Detectability of Dynamical Tidal Effects and the Detection of Gravitational-Wave Transients with LIGO

by

Reed Clasey Essick

B.S., Mechanical Engineering
Washington University in St. Louis, 2011

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Author . . . . Signature redacted
Department of Physics
July 19, 2017

Certified by . . . . Signature redacted
Dr. Erotokritos Katsavounidis
Senior Research Scientist
Thesis Supervisor

Certified by . . . . Signature redacted
Prof. Nevin N. Weinberg
Associate Professor of Physics
Thesis Supervisor

Accepted by . . . . Signature redacted
Prof. Nergis Mavalvala
Curtis and Kathleen Marble Professor of Astrophysics
Associate Head

Detectability of Dynamical Tidal Effects and the Detection of Gravitational-Wave Transients with LIGO

by

Reed Clasey Essick

Submitted to the Department of Physics
on July 19, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Dynamical tidal effects impact the orbital motion of extended bodies, imprinting themselves in several measurable ways. This thesis explores the saturation of weakly nonlinear dynamical tidal interactions within two very different systems: hot Jupiters orbiting main-sequence hosts with radiative cores and compact stellar remnants inspiraling due to gravitational radiation. In addition, it discusses general aspects of detecting Gravitational Waves with ground-based laser interferometers. Data quality and noise reduction along with source parameter estimation, with particular emphasis on localization, are discussed in great detail. Conclusions drawn from statistical ensembles of simulated signals are applied to the first three confirmed detections of Gravitational Waves, all from the coalescence of binary black hole systems.

Thesis Supervisor: Dr. Erotokritos Katsavounidis
Title: Senior Research Scientist

Thesis Supervisor: Prof. Nevin N. Weinberg
Title: Associate Professor of Physics

Acknowledgements

I sincerely thank both my advisers, Erik and Nevin, for their support and guidance throughout my graduate studies. I am also indebted to the rest of my committee for their useful feedback and to my colleagues within the LIGO Lab, the LIGO-Virgo Collaboration, and the MIT Kavli Institute for their companionship and the opportunities working with them has provided.

I would be remiss if I did not also thank my parents, Ray and Jeanna Essick, for their unwavering support throughout my entire life. My friends have also been a welcome refuge during my studies, and in particular I would like to thank the Compact Objects for reminding me of life's true priorities, foremost of which is summer softball.

Portions of this work do not reflect the scientific opinion of the LIGO-Virgo Collaboration and were not reviewed by the collaboration.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the most studied systems in physics is that of two bodies interacting gravitationally. Simple models of these systems in Newtonian gravity are taught to first year students and more complex models are applied to cutting-edge research throughout astronomy and astrophysics [34, 88, 89, 96, 123, 154, 171]. Throughout its long history, the binary problem has always posed a rich opportunity to confront theory with experiment, and I present several such opportunities in this thesis, focusing primarily on the impact of nonlinear tidal dynamics and the general Gravitational Wave (GW) detection problem.

Dynamical tidal effects impact the orbital motion of extended bodies, imprinting themselves in several measurable ways. This thesis explores the saturation of weakly nonlinear dynamical tidal interactions within two very different systems: hot Jupiters orbiting main-sequence hosts with radiative cores and compact stellar remnants inspiraling due to gravitational radiation. In addition, it discusses general aspects of detecting GWs with ground-based laser interferometers. Data quality and noise reduction along with source parameter estimation, with particular emphasis on localization, are discussed in great detail. Conclusions drawn from statistical ensembles of simulated signals are applied to the first three confirmed detections of GWs, all from the coalescence of binary black hole systems.

The remainder of this chapter will review the preliminaries of tidal interactions, binary systems in general, GW sources and their detection within laser interferometers. §2 then discusses nonlinear tidal interactions within both (§2.1) hot Jupiter exoplanetary systems and (§2.2) binary neutron star

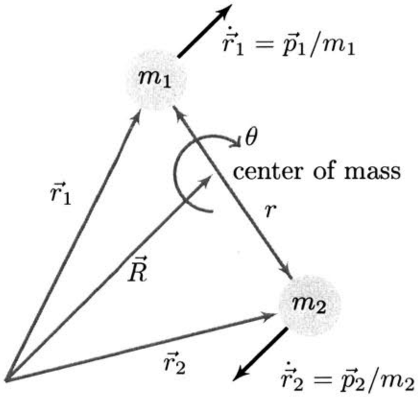


Figure 1-1: Schematic of a general binary, showing multiple sets of coordinates and their relations. We mostly work with the relative separation (r) and angle around the center of mass (θ) of the two components.

coalescences. §3 discusses general aspects of the GW detection problem, decomposing it into (§3.1) trigger generation, (§3.2) data quality and vetting, and (§3.3) parameter estimation with particular emphasis on source localization. These studies are then applied to confirmed detections in §4, namely GW150914 [146], GW151226 [145], GW170104 [151], and the less significant LVT151012 [7].

1.1 Binaries Containing Point Particles

We begin by considering a simple model of two point particles. If these particles interact through an instantaneous central-force potential (Newtonian gravity), this is often called the Kepler problem. The Newtonian solution's breakdown in observed binaries provided some of the first tests of GR [110, 173], and the fully relativistic analog provided the first directly observed GWs [7, 145, 146, 151].

1.1.1 The Kepler Problem

The Kepler problem is most commonly written as

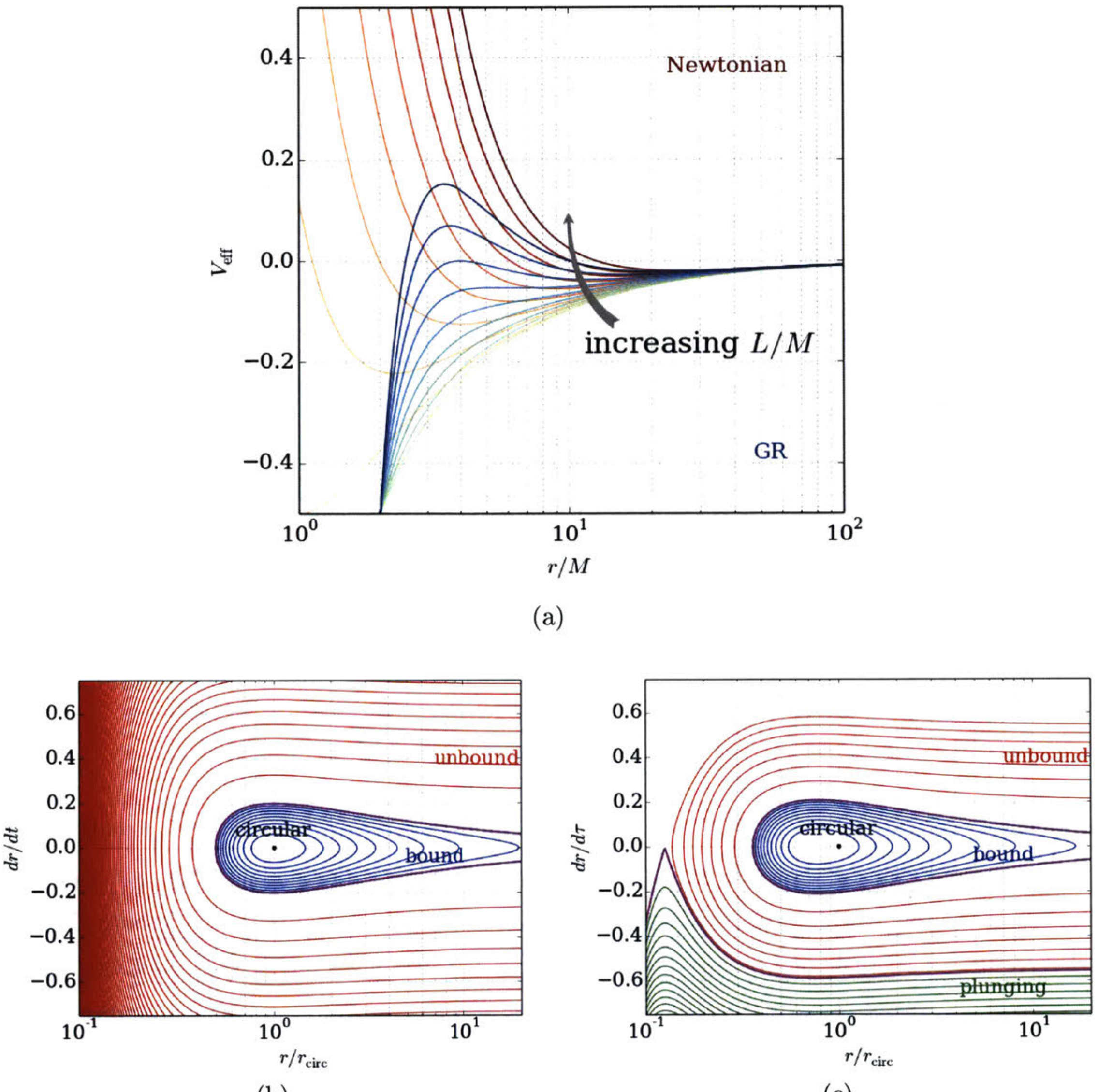
$$
\begin{aligned}
H &= \frac{1}{2m_1}p_1^2 + \frac{1}{2m_2}p_2^2 - \frac{Gm_1m_2}{|\vec{r_1} - \vec{r_2}|} && (1.1) \\
&= \frac{1}{2M}P^2 + \frac{1}{2\mu}p_r^2 - \frac{GM\mu}{r} + \frac{L^2}{2\mu r^2} && (1.2)
\end{aligned}
$$

where $r = |\vec{r_1} - \vec{r_2}|$, $M = m_1 + m_2$, $\mu = m_1m_2/M$, $L = \mu r^2 d\theta/dt$ where θ is an angle describing the position of the components. P is the canonical momentum associated with the system's center of mass (which is often set to zero) and p_r is the canonical momentum associated with r. Fig. 1-1 illustrates this system and our coordinates. From this Hamiltonian, we investigate the dynamics of p_r and r. We define the effective potential (V_{eff}) for the radial coordinate and identify a centrifugal barrier ($L^2/2\mu r^2$) associated with the conserved angular momentum (L). Fig. 1-2 shows the effective potential as a function of angular momentum. This simple model already provides a lot of insight into the system's dynamics, including both bound and unbound orbits, with dynamics described by a 1^{st}-order ordinary differential equation

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 - \frac{GM}{r} + \frac{(L/\mu)^2}{2r^2} = E/\mu \tag{1.3}$$

where $V_{\text{eff}} = -GM/r + (L/\mu)^2/2r^2$. A specific example of common interest involves circular orbits ($dr/dt = 0$), and Eqn. 1.3 then specifies the radius of that orbit as a function of the energy or angular momentum. Circular orbits form a one-parameter family of solutions obtained by minimizing V_{eff} with respect to r; there is a constraint relating E/μ and L/μ. We focus on circular orbits throughout §2. Fig. 1-2 also shows phase portraits of various kinds of orbits, including circular, bound (elliptical orbits), and unbound (hyperbolic orbits) systems. We note that the separatrix between bound and unbound orbits corresponds to $E = 0$ (parabolic orbit).

For point particles, we can include the effects of spin by adding the associated kinetic energies. However, in Newtonian gravity there are no tidal torques on point particles (with vanishingly small radii) and therefore these spins cannot interact. This is not the case with extended objects, as we shall see, and generally does not hold in General Relativity either.



(a)

(b)

(c)

Figure 1-2: (a) Effective potentials for a point-particle in Newtonian gravity and GR. The warm colors represent Newtonian gravity and the cool colors GR, with trends following increasing angular momentum denoted. (b) Phase portrait for r and p_r for Newtonian orbits when $(L/\mu)/M = 5$ in non-dimensionalized units, demonstrating bound and unbound orbits. We note that no trajectory ever reaches $r = 0$ because of the divergent centrifugal barrier. (c) Phase portrait for GR, which resembles the Newtonian case but also includes plunging trajectories that do reach $r = 0$. We omit out-going trajectories corresponding to the plunging orbits. Separatrices in both (b) and (c) are shown in purple.

1.1.2 General Relativity

General Relativity (GR) provides a dynamical theory for the geometry of spacetime as a Riemannian metric.[1] Fundamentally, this metric defines the spacetime interval between two events as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{1.4}$$

The metric ($g_{\mu\nu}$) is a rank-2 symmetric tensor that is the analog of the gravitational potential from Newtonian gravity.[2] The metric encodes the geometry of spacetime, including the curvature, which can be determined through the Riemann tensor[3]

$$R^\rho{}_{\sigma\mu\nu} V^\sigma = [\nabla_\mu, \nabla_\nu] V^\rho \tag{1.5}$$

where the covariant derivative (∇_μ) is defined in terms of the Christoffel connections, in turn defined in terms of the metric

$$\nabla_\mu \xi^\sigma = \partial_\mu \xi^\sigma + \Gamma^\sigma_{\mu\nu} \xi^\nu \quad \Big| \quad \Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu} \right) \tag{1.6}$$

This derivative captures the concept of parallel transport, or moving a vector (ξ^σ) along a possibly curved path in possibly curved spacetime in such a way as to keep it as parallel as possible to itself throughout the process. The Riemann curvature tensor then encapsulates whether parallel transport commutes along two different paths, as it does for flat space ($R^\rho{}_{\sigma\mu\nu} \to 0$). GR also posits that the metric is smooth, meaning it can appear flat at any particular point by an appropriate choice of coordinates. This means that observers confined to a small region of space will not be able to detect the effects of gravity, exactly like the weightlessness experienced by astronauts or other freely-falling observers.

---

[1] We defer to [37, 166] for a more complete discussion of GR and provide only the basics and specific points applicable to our studies here. We also assume units in which $c = G = 1$ throughout.

[2] We assume a basic knowledge of tensors throughout and adopt the Einstein summation convention unless explicitly stated otherwise.

[3] We assumed no torsion, which is the case for metric compatible connections like the Christoffel connection.

Worldlines of free-falling particles follow geodesics defined by parallel transport

$$u^\sigma = \frac{d\xi^\sigma}{d\tau} \tag{1.7}$$

$$u^\mu \nabla_\mu u^\sigma = 0 \tag{1.8}$$

which is the relativistic analog of Newton's first law; object in motion stay in motion unless acted upon by an external force. This also captures the equivalence principle; all freely-falling observers follow the same geodesics and therefore fall at the same rate. We note that this is a radical departure from the Newtonian picture, in which a gravitational force acts on all particles causing them to accelerate. Instead, GR describes gravitation in a force-free paradigm. Particles move along straight lines, generalized to geodesics, and their relative motion is controlled by the curvature of spacetime through $g_{\mu\nu}$ rather than a gravitational force.

GR provides evolution equations for $g_{\mu\nu}$ through the Einstein Field Equations (EFEs)

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu} \tag{1.9}$$

where $T_{\mu\nu}$ is the stress-energy tensor of all matter and fields. $R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}$ is the Ricci Tensor and $R = R^\mu{}_\mu$ the Ricci scalar, both contractions of the Riemann curvature tensor. Λ is a cosmological constant that is often neglected over small scales (e.g. within binary stellar systems). These equations enforce conservation of energy implicitly because $\nabla_\mu G^{\mu\nu} = 0$ and famously describe how matter tells space how to curve while Eqn. 1.8 describes how space tells matter how to move.

Now, a standard analog of the Kepler problem is the (time-like) worldline of a test particle orbiting a circularly symmetric body sourcing a time-independent metric. In this case, the relevant metric was first determined by Schwarzschild [125]

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 \left(d\theta^2 + \sin^2\theta d\phi^2\right) \tag{1.10}$$

In this case, we assume the test particle has a much smaller mass than the central body so that $m_1 + m_2 \approx m_1 = M$ and $\mu \approx m_2$. and the equation of motion for the radial position of the test particle is

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 - \frac{GM}{r} + \frac{(L/\mu)^2}{2r^2} - \frac{GM(L/\mu)^2}{r^3} = E/\mu \tag{1.11}$$

where E/μ and $L/\mu = r^2 d\phi/d\tau$ are constants and assuming the motion is confined to the $\theta = \pi/2$ plane. We notice this looks very similar to the Newtonian case with an additional term in the effective potential $(-GM(L/\mu)^2/r^3)$. This term allows for plunging orbits; when two bodies are close enough the gravitational attraction is strong enough to overcome the centrifugal barrier and the two bodies fall toward one another at an ever increasing rate. This plunging behavior is very important for coalescing compact objects, which eventually fall into each other due to the extra term in V_{eff}. We will discuss such systems in some detail within §2.2, §3, and §4. Fig. 1-2 compares the effective potentials for similar Newtonian and GR systems. It also shows the different phase portraits for the two potentials, and we see the additional plunging orbits present in GR that are not present in Newtonian gravity.

Gravitational Radiation

GR also predicts GWs, which are sourced by accelerated matter or energy. This is typically shown by linearizing the EFEs about a flat background metric, although the choice of background metric is arbitrary. We will follow standard practice and linearize around the Minkowski metric ($\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$) such that

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1.12}$$

We also note that raising and lowering indices of $h_{\mu\nu}$ is done with the unperturbed metric ($h^{\mu}{}_{\nu} = \eta^{\mu\sigma}h_{\sigma\nu}$) because we keep only terms linear in $h_{\mu\nu}$. In vacuum ($T_{\mu\nu} = 0$), the EFEs provide propagation equations for $h_{\mu\nu}$

$$G_{\mu\nu} = \frac{1}{2}\left(\partial_\sigma\partial_\nu h^{\sigma}{}_{\mu} + \partial_\sigma\partial_\mu h^{\sigma}{}_{\nu} - \partial_\mu\partial_\nu h - \partial_\alpha\partial^\alpha h_{\mu\nu} - \eta_{\mu\nu}\partial_\rho\partial_\lambda h^{\rho\lambda} + \eta_{\mu\nu}\partial_\alpha\partial^\alpha h\right) = 0 \tag{1.13}$$

where $h = h^{\mu}{}_{\mu}$. This is obtained directly from the EFEs by using the covariant derivative of the unperturbed metric ($\eta_{\mu\nu}$). By an appropriate choice of coordinates, we can simplify this expression for the *transverse-traceless* (TT) perturbation to

$$\partial_\alpha\partial^\alpha h^{\text{TT}}_{\mu\nu} = 0 \quad \Big| \quad h^{\text{TT}}_{t\mu} = h^{\text{TT}} = 0 \tag{1.14}$$

where $h^{\text{TT}}_{t\mu}$ refers to the components of $h^{\text{TT}}_{\mu\nu}$ with at least one time component. This is a wave equation for a massless particle, which we call a *gravitational wave.* If the wave propagates in the

z-direction, we can write

$$h_{\mu\nu}^{\mathrm{TT}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.15}$$

and identify the two degrees of freedom as the + and × polarizations, named after the ways in which they move test particles (see §1.3). This is commonly called the transverse-traceless gauge because the perturbation's trace vanishes and all changes to the metric are orthogonal to the wave's propagation direction (transverse). There is further gauge freedom in the choice of the x- and y-directions, and the equations of motion are invariant under rotations about the direction of propagation. This choice is often call the polarization angle (ψ) and is important when we project from the wave-frame onto GW detectors.

The TT gauge is particularly important because it is the only gauge invariant portion of the metric perturbation which obeys a wave equation and therefore represents the physical information carried by GWs [67].[4] In fact, we can write the general perturbation as

$$h_{tt} = 2\phi \tag{1.16}$$

$$h_{ti} = \beta_i + \partial_i \gamma \tag{1.17}$$

$$h_{ij} = h_{ij}^{\mathrm{TT}} + \frac{1}{3}\delta_{ij} H + \frac{1}{2}\left(\partial_i \epsilon_j + \partial_j \epsilon_i\right) + \left(\partial_i \partial_j - \frac{1}{3}\delta_{ij}\partial^k \partial_k\right)\lambda \tag{1.18}$$

where $H = \delta^{kl} h_{kl}$, subject to

$$\partial_i \beta_i = 0 \tag{1.19}$$

$$\partial_i \epsilon_i = 0 \tag{1.20}$$

$$\partial_i h_{ij}^{\mathrm{TT}} = 0 \quad \text{(guarantees this is transverse)} \tag{1.21}$$

$$\delta^{ij} h_{ij}^{\mathrm{TT}} = 0 \quad \text{(guarantees this is traceless)} \tag{1.22}$$

along with the boundary conditions γ, ϵ_i, λ, $\partial^k \partial_k \lambda \to 0$ at spatial infinity. If we further parameterize gauge transformations under $(x')^\mu = x^\mu + \zeta^\mu$ such that ($h'_{\mu\nu} = h_{\mu\nu} - \partial_\mu \zeta_\nu - \partial_\nu \zeta_\mu$ and $\zeta_\mu =$

---

[4] Throughout the rest of this section, we will use greek subscripts (α, β, γ, etc) to indicate sums over all spacetime dimensions and latin subscripts (i, j, k, etc) to represent sums over only spatial dimensions.

$(A, B_i + \partial_i C)$), we see that h_{ij}^{TT} is gauge invariant under infinitesimal changes of coordinates

$$\phi \to \phi + \partial_t A \tag{1.23}$$

$$\beta_i \to \beta_i + \partial_t B_i \tag{1.24}$$

$$\gamma \to \gamma - A - \partial_t C \tag{1.25}$$

$$H \to H - 2\partial^k \partial_k C \tag{1.26}$$

$$\lambda \to \lambda - 2C \tag{1.27}$$

$$\epsilon_i \to \epsilon_i - 2B_i \tag{1.28}$$

$$h_{ij}^{\mathrm{TT}} \to h_{ij}^{\mathrm{TT}} \tag{1.29}$$

[67] define four gauge invariant quantities using this decomposition

$$\Phi = -\phi + \partial_t \gamma - \frac{1}{2}\partial_t^2 \lambda \tag{1.30}$$

$$\Theta = \frac{1}{3}\left(H - \partial^k \partial_k \lambda\right) \tag{1.31}$$

$$\Xi_i = \beta_i - \frac{1}{2}\partial_t \epsilon_i \tag{1.32}$$

and h_{ij}^{TT}. They note that only h_{ij}^{TT} obeys a wave equation (Eqn. 2.62–2.65 of [67]). All other gauge invariant, and therefore physically meaningful, parts of the metric perturbation obey Poisson equations and do not contribute dynamically in the radiation field at leading order. There are six gauge invariant degrees of freedom in this decomposition, consistent with the 10 independent components of $h_{\mu\nu}$ and the four arbitrary gauge functions associated with ζ_μ. All but the two degrees of freedom within h_{ij}^{TT} are completely determined by the source via their Poisson equations. To wit, in vacuum we obtain

$$\partial^k \partial_k \Phi = 0 \quad \text{(1 degree of freedom)} \tag{1.33}$$

$$\partial^k \partial_k \Theta = 0 \quad \text{(1 degree of freedom)} \tag{1.34}$$

$$\partial^k \partial_k \Xi_i = 0 \quad \text{(3-1=2 degrees of freedom)} \tag{1.35}$$

$$\partial^\mu \partial_\mu h_{ij}^{\mathrm{TT}} = 0 \quad \text{(6-3-1=2 degrees of freedom)} \tag{1.36}$$

The constraints noted here are $\partial_i \Xi_i = 0$, $\partial_i h_{ij}^{\mathrm{TT}} = 0$, and $h^{\mathrm{TT}} = 0$.

Now, when we consider how GWs are sourced, we define the trace-reversed perturbation $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$ and use the gauge freedom associated coordinate changes in GR to demand $\partial_\mu \bar{h}^\mu{}_\nu = 0$, often called the Lorentz gauge, which is sourced by accelerations in the mass-energy quadrupole at leading order

$$\bar{h}_{ij} = \frac{2G}{r}\frac{d^2}{dt^2}I_{ij}(t-r) \quad \Bigg| \quad I_{ij} = \int d^3y\, y^i y^j T^{tt}(t,y) \tag{1.37}$$

While higher order multipoles also source GWs, we typically consider only the quadrupole's contribution because it dominates for most relevant frequencies in systems of interest. Furthermore, the quadrupole formula assumes relatively slowly moving objects ($v \ll c$), which can break down near the late inspiral of compact objects.

Because of GR's invariance under diffeomorphisms, it is possible to make spacetime appear locally flat at any point in space by an appropriate choice of coordinates. This means the stress-energy associated with perturbations to the metric can be made to vanish at any individual point. Therefore, we define the stress-energy pseudo-tensor ($t_{\mu\nu}$) by averaging over a few wavelengths of the associated perturbation in the radiation zone

$$t_{\mu\nu} = \frac{1}{32\pi G}\left\langle \left(\partial_\mu h_{\rho\sigma}\right)\left(\partial_\nu h^{\rho\sigma}\right)\right\rangle \tag{1.38}$$

The power radiated by a source can be found by integrating the outward flux of the time-like component of this pseudo-tensor across a large sphere at spatial infinity

$$\mathcal{L}_{\mathrm{GW}} = \int d\Omega\, r^2 t_{ti} n^i = \frac{G}{5}\left\langle \frac{d^3\bar{I}_{ij}}{dt^3}\frac{d^3\bar{I}^{ij}}{dt^3}\right\rangle \quad \Bigg| \quad \bar{I}_{ij} = I_{ij} - \frac{1}{3}\delta_{ij}I^k_k \tag{1.39}$$

For two bodies with masses m_1 and m_2, respectively, orbiting their center of mass, separated by a distance R, and with orbital angular velocity Ω_{orb}, this reduces to

$$\mathcal{L}_{\mathrm{GW}} = \frac{32}{5}G^4\frac{m_1^2 m_2^2(m_1+m_2)}{R^5} \tag{1.40}$$

$$= \frac{32}{5}\frac{G^{7/3}\mathcal{M}^{10/3}}{c^5}\Omega_{\mathrm{orb}}^{10/3} \tag{1.41}$$

where the chirp mass ($\mathcal{M}$) is given by $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$; see [113] for more details. Now, equating the energy radiated with the time rate-of-change of the orbital energy ($E_{\mathrm{orb}} =$

$-Gm_1m_2/2R$) yields

$$\Omega_{\text{orb}} = \left(\frac{256}{5}G^{5/3}\mathcal{M}^{5/3}(t_c - t)\right)^{-3/8} \tag{1.42}$$

where t_c is the time at which the point particles coalesce. Physically, this corresponds to a nearly circular orbit of two bodies that radiates energy and angular momentum as GWs. As the system loses energy, the bodies fall toward one another, orbiting faster and radiating at an increasing rate. This will proceed until the particles are so close they plunge toward one another, overcoming the centrifugal barrier associated with their orbital angular momentum, and coalesce into a single object. This process is typically called compact binary coalescence (CBC).

In reality, the maximum frequency is set by the physical extent of the objects or, in the case of black holes, their horizons. This chirping signal is characteristic of CBCs when the bodies are still well separated and is the basis of many searches for GWs (§3.1.1).

1.2 Binaries Containing Extended Bodies

While we can gain a lot of intuition from studying binary systems containing only point particles, real binaries almost always involve extended bodies. Their non-vanishing size produces a gradient in the gravitational force experienced by different parts of each object, which may produce a non-vanishing net torque. This allows components to exchange angular momentum with each other and with the orbit as a whole. We briefly review linear tidal theory here and focus on the implications of nonlinear tidal interactions in subsequent chapters (§2). In particular, we will investigate breakdowns of linear tidal theory and where nonlinearities saturate, with particular emphasis on how these effects may manifest in measurable ways.

1.2.1 Linear Perturbation Theory

Stars are self-gravitating systems composed of gas and plasma. We typically determine their unperturbed configurations using the equations of stellar structure (in geometric units: $G = c = 1$)

$$\frac{dm}{dr} = 4\pi\rho r^2 \tag{1.43}$$

$$\frac{dP}{dr} = -\frac{m\rho}{r^2} \tag{1.44}$$

where m is the mass contained within a radius r, ρ is the stellar density, and P the pressure, along with a description of energy generation and (radiative) transport

$$\frac{dL}{dr} = 4\pi\varepsilon\rho r^2 \tag{1.45}$$

$$4\pi r^2\sigma T^3\frac{dT}{dr} = -\frac{3\kappa\rho L}{16} \tag{1.46}$$

and an equation of state (EOS) describing the matter within the star. Here, L is the luminosity within the star (energy flux), ϵ the local specific energy generation rate, κ the opacity, and σ the Stefan-Boltzmann constant. The analogous equations for a perfect fluid within GR are the Tolman–Oppenheimer–Volkoff (TOV) equations

$$\frac{dm}{dr} = 4\pi r^2\rho \tag{1.47}$$

$$\frac{dP}{dr} = \frac{(\rho + P)(m + 4\pi r^3 P)}{r(r - 2m)} \tag{1.48}$$

which again are closed by the EOS. We note that the quantity m in the TOV equation is *not* the contained mass because the measure does not account for the curved metric. However, it is the *gravitational mass*, which determines the dynamics of orbiting bodies, analogous to the mass parameter in the Schwarzschild metric (Eqn. 1.10). In the rest of this section, we focus on Newtonian gravity.

