling at the Speed of Thought: Einstein and the Quest for Gravitational Waves.* Princeton University Press, 2016.

[30] K. D. Kokkotas. Gravitational wave physics. 2002.

[31] M. Kramer, K. M. Xilouris, D. R. Lorimer, O. Doroshenko, A. Jessner, R. Wielebinski, A. Wolszczan, and F. Camilo. The characteristics of millisecond pulsar emission. i. spectra, pulse shapes, and the beaming fraction. *The Astrophysical Journal*, 501(1):270–285, July 1998.

[32] C. Lammerzahl and V. Perlick. Gravitational waves. `https://www.zarm.uni-bremen.de/fileadmin/user_upload/space_science/gravitational_theory/gravwave.pdf`, 2014.

[33] P. D. Lasky. Gravitational waves from neutron stars: A review. 2015.

[34] J. M. Lattimer and M. Prakash. The equation of state of hot, dense matter and neutron stars. *Physics Reports*, 621:127–164, Mar. 2016.

[35] LIGO Scientific Collaboration. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, 2015.

[36] LIGO Scientific Collaboration, Virgo Collaboration. First all-sky search for continuous gravitational waves from unknown sources in binary systems. *Physical Review D*, 90(6), sep 2014.

[37] LIGO Scientific Collaboration, Virgo Collaboration. All-sky search for periodic gravitational waves in the o1 LIGO data. *Physical Review D*, 96(6), sep 2017.

[38] LIGO Scientific Collaboration, Virgo Collaboration. First narrow-band search for continuous gravitational waves from known pulsars in advanced detector data. *Physical Review D*, 96(12), dec 2017.

[39] LIGO Scientific Collaboration, Virgo Collaboration. Upper limits on gravitational waves from scorpius x-1 from a model-based cross-correlation search in advanced LIGO data. *The Astrophysical Journal*, 847(1):47, sep 2017.

[40] D. R. Lorimer. Binary and millisecond pulsars. *Living Reviews in Relativity*, 1(1), sep 1998.

[41] L. C. Loveridge. Physical and geometric interpretations of the riemann tensor, ricci tensor, and scalar curvature. 2004.

[42] R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs. The Australia Telescope National Facility Pulsar Catalogue. *The Astronomical Journal*, 129(4):1993–2006, apr 2005.

[43] K. T. McDonald. What is the stiffness of spacetime?

[44] G. Meadors. Private communication, 2016.

[45] G. Meadors, E. Goetz, K. Riles, T. Creighton, and F. Robinet. Searches for continuous gravitational waves from scorpius x-1 and XTE j1751-305 in LIGO's sixth science run. *Physical Review D*, 95(4), Feb. 2017.

[46] J. Mizuno, K. Strain, P. Nelson, J. Chen, R. Schilling, A. Rüdiger, W. Winkler, and K. Danzmann. Resonant sideband extraction: a new configuration for interferometric gravitational wave detectors. *Physics Letters A*, 175(5):273–276, apr 1993.

[47] C. Moore, R. Cole, and C. Berry. Gravitational wave detectors and sources. `http://gwplotter.com/`. Accessed: 2019-02-04.

[48] P. L. Nolan et al. FERMILARGE AREA TELESCOPE SECOND SOURCE CATALOG. *The Astrophysical Journal Supplement Series*, 199(2):31, Mar. 2012.

[49] R. Prix. Gravitational waves from spinning neutron stars. In W. Becker, editor, *Neutron Stars and Pulsars*, page 651. Springer-Verlag, 2009.

[50] K. Riles. Gravitational waves: Sources, detectors and searches. 2012.

[51] K. Riles. Recent searches for continuous gravitational waves. *Modern Physics Letters A*, 32(39):1730035, dec 2017.

[52] J. R. Sanders. *Advanced Gravitational Wave Detectors and Detection: Arm Length Stabilization and Directed Searches for Isolated Neutron Stars.* PhD thesis, University of Michigan, 2016. `http://hdl.handle.net/2027.42/120826`.

[53] P. R. Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors (Second Edition).* World Scientific Pub Co Inc, feb 2017.

[54] B. F. Schutz and F. Ricci. Gravitational waves, sources, and detectors. 2010.

[55] The Fermi-LAT collaboration. Fermi Large Area Telescope Third Source Catalog. *The Astrophysical Journal Supplement Series*, 218(2):23, June 2015.

[56] The Fermi-LAT collaboration. Fermi Large Area Telescope Fourth Source Catalog. *arXiv e-prints*, Feb. 2019.

[57] The LIGO Scientific Collaboration. GW170817: Measurements of neutron star radii and equation of state. *Physical Review Letters*, 121(16), Oct. 2018.

[58] The LIGO Scientific Collaboration and the Virgo Collaboration. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs, 2018.

[59] The LIGO Scientific Collaboration and the Virgo Collaboration. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs, 2018.

[60] The LIGO Scientific Collaboration and Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116(6), Feb. 2016.

[61] S. Trembath-Reichert. Automated processing and display of O1 LIGO data for the purpose of CW gravitational wave searches and analysis. April 2016.

[62] J. M. Weisberg and J. H. Taylor. Relativistic binary pulsar b1913+16: Thirty years of observations and analysis, 2004.

[63] G. Weldon and K. Riles. Strain histograms for evaluating continuous gravitational wave candidates. APS April Meeting 2019, Session Y10, 2019.