Given a background stellar model, we determine the behavior of adiabatic Lagrangian fluid displacements (ξ) under the star's restoring forces, which we recognize as an eigenvalue problem. Typically, this involves a decomposition into spherical harmonics ($Y_{lm}(\theta, \phi)$) characterized by their degree (l) and azimuthal number (m). The solutions to the eigenvalue problem are generally categorized by their restoring forces. For non-rotating sun-like stars, p-modes are predominantly radial and are restored by pressure (acoustic waves); g-modes are restored by buoyancy or compositional gradients and their fluid displacements often contain significant transverse components. In fact, there are no $l = 0$ g-modes because there is no horizontal motion (and therefore different fluid strata do not cross). Our studies primarily focus on g-modes because they can have natural frequencies near orbital frequencies for interesting systems, and higher order g-modes have lower natural frequencies (ω_g). Commonly, g-modes have frequencies lower than the fundamental stellar frequency ($\omega_0 = \sqrt{GM/R^3}$) and p-modes have higher frequencies.

The oscillatory region for each g-mode is defined by the star's Brunt-Väisälä frequency (N) [9], which represents the buoyancy frequency for fluid displacements. For sun-like stars, N is linear with the radius near the stellar core because the enclosed mass (and therefore local gravitational acceleration) grows linearly and vanishes at the radiative-convective envelop near the star's surface. g-modes propagate where N is larger than their natural frequency, which implies that high order g-modes with lower natural frequencies propagate deep into the stellar core before N becomes comparable to ω_g. High-order g-modes spectra can be described by WKB scaling relations, which describe their natural frequencies and other parameters as functions of the mode numbers. Generally, these follow the form

$$\omega_g = \alpha \frac{l}{n} \tag{1.49}$$

$$\gamma = \gamma_0 \left(\frac{\omega_0}{\omega_g}\right)^2 \omega_0 \Lambda^2 \quad | \quad \Lambda^2 = l(l+1) \tag{1.50}$$

where γ describes radiative damping within radiatively supported cores of main sequence stars [137]. g-modes excited by the linear tide are launched from radiative-convective interfaces within the star. For solar mass stars, this means they propagate inward toward the stellar core, being geometrically focused to larger amplitudes according to $\xi \sim 1/r^2$. If the modes reach the core before they dissipate, they reflect off the inner Brunt profile. Modes that reflect many times from global standing waves within the resonant cavity defined by N. We focus on these modes.

We can further divide the tide into the equilibrium and dynamical portions. The equilibrium tide is the quasi-static deformation of the star as a whole and is primarily composed of modes with natural frequencies much higher than the orbital frequency as well as responses to non-oscillatory parts of the tidal potential. The dynamical tide is the resonant response of the star and is usually composed of a few high order g-modes. [170] found that the dynamical tide dominates the weakly nonlinear interactions within the star, and therefore we focus only on these modes in §2 [63].

1.2.2 The Tidal Potential

The tidal potential drives the linear modes, and, for low eccentricity orbits, the $l = 2$ spherical harmonic dominates. This driving is encapsulated by the linear overlap integrals between the tidal potential (U) and the mode shapes [170]

$$\begin{aligned} U_\alpha &= -\frac{1}{E_0}\int d^3x\, \rho\xi_\alpha^* \cdot \nabla U && (1.51) \\ &= \frac{M_p}{M}\sum_{lm} W_{lm} I_{\alpha lm}\left(\frac{R}{D(t)}\right)^{l+1} e^{-im\phi(t)} && (1.52) \end{aligned}$$

where

$$\begin{aligned} I_{\alpha lm} &= \frac{1}{Mr^l}\int d^3x\, \rho\xi_\alpha^* \cdot \nabla\left(r^l Y_{lm}\right) \approx \hat{I}_{\alpha lm}\left(\frac{\omega_\alpha}{\omega_0}\right)^{11/6} && (1.53) \\ W_{lm} &= \frac{4\pi}{2l+1} Y_{lm}\left(\theta = \frac{\pi}{2}, \phi = 0\right) && (1.54) \end{aligned}$$

$\hat{I}_{\alpha 2\pm 2} \sim 2.5 \cdot 10^{-3}$ for a sun-like star. For forcing dominated by the $l = 2$, $m = \pm 2$ harmonic

$$U_\alpha = 4\pi^2\sqrt{\frac{3\pi}{10}}\frac{M_p}{M+M_p}\left(\omega_0 P_{\text{orb}}\right)^{-2}\left(\frac{\omega_\alpha}{\omega_0}\right)^{11/6}\hat{I}_{\alpha 2\pm 2} e^{\pm i2\phi} \qquad (1.55)$$

We note that the overlap integral decreases for higher order modes but nonetheless these modes can have sharp enough resonance to be excited to high amplitudes.

Following the phase-space decomposition within [170] (see §2.1), we can immediately solve for the behavior of each mode amplitude (q_α) separately through sets of 1st-order ordinary differential equations

$$\dot{q}_\alpha + \left(i\omega_\alpha + \gamma_\alpha\right) q_\alpha = i\omega_\alpha U_\alpha(t) \qquad (1.56)$$

This is just a set of damped, driven independent simple harmonic oscillators. §2 explores the dynamics of these systems when nonlinear couplings between the modes become important.

1.2.3 Implications for GW Waveforms

Tidal interactions are important for GW sources as well. Generally, the deformability of the constituents modifies the stress-energy tensor and therefore the source term for GWs. This is often described by the component's love number [121, 174], which measures the linear physical response of an object to a tidal field. Within a post-Newtonian (PN) expansion [28], these quasi-static deformations enter at a high PN order but with large enough coefficients that they can still be important during the late inspiral [121].

The tide can also modify the GW signal by dissipating energy. Instead of simply modifying the stress-energy tensor, dissipation of the dynamical tide within the stars implies faster energy loss compared to GW radiation alone and therefore faster orbital evolution. Typically, growth rates for these resonant interactions are small enough that modes cannot absorb large amounts of energy before GW radiation sweeps the orbit through the resonance [81, 92, 93]. This appears to be true for a variety of possible neutron star equations of state and mode types [175, 176]. In §2.2, we describe a nonlinear secular instability that can dissipate energy throughout the inspiral, which may be measurable in astrophysical contexts [62].

1.3 General Problem of GW Detection

Now that we have discussed the basic physics of binaries and how they can source GWs, it is prudent to consider how we can actually measure them. The basic goal for GW detection is to measure the small displacements introduced by the metric perturbations sourced by distant objects. By measuring the separation between a group of neighboring geodesics (called a congruence), we can determine the GW strain.

Let us consider an observer that is initially stationary ($u^{\sigma} = (1,0,0,0)$). Working in the transverse-traceless gage, $\Gamma^{\sigma}_{\mu\nu}u^{\nu} = 0$ because $h^{\mathrm{TT}}_{t\mu} = 0$. We obtain $u^{\mu}\nabla_{\mu}u^{\sigma} = \partial_t u^{\sigma} = 0$, and therefore u^{σ} is constant for all time. This means that stationary observers remain at the same coordinate position regardless of the shape of the perturbation; we can measure the proper separation between neighboring observers within the congruence by measuring the proper distance between fixed coordinate positions. The proper distance between them *can* change and is the basis for interferometric detectors. Fig. 1-3 demonstrates the characteristic pattern introduced by the + and × polarizations.

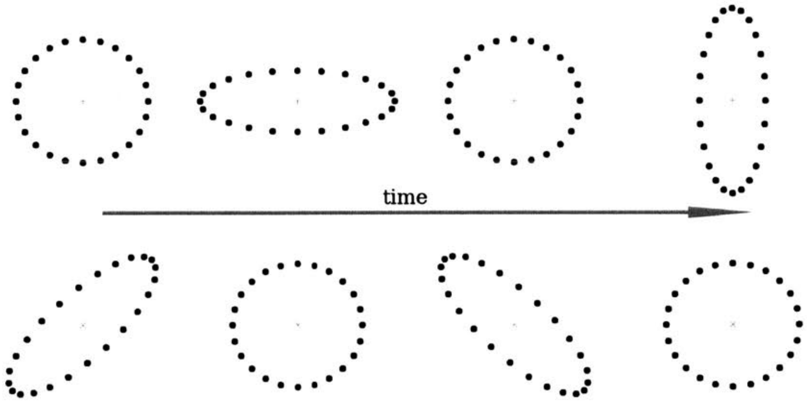


Figure 1-3: The relative change in the proper distance from the origin to a ring of free-falling observers under the action of the (top) + polarization and (bottom) × polarization.

1.3.1 Interferometric Detector Response

Interferometric detectors, sometimes called interferometric observatories (IFOs), measure the proper distance between test particles by recording light's round-trip travel time along null geodesics. To determine the detector response, we must calculate how this round-trip travel time is affected by $h_{\mu\nu}$ [117, 118, 119].

Consider a plane wave propagating in the n_i direction

$$h_{\mu\nu} = h_{\mu\nu}(t - n_i r^i) \tag{1.57}$$

Furthermore, assume one of the detector's arms is pointed in the e_i direction, so that the component of strain along the arm is given by

$$h = h_{\mu\nu} e^\mu e^\nu = h(t - n_i r^i) \tag{1.58}$$

We further consider motion only along the e_i direction, so we are interested in

$$h = h(t - n_i e^i x) = h(t - n_e x) \tag{1.59}$$

where x is the distance along the arm and $n_e \equiv n_i e^i$. Now, we know light will follow a null geodesic,

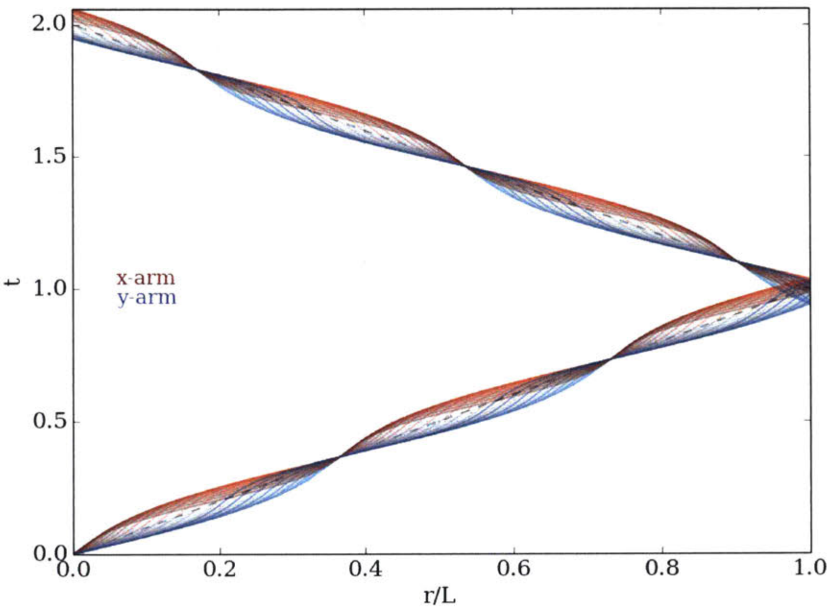


Figure 1-4: Spacetime diagram showing the worldlines of null geodesics along their trips down the interferometer's arms and back. The strain is greatly exaggerated to demonstrate the effect, and we depict multiple GW phases with different shades of (x-arm) red and (y-arm) blue. Note, we plot the coordinate position, not the physical separation between points. Points where all worldlines cross correspond to complete cycles of the monochromatic GW strain.

so that $dt^2 = (1+h)dx^2$. Furthermore, we assume there is a single Fourier component of the GW strain, $h = h_0 e^{-i\omega(t-n_e x)}$. Fig. 1-4 demonstrates the null geodesics in coordinate position (not physical separation) as the light travels down the detector's arms and back.

Consider the outbound trip. In this case, $dx/dt > 0$ and

$$dt = \sqrt{1+h}\,dx = (1 + \frac{1}{2}h + O(h^2))dx \tag{1.60}$$

We define $d\tau = dt - dx$ and note that τ denotes the phase of light as it propagates down the arm. Worldlines of constant τ represent worldlines of constant phase. We assume $h \ll 1$, and obtain

$$\int_{t_0-x_0}^{T+t_0-(L+x_0)} d\tau\, e^{i\omega\tau} = \frac{h_0}{2} \int_{x_0}^{L+x_0} dx\, e^{-i\omega(1-n_e)x} \tag{1.61}$$

$$\frac{e^{i\omega(t_0-x_0)}}{i\omega}\left(e^{i\omega(T-L)} - 1\right) = \frac{h_0}{2}\frac{e^{-i\omega(1-n_e)x_0}}{-i\omega(1-n_e)}\left(e^{-i\omega(1-n_e)L} - 1\right) \tag{1.62}$$

For the return trip, we know $dx/dt < 0$ and define $d\tau' = dt + dx$. Again, τ' is the phase of the light on the return path. This yields

$$\int_{T+t_0+L+x_0}^{T'+t_0+x_0} d\tau'\, e^{i\omega\tau'} = -\frac{h_0}{2} \int_{L+x_0}^{x_0} dx\, e^{i\omega(1+n_e)x} \tag{1.63}$$

$$\frac{e^{i\omega(t_0+x_0)}}{i\omega}\left(e^{i\omega T'} - e^{i\omega(T+L)}\right) = -\frac{h_0}{2}\frac{e^{i\omega(1+n_e)x_0}}{i\omega(1+n_e)}\left(1 - e^{i\omega(1+n_e)L}\right) \tag{1.64}$$

We are interested in the change in round-trip time relative to the unperturbed state ($(T'-2L)/2L \sim h_0$), and therefore write

$$e^{i\omega(T'-2L)} - e^{i\omega(T-L)} = -\frac{h_0}{2}e^{-i\omega(t_0+x_0+2L)}\frac{e^{i\omega(1+n_e)x_0}}{1+n_e}\left(1 - e^{i\omega(1+n_e)L}\right) \tag{1.65}$$

$$e^{i\omega(T-L)} = 1 - \frac{h_0}{2}e^{-i\omega(t_0-x_0)}\frac{e^{-i\omega(1-n_e)x_0}}{1-n_e}\left(e^{-i\omega(1-n_e)L} - 1\right) \tag{1.66}$$

which yields

$$e^{i\omega(T'-2L)} - 1 = \frac{h_0}{2}\left(e^{-i\omega(t_0-x_0+(1-n_e)x_0)}\left(\frac{1-e^{-i\omega(1-n_e)L}}{1-n_e}\right)\right. \tag{1.67}$$

$$\left. -e^{-i\omega(t_0+x_0+2L-(1+n_e)x_0}\left(\frac{1-e^{i\omega(1+n_e)L}}{1+n_e}\right)\right) \tag{1.68}$$

$$= \frac{h_0}{2}e^{-i\omega(t_0-n_e x_0)}\left(\frac{1-e^{-i\omega(1-n_e)L}}{1-n_e} - e^{-2i\omega L}\frac{1-e^{i\omega(1+n_e)L}}{1+n_e}\right) \tag{1.69}$$

We now assume $T' - 2L \ll 1/\omega$ and obtain

$$\frac{T'-2L}{2L} = \frac{h_0 e^{-i\omega(t_0-n_e x_0)}}{4i\omega L}\left(\frac{1-e^{-i\omega(1-n_e)L}}{1-n_e} - e^{-2i\omega L}\frac{1-e^{i\omega(1+n_e)L}}{1+n_e}\right) \tag{1.70}$$

We recognize $h_0 e^{-i\omega(t_0-n_e x_0)}$ as the strain projected along the arm at the spacetime coordinates corresponding to the start of the round-trip, and can therefore express this as

$$\frac{\delta T}{2L} = \frac{1}{4i\omega L}\left(\frac{1-e^{-i\omega(1-n_e)L}}{1-n_e} - e^{-2i\omega L}\frac{1-e^{i\omega(1+n_e)L}}{1+n_e}\right) e_i e_j h^{ij} \tag{1.71}$$

This yields the transfer function from the astrophysical signal to the observed fractional change in round-trip travel time down a single arm of the detector. Furthermore, if we define

$$D(\omega, n_e) \equiv \frac{1}{4i\omega L}\left(\frac{1-e^{-i\omega(1-n_e)L}}{1-n_e} - e^{-2i\omega L}\frac{1-e^{i\omega(1+n_e)L}}{1+n_e}\right), \tag{1.72}$$

we can compactly report the output of an interferometer as

$$\delta V = \left(D(\omega, n_\alpha e_x^\alpha)e_x^i e_x^j - D(\omega, n_\beta e_y^\beta)e_y^i e_y^j\right) h_{ij} \tag{1.73}$$

$$= D^{ij} h_{ij} \tag{1.74}$$

We note that, while a single arm is sensitive to GW strains, by comparing the difference between two separate arms we can reject many sources of noise and produce a much more sensitive device. This is why IFOs contain multiple arms and why our read-out measures the difference between the x- and y-arms.

This yields standard expressions for the antenna patterns as follows

$$F_+ = D_{ij}\epsilon_+^{ij} \tag{1.75}$$

$$F_\times = D_{ij}\epsilon_\times^{ij} \tag{1.76}$$

with ϵ_+ and $\epsilon_\times$ defining the polarization tensors for the + and × polarizations, respectively (Eqn. 1.15). This gives us a coordinate-independent expression for the antenna responses which includes frequency-dependent effects for a Michelson interferometer.

We note that with some elbow grease, we can massage the expression for $D(\omega, n_e)$ into

$$D(\omega, n_e) = \frac{e^{-i\omega L}}{2\left(1-n_e^2\right)}\left(\mathrm{sinc}(\omega L) - n_e^2\,\mathrm{sinc}(\omega n_e L) - \frac{i n_e}{\omega L}\left(\cos(\omega L) - \cos(\omega n_e L)\right)\right) \tag{1.77}$$

From this, we immediately see that $|D(\omega, n_e)| = |D(\omega, -n_e)|$, as required by time-reversal symmetry. We also note that the relevant frequency for these effects corresponds to the unperturbed round-trip travel time, or the *free spectral range* ($f_{\mathrm{fsr}} = c/2L$). Fig. 1-5 and Fig. 1-6 show the antenna response's behavior as a function of frequency and direction.

1.3.2 Pragmatic Issues with Ground-Based Detectors

Although we model our detectors as sets of freely falling observers, in reality our instruments are fixed to the Earth. This means the observers are not truly free-falling, at least not in the vertical direction, but instead are suspended from multiple pendula in order to isolate them from ground motion. These pendula respond as freely-falling obsevers in the horizontal plane as long as the GW frequencies are larger than the pendulum's natural frequency. In fact, noise mitigation and isolation from the environment constitute a significant part of the detector's subsystems, including both active and passive control loops [138, 143]. This also means that, while models of IFO responses may include only a few degrees of freedom (e.g. the lengths of the arms), in reality there are hundreds of thousands of degrees of freedom associated with the positions and velocities of all the suspension systems and surrounding instruments.

However, our isolation is never perfect and environmental disturbances or noise transients within control loops can couple to the measurement of $h(t)$, generating noise artifacts. The scale of the detectors can often make it difficult to diagnose the cause of these artifacts by inspection, and

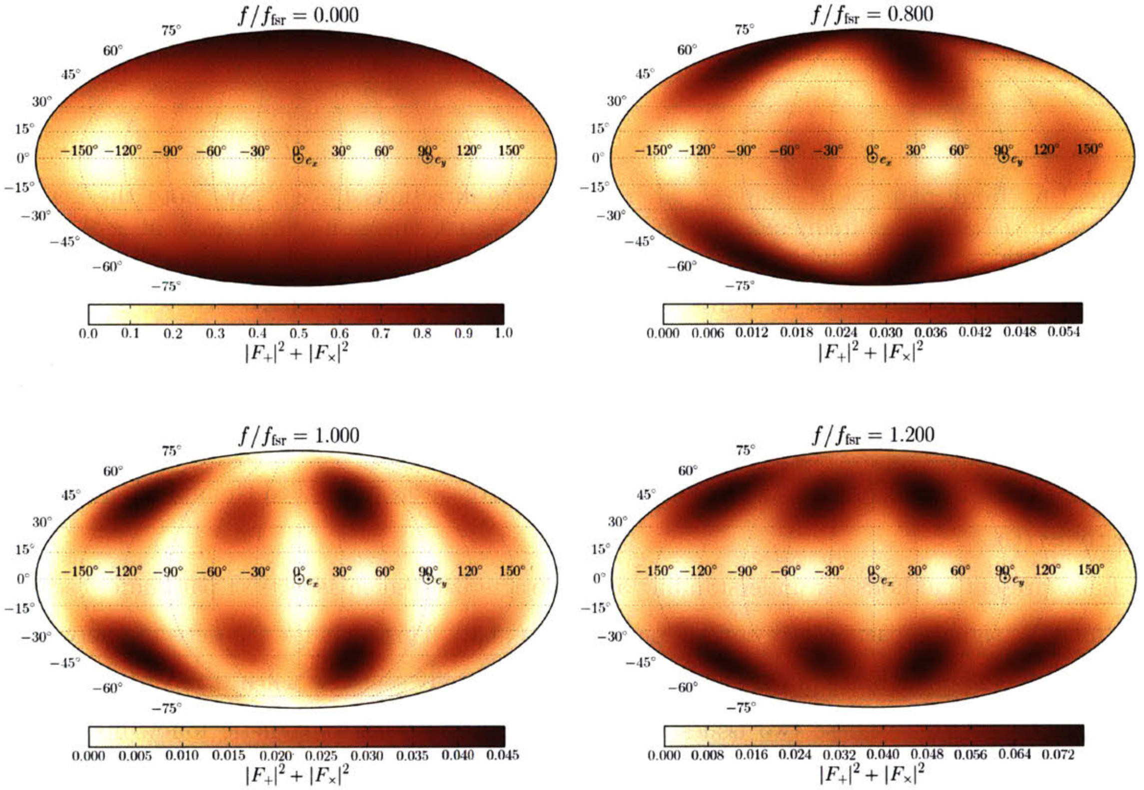


Figure 1-5: At frequencies small compared to f_{fsr}, we notice little difference in the overall shape of the antenna response's magnitude, although the phase does change. The predominant change is a decrease in the magnitude (Fig. 1-6). However, when we approach (and exceed) f_{fsr}, there are large changes to the directional sensitivity. In particular, the maximum in the detector response directly overhead ($\theta = 0$) becomes a zero.

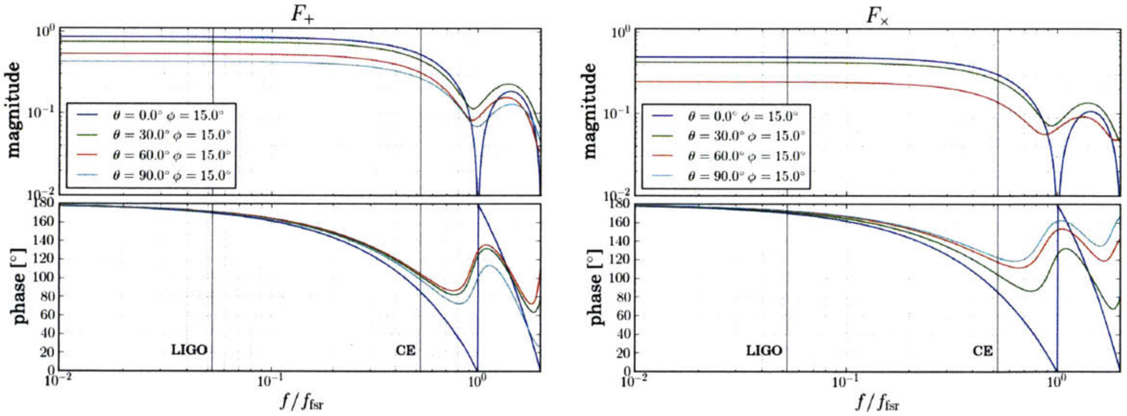


Figure 1-6: Bode plots of F_+ and $F_\times$ at a few source directions (θ, ϕ). We note the general decrease in $|F|$ at higher frequencies as well as the significant change in phase, even at $f \lesssim f_{\rm fsr}/2$. The annotations show the dynamical frequency of a $2M_\odot$ neutron star with a 12 km radius ($f_{\rm ns} = \frac{1}{2\pi}\sqrt{\frac{GM}{R^3}} = 1972$ Hz) normalized by $f_{\rm fsr}$ for LIGO [138] and CE [139], respectively.

we instead employ statistical tools to find correlations between auxiliary degrees of freedom and artifacts in $h(t)$. This thesis discusses solutions for measuring robust statistical correlations in the presence of non-Gaussian, non-stationary noise and how that can be applied to improve a search's sensitivity (§3.2).

I also present methods for understanding and characterizing some aspects of parameter estimation for GW signals, particularly source localization (§3.3). This includes the detector network's exposure, meaning the parts of the sky to which it was actually sensitive when it recorded science-quality data (§3.5.2). These two issues have implications for electromagnetic follow-up of GW candidates.

Although the results presented here represent general trends applicable to arbitrary networks of detectors, we focus on networks consisting of the advanced LIGO Livingston (LLO) and Hanford (LHO) detectors [138], located in Livingston, LA and Hanford, WA, respectively, and advanced Virgo [153], located near Pisa, Italy. While their technical details differ (e.g. the LIGOs have 4-km arms and Virgo has 3-km arms), they all are km-scale Michelson interferometers with Fabry-Perot resonating cavities spanning their arm lengths. These machines have reached the greatest sensitivity to displacements ($\sim 10^{-18}$m) and correspondingly small strains ($\sim 10^{-21}$) ever achieved. For a $1.4M_\odot$-$1.4M_\odot$ binary neutron star system (BNS) this corresponds to sky-averaged detection ranges of 200 Mpc and 130 Mpc for LIGO and Virgo, respectively, at design sensitivity [8].

Chapter 2

Nonlinear Tides

As described in §1.1, binaries containing point particles constitute a wealth of physics. However, real binaries involve extended bodies and we must consider their tidal interactions as well (§1.2). In this chapter, I describe the impact of nonlinear dynamics excited by tidal interactions, focusing first on (§2.1) resonant interactions within exoplanet systems involving Jupiter-mass planets in tight orbits around sun-like stars and (§2.2) secular instabilities within neutron stars coalescing due to the emission of GWs. Results presented in §2.1 are primarily taken from [63] while §2.2 is based primarily on [62].

In many situations, linear tidal theory is insufficient and can be shown to be unstable to small perturbations. These instabilities are present because the linear equations of motion do not represent the entire Hamiltonian and there are *nonlinear* interactions between the eigenmodes within extended bodies. In particular, it is natural to write the Hamiltonian as a power series in the time-dependent amplitudes of linear eigenmodes, sometimes called a Galerkin decomposition:

$$H_* = \sum_{\alpha} \left(\dot{q}_\alpha^2 + \omega_\alpha^2 q_\alpha^2 \right) + \sum_{\alpha,\beta,\gamma} \kappa_{\alpha\beta\gamma} q_\alpha q_\beta q_\gamma + \cdots \tag{2.1}$$

where each mode's amplitude ($q_\alpha(t)$) is a generalized coordinate [161]. We focus on weakly nonlinear systems in which only the leading-order nonlinear interactions are important. This means we neglect all higher-order terms and include only the *3-mode* interactions with coupling coefficients $\kappa_{\alpha\beta\gamma}$. Substantial work has been devoted to computing these 3-mode coupling coefficients [164, 168, 169, 170], and we will assume they are known and only examine the resulting dynamics of networks of nonlinearly coupled modes. Specifically, we study nonlinear interactions between stellar eigenmodes driven by tidal potentials in close binaries, focusing on hot Jupiters orbiting main sequence stars and binary systems of compact stellar remnants, such as neutron stars.

2.1 Hot Jupiters

The term *hot Jupiter* refers to a gaseous giant planet orbiting very close to its host star [107]. These exoplanets were some of the first to be discovered because their large masses and radii as well as their short orbital periods generate strong signals in both radial velocity and photometric-transit measurements [154]. To date, roughly 61 hot Jupiters with $M_p \sin i \geq M_J$ and $P_{\text{orb}} \leq 3$ days have been discovered [1], and approximately 27 of these planets orbit solar-type stars (0.8-$1.1 M_\odot$). We focus on the latter, specifically modeling the interior g-modes of sun-like hosts and the couplings between them. We find that resonant nonlinear interactions produce instabilities that can dramatically affect the orbital dynamics of these systems.

The tide raised by a hot Jupiter excites large amplitude waves within its host star. These waves transfer energy and angular momentum from the orbit to the star and as a result the planet gradually spirals inward. The rate of orbital decay is determined by the efficiency of tidal dissipation and depends on the amplitude of the waves as well as the effectiveness of frictional processes within the star.

Tidal dissipation is often parametrized by the stellar tidal quality factor Q'_*, where larger Q'_* implies less dissipation. Perhaps the best constraints on Q'_* for solar-type stars come from the observed circularization rate of solar-type binaries, which yield $Q'_* \sim 10^6$ [102]. However, because Q'_* is not a fundamental property of the star (it depends on the shape and size of the orbit and the mass of the perturber), this result does not necessarily imply $Q'_* \sim 10^6$ for hot Jupiter systems. There have been a number of efforts to measure Q'_* from statistical modeling of the observed sample of hot Jupiters (see [107] for a review). [112] finds that the distribution favors $Q'_* \gtrsim 10^7$ for a specific set of assumptions about the initial conditions. [84] finds $10^6 \lesssim Q'_* \lesssim 10^9$ and notes that it is difficult to obtain tighter constraints because of the limited sample size and uncertainties in the initial period distribution and stellar age. Although there are no direct observational measurements of Q'_* from individual hot Jupiter systems (e.g. from the detection of orbital decay), [83] argues that the distribution shows evidence for ongoing removal and destruction by tides. In addition, [136] proposes that the observed dearth of close-in planets around fast-rotating stars [101] can be attributed to tidal ingestion of giant planets.

We study the orbital evolution of hot Jupiters due to the excitation and damping of tidally driven g-modes within solar-type host stars. Linearly resonant g-modes (the dynamical tide) are driven to such large amplitudes in the stellar core that they excite a sea of other g-modes through weakly nonlinear interactions ([20, 21]; [170], hereafter WAQB). These secondary waves can have much shorter wavelengths than the primary wave and, as a result, they can have much larger damping rates (due to radiative diffusion). Systems in which nonlinear interactions are important may therefore dissipate tidal energy much more rapidly than the linear theory estimates. By solving the dynamics of large networks of nonlinearly coupled modes, we show that the nonlinear dissipation rate of the dynamical tide is several orders of magnitude larger than the linear dissipation rate $(Q'_*)_{\rm lin} \sim 10^8 - 10^{10}$; [74, 108, 137]. We find stellar tidal quality factors $Q'_* \simeq 10^5 - 10^6$ for systems with planet mass $M_p \gtrsim 0.5 M_{\rm J}$ and orbital period $P \lesssim 2$ days, which implies that such systems decay on timescales that are small compared to the main-sequence lifetime of their hosts. According to our results, there are $\simeq 10$ currently known exoplanetary systems, including WASP-19b and HAT-P-36-b, with orbital decay timescales shorter than a Gyr. Rapid, tidally induced orbital decay may explain the observed paucity of planets with $M_p \gtrsim M_{\rm J}$ and $P < 2$ days around solar-type hosts and could generate detectable transit-timing variations in the near future.

For a planet with mass $M_p \gtrsim 3M_{\rm J}(P/{\rm day})^{-0.1}$ orbiting a solar-type star, the primary wave reaches such large amplitudes near the stellar center that it overturns the background stratification and breaks [19, 20]. In this *strongly* nonlinear regime, the primary wave deposits nearly all of its energy and angular momentum in a single group travel time through the star. The tidal dissipation rate therefore equals the energy flux of the initial, linearly driven primary wave. The three-dimensional numerical simulations of wave breaking by [19] yield $Q'_* \simeq 10^5 (P/1 \text{ day})^{2.8}$ for $M_p \gtrsim 3M_{\rm J}$ and a solar-type star. This corresponds to an inspiral time of ≈ 1 Gyr for a $3M_{\rm J}$ planet in a 2 day orbit.

For a planet with mass $0.5 \lesssim M_p/M_{\rm J} \lesssim 3$, the primary wave does not break. Nonetheless, it is sufficiently nonlinear that it excites many secondary waves near the stellar center. In this *weakly* nonlinear regime, the primary wave only deposits a fraction of its energy and angular momentum in a single group travel time. The value of that fraction, which determines the rate of tidal dissipation, depends on the detailed interaction between the primary wave and the sea of secondary waves. We calculate this interaction (and its saturation) in the weakly nonlinear regime. Similar types of analyses have been carried out in the context of the r-mode instability in spinning neutron stars [29, 32].

We begin with an example of resonant parametric instability (§2.1.1) before moving on to our specific application within main-sequence host stars in hot Jupiter systems.

2.1.1 The Mathieu Equation

The Mathieu equation provides a simple model that demonstrates parametric nonlinear instabilities [100]

$$\ddot{q} + \gamma\dot{q} + \left(\omega^2 - \kappa\cos t\right) q = 0 \tag{2.2}$$

Classically, the Mathieu equation does not contain a dissipative term ($\gamma\dot{q}$), but we include damping because of the obvious connection to stellar eigenmodes. We note that as $\kappa \to 0$, the system becomes a standard damped harmonic oscillator and is well understood; all solutions are stable and decay exponentially. However, the *parametric forcing* present when $\kappa \neq 0$ can destabilize the system and cause it to grow without bound. To find what conditions lead to this instability, we make a few simplifying assumptions: weak parametric driving ($\kappa = K\epsilon$), weak damping ($\gamma = \Gamma\epsilon^2$), near resonant forcing ($\omega^2 = 1 + W^2\epsilon^2$), and that there are two relevant time-scales related by $\tau \sim \epsilon^2 t$. This allows us to use multi-scale analysis techniques, sometimes called *two-timing* [23, 133].

In particular, we assume that the solution is a function of two time-scales

$$q = q(t,\tau) \Rightarrow \dot{q} = \frac{\partial q}{\partial t} + \frac{\partial \tau}{\partial t}\frac{\partial q}{\partial \tau} = q_t + \epsilon^2 q_\tau \tag{2.3}$$

We furthermore expand q as a power series in ϵ

$$q = q^{(0)} + \epsilon q^{(1)} + \epsilon^2 q^{(2)} + \cdots , \tag{2.4}$$

insert this ansatz into the Mathieu equation, and group similar powers of ϵ. This produces

$$\begin{aligned}
\epsilon^0 &: q^{(0)}_{tt} + q^{(0)} = 0 && (2.5)\\
\epsilon^1 &: q^{(1)}_{tt} + q^{(1)} = q^{(0)} K \cos t && (2.6)\\
\epsilon^2 &: q^{(2)}_{tt} + q^{(2)} = W^2 q^{(0)} - 2q^{(0)}_{t\tau} - \Gamma q^{(0)} + q^{(1)} K \cos t && (2.7)
\end{aligned}$$

and so on. By iteratively solving these equations, we find

$$\begin{aligned}
\epsilon^0 &: q^{(0)} = Ae^{it} + Be^{-it} && (2.8)\\
\epsilon^1 &: q^{(1)} = Ce^{it} + De^{-it} + \frac{K}{2}\left(A + B - \frac{A}{3}e^{2it} - \frac{B}{3}e^{-2it}\right) && (2.9)
\end{aligned}$$

which implies the ϵ^2 terms demonstrate resonant driving

$$\begin{aligned}
q^{(2)}_{tt} + q^{(2)} =& -2\left(iA_\tau e^{it} - iB_\tau e^{-it}\right) - \Gamma\left(iAe^{it} - iBe^{-it}\right) + W^2\left(Ae^{it} + Be^{-it}\right)\\
&+ \frac{K}{2}\left(e^{it} + e^{-it}\right)\left(Ce^{it} + De^{-it} + \frac{K}{2}\left(A + B - \frac{A}{3}e^{+2it} - \frac{B}{3}e^{-2it}\right)\right)\\
=& \left(-2iA_\tau - \Gamma iA + W^2 A + \frac{K^2}{4}\left(B + \frac{2}{3}A\right)\right)e^{it}\\
&+ \left(2iB_\tau + \Gamma iB + W^2 B + \frac{K^2}{4}\left(A + \frac{2}{3}B\right)\right)e^{-it}\\
&+ \text{nonresonant terms} \qquad (2.10)
\end{aligned}$$

Because resonant driving would cause $q^{(2)}$ to diverge, we require all resonant forcing to vanish and obtain

$$\begin{aligned}
2A_\tau &= -\Gamma A - i\left(W^2 A + \frac{K^2}{4}\left(B + \frac{2}{3}A\right)\right) && (2.11)\\
2B_\tau &= -\Gamma B + i\left(W^2 B + \frac{K^2}{4}\left(A + \frac{2}{3}B\right)\right) && (2.12)
\end{aligned}$$

This is just a linear eigenvalue problem for the coupled evolution of A and B. Assuming A, $B \propto e^{s\tau}$, we find

$$s = \frac{-\Gamma \pm \sqrt{\frac{5}{144}K^4 - \frac{1}{3}K^2 W^2 - W^2}}{2} \qquad (2.13)$$

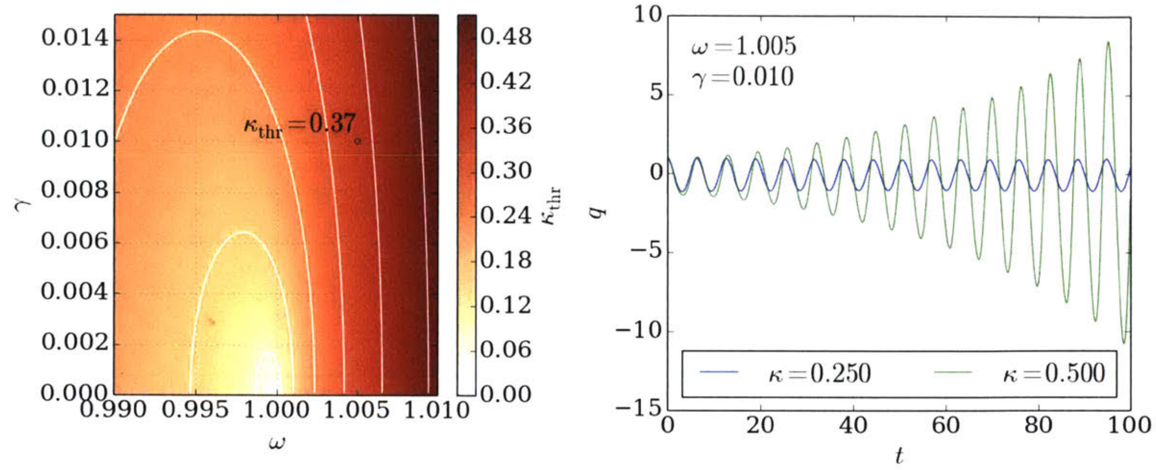


Figure 2-1: (left) Threshold coupling strengths to induce instability as a function of the Mathieu equation's parameters. (right) Time evolution of similar systems below and above the instability threshold. We see that both larger damping rates and nonresonant forcing raise the stability threshold.

We note that when $K \to 0$, we can set $W^2 \to 0$ without loss of generality and obtain the expected result $s = -\Gamma/2$; the solution decays exponentially as $q^{(0)} \propto e^{-\gamma t/2}$. However, as $K \to \infty$, we obtain solutions with $s > 0$ implying exponential growth of q. Note also that the threshold coupling necessary for instability is larger further from resonance (larger W^2). To wit, the threshold coupling is

$$K_{\rm thr}^2 = \frac{72}{15}\left(W^2 + \sqrt{\frac{9}{4}W^4 + \frac{5}{4}\Gamma^2}\right) \tag{2.14}$$

We therefore conclude that, near resonance, parametric forcing can cause exponential growth. Fig. 2-1 demonstrates how the instability threshold depends on the system's parameters. A similar resonant instability manifests in 3-mode interactions between g-modes within sun-like stars.

2.1.2 Resonant 3-mode Parametric Instabilities

We analyze eigenmodes within sun-like stars using a slightly different representation of the amplitude equations, sometimes called the *phase-space* representation (see WAQB). While still based on a decomposition of the Lagrangian fluid displacements into eigenmodes, we assign a separate amplitude to the positive and negative frequency components of each mode. Therefore, we refer to each mode q_α via a combination of radial number (n), angular numbers (l and m), and frequency

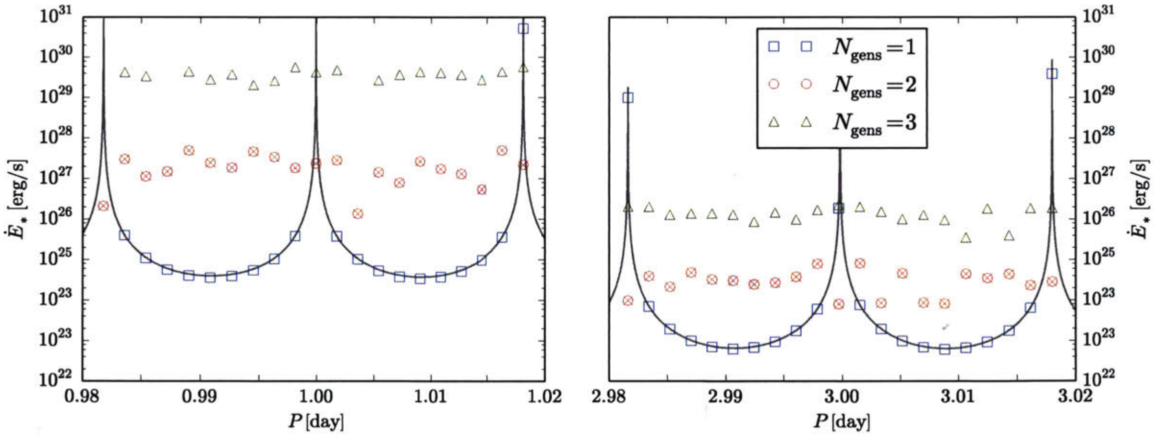


Figure 2-2: Resonance diagrams showing the dissipation within the host star as a function of the number of generations included in the network for a Jupiter-mass companion in a (left) 1 day and (right) 3 day orbit around a sun-like host. Grey lines represent the analytic prediction for the linear response alone.

sign (±). The fluid motion is then written as a sum over these amplitudes via

$$\begin{bmatrix} \vec{\xi}(\vec{x},t) \\ \dot{\vec{\xi}}(\vec{x},t) \end{bmatrix} = \sum_\alpha q_\alpha(t) \begin{bmatrix} \vec{\xi}(\vec{x}) \\ -i\omega_\alpha \vec{\xi}(\vec{x}) \end{bmatrix} \tag{2.15}$$

where the sum over α extends over all n, l, m, and frequency signs. We note that q_α can be complex, in general.

This representation yields the following coupled set of 1^{st}-order ordinary differential amplitude equations

$$\dot{q}_\alpha + (\gamma_\alpha + i\omega_\alpha) q_\alpha = i\omega_\alpha U_\alpha(t) + i\omega_\alpha \sum_{\beta,\epsilon} \kappa_{\alpha\beta\epsilon} q^*_\beta q^*_\epsilon \tag{2.16}$$

where (*) denotes complex conjugation [63]. We neglect nonlinear tidal forcing ($U_{\alpha\beta}$) and 3-mode couplings between the dynamical tide and the equilibrium tide. These two terms strongly cancel and therefore are not expected to effect the system's dynamics (see WAQB and [164]). This network of weakly nonlinear oscillators exhibits resonant instabilities similar to those observed in the Mathieu equation (§2.1.1) and we study the dynamics of these networks. We note that U_α represents the generalized force applied to q_α by the tidal potential and γ_α represents a dissipative interaction. For g-modes within sun-like stars, the dominant source of damping is believed to be radiative diffusion [137].

The coupling coefficients (κ) represent the inner product of three spatial mode shapes integrated against the background stellar structure (see WAQB). Because we decompose the fluid motion into spherical harmonics, this imposes angular selection rules. Specifically, we require $m_\alpha + m_\beta + m_\epsilon = 0$ and $|l_\alpha - l_\beta| \leq l_\epsilon \leq l_\alpha + l_\beta$ (the triangle rule). What's more, there are radial resonance conditions. If the modes oscillate rapidly relative to each other, their product will oscillate and the integrand will not accumulate significantly. However, if the difference of the daughter radial mode numbers is less than the parent's, the integrand can accumulate through a process similar to beating harmonics (Appendix B.1 of WAQB). This imposes relatively strict selection rules on which (n, l, m) couple strongly.

Furthermore, most of the coupling occurs in the stellar core where the mode amplitudes are geometrically focused to large values. Beyond that, coupling happens near the parent's inner turning point, where the parent mode shape is evanescent (non-oscillatory) and the product of the daughter modes' oscillations can produce a slowly varying integrand. We note that a linear Brunt profile in the region near the stellar core is important for the latter so that lower frequency daughters still oscillate at the parent's inner turning point.

These networks of modes represent the fluid motion within the host star and model how it is excited by its companion. Fig. 2-2 shows resonance diagrams for the energy dissipated by networks of modes within the host (see §2.1.3). Indeed, we see a large increase in dissipation ($\dot{E}$) when we include multiple generations of nonlinearly coupled modes. We now analyze the stability of these modes and the growth rates of associated instabilities.

3-mode Stability and Growth Rates

We begin by analyzing the stability of the linear solution, focusing particularly on the dynamical tide (§1.2). Fig. 2-3 shows the cumulative number of unstable 3-mode pairs in a sun-like host with a 0.1 M_J companion in a $\sim$ 3 day orbit; there are nearly 10^3 unstable daughter pairs and many more unstable granddaughters (§2.1.3).

For simplicity, assume the dynamical tide is composed of a single resonant g-mode. We focus on circular orbits so that $U_\alpha(t) = U_\alpha e^{-im_\alpha \Omega t}$, where $\Omega = 2\pi/P$ is the orbital angular frequency

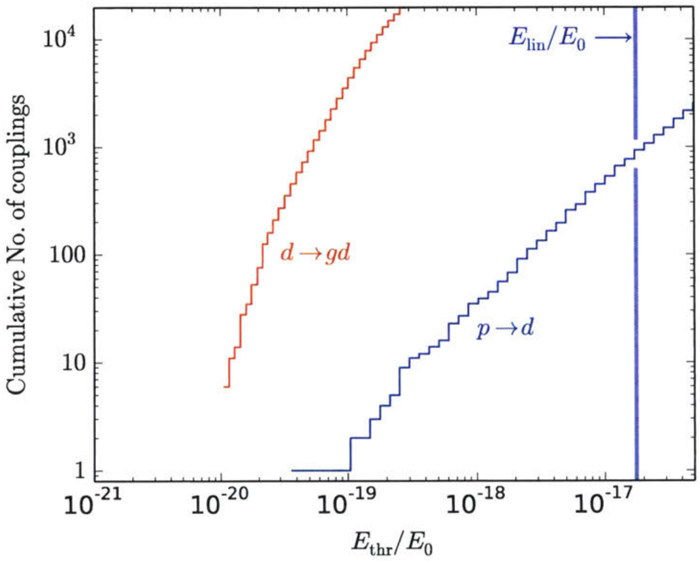


Figure 2-3: The cumulative number of unstable parent-daughter ($p \to d$) pairs of modes and daughter-granddaughter ($d \to gd$) pairs as a function of the instability threshold for a sun-like host with a $0.1M_J$ companion in a ~ 3 day orbit (about half-way between a resonance peak and trough).

and m_α is q_α's azimuthal number. The steady-state linear solution is then

$$q_\alpha = \frac{i\omega_\alpha U_\alpha}{\gamma_\alpha + i(\omega_\alpha - m_\alpha\Omega)} e^{-im_\alpha\Omega t}. \tag{2.17}$$

We now examine the stability of this solution assuming a single pair of *daughter* modes is coupled to the *parent*. The daughters' equations of motion, assuming $\beta \neq \epsilon$, become

$$\dot{q}_\beta + (\gamma_\beta + i\omega_\beta)q_\beta = 2i\omega_\beta\kappa_{\alpha\beta\epsilon}q_\alpha^* q_\epsilon^* \tag{2.18}$$

and analogously for $\beta \leftrightarrow \epsilon$. Here, we've assumed $q_\beta,\ q_\epsilon \ll q_\alpha$ and that q_α is exactly the linear equilibrium result. We have also neglected the linear forcing for the daughter mode ($U_{\beta,\epsilon}$) because it will be highly nonresonant and therefore should not contribute to the dynamics meaningfully. Furthermore, if the daughter mode is self-coupled ($\beta = \epsilon$), we can immediately rearrange this to obtain a modified Mathieu equation. We therefore expect parametric forcing to produce instabilities near resonance.

Writing each mode amplitude as $x = qe^{-i(\omega-\Delta)t}$, we can determine the resonance conditions and stability of this system. Importantly, we find that the time-dependence drops out of the system completely if $m_\alpha\Omega + \omega_\beta + \omega_\epsilon = \Delta_\beta + \Delta_\epsilon$; minimizing $\Delta_\beta + \Delta_\epsilon$ turns out to be the resonance condition for the instability. This yields

$$\dot{x}_\beta + (\gamma_\beta + i\Delta_\beta)x_\beta = 2i\omega_\beta\kappa_{\alpha\beta\epsilon}\left(\frac{-i\omega_\alpha U_\alpha}{\gamma_\alpha - i(\omega_\alpha - m_\alpha\Omega)}\right)x_\epsilon^* \tag{2.19}$$

and analogously for $\beta \leftrightarrow \epsilon$. Again, we obtain the solution as an eigenvalue assuming $x_\beta,\ x_\epsilon \sim e^{st}$ (which renders the problem algebraic), where s is implicitly determined by

$$(s + \gamma_\beta + i\Delta_\beta)(s + \gamma_\epsilon - i\Delta_\epsilon) = 4\omega_\beta\omega_\epsilon\kappa_{\alpha\beta\epsilon}^2 A_\alpha^2 \quad \Bigg| \quad A_\alpha^2 = \frac{\omega_\alpha^2 U_\alpha^2}{\gamma_\alpha^2 + (\omega_\alpha - m_\alpha\Omega)^2} \tag{2.20}$$

We note that as $A_\alpha \to \infty$, the system becomes unstable and $\mathcal{R}\{s\} \to 2\sqrt{\omega_\beta\omega_\epsilon}\kappa_{\alpha\beta\epsilon}|A_\alpha|$, meaning that the growth rate of the instability is proportional to the amplitude of the parent. From this, we can also determine the threshold amplitude for stability.

$$(A_\alpha)^2_{\text{thr}} = \frac{\gamma_\beta\gamma_\epsilon + \Delta_\beta\Delta_\epsilon}{4\omega_\beta\omega_\epsilon\kappa_{\alpha\beta\epsilon}^2} = \frac{\gamma_\beta\gamma_\epsilon}{4\omega_\beta\omega_\epsilon\kappa_{\alpha\beta\epsilon}^2}\left(1 + \left(\frac{m_\alpha\Omega + \omega_\beta + \omega_\epsilon}{\gamma_\beta + \gamma_\epsilon}\right)^2\right) \tag{2.21}$$

where we've used the vanishing imaginary part from Eqn. 2.20 in the last step to determine $\Delta_\beta \gamma_\epsilon = \Delta_\epsilon \gamma_\beta$. From this, we clearly see the competition between resonance ($\Delta_\beta \Delta_\epsilon$) and damping ($\gamma_\beta \gamma_\epsilon$) which determines stability, similar to what was seen in §2.1.1. We can therefore determine whether a 3-mode triple will be unstable by comparing the parent's linear amplitude to this threshold.

Collective Stability and Growth Rates

Of course, real stars contain more than three eigenmodes, and those modes are coupled through a complicated network. For high-order g-modes, there may be many modes with similar properties (ω, γ, etc) corresponding to neighboring radial numbers (n). Therefore, if one of them participates in an unstable 3-mode triple, it is likely that all of them will participate in an unstable triple. In general, we should consider the stability of the collection of g-modes rather than restricting ourselves to the individual 3-mode sets. We investigate various different coupling scenarios, finding similar behavior for all of them. Larger networks of coupled modes can often have faster growth rates than 3-mode pairs along with suppressed stability thresholds. Nonetheless, we find that a few highly resonant 3-mode pairs generally dominate the long-term behavior of the system (Fig. 2-5).

The general behavior of a set of N collectively coupled modes resembles the behavior of a single unstable self-coupled mode with the coupling coefficient enhanced by a factor of N (lower A_{thr} and larger growth rate). This is because $\dot{q} \sim \sum_\beta \kappa q_p^* q_\beta^* \sim N \kappa q_p^* q^*$ if all daughters are similar. Appendix A provides more details on how this behavior can be derived rigorously.

Effects of Multiple Parents

If we linearize around the linear-equilibrium solution involving multiple parents, the equations of motion for the daughter modes become

$$\partial_t q_\beta + (i\omega_\beta + \gamma_\beta) = 2i\omega_\beta q_\epsilon^* \sum_{p \,\in\, \mathrm{parents}} \kappa_{p\beta\epsilon} q_p^* \tag{2.22}$$

We can analyze this system as if there is a single parent with complex amplitude

$$\kappa q^* = \sum_{p \,\in\, \mathrm{parents}} \kappa_{p\beta\epsilon} q_p^*. \tag{2.23}$$

We note the possibility for parent modes to interfere with one another when driving daughter modes, possibly rendering daughters stable under multi-parent driving when they were unstable to any individual parent. Most notably, if the parents are regularly spaced in frequency and driven at the midpoint between resonance peaks, there can be strong destructive interference. This is because each parent is paired with a partner on the opposite side of the driving frequency, and each pair consists of parents oscillating with nearly opposite phase. This narrow trap in the resonance troughs is readily apparent at orbital periods above 4 days for a solar-mass host of a Jupiter-mass companion. However, we did not observe significant trapping below $\sim$ 4 day orbital periods, where we focus our attention for this study. This may be due to the asymmetric spacing of resonances, which will destroy this near perfect cancellation or the amplitudes being large enough to overcome any cancellation that was present. In the hot Jupiter context, this issue is probably only of theoretical interest because the orbital evolution time scales are $\gtrsim 10^{11}$ yr for $P \gtrsim 3$ day, even for massive companions.

Nonlinear Equilibria

Now that we have shown how and when sets of modes can become unstable, we consider whether they approach new steady states. As it turns out, we can solve for the simple 3-mode system's non-linear equilibrium. Because the intuition gained is useful, we reproduce the solution here beginning with the equations of motion for our transformed coordinates $x = qe^{-i(\omega-\Delta)t}$

$$\begin{aligned}
\dot{x}_\alpha + (i\Delta_\alpha + \gamma_\alpha)x_\alpha =& i\omega_\alpha U_\alpha e^{-i(m_\alpha\Omega-\omega_\alpha+\Delta_\alpha)t} \\
&+ 2i\omega_\alpha \kappa_{\alpha\beta\epsilon} x^*_\beta x^*_\epsilon e^{i(\omega_\alpha+\omega_b eta+\omega_\epsilon-\Delta_\alpha-\Delta_\beta-\Delta_\epsilon)t} && (2.24) \\
\dot{x}_\beta + (i\Delta_\beta + \gamma_\beta)x_\beta =& 2i\omega_\beta \kappa_{\alpha\beta\epsilon} x^*_\alpha x^*_\epsilon e^{i(\omega_\alpha+\omega_\beta+\omega_\epsilon-\Delta_\alpha-\Delta_\beta-\Delta_\epsilon)t} && (2.25) \\
\dot{x}_\epsilon + (i\Delta_\epsilon + \gamma_\epsilon)x_\epsilon =& 2i\omega_\epsilon \kappa_{\alpha\beta\epsilon} x^*_\alpha x^*_\beta e^{i(\omega_\alpha+\omega_\beta+\omega_\epsilon-\Delta_\alpha-\Delta_\beta-\Delta_\epsilon)t} && (2.26)
\end{aligned}$$

Again, we can remove all time dependence from these equations by setting

$$m_\alpha\Omega = \omega_\alpha - \Delta_\alpha = \Delta_\beta + \Delta_\epsilon - \omega_\beta - \omega_\epsilon \tag{2.27}$$

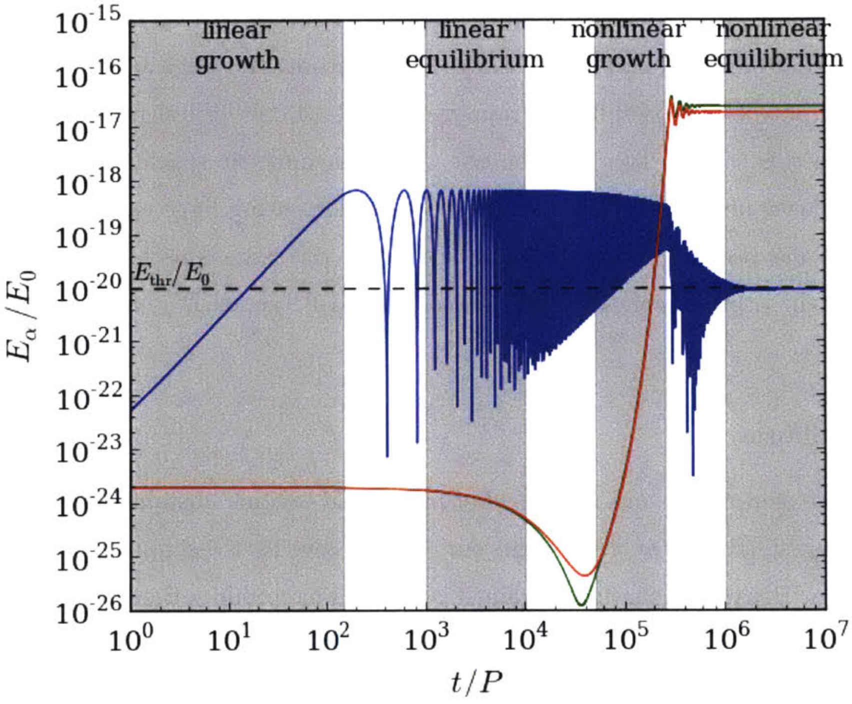


Figure 2-4: 3-mode system's evolution toward nonlinear equilibrium starting from randomized initial conditions. We note that there is a clear separation of time-scales between the linear growth rates, 3-mode growth rates, and linear damping rates. We note that the parent's amplitude oscillates as the homogeneous solutions decay and it approaches the linear equilibrium. Only once it spends a significant amount of time above $E_{\rm thr}$ do the daughter modes become unstable.

and assuming constant amplitudes ($\partial_t \rightarrow 0$). This renders the problem algebraic. We summarize the main results below and refer to Appendix B of [63] for a more complete derivation.

$$\Delta_\beta \gamma_\epsilon = \Delta_\epsilon \gamma_\beta \tag{2.28}$$

$$\frac{\Delta_\beta \Delta_\epsilon + \gamma_\epsilon \gamma_\beta}{4\omega_\beta \omega_\epsilon \kappa_{\alpha\beta\epsilon}^2} = A_\alpha^2 = (A_\alpha)_{\text{thr}}^2 \tag{2.29}$$

$$\left(\frac{A_\beta}{A_\epsilon}\right)^2 = \frac{\gamma_\epsilon \omega_\beta}{\gamma_\beta \omega_\epsilon} \tag{2.30}$$

$$\delta = \delta_\alpha + \delta_\beta + \delta_\epsilon \tag{2.31}$$

$$\gamma_\beta \frac{A_\beta}{A_\epsilon} = 2\omega_\beta \kappa_{\alpha\beta\epsilon} A_\alpha \sin\delta \tag{2.32}$$

$$\Delta_\beta \frac{A_\beta}{A_\epsilon} = 2\omega_\beta \kappa_{\alpha\beta\epsilon} A_\alpha \cos\delta \tag{2.33}$$

$$A_\beta A_\epsilon = \left(\frac{A_\alpha}{2\omega_\alpha \kappa_{\alpha\beta\epsilon}}\right)\left(\Delta_\alpha \cos\delta + \gamma_\alpha \sin\delta \pm \sqrt{(\Delta_\alpha \cos\delta + \gamma_\alpha \sin\delta)^2 + \frac{\Delta_\alpha^2 + \gamma_\alpha^2}{A_\alpha^2}\left(\frac{\omega_\alpha^2 U_\alpha^2}{\Delta_\alpha^2 + \gamma_\alpha^2} - A_\alpha^2\right)}\right) \tag{2.34}$$

where $x = Ae^{i\delta}$ and the sign in Eqn. 2.34 depends on the sign of $\omega_\alpha \kappa_{\alpha\beta\epsilon}$. The key bits of insight from this are that the parent amplitude drops to the instability threshold and that, although it is not immediately obvious, the daughter equilibrium amplitudes are typically larger than the parent's linear amplitude. This means there is an inversion of the mode amplitudes; the parent decays as it pumps energy into the daughter modes. Furthermore, this fixed point is stable and perturbations are typically damped out on time-scales set by the linear damping rates ($\sim \gamma^{-1}$), which are long compared to dynamical time-scales set by the 3-mode growth rate ($\sim (\omega\kappa A)^{-1}$). Fig. 2-4 demonstrates these features.

Just as for stability, we should consider what the eventual steady-state will be for a large collection of unstable modes instead of an isolated 3-mode triple. Generally, if collective modes are unstable, they will have faster growth rates than 3-mode instabilities. Therefore, collective instabilities will dominate the dynamics over short time-scales. We expect the parent amplitude to decrease to the threshold energy of the associated instability that currently dominates the dynamics. Typically, there are a few extremely resonant 3-mode triples corresponding to very small A_{thr}, much smaller than the collective modes. This means that the 3-mode instabilities will still be excited even when the collective mode has suppressed the parent's amplitude; they will slowly grow until they

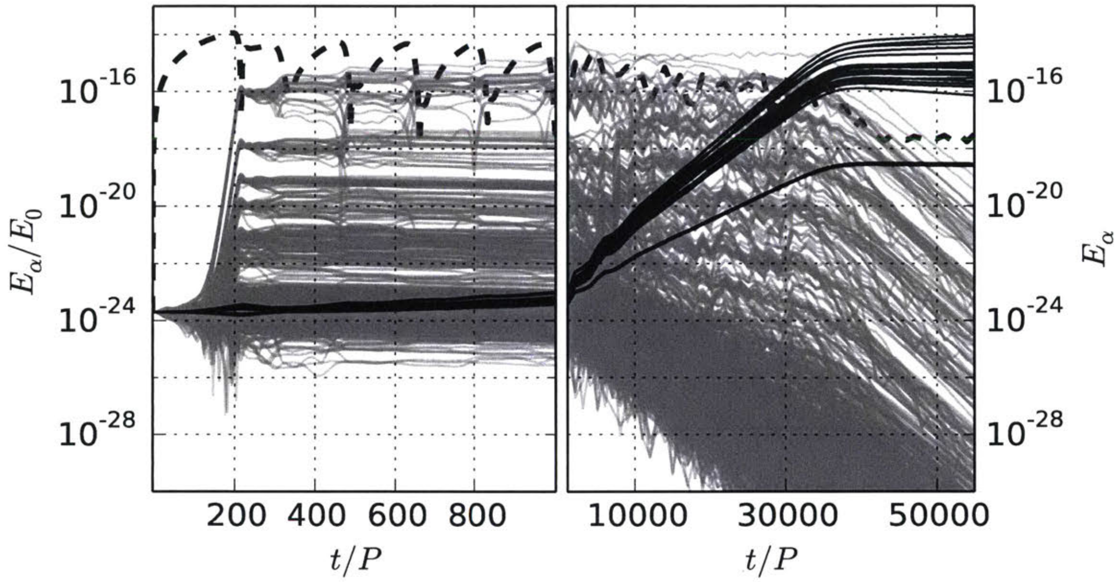


Figure 2-5: 3-mode and collective instabilities' evolution toward nonlinear equilibrium. (left) Even though the collective mode's (grey) growth rate is *much* faster than the 3-mode (black) growth rate, the 3-mode equilibrium eventually suppresses the parent (dashed) amplitude enough to render the collective mode stable, thereby dominating the long-term steady state (right).

eventually surpass the collective mode's amplitude and drive the parent to even lower amplitudes. At this point, the collective mode becomes stable and will decay. The system's long-term behavior will therefore be dominated by the lowest threshold 3-mode pair, which is demonstrated in Fig. 2-5.

This intuition holds when there is only one generation of parent-child couplings. As we will see, the intuition is still valuable but the steady state changes dramatically when more generations are introduced.

2.1.3 Dynamics of Large Networks

Based on the observation that the 3-mode triple with the lowest E_{thr} is expected to dominate the long term dynamics in simple systems, we compose more complicated networks by adding in couplings in order of increasing E_{thr}. In this way, we approximate the dynamics of the full star by simulating only those modes we expect to be most dynamically relevant. Throughout this study, we ensured our numerical searches for low E_{thr} triples spanned enough modes that the exact boundary of the search did not affect the results.

From Eqn. 2.21, we see there is a natural competition between the detuning and the damping when minimizing E_{thr}. Typically, to reduce the detuning we must choose higher l modes. These have larger damping rates, and therefore some balance must be reached depending on the particular stellar model. We find that solar-mass hosts support g-modes spaced closely enough that small detunings are possible even with low l. However, lower mass stars have coarser mode spacing and lower damping rates, meaning higher l modes can be important (as high as $l \gtrsim 10$). Nonetheless, when we reach several generations into the network (grandchildren, great grandchildren, etc), we find that lower l are again favored, although there remains a larger tail to higher l for low mass hosts than for solar-mass hosts.

The intuition gained from simple 3-mode systems is relevant for more complicated networks as well. In particular, the inversion of the parent and daughter amplitudes is key. When the daughters are driven to large amplitudes, they excite their own children (granddaughters). When the granddaughters reach large amplitudes, they siphon energy from the daughters, causing the daughters' amplitudes to collapse. This allows the parent to recover and causes the granddaughters to decay. When the parent has recovered, the entire cycle restarts and we observe a sawtooth pattern in the amplitudes of these modes. What's more, adding additional generations does not significantly affect the global dynamics of the system; they simply cause the granddaughter to decay more rapidly. This is reminiscent of a turbulent cascade from long wavelength perturbations to smaller length scales, where global behavior is insensitive to the precise micro-physics of the smallest length scale. What's more, we find that the global dissipation is insensitive to the linear damping rates of the youngest generation (as long as those modes remain unstable); it is instead the nonlinear growth rate that drives the dissipation time-scale.

Networks with several generations quickly establish a steady-state solution, which often appears chaotic (Fig. 2-6). The individual trajectories of separate modes depend on the initial conditions and the network's structure, but global properties of the steady state solution are robust. For instance, $\dot{E}$ may be the same if we have 100 modes dissipating an equal amount each, or only 10 modes but each dissipating 10 times as much. Fig. 2-7 shows the convergence of $\dot{E}$ with many different network structures, numbers of modes, and numbers of generations.

A reasonable question is what sets the scale for the global dissipation and how many generations are involved in the cascade. A simple estimate of $\dot{E} \sim 2\Gamma_{3\mathrm{mode}} E_{\mathrm{break}}$ gives a reasonable approximation for the dissipation, and at least approximately captures the correct scaling (see Appendix B for more detail). We see that $\dot{E}$ is independent of the linear damping rates ($\Gamma_{\mathrm{3mode}} \sim 2\omega\kappa A_{\mathrm{parent}}$) and is naturally associated with the cyclic growth and breaking of g-modes. This occurs when the fluid displacement is of the same size as the wavelength ($\xi_r k_r \sim 1$), and the g-mode overturns the density stratification within the star. We note that E_{break} depends on the mode's angular number (l), and with all else equal we observe roughly the correct scaling when lower host mass networks naturally select larger l modes (higher l modes penetrate deeper into the stellar core, are more geometrically focused, and break at lower amplitudes).

We can also estimate how many generations would participate in the cascade based on the WKB scaling relations for the mode and coupling parameters. For stability thresholds dominated by the detuning ($\Delta \sim \omega/n$), $E_{\mathrm{thr}} \sim \omega^6$ and daughters (with $\omega \sim \omega_{\mathrm{parent}}/2$) are even more unstable than their parents. This means that unstable parents will almost certainly excite unstable children. However, for damping dominated stability (modes are spaced closely enough that the detuning almost always vanishes); we have $E_{\mathrm{thr}} \sim \omega^{-2}$ and daughters are *more* stable than their parents. This turn-over typically happens after $\sim$ 5 or more generations for stellar-mass hosts, but we find we can model the global dissipation with only 3 generations (linear parents, daughters, and granddaughters).

Because there is no great change between granddaughter networks and larger networks, we define reference networks with up to granddaughters for computational reasons. We use these throughout our study to model the dissipation from larger network in a tractable way. Fig. 2-7 shows how our reference networks compare to others.

2.1.4 Sun-like Hosts

We now focus on the implications of these resonant 3-mode instabilities within sun-like stellar hosts. Specifically, we study main sequence stars with radiative cores and convective envelopes ($M_* \in [0.5, 1.1]M_\odot$). Because we are interested in the system's long-term orbital evolution, we compute time-averaged statistics by sampling over several orbital periods spanning at least one resonance peak.

$$\langle X \rangle = \frac{\int dt\, X}{\int dt} = \frac{\int dP\, \dot{P}^{-1} X}{\int dP\, \dot{P}^{-1}} \tag{2.35}$$

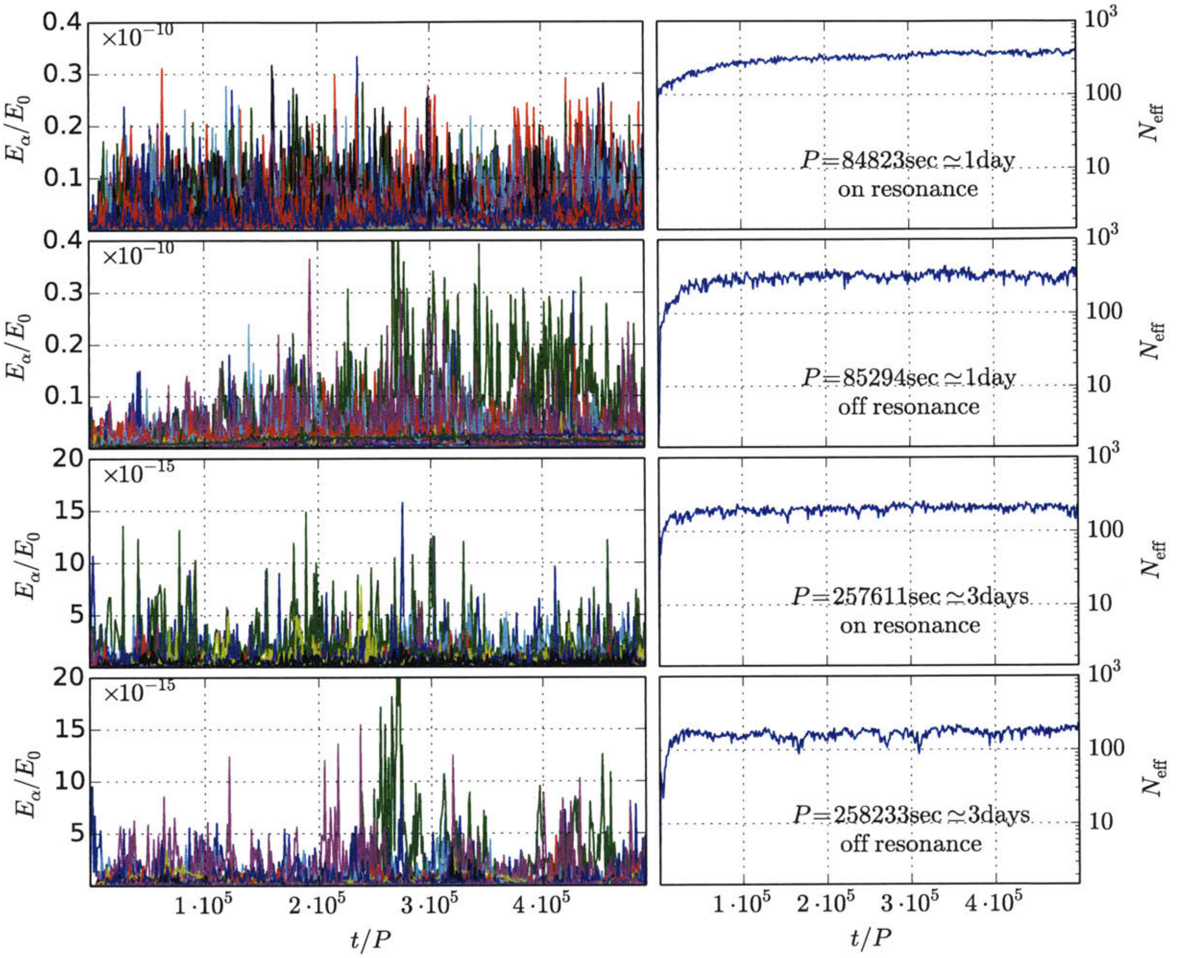


Figure 2-6: (left) Mode amplitudes and (right) effective number of modes contributing to dissipation as a function of time. We observe chaotic cycles for individual mode amplitudes, but the networks' overall behaviors saturate relatively quickly.

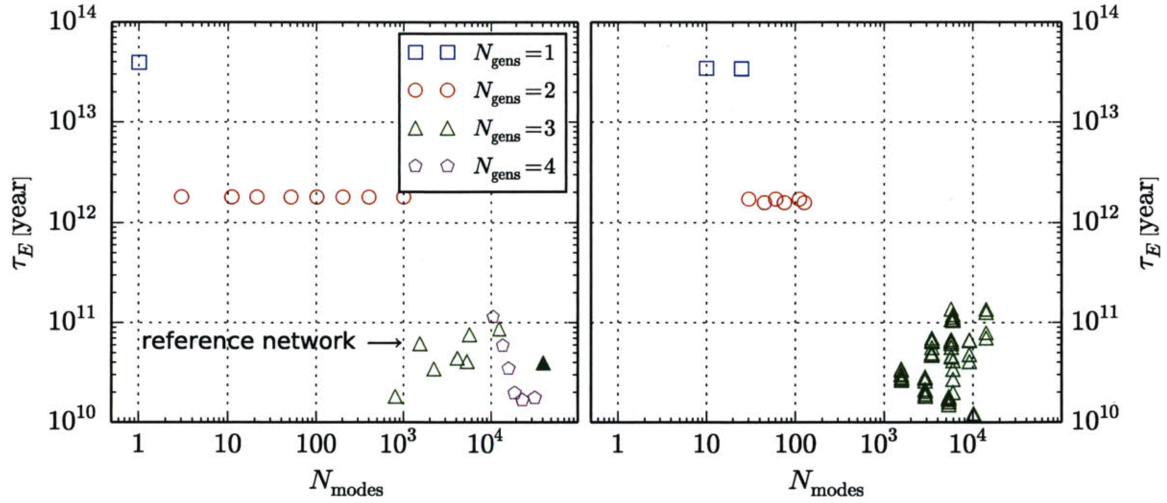


Figure 2-7: Saturation of network dissipation as a function of N_{mode} and N_{gen} for a Jupiter mass companion in a 3 day orbit (257928 sec, chosen to be approximately half way between a resonance trough and peak) around a solar mass host with (left) a single parent and (right) multiple parents. Although subject to some modeling uncertainties, we observe saturation after only 3 generations, independent of the number of parents.

Furthermore, we can estimate $\dot{P}$ from our numerical simulations of large networks. From energy balance, we expect

$$\frac{d}{dt}\left(E_{\text{orb}} + E_{\text{int}} + E_{\text{rot}} + E_{\text{modes}}\right) = -2\sum_{\alpha}\gamma_\alpha E_\alpha \tag{2.36}$$

where E_{orb}, E_{int}, E_{rot}, and E_{modes} are the energies associated with the orbit, the tidal interaction, rotation of the star/planet, and the stellar eigenmodes, respectively. From this, we can determine $\dot{P}$ via

$$\dot{P} = -2\left(\frac{dE_{\text{orb}}}{dp} + \frac{dE_{\text{int}}}{dP} + \frac{dE_{\text{rot}}}{dP} + \frac{dE_{\text{modes}}}{dP}\right)^{-1}\sum_{\alpha}\gamma_\alpha E_\alpha \tag{2.37}$$

$$\approx -2\left(\frac{dE_{\text{orb}}}{dP}\right)^{-1}\sum_{\alpha}\gamma_\alpha E_\alpha \tag{2.38}$$

$$\propto \frac{P}{E_{\text{orb}}}\sum_{\alpha}\gamma_\alpha E_\alpha \tag{2.39}$$

We recognize that $|dE_{\text{orb}}/dP| \gg |dE_{\text{int}}/dP|$, $|dE_{\text{rot}}/dP|$, $|dE_{\text{modes}}/dP|$ because E_{orb} is so much larger than any other energy scale that even small changes in it dominates over other effects. From this we obtain our estimate

$$\langle X\rangle \approx \frac{\sum_P \left(\frac{P}{E_{\rm orb}}\sum_\alpha \gamma_\alpha E_\alpha\right)^{-1} X}{\sum_P \left(\frac{P}{E_{\rm orb}}\sum_\alpha \gamma_\alpha E_\alpha\right)^{-1}} \tag{2.40}$$

We use this to determine the time-averaged orbital decay time-scale by marginalizing over the neighborhood of a resonance peak. Specifically, we compute

$$\langle \tau_E\rangle = \left\langle \frac{E_{\rm orb}}{2\sum_\alpha \gamma_\alpha E_\alpha}\right\rangle . \tag{2.41}$$

This timescale is often reported in terms of the star's tidal quality factor (Q'_*), which is defined as [73, 84]

$$Q'_* = 7.5\times 10^6 \left(\frac{\langle\tau_E\rangle}{\rm Gyr}\right)\left(\frac{M_p}{M_J}\right)\left(\frac{P}{\rm day}\right)^{-13/3} . \tag{2.42}$$

Theoretical estimates for Q'_* range from 10^5–10^{12} in the literature.

Analytic Approximations

The linear dissipation rate of individual resonant modes is $\dot{E}_{\rm lin} \simeq 2\gamma_\alpha E_{\rm lin}$, where $E_{\rm lin}$ is the linear equilibrium energy of mode α. After summing over many parents near the resonance, using the WKB estimates for the damping and forcing coefficients, and time-averaging, we find

$$\langle \tau_E\rangle_{\rm lin} \simeq 1.4\times 10^{12}\left(\frac{M}{M_\odot}\right)^{-5/6}\left(\frac{R}{R_\odot}\right)^{11/2}\left(\frac{M_p}{M_j}\right)^{-1}\left(\frac{P}{\rm day}\right)^{3} {\rm yr} \tag{2.43}$$

We can also estimate the nonlinear dissipation rate of networks consisting of only parents and daughters (but not granddaughters, etc). This is because the dissipation in that case is dominated by the single daughter pair (β, ϵ) with the lowest instability threshold $E_{\rm thr}$. For the parameters of a typical hot Jupiter system, the nonlinear equilibrium energy of such a daughter pair is $E_{\beta,\epsilon} \simeq |U_\alpha/2\kappa_{\alpha\beta\epsilon}|E_0$ [170]. The total dissipation rate of the system is approximately the dissipation due to these two daughters $\dot{E}_{\text{p-d}} \simeq 2\times 2\gamma_{\beta,\epsilon}E_{\beta,\epsilon}$. There is a small correction to this because the lowest $E_{\rm thr}$ daughters have slightly different parameters and therefore do not sit at exactly the same amplitudes. After accounting for this small correction, we find

$$\langle \tau_E\rangle_{\text{p-d}} \simeq 2.0\times 10^{11}\left(\frac{\Lambda^2_{\beta,\epsilon}}{2}\right)^{-1}\left(\frac{T_{\alpha\beta\epsilon}}{0.2}\right)\left(\frac{M}{M_\odot}\right)^{-11/6}\left(\frac{R}{R_\odot}\right)^{11/2}\left(\frac{P}{\rm day}\right)^{19/6} {\rm yr} \tag{2.44}$$

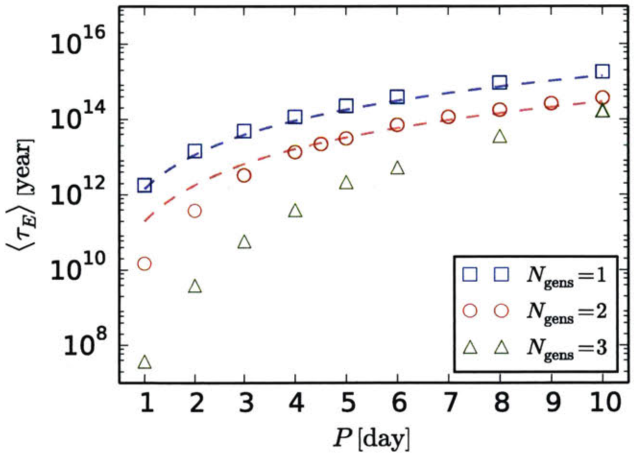


Figure 2-8: Analytic approximations for orbital decay time-scales along with numerical results. The disagreement between analytic and numeric estimates for $N_{\text{gens}} = 1$ is due to our assumption that $l = 1$ for all daughter modes, which breaks down at low P.

where $\Lambda^2 = l(l+1)$. Here we assumed $l_{\beta,\epsilon} = 1$, which is representative of the typical lowest E_{thr} daughters for $P \gtrsim 2$ days within solar-mass hosts. We find good agreement between parent-daughter network simulations that include many daughters and this analytic estimate. However, at these shorter periods, the available daughter modes are spaced further apart in frequency and the lowest E_{thr} pair may be pushed to $l_{\beta,\epsilon} > 1$. This causes the small discrepancy between the numeric and analytic estimates seen in Fig. 2-8; large l modes dissipate more energy and reduce $\langle \tau_E \rangle$.

Numerical Results

We use our reference networks to scan large parts of parameters space, averaging over the dissipation across several resonance peaks near each orbital period. This allows us to numerically determine $\langle \tau_E \rangle$, which we show in Fig. 2-9, and we immediately notice a few clear trends. First, there is a sharp cut-off at low M_p, but we see this near $M_p \sim 0.5 M_J$ rather than $\sim 3 M_J$, as was predicted in

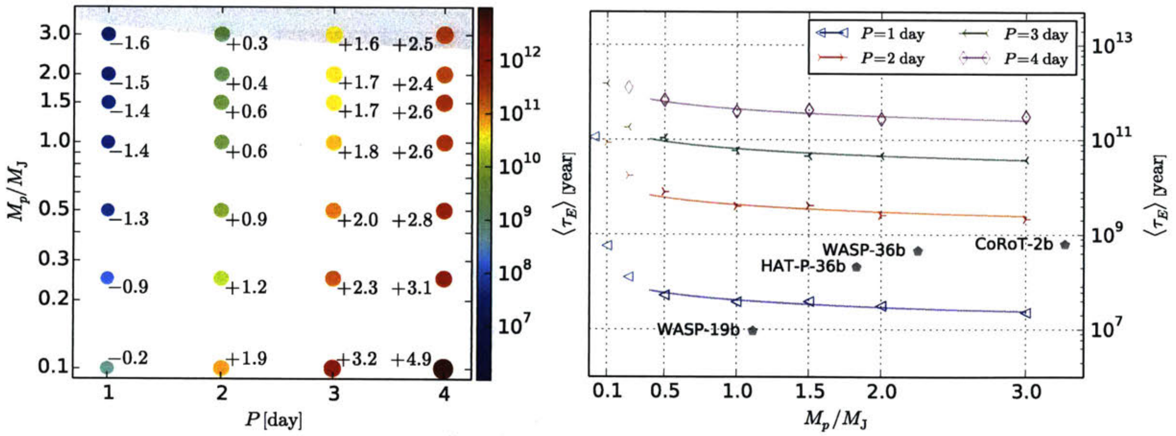


Figure 2-9: Several representations of the measured $\langle\tau_E\rangle$ and our fitting function for a solar mass host. The simple power law in Eqn. 2.45 captures the behavior well above $M_p \sim 0.5M_J$ for all periods considered.

the literature [19, 20, 21]. We also find a relatively simple fitting formula

$$\langle\tau_E\rangle = 4.1 \times 10^7 \left(\frac{M_p}{M_J}\right)^{-0.5} \left(\frac{P}{\text{day}}\right)^{6.7} \text{ years} \tag{2.45}$$

$$Q'_* = 3.0 \times 10^5 \left(\frac{M_p}{M_j}\right)^{0.5} \left(\frac{P}{\text{day}}\right)^{2.4} \tag{2.46}$$

which reasonably approximates the numerical results, but should not be trusted beyond a factor of ~ 2 due to both statistical and systematic modeling uncertainties. We note a relatively weak dependence on M_p once these instabilities become active and a very strong dependence on P. This means closer orbits, almost independent of M_p, are more promising for detecting tidally driven orbital decay.

Decay Timescales for Known Systems

Based on the Exoplanet Orbit Database [1], we studied 11 known planets orbiting approximately solar-type stars ($M = 1.0 \pm 0.1 M_\odot$ and $T_{\text{eff}} \simeq 5500$ K) with decay times $\langle\tau_E\rangle < 1$ Gyr according to our results. Of these, 7 have expected decay times $\langle\tau_E\rangle \lesssim 0.3$ Gyr. In order of increasing P, they are WASP-19b, TrES-3b, HAT-P-36b, WASP-77Ab, WASP-4b, WASP-36b, and WASP-46b. Fig. 2-10 shows these planets on the $M_p \sin i$–P plane, with $\langle\tau_E\rangle$ labeled for each system and a contour of $\langle\tau_E\rangle$ superimposed. These planets all have $M_p \sin i > M_{\text{J}}$, $P < 2.0$ days, and eccentricities consistent

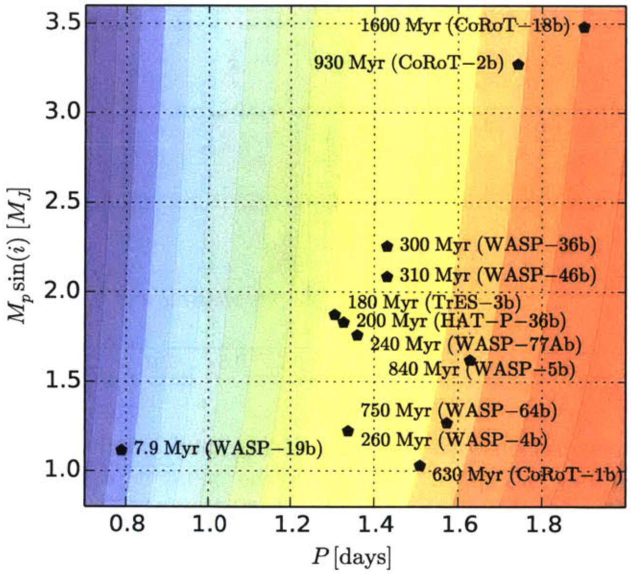


Figure 2-10: Known systems orbiting sun-like hosts and our associated predictions of the orbital decay rates. WASP-19b is a notable outlier, with an extremely short decay time-scale due to it's extremely short period (< 1day).

with or very close to zero. Since these are all transiting systems, $M_p \sin i \simeq M_p$ and the reported errors in the measured mass are typically $\lesssim 0.1 M_{\rm J}$.

We note that two of the planets (CoRoT-2b and CoRoT-18b) have masses $M_p \sin i > 3M_{\rm J}$. This suggests that they are in the strongly nonlinear regime where the parent wave breaks within the stellar core [19, 20, 21].

Of the 11 planets with $\langle \tau_E \rangle < 1$ Gyr, there are five for which studies report at least some constraint on the age of the system. In three of these, the age uncertainties are sufficiently large that the systems might be relatively young, i.e. ~ 1 Gyr (WASP-64b, WASP-5b, CoRoT-2b: $1.2^{+1.2}_{-0.7}$, $5.4^{+4.4}_{-4.3}$, $2.7^{+3.2}_{-2.7}$, respectively). However, WASP-4b and WASP-19b appear to be older systems: 7.0 ± 2.9 and $10.2^{+3.0}_{-3.8}$Gyr, respectively. Assuming that the planets arrived close to their current orbits when their host stars first formed, old stellar ages seem to be in tension with the small $\langle \tau_E \rangle$ we predict, especially in the case of WASP-19b. If our results are correct, then perhaps these planets were scattered into their current orbits well after the stars formed or they just happened to initially reside at separations with decay time-scales very close to their current ages.

Table 2.1: Orbital parameters for known systems and a summary of predictions for the orbital decay time-scales and change in orbital parameters after 10 years.

	WASP-19b	HAT-P-36b	WASP-36b	CoRoT-2b
$M[M_\odot]$	0.930 ± 0.02	1.022 ± 0.049	1.020 ± 0.032	0.970 ± 0.06
$R[R_\odot]$	0.990 ± 0.02	1.096 ± 0.056	0.943 ± 0.019	0.902 ± 0.018
$M_p \sin i\ [M_J]$	1.114 ± 0.039	1.83 ± 0.1	2.255 ± 0.089	3.27 ± 0.171
P[sec]	68155.776 ± 0.026	114682.78 ± 0.26	132828.36 ± 0.23	150594.64 ± 0.86
eccentricity	$0.0046^{+0.0044}_{-0.0028}$	0.063 ± 0.032	0	$0.0143^{+0.0077}_{-0.0076}$
age[Gyr]	$10.2^{+3.0}_{-3.8}$	null	null	$2.7^{+3.2}_{-2.7}$
$\langle \tau_E \rangle$ [Myr]	9.2 ± 0.128	205 ± 3.9	454 ± 15.7	623 ± 27.6
min τ_E[Myr]	6.3	84.7	214	241
max τ_E[Myr]	12.4	311.	853	1150
$T_{\rm shift}$[sec]	257 ± 3.6	11.5 ± 0.22	5.2 ± 0.18	3.80 ± 0.17
min $T_{\rm shift}$ [sec]	191	7.6	2.8	2.1
max $T_{\rm shift}$ [sec]	375	27.9	11.	9.8

Several recent papers consider the prospects for the direct detection of orbital decay of individual planets by measuring transit timing variations (TTVs) over long time baselines ($\Delta t \gtrsim 5$ year, see [25, 70, 160, 167]). In order to evaluate this possibility, we simulated four known systems spanning a variety of companion masses and orbital periods (but each with a solar-type host[1]), calculating their tidally induced TTV ($T_{\rm shift}$) as a function of Δt (see [25] for a derivation)

$$T_{\rm shift} \approx \frac{1}{2}\frac{\dot{\Omega}}{\Omega}\left(\Delta t\right)^2 = \frac{3}{4\tau_E}\left(\Delta t\right)^2 . \tag{2.47}$$

In order of increasing orbital period, we analyze WASP-19b [76, 78, 105], HAT-P-36b [17], WASP-36b [129], and CoRoT-2b [15, 72].

To calculate the orbital decay rate of these systems, we simulate a small range of orbital periods centered on each system's measured period. We then compute the time-averaged decay rates. We do this in order to mitigate any differences between the resonances of our stellar models and the actual resonances of the stellar hosts. Furthermore, this allows us to compute a minimum and maximum expected decay rate, corresponding to the troughs and peaks of the resonances, respectively.

Table 2.1 lists $\langle \tau_E \rangle$ as well as the minimum and maximum τ_E. The $\langle \tau_E \rangle$ of the four systems ranges from about 10 Myr (WASP-19b) to 600 Myr (CoRoT-2b), while the minimum (maximum)

---

[1] This requirement is why we do not consider WASP-18b, which was analyzed in [25].

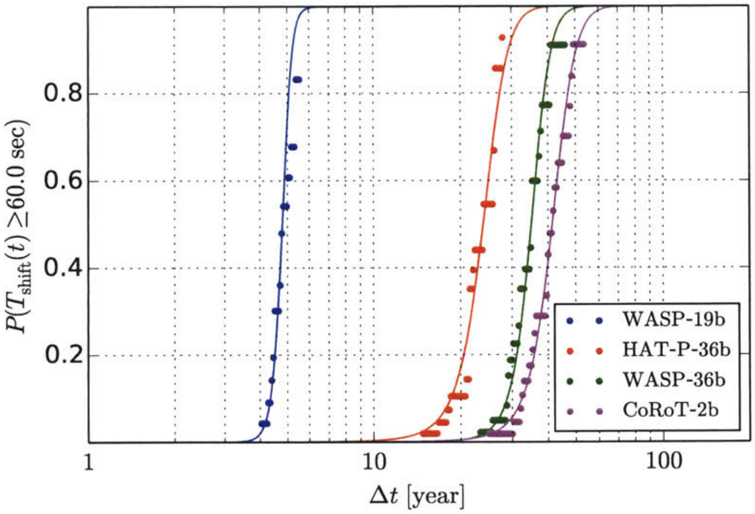


Figure 2-11: Probabilities of T_{shift} being large enough to be detectable after a given baseline.

τ_E is approximately two times smaller (larger). WASP-19b has by far the shortest decay time owing to its extremely short orbital period (18.9 hours).

Table 2.1 also lists the systems' average, minimum, and maximum T_{shift} after ten years of evolution. These provide an estimate of the magnitude of the tidally induced deviations we would expect to observe from these systems over the next ten years.

We quantify these effects further in Fig. 2-11, which shows the cumulative probability of observing tide-induced deviations as a function of time. We choose a detection threshold of $(T_{\mathrm{shift}})_{\mathrm{thr}} = 60$ sec based on the expected uncertainties in TTVs [71, 167]; different choices will scale Δt through Eqn. 2.47. According to our results, WASP-19b should produce a detectable T_{shift} in the very near future, with a $\approx 50\%$ chance of observing a deviation now given the current baseline [78] and a high likelihood of detection after only a few more years. It will take considerably longer before detections are possible in the other three systems.

We note that even if, for some reason, our calculations overestimate the dissipation rate by an order of magnitude, the curves in Fig. 2-11 would only shift to the right by a factor of $\sqrt{10} \sim 3$. Finally, as [167] points out, the Applegate effect[2] could produce T_{shift} values that are comparable to the tidally induced values and distinguishing the two may not be simple.

---

[2] Long-term cyclic changes in the star's oblateness associated with stellar activity, which modify the gravitational potential and may induce TTVs of similar magnitude to those predicted here.

2.2 Neutron Star Binaries

§2.1 discussed the impact of resonant nonlinear interactions excited by the tide within sun-like stars. We now discuss the impact of weakly nonlinear interactions within neutron stars in close binaries with other compact objects. The dynamics of these systems are dominated by the emission of GWs, but compact, degenerate stars can support secular instabilities which may be relevant throughout orbital frequencies to which ground-based GW detectors are sensitive. We now consider a phenomenological model for these instabilities and estimate how their saturation impacts GW measurements.

In the coming years, LIGO expects to detect GWs from neutron stars (NSs) in coalescing binaries. Although a NS can be treated as a point particle (PP) to a first approximation, at some level tides will modify the rate of inspiral and thus the GW signal. The impact of the tidal effects are, however, uncertain. In part this is due to uncertainties in the NS equation of state, and indeed there is hope that GW observations will eventually provide precise constraints on the equation of state [11, 50, 52, 79, 91, 120]. In addition, there are uncertainties in the tidal fluid dynamics both near the merger when matter and GR effects are strong [64, 121, 174] and during the long inspiral phase when the tide is weakly nonlinear [164, 168, 169].

Many previous studies considered the impact of the linear tide, implicitly assuming that nonlinear effects are negligible at GW frequencies below $f \approx 400\,\mathrm{Hz}$. These include studies of the linear equilibrium tide [11, 50, 52, 79, 91, 120] and the linear dynamical tide in non-rotating NSs [80, 92, 122, 132, 175] and rotating NSs [68, 81, 93]. The equilibrium and dynamical tide refer, respectively, to the quasi-static and resonant response of a star to a tidal field (see, e.g. [107]). Typically these studies conclude that linear tidal effects will be difficult to measure with current instruments without a gold-plated detection (signal-to-noise ratios $\gtrsim 50$; [120]) or stacked data from dozens of marginal events [11, 52, 91]. Moreover, because they find that tidal effects only become significant during the late inspiral, there are proposals to test vacuum GR using waveforms from NS systems at $f \lesssim 400$ Hz [10].

Recently, it has been suggested that the tide is subject to a weakly nonlinear fluid instability during the early inspiral ([169] WAB, [164] VZH, and [168] W16). The instability involves a nonresonant coupling between the quasi-static equilibrium tide, pressure supported p-modes, and buoyancy (i.e. gravity) supported g-modes. Typically, modes first become unstable at $f \approx 50\,\mathrm{Hz}$

and are driven thereafter to potentially large amplitudes. This continuous transfer of energy from the orbit into the modes increases the rate of inspiral and induces an ever-growing phase shift relative to the PP waveform. Although there has been disagreement in the literature about the magnitude of the growth rates, all studies of p-g coupling predict an instability. Furthermore, W16 find that non-static tidal effects (e.g. compressibility) enhance the growth rates, enabling a very large number of modes to reach significant amplitudes well before the binary merges.

Studies of the p-g instability have mainly focused on calculating the instability threshold and growth rates; they have not attempted to study its saturation in any detail. As a result, we do not know the rate at which the instability extracts energy from the orbit and thus we cannot say precisely how it will impact the GW signal. We construct a parametrized model of the saturation and explore the instability's impact as a function of the model parameters. It is worth emphasizing, however, that although we believe our saturation model adequately captures the range of possibilities, without a proper saturation study we cannot be certain. We find that for plausible assumptions about the saturation, current gravitational wave detectors might miss $> 90\%$ of events if only PP waveforms are used. Parameters such as the chirp mass, component masses, and luminosity distance might also be significantly biased. On the other hand, we find that relatively simple modifications to the point particle waveform can alleviate these problems and enhance the science that emerges from the detection of binary neutron stars.

2.2.1 Nonresonant p-g Instability

As the NS inspirals and the amplitude of its tidal deformation increases, the tidal flow becomes susceptible to nonlinear fluid instabilities. These will initially manifest as weakly nonlinear interactions between the tide and internal oscillation modes of the star. WAB applied the formalism developed in [170] to determine the influence of such nonlinear interactions on the inspiral of NS binaries. This revealed a new form of nonlinear instability in which the tide excites a high-frequency p-mode coupled to a low-frequency g-mode. Because the p-mode's (linear eigen-)frequency is much higher than the tidal frequency, the p-g pair is not resonant with the tide. This form of three-wave interaction is therefore very different from the well-known resonant parametric instability in which the tide excites a pair of g-modes whose frequencies approximately sum to the tidal frequency, such as those discussed in §2.1.[3]

In analyzing the growth rates of the p-g instability, WAB considered only three-wave interactions between the tide, a p-mode, and a g-mode. VZH showed that four-wave interactions between the tide (twice) and two g-modes enter the analysis at the same order as the three-wave interactions. They found that the four-wave interactions significantly cancel the three-wave interactions and concluded that although the $m = \pm 2$ component of the equilibrium tide can be p-g unstable, the growth rates are too small to influence the inspiral in a measurable way.

However, the analysis in VZH assumes that the equilibrium tide is incompressible. Although that is the case for the static equilibrium tide (the $m = 0$ component), the non-static equilibrium tide ($m \pm 2$) is compressible. W16 accounted for this compressibility and found that it undoes the cancellation between the three- and four-wave interactions, yielding rapid p-g growth rates even during the early inspiral. Specifically, W16 found that thc instability turns on at gravitational wave frequencies

$$f_i \simeq 45 \left(\frac{\omega_g}{10^{-4} \lambda \omega_0} \right)^{1/2} \text{ Hz}, \tag{2.48}$$

where ω_g is the g-mode's linear eigen-frequency, $\omega_0 = (GM/R^3)^{1/2}$ is the dynamical frequency of a NS with radius R and mass M, and $\lambda \sim 0.1$–1 is a slowly undulating function of binary separation that depends on how close the (quasi-static) equilibrium tide is to a resonance (see Fig. 9 in W16).

From Eqn. 2.48, we see that low frequency (i.e. high order) g-modes become unstable first. However, it is not clear what sets the minimum ω_g. The maximum ω_g is determined by the magnitude of the NS buoyancy frequency $\sim \omega_0/10$. W16 showed that, for $\omega_g \gtrsim 10^{-4}\omega_0$ (which corresponds to $\ell = 2$ g-modes with radial order $n \lesssim 10^3$), linear damping of the modes does not modify the instability threshold nor the growth rates. However, it is possible that other physical effects will limit the minimum ω_g (e.g. magnetic fields). As we describe in §2.2.2, our saturation model therefore includes a parameter that accounts for the uncertainty in f_i.

Once unstable, the coupled p-modes and g-modes are continuously driven by the tide and their energy grows at

$$\Gamma \approx 2\lambda\epsilon\omega_0 \simeq 20\lambda \left(\frac{M_2}{M_1 + M_2} \right) \left(\frac{f}{100\,\text{Hz}} \right)^2 \text{ Hz}, \tag{2.49}$$

where $\epsilon = (M_2/M_1)(R_1/D)^3$ is the tidal amplitude parameter due to mass M_2 acting on mass M_1

---

[3] WAB showed that, although some g-modes are also susceptible to the resonant parametric instability during the inspiral, their growth rates are too small to influence the GW signal.

and we assume $\omega_0 = 10^4\,\mathrm{rad\ s}^{-1}$ (c.f. Eqn. 112 in W16; here we include an additional factor of 2 to yield the growth rate of the energy rather than the amplitude). This equation is valid regardless of the relative size of the objects (i.e. both $M_1 > M_2$ and $M_1 < M_2$). Note that Γ is independent of ω_g, unlike f_i. Because the modes have enough time to grow by many tens of e-foldings before the binary merges (see W16 Section 5.4), eventually they reach such large energies that their growth saturates due to nonlinear damping (i.e. by exciting secondary waves through nonlinear wave-wave interactions). At saturation, there is a balance between continuous driving by the tide and decay through nonlinear damping. This suggests that the excited modes will continuously dissipate orbital energy at a rate

$$\dot{E}_{NL} \approx \Gamma N E_{\rm sat}, \tag{2.50}$$

where N is the number of independently unstable modes. The value of N is uncertain, but because the modes do not need to be resonant, $N \sim 10^3$–10^4 is possible based on the modes' typical radial order and angular degree ($n \sim 1000$, $\ell \sim$ few).

Given $\dot{E}_{NL}$, we can calculate the cumulative phase shift of the GW signal relative to the point-particle (PP) signal

$$\Delta\phi(f) \approx 2\pi \int_{f_i}^{f} \frac{\dot{E}_{NL}}{\dot{E}_{\rm GW}} \tau df, \tag{2.51}$$

where $\dot{E}_{\rm GW}$ is the GW luminosity, $\tau = f/\dot{f}$ is the inspiral time scale (both of which are dominated by the leading order quadrupole formula for two point masses [113]), and $\dot{f}$ is the rate at which the gravitational-wave frequency increases with time. Note that if the binary contains two NSs, the instability manifests in each star separately and their individual $\dot{E}_{\rm NL}$ add to the system's total $\Delta\phi$.

In general, $E_{\rm sat}$ will be a complicated function of Γ, N, the properties of the unstable modes, the NS structure, and the equation of state. Calculating $E_{\rm sat}$ is therefore challenging. Nonetheless, we might expect wave breaking to set an approximate upper bound. At wave breaking, a g-mode overturns the local stratification and a p-mode induces order unity density perturbations. WAB show that g-modes in a NS break at an energy

$$E_{\rm break} \sim 10^{-8} \left(\frac{\omega_g}{10^{-4}\Lambda_g\omega_0}\right)^2 \left(\frac{r}{R}\right)^2 E_0, \tag{2.52}$$

where r is the radial location within the star at which the breaking occurs, $\Lambda_g = \ell_g(\ell_g+1)$, and $E_0 = GM^2/R$. This is lower than the energy at which the p-modes break and thus the g-modes probably determine $E_{\rm sat}$ for the p-g instability. Although we use $E_{\rm break} \sim 10^{-8} E_0$ as a reference value throughout our study, note that if the g-modes break at $r \ll R$ the actual value will be much smaller.

These considerations motivate the ansatz $E_{\rm sat} = \beta E_{\rm break}$, where $\beta \to 1$ corresponds to saturation at the g-mode wave breaking energy. Observations of g-mode instabilities in the ocean, the atmosphere, and laboratory experiments often find that saturation indeed occurs by wave breaking (see the review by Staquet and Sommeria [131]). Numerical studies of the dynamical tide in hot Jupiter systems find that g-modes driven by the parametric instability also saturate at energies $E_{\rm sat} \sim E_{\rm break}$ [19, 20, 63]. This suggests that perhaps $\beta \sim 1$ for the p-g instability as well.

To summarize, $\dot{E}_{NL}$ and therefore $\Delta\phi$ are poorly constrained because of uncertainties in the minimum ω_g, the number of unstable modes N, and the saturation energy $E_{\rm sat}$ (or equivalently, β). We note that this instability may be relevant for any degenerate star, including both white dwarfs and NSs. This is because degenerate stars have nearly constant densities within their large cores and this supports significant overlap between the regions where p-modes and g-modes propagate, thereby allowing large coupling coefficients between the linear tide, p-modes, and g-modes. We focus on NSs because their compactness allows them to reach orbital frequencies relevant for ground-based GW detectors [8].

2.2.2 Phenomenological Model

While the saturation of the p-g instability is likely to be a complicated process, we construct a relatively simple model. Given Eqn. 2.49 and 2.50, we model the saturation with three parameters (A, f_0, n) such that

$$\dot{E}_{\rm NL} \propto \lambda f^2 N E_{\rm sat} \propto A f^{n+2} \Theta\left(f - f_0\right), \tag{2.53}$$

where Θ is the Heaviside function. The model assumes that $\beta N \lambda \propto f^n$ for $f > f_0$. The parameters A and n determine the overall amplitude and frequency dependence of $\dot{E}_{\rm NL}$ while f_0 is the frequency at which the modes reach saturation. By allowing A, f_0, and n to vary, we can account for the uncertainties in f_i, λ, N, and $E_{\rm sat}$.

We compute the tidal phase shift $\Delta\phi(f)$ relative to the non-spinning PP model using a zeroth order post-Newtonian (PN) expansion. We expect that higher order PN terms will simply add to the PP result without significantly modifying the effects from nonlinear tidal interactions. Moreover, any correction from higher order PN terms will be small compared to the zeroth order term since the phase shift accumulates predominantly at low frequencies ($f \lesssim 100\,\mathrm{Hz}$).

We assume a circular, quasi-Keplerian orbit that loses energy due to gravitational radiation and dissipative tidal interactions (between star 1 and star 2)

$$\dot{E}_{\rm orb} = -\dot{E}_{\rm GW} - \dot{E}_1 - \dot{E}_2, \tag{2.54}$$

where

$$\dot{E}_{\rm orb} = -\frac{G^{2/3}\pi^{2/3}\mathcal{M}^{5/3}\dot{f}}{3f^{1/3}}, \tag{2.55}$$

$\mathcal{M} = (M_1 M_2)^{3/5}/(M_1+M_2)^{1/5}$ is the chirp mass, $f = \Omega/\pi$ is the GW frequency, $\Omega = [G(M_1 + M_2)/a^3]^{1/2}$ is the Keplerian frequency, and (see [113])

$$\dot{E}_{\rm GW} = \frac{32\pi^{10/3}}{5}\frac{G^{7/3}\mathcal{M}^{10/3}}{c^5}f^{10/3}. \tag{2.56}$$

We model the dissipation due to the tide raised in M_1 by M_2 as

$$\dot{E}_1 = \Gamma_1 N_1 E_{\rm sat,1} \tag{2.57}$$

(and similarly for the tide raised in M_2 if both objects are NSs), where Γ is the growth rate of the instability, N is the number of unstable modes, and $E_{\rm sat}$ is the energy at which the unstable modes saturate. We have

$$\Gamma_1 = 2\lambda_1\epsilon_1\omega_{0,1} = 2\lambda_1\frac{M_2}{M_1}\left(\frac{R_1}{D}\right)^3\omega_{0,1}, \tag{2.58}$$

$$E_{\rm sat,1} = \beta_1 E_{\rm break,1} = \beta_1\left(\frac{\omega_{g,1}}{\Lambda_{g,1}\omega_{0,1}}\right)^2 E_{0,1}, \tag{2.59}$$

where $\omega_0^2 = GM/R^3$ and $E_0 = GM^2/R$. Thus,

$$\dot{E}_1 = 2\pi^2\frac{M_1M_2}{M_1+M_2}(GM_1)^{2/3} \times \left[\omega_{0,1}^{-1/3}\left(\frac{\omega_{g,1}}{\Lambda_{g,1}\omega_{0,1}}\right)^2\beta_1 N_1\lambda_1\right] f^2. \tag{2.60}$$

As the orbit decays, the fraction of the breaking amplitude at which the instability saturates (β) may increase and there may be more unstable modes (N). Therefore, we expect these parameters to vary with frequency and for simplicity we assume

$$\beta_1 N_1 \lambda_1 = [\beta_1 N_1 \lambda_1]_{\text{ref}} \left(\frac{f}{f_{\text{ref}}}\right)^{n_1} \Theta_1, \tag{2.61}$$

i.e., a power law dependence with a sudden onset of the dissipation at $f = f_{0,1}$ as captured by the Heaviside function $\Theta_1 = \Theta(f - f_{0,1})$ and motivated by the rapid growth rates relative to the inspiral rate as described in §5.4 of W16. We define the magnitude of $\beta_1 N_1 \lambda_1$ relative to the value at an arbitrary reference frequency f_{ref}. Throughout our study we set $f_{\text{ref}} = 100$ Hz (for both star 1 and star 2). Then

$$\dot{E}_1 = \frac{(2GM_1)^{2/3} M_1 M_2}{M_1 + M_2} (\pi f_{\text{ref}})^{5/3} A_1 \left(\frac{f}{f_{\text{ref}}}\right)^{2+n_1} \Theta_1, \tag{2.62}$$

where

$$\begin{aligned} A_1 &= \left(\frac{2\pi f_{\text{ref}}}{\omega_{0,1}}\right)^{1/3} \left(\frac{\omega_{g,1}}{\Lambda_{g,1}\omega_{0,1}}\right)^2 [\beta_1 N_1 \lambda_1]_{\text{ref}} \\ &\simeq 4 \times 10^{-9} \left(\frac{\omega_{g,1}}{10^{-4}\Lambda_{g,1}\omega_{0,1}}\right)^2 [\beta_1 N_1 \lambda_1]_{\text{ref}} \end{aligned} \tag{2.63}$$

is a dimensionless amplitude parameter that depends on the equation of state and how the instability saturates. The three parameters of our saturation model are therefore A_1, n_1, and $f_{0,1}$ for star 1 and similarly for star 2.

Eqn. 2.54 then implies

$$\dot{f} = 3\pi f_{\text{ref}}^2 x^{7/3} \left[B x^{4/3} + C_1 x^{n_1} + C_2 x^{n_2}\right] \tag{2.64}$$

where $x = f/f_{\text{ref}}$,

$$B = \frac{32}{5} \left(\frac{G\mathcal{M}\pi f_{\text{ref}}}{c^3}\right)^{5/3}, \tag{2.65}$$

$$C_1 = \left(\frac{2M_1}{M_1 + M_2}\right)^{2/3} A_1 \Theta_1, \tag{2.66}$$

and similarly for C_2. The phase of the GW signal $d\phi = 2\pi f dt = 2\pi f df/\dot{f}$ and

$$\phi(f) = \frac{2}{3} \int_0^{f/f_{\rm ref}} \frac{x^{-4/3}dx}{Bx^{4/3} + C_1 x^{n_1} + C_2 x^{n_2}}. \tag{2.67}$$

For typical NS parameters, $B \sim 10^{-4}$ and $B \gg C_{1,2} \approx A_{1,2}$ as long as

$$[\beta N \lambda]_{\rm ref} \ll 10^5 \left(\frac{10^{-4}\Lambda_g \omega_0}{\omega_g}\right)^2, \tag{2.68}$$

which we expect to be satisfied. Thus, the orbital decay due to gravitational radiation always strongly dominates and we can expand the $\phi(f)$ integrand as a power series. The phase shift relative to the PP waveform is therefore

$$\begin{aligned}\Delta\phi(f) &\simeq -\frac{2}{3B^2} \int_0^{f/f_{\rm ref}} dx \left[C_1 x^{n_1-4} + C_2 x^{n_2-4}\right] \\ &\simeq 0.4 \left(\frac{\mathcal{M}}{1.2 M_\odot}\right)^{-10/3} \left(\frac{C_{1,2}}{10^{-8}}\right) \left[\frac{x_0^{n-3} - x^{n-3}}{n-3}\right] \text{ rad},\end{aligned} \tag{2.69}$$

where in the second line $x_0 = f_0/f_{\rm ref}$, and we assumed $n < 3$ and $M_1 = M_2$. The phase shift is negative, which means that the orbit reaches a given frequency in fewer orbits than in the PP model.

The total phase shift accumulated by the time the NS merges is $\Delta\phi(f \gg f_0) \propto A f_0^{n-3}$. Because the growth rates are large compared to the inspiral time, $f_0 \simeq f_i$ and thus $\Delta\phi(f \gg f_0) \propto \omega_g^{(n+1)/2}$, assuming $[\beta N \lambda]_{\rm ref}$ is independent of ω_g. Because we expect $n > -1$, we see that unstable modes with *larger* ω_g contribute more to $\Delta\phi$ at merger (as long as ω_g is sufficiently small that the modes reach saturation before the merger). This is because modes with smaller ω_g have smaller $E_{\rm break}$ (Eqn. 2.52) and thus contribute less to the total $\dot{E}_{\rm NL}$ despite being unstable earlier in the inspiral (Eqn. 2.48).

The phase shift depends on the component masses as $\Delta\phi \propto (1+q)^{-2/3}\mathcal{M}^{-10/3}$, where $q = M_2/M_1$ is the mass ratio.[4] Highly asymmetric systems, such as NS-black hole (BH) binaries, therefore have much smaller $\Delta\phi$, all else being equal. This is because NS-BH orbits decay faster

---

[4] Normally, we only consider $q \leq 1$ because of a symmetry under the interchange $M_1 \leftrightarrow M_2$, but this is not the case for $\Delta\phi$ caused by only the tide in M_1 raised by M_2. If we included the phase shift induced by both the tide in M_1 raised by M_2 and vice versa, as we do later, the symmetry is restored.

and there is less time for the nonlinear tidal effects to accumulate during the early inspiral. For example, $\Delta\phi$ is approximately 100 times smaller for a NS-BH binary with a $1.4M_\odot$ NS and a $10M_\odot$ BH compared to an NS-NS binary with $M_1 = M_2 = 1.4M_\odot$ (accounting for the $\Delta\phi$ due to both NSs). As we describe below, we expect $A \lesssim 10^{-6}$ and a NS-BH binary has $\Delta\phi \lesssim 1$ rad. We show in §2.2.4 that such a phase shift is at the margins of detectability.

In our analysis, we consider values of A in the range $10^{-9} \lesssim A \lesssim 10^{-6}$. From Eqn. 2.63 we see that $A \sim 10^{-6}$ corresponds to, e.g. $N \sim 10$ ($\sim 10^3$) modes with $\omega_g/\omega_0 \sim 10^{-3}$ ($\sim 10^{-4}$) each saturating near their wave breaking energy $\beta \sim 0.1$–1. These values of N are based on the radial and angular orders of such modes ($n \sim 100$–1000 and $\ell \approx$ few). We therefore do not expect A to be much larger than 10^{-6}. Regarding the low end of our A range, we will show that for $A \lesssim 10^{-8}$ the phase shift is too small to be detectable.

Because we do not expect $\dot{E}_{\rm NL}$ to be a particularly strong function of f, we consider values for n in the range $0 \le n \le 2$. As the binary separation decreases, higher frequency modes become unstable (Eqn. 2.48), which suggests that N and perhaps $E_{\rm sat}$ increase with f, implying $n > 0$. Finally, the rapid growth rates suggest that f_0 is close to f_i. We therefore consider values in the range $30 \lesssim f_0 \lesssim 80\,{\rm Hz}$.

The saturation parameters may depend on the stellar structure and thus the component masses. We therefore allow each star in a binary NS system to have its own A, n, and f_0. Following previous work [52], we expand all the parameters around a reference mass. To wit,

$$A(M_i) \equiv A^{(0)} + A^{(1)}(M_i - 1.4M_\odot) + \cdots . \tag{2.70}$$

In our analyses, we keep only the zeroth and first order terms and, although we marginalize over both orders, we focus on the zeroth order terms throughout this study (dropping the superscript unless otherwise indicated). For simplicity we consider only the mass dependence; future improvements to the model might allow for dependencies on other stellar parameters (e.g. spin and magnetic fields).

In Fig. 2-12 we show the time domain waveform, with and without the nonlinear corrections to the PP solution, assuming a non-spinning $1.4M_\odot$–$1.4M_\odot$ binary NS system. Although both waveforms are identical at early times ($f < f_0$), a phase shift accumulates throughout the inspiral. For these parameters, the cumulative phase shift at $f \simeq 200$Hz is $\Delta\phi \simeq 200°$. As we show in §2.2.4, the nonlinear tidal effects begin to be detectable at such phase shifts.

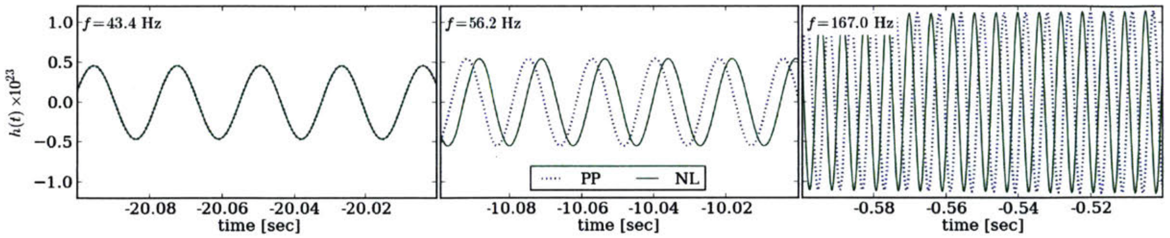


Figure 2-12: Time-domain waveforms with (green solid) and without (blue dashed) $\Delta\phi$ when $n = 0$, $f_0 = 50$ Hz, and $A = 4 \times 10^{-8}$.

In Fig. 2-13 we show $\Delta\phi(f)$ for a range of A, f_0, and n. Large A implies large total phase shift $\Delta\phi(f \to \infty)$ whereas large f_0 or n imply the opposite. We also see that although $\Delta\phi(f \to \infty)$ depends on all three parameters, the slope is mostly determined by A and n. Moreover, because we expect $n < 3$, $\Delta\phi$ accumulates most rapidly at low frequencies and asymptotes to a constant value at large frequencies. Since the PP models can account for a constant overall phase shift, detecting the nonlinear tidal effects depends primarily on the low-frequency sensitivity of the detectors.

Assuming a parametrized post-Einsteinian formalism, [43] studies modifications to PP GR waveforms that are, in some ways, similar to ours. In particular, they assume a power-law form for the phase shift, $\Delta\phi(f) \sim Af^n$, and explore a range of power-law amplitudes and exponents. However, they do not include a turn-on frequency f_0. Furthermore, they focus on high frequencies because they find that solar-system tests are more sensitive to deviations from GR than GW measurements at low frequencies. Nonetheless, their conclusions are consistent with ours to the extent that they can be compared.

2.2.3 Bayesian Inference

We use Bayesian methods to assess how our model of the nonlinear tidal effects impacts the GW data analysis. Specifically, we use Nested Sampling [127, 163] within LALInference [162] to compute posterior distributions and the evidence. In the most general PP case, the GW signal emitted by a binary in a circular orbit depends on 15 parameters, including the two component masses, source location, orientation, distance, and 6 degrees of freedom for the two spins. We collectively refer to the unknown parameters as $\vec{\theta}$. In a Bayesian framework, the evidence Z of data d given a model $\mathcal{H}$

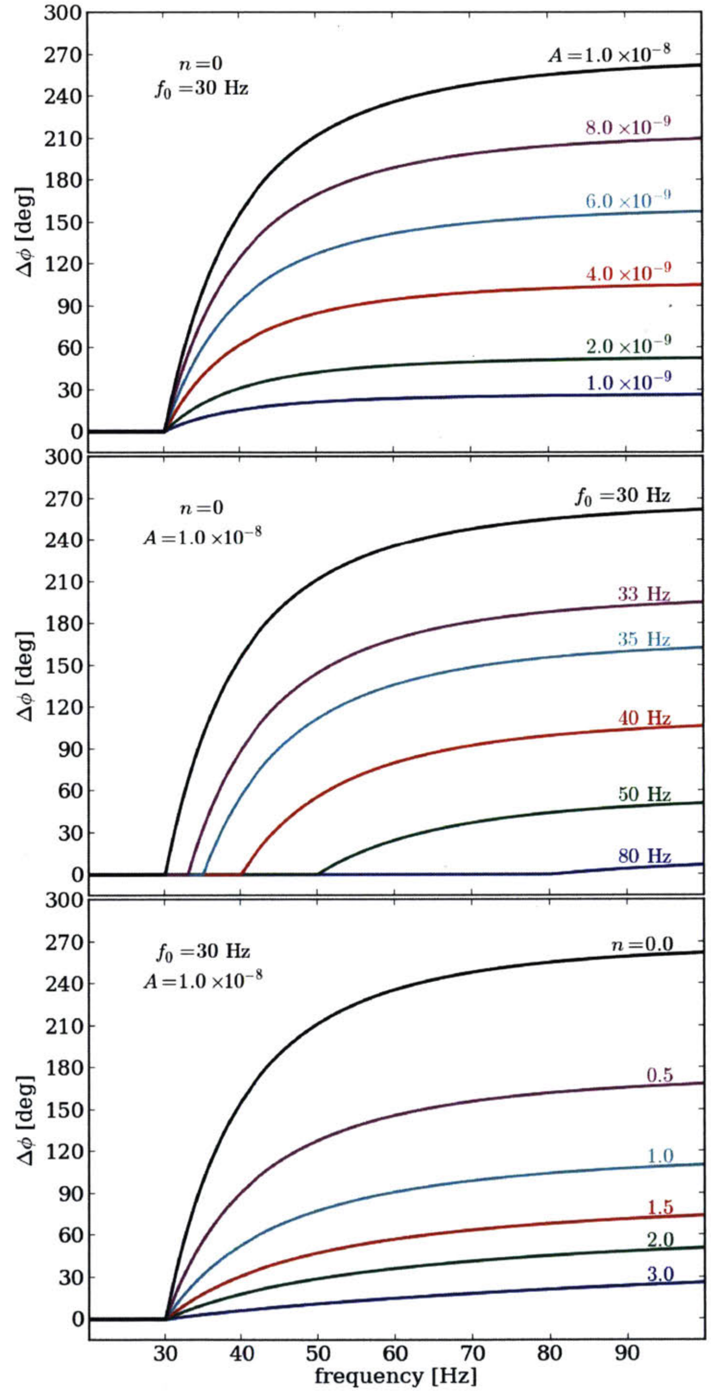


Figure 2-13: Frequency-domain representation of $\Delta\phi$ as a function of (top) A, (middle) f_0, and (bottom) n.

is

$$Z \equiv p(d|\mathcal{H}) = \int d\vec{\theta}\, p(d|\vec{\theta}, \mathcal{H}) p(\vec{\theta}|\mathcal{H}), \tag{2.71}$$

where the first term in the integral is the likelihood and the second is the prior, both of which depend on the model. The multidimensional posterior distribution for $\vec{\theta}$ can be written using Bayes' theorem as

$$p(\vec{\theta}|d, \mathcal{H}) = \frac{p(d|\vec{\theta}, \mathcal{H}) p(\vec{\theta}|\mathcal{H})}{Z}. \tag{2.72}$$

Furthermore, if two (or more) competing models are available, odds ratios between pairs of models can be calculated as

$$\mathcal{O}_B^A = \frac{p(\mathcal{H}_A|d)}{p(\mathcal{H}_B|d)} = \frac{p(\mathcal{H}_A) Z_A}{p(\mathcal{H}_B) Z_B}, \tag{2.73}$$

where the ratio of priors reflects the initial relative belief in each model. We assume that no model is preferred a priori and therefore $\mathcal{O}_B^A \to Z_A/Z_B$.

When the gravitational waveform's shape is known a priori, we use templates to represent the expected signal. These templates are parameterized by $\vec{\theta}$ and form a manifold onto which we project the data. By measuring how well different points on the manifold match the data, we construct posterior distributions for each signal parameter. This is effectively what is done within Eqn. 2.71 and 2.72. However, if the manifold does not accurately capture the full range of possible signals, biases may be introduced. Furthermore, if no point on the manifold represents the data well, we may not be able to recover the signal at all (small Z). This effect, commonly referred to as template mismatch, can occur if the phase shift introduced by nonlinear tides is sufficiently large and neglected.

In what follows, we consider two models: $\mathcal{H}_{PP}$ treats the two objects as point particles, whereas $\mathcal{H}_{NL}$ includes nonlinear tidal effects. The $\mathcal{H}_{PP}$ model uses a simple inspiral-only analytic approximant (TaylorF2) [34]. The $\mathcal{H}_{NL}$ model augments the TaylorF2 phase evolution with a tide-induced phase evolution given by Eqn. 2.69.

We focus on a single, optimally oriented, non-spinning[5] binary NS system, analyzed at distances corresponding to network signal-to-noise ratios (ρ_{net}) near 12, 25 and 50. These roughly correspond to marginal, confident, and gold-plated detections, respectively. We also neglect linear tides, which we expect to decouple from the nonlinear tidal effects because the former are significant at high

---

[5] We briefly consider aligned spins in §2.2.4.

Table 2.2: Prior distributions for nonlinear model parameters.

	Parameter	Minimum	Maximum	Distribution
PP	M_1	$1M_\odot$	$10M_\odot$	$dN \propto dM_1$
	M_2	$1M_\odot$	$10M_\odot$	$dN \propto dM_2$
	D_L	0 Mpc	300 Mpc	$dN \propto D_L^2 dD_L$
	$\cos\theta_{jn}$	-1	1	$dN \propto d\cos\theta_{jn}$
	α	0	2π	$dN \propto d\alpha$
	$\cos\delta$	-1	1	$dN \propto d\cos\delta$
NL	$A(1.4M_\odot)$	10^{-10}	10^{-5}	$dN \propto d\log A$
	$\frac{1}{A}dA/dm(1.4M_\odot)$	$-1M_\odot^{-1}$	$1M_\odot^{-1}$	$dN \propto d(\log A/dm)$
	$f_0(1.4M_\odot)$	10 Hz	100 Hz	$dN \propto df_0$
	$df_0/dm(1.4M_\odot)$	-10 Hz/$M_\odot$	10 Hz/$M_\odot$	$dN \propto d(df_0/dm)$
	$n(1.4M_\odot)$	-1	3	$dN \propto dn$
	$dn/dm(1.4M_\odot)$	$-1M_\odot^{-1}$	$1M_\odot^{-1}$	$dN \propto d(dn/dm)$

frequencies while the latter are most significant at relatively low frequencies (see Fig. 2-13). We include the LIGO Hanford and Livingston detectors [138] in addition to Virgo [153], assuming expected sensitivities for the second observing run (O2) [8]. While these were not realized exactly, they should approximate the relative sensitivities of the detectors. Because detections will be driven by the two LIGO instruments, which are expected to be more sensitive than Virgo, we place our signal directly overhead North America [8]. Virgo will mostly just improve localization through triangulation, although it could also help constrain intrinsic parameters for loud, precessing systems through improved polarization constraints. Finally, we use a zero-noise realization for our simulations, which is equivalent to taking the expected value of the evidence and posterior distributions from many noise realizations [157]. Details of our priors on all parameters are provided in Table 2.2.

2.2.4 Detectability and Biases when Nonlinear Tides are Neglected

We begin by investigating the impact of neglecting nonlinear tidal effects. We do this by injecting signals that include the tide-induced phase shift but then fit the data using only the PP waveforms. This causes significant template mismatch if the tidal effects are large, impairing our ability to detect events and biasing the inferred parameters.

Detectability and bias are related but subtly different [49, 97]. For example, the best fit may not be very good but nevertheless remain near the true parameters (i.e. unbiased but impaired detection). Alternatively, we may be able to find a good fit but only with parameters that are far from the true values (i.e. biased but unimpaired detection). Depending on the magnitude of $\Delta\phi$ and its frequency evolution, we observe one or both effects.

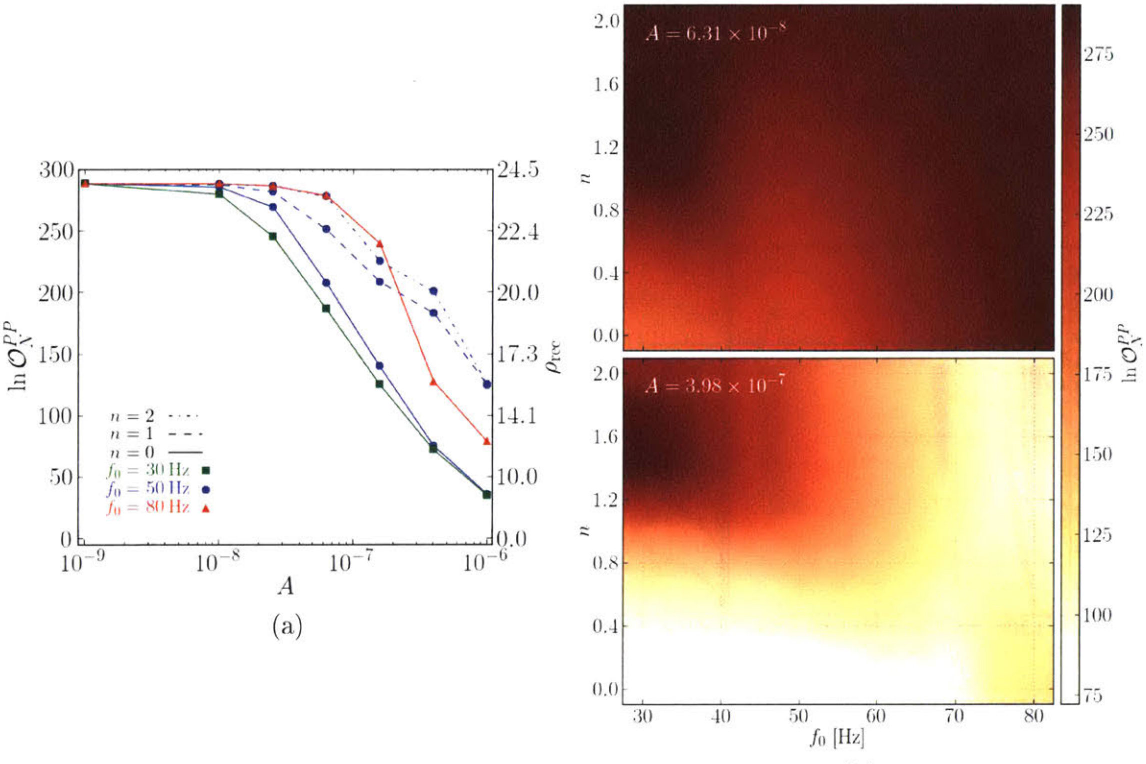
Detectability

As $\Delta\phi$ increases, the template mismatch worsens. We generally find that when $A \gtrsim 10^{-8}$ nonlinear tidal effects begin to be noticeable for current detector sensitivities. From Fig. 2-13, we see that this corresponds to $\Delta\phi \gtrsim 1$ rad, which is similar to other estimates of the minimum measurable $\Delta\phi$ [18, 48]. In terms of the saturation model (see Eqn. 2.63), $A \sim 10^{-8}$ corresponds to, e.g., $N \sim 10$ unstable modes with $\omega_g \sim 10^{-4}\omega_0$ saturating at $E_{\mathrm{sat}} \sim E_{\mathrm{break}}$ or equivalently $N \sim 10^3$ such modes saturating at $E_{\mathrm{sat}} \sim 0.01E_{\mathrm{break}}$.

We illustrate this result in Fig. 2-14 for signals that include nonlinear tidal effects injected with signal-to-noise ratio $\rho_{\mathrm{net}} \simeq 25$. We show the odds ratio $\mathcal{O}_N^{PP}$ of a PP waveform model relative to pure Gaussian noise as a function of A for different values of n and f_0. For small A, $\mathcal{O}_N^{PP}$ plateaus at large values because the PP signal model matches the data well. However, as A increases, the PP model matches the data less and less, thereby decreasing the evidence for the existence of a signal. $\mathcal{O}_N^{PP}$ can be mapped into the recovered ρ_{net} (called ρ_{rec}), and we see that for $A \sim 10^{-6}$ more than half of the signal is lost ($\rho_{\mathrm{rec}} < \rho_{\mathrm{net}}/2$). In that case, the horizon distance is halved and we miss approximately $1 - (1/2)^3 \simeq 90\%$ of NS merger events. For $\rho_{\mathrm{net}} \simeq 12$, extreme values of A can produce $\mathcal{O}_N^{PP} < 1$, which implies that Gaussian noise alone is preferred over the PP signal model even though we use a zero-noise realization.

We injected similar signals with three different ρ_{net} (12, 25, and 50). We only show the results for $\rho_{\mathrm{net}} \simeq 25$ because we find that all ρ_{net} yield very similar results modulo the usual broadening of posteriors associated with lower ρ_{net} signals. For example, all ρ_{net} produce nearly identically shaped $\mathcal{O}_N^{PP}$ curves and simply scale $\mathcal{O}_N^{PP}$ up or down. Signal loss due to template mismatch produces this behavior because we lose a fixed fraction of the inner product between the template and the data regardless of the overall amplitude.

As Fig. 2-14 shows, the decrease of $\mathcal{O}_N^{PP}$ with increasing A depends only mildly on n and f_0. We can see this in more detail in Fig. 2-14(b), which shows $\mathcal{O}_N^{PP}$ as a function of n and f_0 for two values of A. Typically, small n imply more rapid accumulation of phase shift and small f_0 imply



(a)

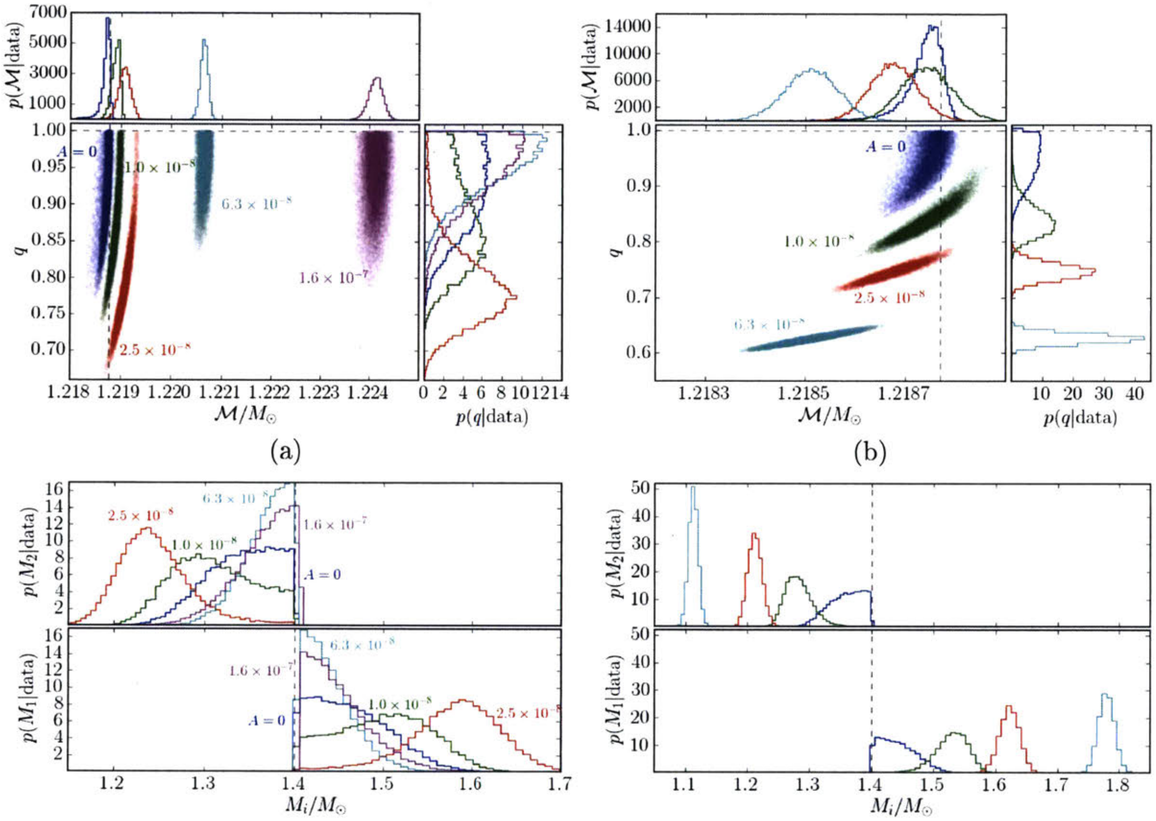
(b)

Figure 2-14: $\ln O_N^{PP}$ for a variety of NL parameters when $\rho_{\rm net} \approx 25$. Although the details depend on n and f_0, there is a clear trend to lower $\ln O_N^{PP}$ with larger A.

more total phase shift, both of which produce larger template mismatch and lower $\mathcal{O}_N^{PP}$. We also see that $\mathcal{O}_N^{PP}$ depends more strongly on A for high f_0 injections than for low f_0 injections.

Effects of spin

We also briefly investigated the effects of spins with TaylorF2 approximants. These signals allow the components to spin either aligned or anti-aligned with the orbital angular momentum, and therefore do not include precession effects. Spins can change the waveform's duration, which may be confused with the analogous effect from nonlinear tidal interactions. Searches often use TaylorF2 for low-mass systems involving NSs and restrict themselves to only relatively small spins (dimensionless spin parameters $|\chi_{1,2}| \leq 0.05$; [144]). We performed a grid-based calculation to determine the possible improvements in detectability provided by spins up to $|\chi_{1,2}| \leq 0.1$. We find that including spins



(c) (d)

Figure 2-15: Biases in $\mathcal{M}$, q, M_1, and M_2. (a) and (c): $n = 0$, $f_0 = 50$ Hz, $\rho_{\text{net}} \approx 25$. (b) and (d): $n = 2$, $f_0 = 50$ Hz, $\rho_{\text{net}} \approx 50$. We see that the component masses can be significantly biased, but they return to the injected values when $A \to \infty$, although the posteriors widen with the associated loss of ρ_{net}.

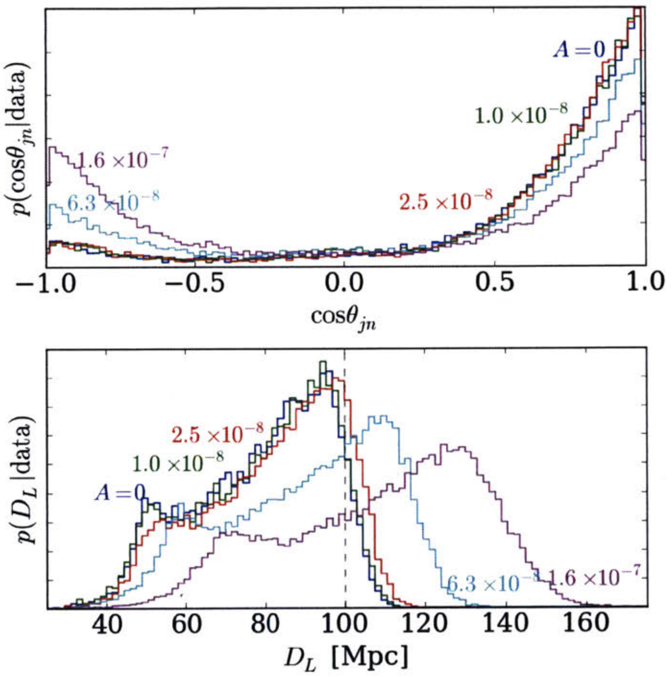


Figure 2-16: Biases in estimated D_L, and $\cos\theta_{jn}$ when NL tidal effects are present but neglected with $n = 0$, $f_0 = 50$ Hz, and $\rho_{\text{net}} \approx 25$. D_L is biases to larger values because $\mathcal{M}$ is biases to larger values, while the posterior for $\cos\theta_{jn}$ is predominantly affected by the loss of ρ_{net}.

only marginally increases $\rho_{\mathrm{rec}}/\rho_{\mathrm{net}}$ (e.g. from 0.30 to 0.34 for $A = 10^{-6}$). The slight improvement is likely due to spins compensating somewhat at high frequencies for the biases in chirp mass (see §2.2.4) induced at low frequencies by the NL effects. Although we did not fully explore the effect of spins, our analysis suggests that measurements of the spin may be biased, which could have implications for population synthesis inferences [142].

Full spinning waveforms may increase the match further, but it is unlikely that they will recover a significant fraction of the lost ρ_{net}. We conclude that spin may be important for studies of populations of marginally detectable sources with marginally relevant values of A. When A is large, we see a dramatic reduction in our ability to recover signals even when using spinning PP waveforms.

Biases

When A is small, PP models fit the true waveform well and the posterior distributions are centered on the true values. At $A \sim 10^{-8}$ we begin to observe biases in the recovered parameters even though $\mathcal{O}_N^{PP}$ has decreased by only a few percent. This is sometimes called a *stealth bias* [43, 159]. Fig. 2-15 shows the joint and marginal posterior distributions of the chirp mass $\mathcal{M}$ and the mass ratio ($q = M_2/M_1$) as a function of A for $n = 0$, $f_0 = 50$ Hz with $\rho_{\mathrm{net}} \simeq 25$. Here and throughout the rest of this study, we follow the standard convention $M_1 \geq M_2$ so that $0 < q \leq 1$, reflecting a symmetry under the interchange $M_1 \leftrightarrow M_2$. $\mathcal{M}$ is measured particularly well because it dominates the frequency evolution of inspirals [113]. We observe a clear bias in $\mathcal{M}$ as A increases. This is because larger A imply faster orbital decay, which can be confused with heavier systems. Even at $A = 10^{-8}$, we observe a statistically significant bias in $\mathcal{M}$ even though $\mathcal{O}_N^{PP}$ is essentially identical to the $A = 0$ result. Therefore, nonlinear tidal effects can bias parameter estimation even before they impact detection. However, we note that although the bias in $\mathcal{M}$ can be much larger than the statistical uncertainty, in absolute terms it remains small ($\lesssim 1\%$) even for large values of A.

Nonlinear tides also introduce biases in the mass ratio q, particularly when the impact on detectability is marginal. For $A \lesssim 5 \times 10^{-8}$, q is biased toward more asymmetric component masses. This is because asymmetric systems also decay faster. In fact, for large f_0, q is biased so much that $\mathcal{M}$ is inferred to be *smaller* than it really is. For our $1.4M_\odot$-$1.4M_\odot$ system, we find that at $A \sim$ few $\times 10^{-8}$ the larger mass may be inferred to be as much as $1.8M_\odot$ and the smaller mass as little as $1.1M_\odot$. For different values of n, the bias in q can be even more extreme than this. Although we are not likely to mis-classify a NS-NS binary as a NS-BH system for canonical $1.4M_{\odot}$-$1.4M_{\odot}$ systems, there might be some confusion for masses near the maximal NS mass.

As Fig. 2-15 shows, the bias in q is large for intermediate values of $A \sim 10^{-8}$ but small for $A \ll 10^{-8}$ and $A \gg 10^{-8}$. By contrast, we find that the bias in $\mathcal{M}$ increases nearly monotonically with A. Apparently, for $A \lesssim 10^{-8}$, which corresponds to $\Delta\phi \lesssim 1$ radian, the PP model can still approximate the data reasonably well, but only with a substantially biased q. We find that this trend holds for all values of f_0 and n. However, for $A \gg 10^{-8}$, $\mathcal{O}_N^{PP}$ decreases significantly and even though no set of PP parameters captures the data well, the true parameter values again offer the best fit (with the exception of $\mathcal{M}$, which remains biased at large A).

Despite the potential for biases, the posteriors for the component masses M_1 and M_2 almost always have some support near the true value, even if it corresponds to a long tail relative to the mode of the distribution. We also find that heavier systems with larger $\mathcal{M}$ (including NS-BH systems) are less biased by NL effects because $\Delta\phi \propto \mathcal{M}^{-10/3}$. Such systems have smaller $\Delta\phi$ because they decay faster and spend less time in the slow inspiral phase where nonlinear tides make their greatest impact. Therefore, for the same A, the posteriors and odds ratios of NS-BH systems more closely resemble the PP model.

Nonlinear tides can also bias the luminosity distance (D_L). In Fig. 2-16 we show the posterior distributions of D_L and orbital inclination (θ_{jn}; the angle between the system's total angular momentum and the line of sight to the source). As we showed above, the PP model compensates for larger A by increasing $\mathcal{M}$. However, systems with larger $\mathcal{M}$ are intrinsically more luminous and therefore are inferred to come from larger D_L. Despite the bias, we find that the posterior distribution of D_L is broad enough to cover the true value for our injections.

The other extrinsic parameters, such as θ_{jn} and source position, are unbiased by nonlinear tides. This is because the phase shift affects both polarizations equally and these other extrinsic parameters depend primarily on the ratio of the two polarizations. Although not biased, the decrease in $\mathcal{O}_N^{PP}$ with increasing A does broaden the posteriors of all extrinsic parameters.

2.2.5 Measurability and Model Selection with Nonlinear Tides

Having quantified the impact of neglecting nonlinear tidal effects, we now consider how well they can be measured. We evaluate the statistical evidence for their existence and assess how well we can constrain the nonlinear tide parameters from the data. To do this, we repeat the simulations in §2.2.4 but now use a model that *does* include the nonlinear effects when recovering the signal. We thereby obtain posterior distributions for A, n, and f_0 as well as odds ratios $\mathcal{O}_N^{NL}$ comparing the nonlinear tide model to Gaussian noise.

Model Selection

By computing both $\mathcal{O}_N^{PP}$ and $\mathcal{O}_N^{NL}$, we obtain an odds ratio comparing the two signal models $\ln \mathcal{O}_{PP}^{NL} = \ln \mathcal{O}_N^{NL} - \ln \mathcal{O}_N^{PP}$. This provides a statistical measure of the evidence for each model. If $\mathcal{O}_{PP}^{NL}$ is large, the nonlinear (NL) model is favored.

In Fig. 2-17 we show $\mathcal{O}_{PP}^{NL}$ as a function of A. For $A \lesssim 10^{-8}$, we find $\mathcal{O}_{PP}^{NL} < 1$ which implies that the model neglecting nonlinear tides is favored. This is due to Occam's razor, which penalizes the more complicated models that include nonlinear tides because they do not match the data significantly better than the simpler models that ignore them. Typically, the Occam factor corresponds to $\ln \mathcal{O}_{PP}^{NL} \sim -0.1$ and is not strongly dependent on ρ_{net}. This agrees with our intuition from the Laplace approximation, valid in the limit of large ρ_{net}. In this limit, we assume the likelihood is well modeled by a Gaussian in the model parameters and the prior is nearly constant over the Gaussian's support.

$$Z = \int d^D\theta \, e^{-\frac{1}{2}(\theta-\theta_{\text{MAP}})^T\Gamma(\theta-\theta_{\text{MAP}})} p(\theta) \tag{2.74}$$

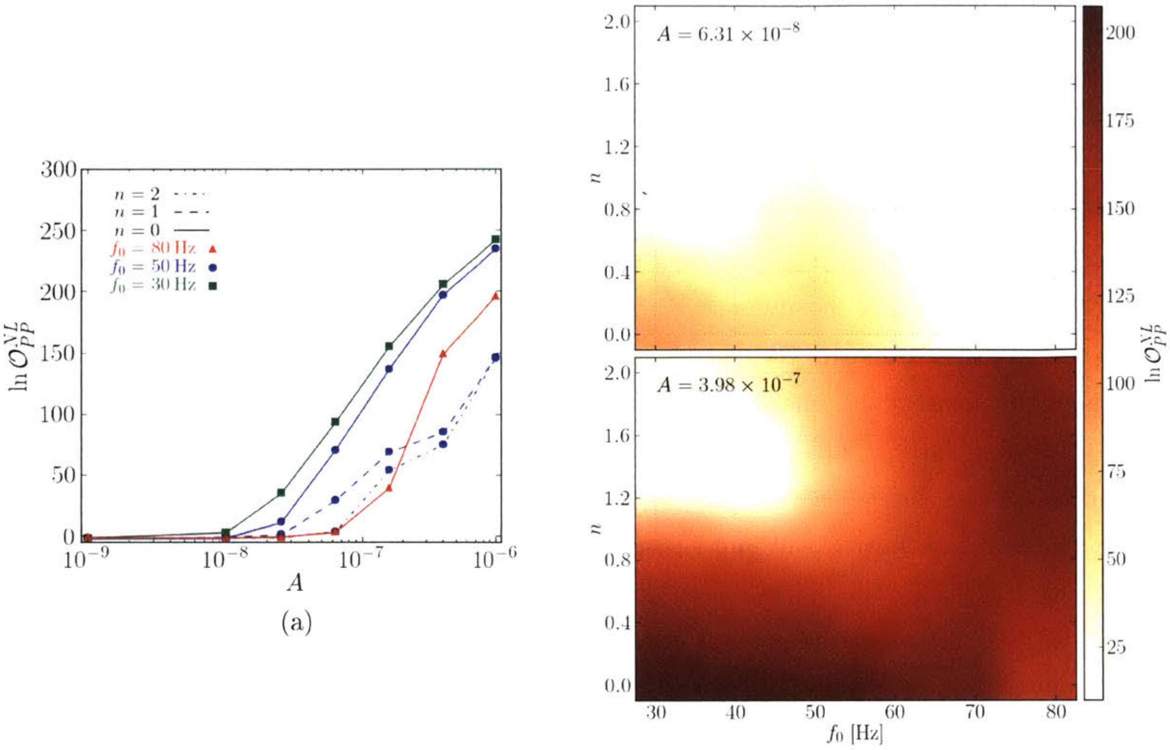
$$\sim e^{-\frac{1}{2}\theta_{\text{MAP}}^T\Gamma\theta_{\text{MAP}}} p(\theta_{\text{MAP}}) (2\pi)^{D/2} \sqrt{\det|\Gamma^{-1}|} \tag{2.75}$$

$$\sim \frac{\mathcal{L}_{\text{max}}}{V_{\text{prior}}} \sqrt{\det|\text{Cov}|} \tag{2.76}$$

We expect det|Cov| to be independent of the NL model parameters when they are too small to measure, meaning the odds ratio is dominated by the ratio of the prior volumes ($O_{PP}^{NL} \sim V_{\text{prior}}^{PP}/V_{\text{prior}}^{NL} < 1$). Because the NL model has a larger prior volume (more parameters), the PP model is preferred over it.[6] For $\rho_{\text{net}} \simeq 25$, $\ln \mathcal{O}_{PP}^{NL} \sim -0.1$ corresponds to less than 0.05% of $\ln \mathcal{O}_N^{PP}$. However, when $A \gtrsim 10^{-8}$, the NL models are strongly favored. Comparing with Fig. 2-17, we see that $A \sim 10^{-8}$ is also where $\mathcal{O}_N^{PP}$ begins to decrease. This is not a coincidence. The NL models are able to reconstruct the signal equally well regardless of A and thus $\mathcal{O}_N^{NL} \simeq$ constant. Therefore, $\ln \mathcal{O}_{PP}^{NL} \simeq$ constant $- \ln \mathcal{O}_N^{PP}$ and the critical values of A for detectability and model selection are

---

[6] Note, the assumption that det|Cov| is independent of the NL parameters can break down when the NL parameters are measurable.



(a)

(b)

Figure 2-17: $\ln O_{PP}^{NL}$ for various NL parameters when $\rho_{\rm net} \approx 25$. We note that $\ln O_{PP}^{NL} < 0$ but $|\ln O_{PP}^{NL}| \ll |\ln O_{N}^{PP}|$ when $A \to 0$, meaning we lose some evidence by expanding the model when NL effects are unmeasurable, but it is a small effect.

the same. Fig. 2-17(b) shows that the trend continues as a function of n and f_0 as well; Fig. 2-14(b) and 2-17(b) are inverses. Therefore, the regions of parameter space where the PP models fail correspond to the regions where the models with nonlinear tides are most favored. It also means that we can recover all of the $\rho_{\rm net}$ that is lost when nonlinear tides are neglected by using a more complete model.

Tests of GR

While it is clear that we can distinguish NL models from PP models for large A, it is also interesting to consider whether we can detect deviations from the PP model without the correct alternative model. Test Infrastructure for GEneral Relativity (TIGER) [10, 94, 95] is designed to answer exactly this question and computes odds ratios between the PP model and generic deviations from vacuum

GR ($\mathcal{O}^{!PP}_{PP}$). It does so by allowing the PN coefficients to vary away from their GR predictions and computing the evidence for the modified models. Furthermore, TIGER is agnostic about the effects of linear tides and only considers $f \lesssim 400$Hz [11]. In this way, it focuses on the early inspiral alone, during which the PP model is expected to be correct. We used TIGER to analyze a single injection ($A = 1.6 \times 10^{-7}$, $f_0 = 50\,\mathrm{Hz}$, $n = 2$) and observed large evidence for models allowing the first four PN coefficients to vary. They correspond to $\ln \mathcal{O}^{!PP}_{PP} \simeq 45$ when $\rho_{\mathrm{net}} \simeq 25$ and there is strong evidence in favor of the alternative hypothesis. By comparison, when we use the NL model rather than TIGER to recover the same injection we find $\ln \mathcal{O}^{NL}_{PP} \simeq 53$. We also note that $n = 2$ corresponds to some of the smaller $\mathcal{O}^{NL}_{PP}$ observed; other parameters are likely to produce even larger evidence in favor of TIGER's alternative hypothesis.

Various studies have shown TIGER to be insensitive to most uncertainties associated with compact binary coalescences and interferometric observatories (e.g. linear tides and calibration uncertainties; [12, 10]). However, we find that nonlinear tide effects, if large and ignored, can fool the TIGER machinery and suggest that GR is not the correct theory of gravity when, in fact, we have simply neglected relevant physics within the NSs. To our knowledge, this is the first example of an effect that, if ignored, could fool TIGER. This therefore emphasizes the implicit assumption within the TIGER analysis that all relevant physics has already been included in the model.

Measurability

We found that neglecting nonlinear tides when $A \gtrsim 10^{-8}$ can significantly hamper detection and bias parameter estimation. Conversely, we found that if $A \gtrsim 10^{-8}$, there will be strong statistical evidence for nonlinear tides. We now consider how well we can measure the nonlinear parameters with data from a single event.

We first evaluate what upper bound on A is achieved when nonlinear effects are extremely small (i.e. for injected signals with $A \to 0$). In Fig. 2-18 we show the posterior distributions of A for different values of ρ_{net} assuming a uniform prior for $\log A$. We find that the upper bound is near $A \sim 10^{-8}$, with a slight decrease with increasing ρ_{net}. This is not surprising given that at this A the tidal effects begin to be noticeable.

In Fig. 2-19 we show the marginal posterior distributions for A, n, and f_0 for injections at $\rho_{\mathrm{net}} \simeq 25$. When $A \lesssim 10^{-8}$, we cannot measure n or f_0. However, for $A \gtrsim 10^{-8}$, we can measure both n and f_0 to relatively high precision even at $\rho_{\mathrm{net}} \sim 12$. Typically, we measure A and f_0

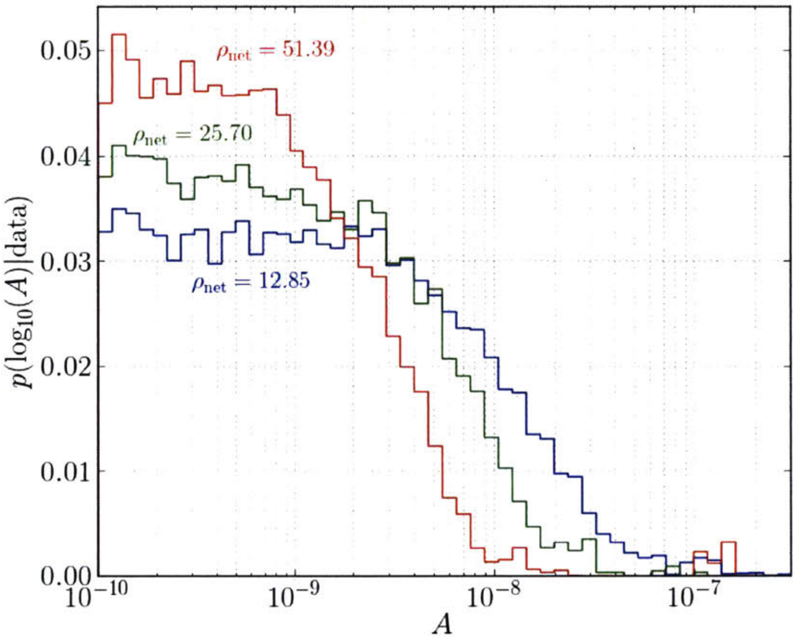


Figure 2-18: The NL amplitude (A) is constrained to $\lesssim 10^{-8}$ if NL effects are small. The constraint tightens slightly with increased $\rho_{\rm net}$.

comparably, based on a comparison of the Kullback–Leibler divergence [90, 109] from the prior to the posterior and the entropy of the posteriors. Measuring n, however, requires either larger A or $\rho_{\rm net}$.

There are also degeneracies among many of the parameters in our model. The strongest degeneracy is between A and $\mathcal{M}$, which we show in Fig. 2-21. When $A \sim 10^{-8}$ and nonlinear tides are marginally detectable we find a negative correlation between $\mathcal{M}$ and A (larger $\mathcal{M}$ favor smaller A and vice versa). This is because a bias toward larger $\mathcal{M}$ shortens the inspiral and thereby mimics the effects of the nonlinear tide. When $A \gtrsim 10^{-8}$, the degeneracy between $\mathcal{M}$ and A is present but truncated because $A \ll 10^{-8}$ is ruled out. However, the trends are somewhat different when $n = 2$; the correlations can change their behavior depending on the values of the injected parameters. Fig. 2-15(b) shows the joint and marginal distributions for $\mathcal{M}$ and q for injections with $n = 2$, $f_0 = 50\,\text{Hz}$, and $\rho_{\rm net} \simeq 50$. Unlike in Fig. 2-15(a) where $\mathcal{M}$ is biased to larger values as A increases, here we see that $\mathcal{M}$ is biased to smaller values as A increases. This is because the bias in q is much stronger and pushes the posterior backward along the degeneracy between $\mathcal{M}$ and

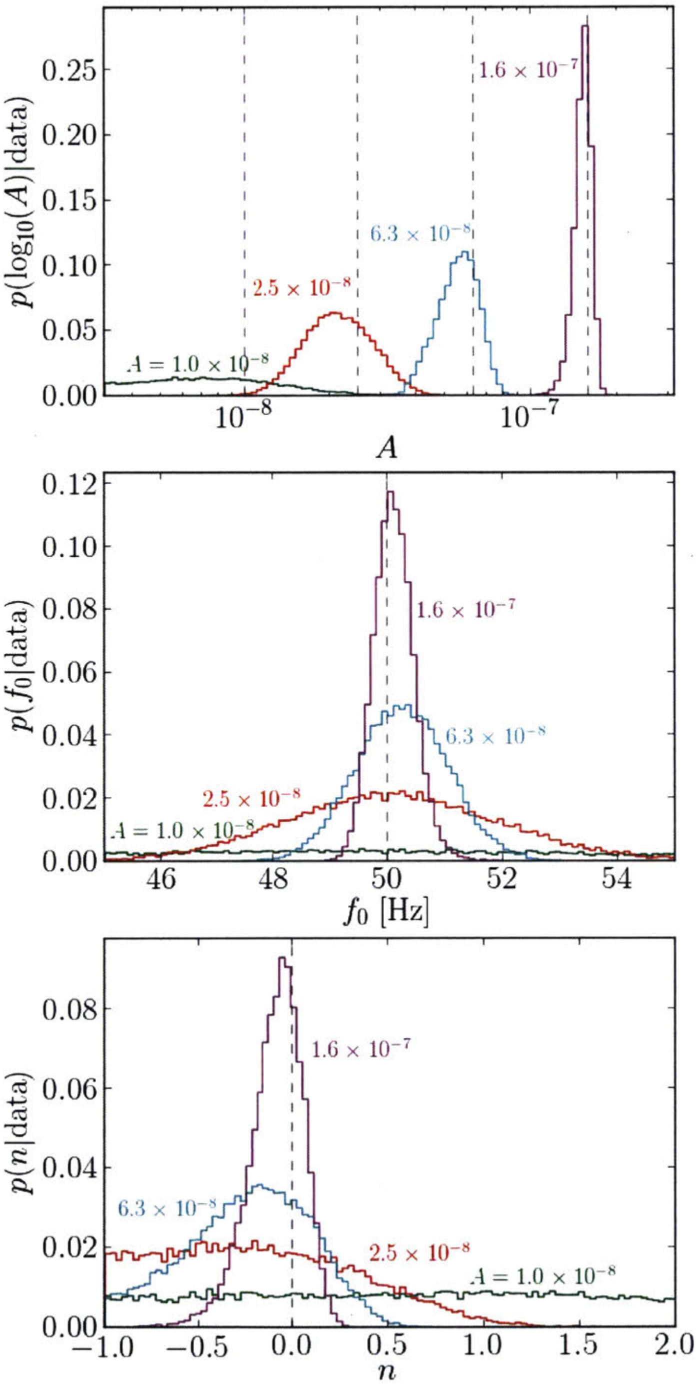


Figure 2-19: Posterior distributions for NL parameters as a function of A when $n = 0$, $f_0 = 50$ Hz, and $\rho_{\text{net}} \approx 25$. Typically, we can measure A and f_0 well, but measurements of n require higher ρ_{net} to achieve the same precision.

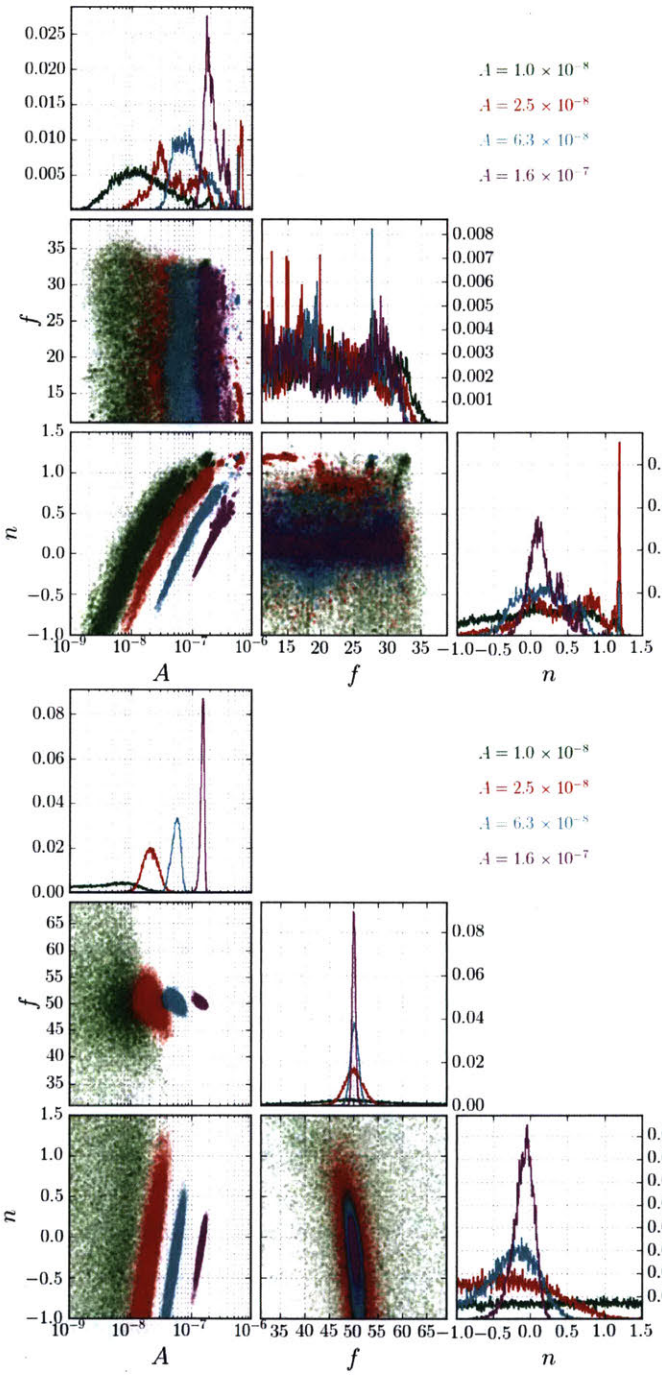
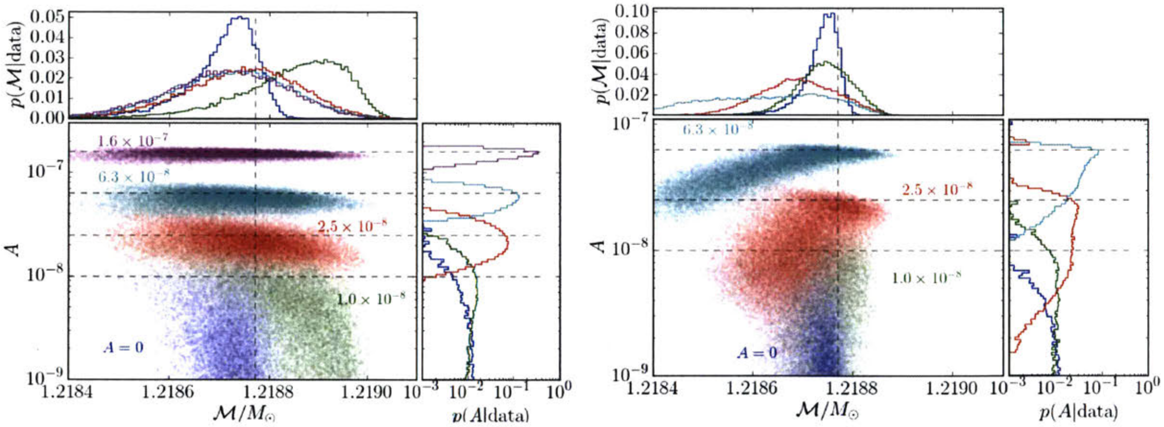


Figure 2-20: Correlations between NL parameters for (top) n=0, f_0=30 Hz, $\rho_{\text{net}} \approx 25$ and (bottom) n=0, f_0=50 Hz, $\rho_{\text{net}} \approx 25$. We notice that there can be strong correlations between A and n, which depends strongly on f_0. The general shape of the correlations are not strongly dependent on n or A.



(a)

(b)

Figure 2-21: NL and GR parameters are correlated, and the strongest correlation is between A and $\mathcal{M}$. We note that the correlation coefficient can be both positive and negative, depending on the NL parameters. (a): $n = 0$, $f_0 = 50$ Hz, $\rho_{\text{net}} \approx 25$. (b): $n = 2$, $f_0 = 50$ Hz, $\rho_{\text{net}} \approx 50$.

q [16]. Fig. 2-21(b) shows the correlation between $\mathcal{M}$ and A when $n = 2$. We see that it is reversed relative to the $n = 0$ case shown in Fig. 2-21(a). In particular, larger A imply larger $\mathcal{M}$. This is because at smaller A, the model compensates with a more asymmetric q and a decrease in $\mathcal{M}$.

We also find degeneracies between the nonlinear tidal parameters. A has a strong positive correlation with n because $\Delta\phi \propto A/(n-3)$. Interestingly, this forces A to have a weak negative correlation with f_0 because larger A imply larger n, which then requires a smaller f_0 to maintain roughly the same $\Delta\phi$. However, this correlation breaks down for large f_0, because $\Delta\phi$ then depends more strongly on f_0, weakening the correlation between n and A and strengthening the correlation between f_0 and A.

So far, we have focused on only the leading order terms in our Taylor expansions of A, f_0, and n (see Eqn. 2.70). However, our reconstructions also sampled the first order terms. We do not find any strong correlations between the zeroth and first order terms. Nonetheless, while most marginal posterior distributions for the first order terms are completely unconstrained, occasionally we observe weak constraints on df_0/dm near the boundary of its prior. If this result holds more generally and we are able to measure f_0 as a function of component mass from a series of detections, we may be able to use f_0 to make cosmological measurements using GWs alone. This is because f_0 provides an intrinsic frequency scale that gives a handle on the redshift of the otherwise conformal inspiral [99, 104]. Indeed, if we can measure f_0 as a function of mass, we may extract both the redshift and the luminosity distance directly from the GW signal without recourse to an electromagnetic counterpart. Similar approaches already exist in the literature including when one knows the NS equation of state [53, 104], when the post-merger signal is observed [103], when the shape of the NS mass distribution is known [134, 135], and when no electromagnetic counterpart is found but there is a reliable galaxy catalog [51]. Further studies will be needed, however, to test the usefulness of f_0 and to evaluate the robustness of our saturation model.

We also carried out analyses in which we allow each body to have independent values of A, f_0, and n as an alternative to the Taylor series expansions in component mass. Because there is a relatively weak dependence on mass ratio q in the phase shift and because binary NS systems should have $q \sim 1$, we find a strong degeneracy between A_1 and A_2. Generally, the posterior supports large A for one mass and small A for the other, disfavoring nearly equal A for both masses (even if the masses are similar). The Taylor expansion approach, by contrast, ensures similar values of A for similar mass NSs. We therefore consider it a better method. Most important, the weak constraints placed on the first order terms suggests that we capture most of the nonlinear tidal effects with just the zeroth order terms.

Chapter 3

Detection of Gravitational-Wave Transients

Having considered the effects of nonlinear interactions within tidal systems in §2, including the possible impact on GWs observed from binary neutron star coalescences (§2.2), it behooves us to consider GW detection itself in greater detail. This chapter focuses on many, but not all, aspects of GW transient detection using ground-based laser interferometers and is largely based on [22, 27, 40, 59, 60, 165], among other work. We apply the conclusions from these studies to actual detections in §4.

Ground-based GW detectors are complex machines with hundreds of thousands of degrees of freedom. While basic interferometric readout may be described by only a handful of these, the supporting subsystems introduce many more associated with the position and movement of the device as well as diagnostic monitors of the surrounding environment. The basic idea, however, is to isolate test masses by suspending them from several pendula, which leaves them free to move over small distances in the horizontal plane. These test masses then approximate freely falling observers and we measure the separation between elements of this congruence with light, replicating the theoretical set-up introduced in §1.3.

In general, GW detectors output a lot of information beyond just calibrated strain time-series ($h(t)$). Using all available information, analysts then try to answer three basic questions: which bits of data are likely to contain GWs, which bits of data are valid and/or unlikely to contain noise

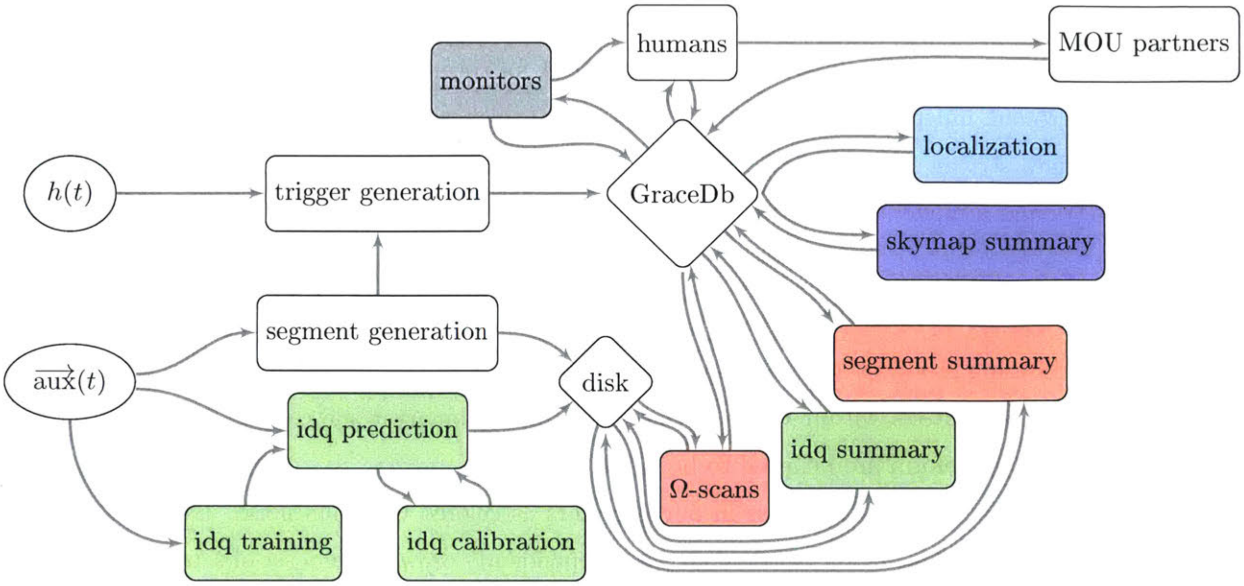


Figure 3-1: Simplified work-flow of online searches, data quality, and follow-up. Details of the shaded items are discussed in (green) §3.2.2, (cyan) §3.3, (red) §3.6.1, (blue) §3.6.2, and (grey) §3.6.3.

artifacts from terrestrial sources, and, assuming there is a GW in the data, what is the spectral structure of that GW. I will refer to these as trigger generation, data quality, and parameter estimation, respectively.

Fig. 3-1 shows a simplified schematic of the entire GW detection process, from data acquisition to completed inference and dissemination to the outside world. We discuss trigger generation in §3.1, data quality and vetting in §3.2, parameter estimation with emphasis on source localization in §3.3, and general aspects of monitoring and infrastructure in §3.6. We also describe a few pragmatic issues with IFOs in §3.4 and §3.5, frequency dependence of the antenna responses and the impact of diurnal cycles within the detectors, respectively.

3.1 Trigger Generation

There are many approaches to detecting gravitational waves, but I will focus on two of the most common for short duration transients. These algorithms statistically determine when GWs are likely to be present in a data stream. Matched filtering (also called Wiener filtering [156, 172]) assumes the signal's frequency evolution is known and uses this prediction to reject noise. A coherent likelihood approach can use multiple detectors to look for correlated excess power inconsistent with the background, independent of the signal's (unknown) spectral features. Both attempt to compress the full data stream into sufficient statistics that maximize a likelihood ratio comparing a model with GWs and a model with pure Gaussian noise.

3.1.1 Matched Filtering

Matched filtering produces an optimal search statistic for known signal waveforms in pure Gaussian noise. It is ubiquitous within GW searches for compact binary coalescences, where theoretical models of the waveform are believed to be reliable.

We begin by modeling the noise in our detectors as colored Gaussian noise. By this we mean that frequency bins in the Fourier domain are independently drawn from a set of separate Gaussian distributions. Usually, we measure the standard deviation of the amplitude and assume a random phase for each bin. This means we can write the probability of a particular noise realization n_i in the frequency domain as

$$\ln p(n) = -\frac{1}{2}\sum_i \frac{|n_i|^2}{\sigma_i^2} \sim -2\int_0^\infty df\, \frac{|n(f)|^2}{S_n(f)} \tag{3.1}$$

where the explicit sum over discrete frequencies extends over both positive and negative frequencies and S_n is the one-sided power spectral density (PSD).

Now, we assume that the noise is additive so that, in the presence of a signal, the detector output is $d = Ah + n$, meaning we can write the probability of observing a signal of known shape (h) with amplitude A as

$$\ln p(A) = \ln p(d - Ah) = -2\int df \frac{|d - Ah|^2}{S_n} \tag{3.2}$$

$$= -2\int df \frac{|d|^2 + |A|^2|h|^2 - d^*Ah - dA^*h^*}{S_n} \tag{3.3}$$

We can then maximize this with respect to A and obtain

$$A_{\text{max}} = \frac{1}{2}\left(\frac{\int df \frac{h^* d + d^* h}{S_n}}{\int df \frac{|h|^2}{S_n}}\right) \tag{3.4}$$

$$\ln p(A_{\text{max}}) - \ln p(A = 0) = \frac{1}{2}\rho^2 \tag{3.5}$$

where

$$\rho^2 = \frac{\left|\int df \frac{h^* d + d^* h}{S_n}\right|^2}{\int df \frac{|h|^2}{S_n}} \tag{3.6}$$

is the matched filter signal-to-noise ratio.[1] We note that this is a sufficient statistic in pure Gaussian noise, and all searches that use predicted waveforms implement some form of matched filtering.

Weiner filters [156, 172] also provide a way to compute ρ in the time domain, allowing analysts to filter the data optimally in real time (ρ is linear in the signal amplitude). Fig. 3-2 demonstrates this for a simple chirping sine-Gaussian signal buried in white noise. We see that there is a clear peak in $\rho(t)$ corresponding to the location of the hidden signal, even though the signal's time-domain amplitude is much less than the noise.

We should note that GW detectors do not exhibit pure Gaussian noise, but instead also contain non-Gaussian noise artifacts (glitches). This fact motivates most searches to include other detection statistics, usually some form of signal consistency check or χ^2 statistic that compares the predicted waveform's distribution of power with what was observed in the data [35, 116, 155]. We will not discuss signal consistency checks of this kind in any great depth, but we *will* discuss non-Gaussian artifacts in detail in §3.2.

3.1.2 Coherent Likelihood

While matched filtering techniques are often used to examine a single data stream at a time, we can use multiple detectors to reject noise. In this way, we can perform a coherent search for GWs in multiple strain measurements simultaneously. Typically, these searches are used to identify unmodeled GW signals that may otherwise appear as excess noise, commonly referred to as bursts.[2]

We begin with the antenna patterns for a detector (§1.3.1), which are transfer functions from

---

[1] The value of the denominator is somewhat arbitrary and is usually normalized to unity.

[2] One can conduct a single-interferometer burst search as well, but we focus on searches using at least two detectors in order to reduce the non-Gaussian noise background through coincidence requirements.

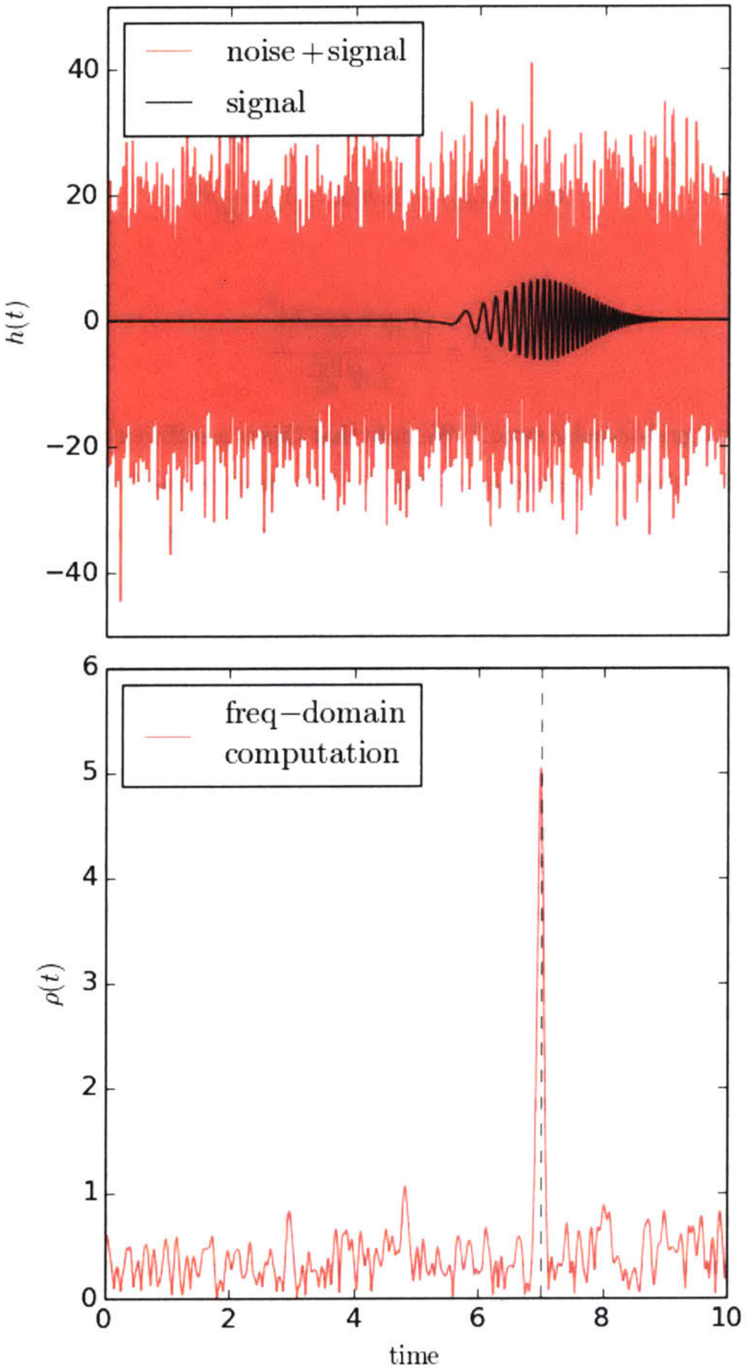


Figure 3-2: Demonstration of a matched filter. By convolving the known waveform, weighted by the noise spectral density, through noisy data we can clearly identify a strong peak in ρ that is well localized around a signal that is completely hidden by the noise.

the astrophysical strain into the detector output. Again, assuming additive noise and antenna responses $F_{\alpha j}$, we write

$$d_\alpha = n_\alpha + \sum_j F_{\alpha j} h_j \tag{3.7}$$

where α indexes the detectors and j indexes the polarizations (for GR, $j \in \{+, \times\}$). We can then write the probability that there is a signal in the data from multiple detectors as

$$\begin{aligned}
\ln p &= -2 \int df \sum_\alpha \frac{(d_\alpha - F_{\alpha j} h_j)^* (d_\alpha - F_{\alpha i} h_i)}{S_\alpha} && (3.8)\\
&= -2 \int df \sum_\alpha \frac{|d_\alpha|^2}{S_\alpha} - \sum_{\alpha,j} \frac{d_\alpha F^*_{\alpha j} h^*_j + d^*_\alpha F_{\alpha j} h_j}{S_\alpha} + \sum_{\alpha,j,k} \frac{h_k F_{\alpha k} F^*_{\alpha j} h^*_j}{S_\alpha} && (3.9)\\
&= -2 \int df \sum_\alpha \frac{|d_\alpha|^2}{S_\alpha} - \sum_{\alpha,j} d_\alpha B_{\alpha j} h^*_j + d^*_\alpha B^*_{\alpha j} h_j + \sum_{j,k} h_k A_{jk} h^*_j && (3.10)
\end{aligned}$$

where

$$\begin{aligned}
A_{kj} &= \sum_\alpha \frac{F_{\alpha k} F^*_{\alpha j}}{S_\alpha} && (3.11)\\
B_{\alpha j} &= \frac{F^*_{\alpha j}}{S_\alpha} && (3.12)
\end{aligned}$$

Now, for un-modeled signals we maximize this over each strain component $h_j(f)$, which yields an estimator for the strain and a corresponding likelihood[3]

$$\begin{aligned}
\hat{h}_j &= \sum_{\alpha,k} \left(A^{-1}\right)_{jk} B_{\alpha k} d_\alpha && (3.13)\\
\ln p(\hat{h}) - \ln p(\hat{h} = 0) &= 2 \int df \sum_{\alpha,\beta,j,k} d^*_\alpha B^*_{\alpha j} \left(A^{-1}\right)_{jk} B_{\beta k} d_\beta && (3.14)
\end{aligned}$$

and we note that the likelihood is again quadratic in the data.

This result simultaneously fits the signal in multiple detectors as a function of the location on the sky. We note that these solutions involve an inversion of the antenna patterns, which can lead to singularities and degeneracies. In general, if we have equal numbers of detectors and polarizations there is a degeneracy across the entire sky (no source location is preferred over another). Additional

---

[3] We can also marginalize over all h_j with a suitable prior to compute the Bayesian evidence for a GW being present.

assumptions must be made, and we will discuss a few in more detail in §3.3. A common one is motivated by the near alignment of the Livingston and Hanford detectors. Throughout most of the sky, the combination of these detectors is effectively sensitive to only a single polarization, and therefore we often assume there is only one polarization present [86, 87].

For two interferometers and a single effective polarization, the location on the sky maps into a time-delay between when the signal arrives at each detector. We can then search for the optimal time-delay by sliding data from one interferometer past data from the others, as is depicted in Fig. 3-3. When the time-delay cancels the physical one introduced by the direction to the source relative to the detectors, we see a strong peak in the coherent likelihood. Again, this constitutes a sufficient statistic in pure Gaussian noise.

Although a coherent likelihood search is typically less sensitive than an analogous matched filter search, it is generally more robust against modeling uncertainties. Therefore, burst searches typically target poorly modeled or unanticipated signals. Furthermore, like matched filtering, burst searches typically include other search statistics to handle non-Gaussian noise artifacts. Generally, searches produce something like a signal-to-noise ratio (or coherent likelihood, sometimes called the correlated energy) and a measure of how similar the waveforms are within each detector. Because they search for signals with unknown waveforms, they cannot use signal consistency checks (with few exceptions).

3.1.3 Statistical Decision Theory

Both matched filtering and coherent likelihood techniques provide a set of detection statistics and an intuitive way of ranking events. However, we can more generally determine the optimal way to rank events given an arbitrary number of statistics. The likelihood ratio provides such a ranking. The likelihood ratio compares the probabilities of the data being due to one of two models. Assuming we know the distributions over detection statistics for each model ($p(\text{data}|\text{signal})$ and $p(\text{data}|\text{noise})$, respectively), the Neyman–Pearson lemma [106] proves the optimality of the likelihood ratio, defined as

$$\Lambda = \frac{p(\text{data}|\text{signal})}{p(\text{data}|\text{noise})} \tag{3.15}$$

Λ naturally ranks signal-like events higher than noise-like events, and tells us which parts of parameter space to search first for signals. In this way, we naturally define the optimal detection efficiency

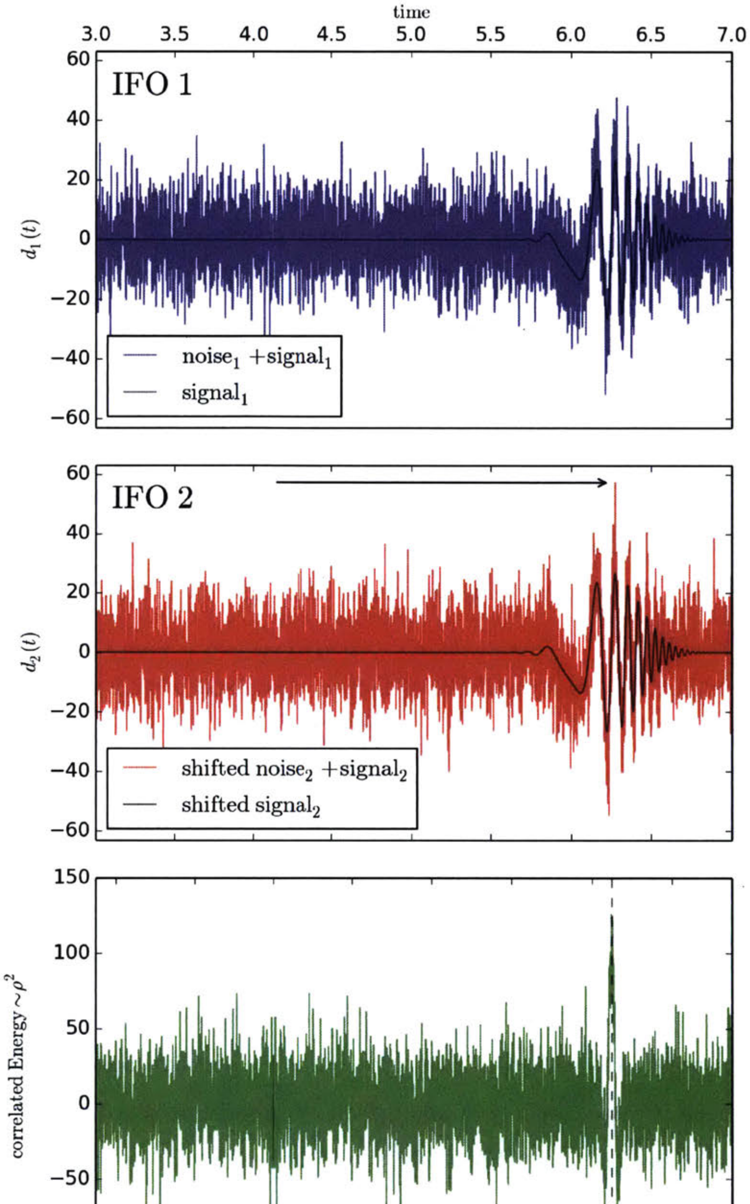


Figure 3-3: Demonstration of an un-modeled search using the coherent likelihood (correlated Energy$\sim \rho^2$). By sliding data from one detector past the other and computing the coherent likelihood, assuming a single polarization and isotropic antenna patterns, we identify a peak corresponding to when the signals are aligned. Physically, this time-shift is associated with the source's location relative to the line-of-sight between the two detectors (§3.3.1).

and false alarm probability (FAP) of the search as integrals over the D-dimensional parameter space defined by a set of statistics (θ) (see, for example [2, 26, 36, 98])

$$\text{detection efficiency}: P(\Lambda \geq \alpha|\text{signal}) = \int\limits_{\Lambda\geq\alpha} d^D\theta\, p(\theta|\text{signal}) \tag{3.16}$$

$$\text{false alarm probability}: P(\Lambda \geq \alpha|\text{noise}) = \int\limits_{\Lambda\geq\alpha} d^D\theta\, p(\theta|\text{noise}) \tag{3.17}$$

Fig. 3-4 shows a 2-dimensional example. In particular, Fig. 3-4 shows the Receiver Operating Characteristic (ROC) curve generated by plotting $P(\Lambda \geq \alpha|\text{signal})$ against $P(\Lambda \geq \alpha|\text{noise})$ for all α. ROC curves define what fraction of signals can be detected at each FAP; different detection schemes can produce different ROC curves. Ranking by the likelihood ratio maximizes the area under the ROC curve, and generally ROC curves which are further toward the "top left" corner are better. Fig. 3-4 shows that, if we are willing to accept a 1% FAP, we can detect ~ 25% of signals for these particular distributions.

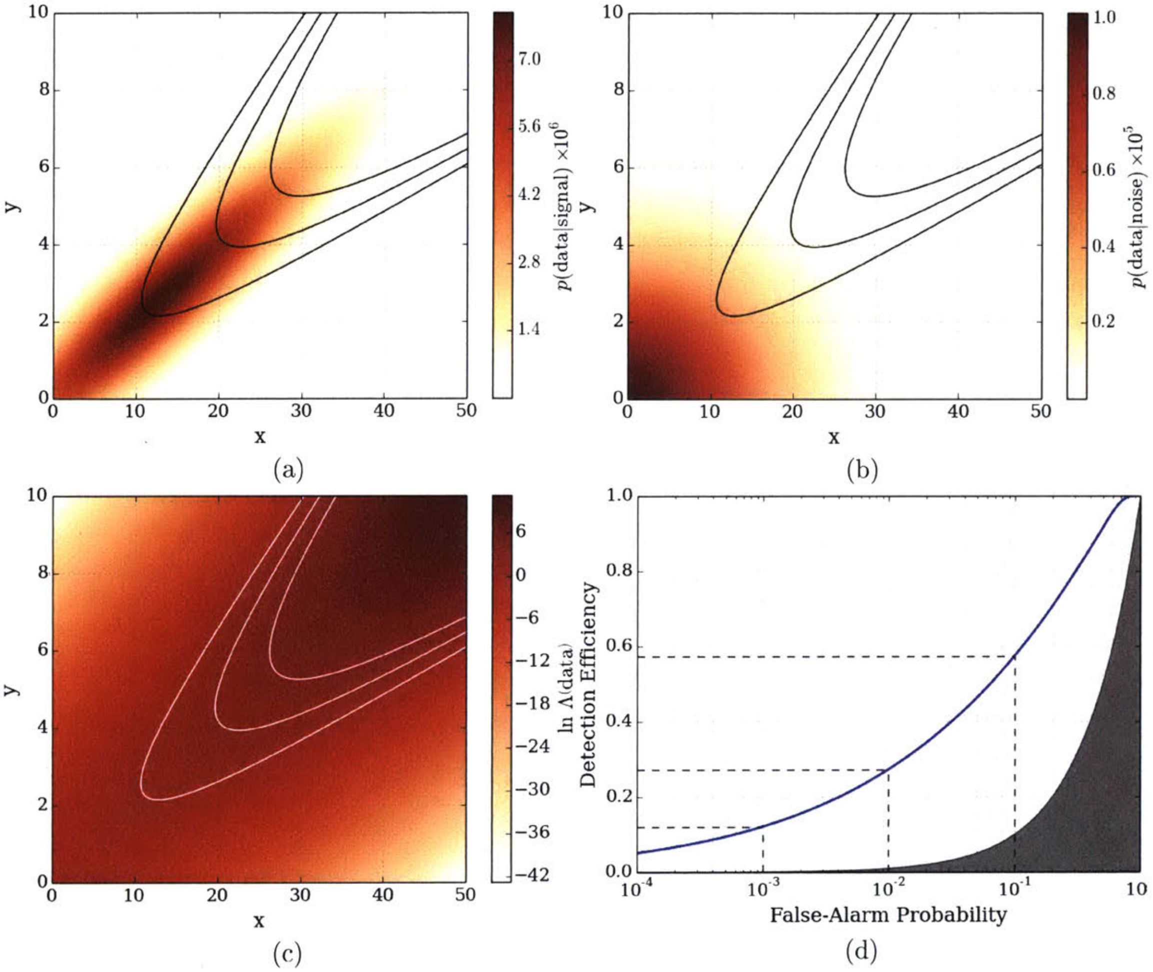
We note that the odds ratio $(\ln O)$ and Bayes factors discussed in §2.2 are essentially forms of the likelihood ratio. These include the application of priors, which do not depend on the data, and therefore do not affect the ordinal ranking assigned to a set of candidates. Therefore, both Λ and $\ln O$ are affine parameters for the same optimal ROC curve.

3.2 Data Quality and Vetting

While our example trigger generation statistics are sufficient in Gaussian noise, real detectors do not produce pure Gaussian noise. Instead, they also produce non-Gaussian noise transients (glitches). Significant effort has been invested in search algorithms to determine search statistics that are robust against non-Gaussian backgrounds. Similarly, significant effort has been invested in identifying glitches based off auxiliary data which is insensitive to GWs but nominally more sensitive to the glitches' source. Historically, analysts have often sought one-to-one deterministic mappings from cause to effect based primarily on their expertise with the detectors. While this has been successful in some cases,[4] real interferometers are complicated enough that these efforts are far from comprehensive. Instead, we often rely on algorithmic approaches to determine correlations between

---

[4] An example is described in the appendix of [143].



(a) (b) (c) (d)

Figure 3-4: Example (a) signal and (b) noise distributions of detection statistics x and y, along with the corresponding (c) likelihood ratio (Λ) and (d) ROC curve. Contours in (a), (b), and (c) demonstrate the decision surfaces corresponding to the FAPs called out in (d). The shaded region in (d) corresponds to classifiers that perform worse than flipping a fair coin, and we see that Λ significantly out-performs this.

auxiliary degrees of freedom and glitches. In this sense, data quality becomes another instance of statistical decision theory with a different target: glitches and other noise artifacts instead of GWs.

The statistical approach immediately runs into problems associated with scale. We record $\sim 200,000$ auxiliary degrees of freedom associated with the detector's state at a variety of sample rates. Within each of these channels, there may be several populations of artifacts characterized by different distributions over amplitude, frequency, duration, etc, some of which may be correlated with $h(t)$ and some of which may not. To further complicate matters, the correlations are often non-stationary. A good witness channel yesterday is not guaranteed to be a good witness today. These issues strongly motivate the need for reliable, scalable correlation machines which can be run repeatedly to track non-stationarity. In a broad sense, we utilize machine learning to infer important correlations from vectorized representations of the auxiliary state surrounding glitches.

§3.2.1 discusses several algorithms used to infer robust statistical correlations between auxiliary witnesses and $h(t)$ noise artifacts. Their performance is characterized in offline/acausal applications before being applied to low-latency searches in §3.2.2.

3.2.1 Algorithmic Development

Interferometric detectors record their physical environment and detailed interferometry status through thousands of auxiliary channels that present no or negligible coupling to GWs. Information from these channels presents an important handle for understanding (and fixing) the sources of noise in the instruments, reducing the background and ultimately establishing confidence in detections. The problem of identifying and automating the use of information from auxiliary channels is long standing within the GW data analysis community [41, 42, 46].

A number of statistical quantities have been developed [128] in order to help characterize the performance of a particular auxiliary channel or veto strategy, such as *glitch detection efficiency*: the fraction of $h(t)$ glitches removed, *use percentage*: the fraction of auxiliary glitches which can be associated with a $h(t)$ glitch [46, 82], *deadtime*: the effective fraction of analysis livetime removed when applying the veto strategy, and *Poisson significance*: the statistical significance of a measured correlation between auxiliary and $h(t)$ glitches assuming random coincidence [85]. These veto metrics are most appropriate for a simple veto strategy, such as a coincidence between an auxiliary and $h(t)$ glitch within a short specified time window. Expansions on this approach include making use of our knowledge of the instrument to anticipate when noise coupling between the auxiliary and $h(t)$ channels is strongest and/or consistent with observation [13, 14]. The use of machine learning algorithms to digest the large amounts of auxiliary channel information and predict $h(t)$ glitches is also an active area of study [27]. Accidental transient noise coincidence across multiple instruments is still a dominant source of background in astrophysical searches, especially in searches for un-modeled transients.

All our algorithms assume some level of stationarity within the data. This assumption, even if implicit, allows us to measure correlations using samples from relatively long time periods (much longer than the typical duration between noise transients). What's more, we typically assume there are several populations of transients in each auxiliary channel, some of which are correlated with $h(t)$ and some of which are not. However, we often assume that these populations can be reasonably well separated with a few thresholds on transient parameters.

Because of the classification problem's scale, a major task is the automatic rejection of irrelevant or redundant information. One such approach, commonly referred to as *Ordered Veto List* (OVL) [59], does this by hierarchically considering pair-wise correlations between $h(t)$ and a single auxiliary channel. Similar approaches are also implemented in [130] and [82]; OVL supports those algorithms through a choice of ranking statistic. The basic concept is illustrated in Fig. 3-5. We define a set of veto configurations, or recipes for how to reject transients in $h(t)$ based on auxiliary glitches. OVL uses a threshold on the auxiliary glitch's amplitude (technically, a measure of the significance of the glitch in white Gaussian noise) and a time window surrounding the auxiliary glitch's central time. For a single auxiliary channel, OVL constructs segments based on the union of all time windows surrounding sufficiently loud auxiliary triggers and vetoes any $h(t)$ transients that fall within those segments. This procedure is carried out hierarchically through a list of veto configurations with only data that has *not* already been vetoed available to subsequent configurations. The algorithm proceeds iteratively, ranking each veto configuration based on its performance and then re-ordering the configurations before the next iteration. This process is repeated several times until the list settles into a reasonable state that does not change much from iteration to iteration.

OVL supports three ranking metrics when ordering veto configurations: efficiency over deadtime (ϵ/f), use-percentage ($\%_{\mathrm{use}}$), and Poisson significance (p_{poisson}). ϵ/f is the fraction of remaining $h(t)$ transients removed by this configuration divided by the fraction of remaining time removed by this configuration. We note that this is similar to a likelihood ratio with the efficiency approximating

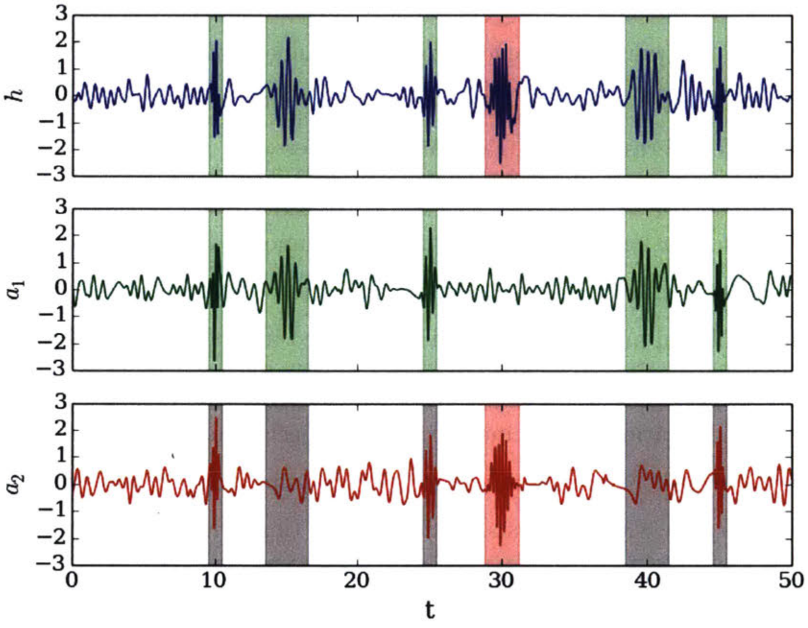


Figure 3-5: Demonstration of the hierarchical application of veto configurations implemented within OVL. The green segments are excluded first, removing some redundant vetoes that may have been generated by the red segments.

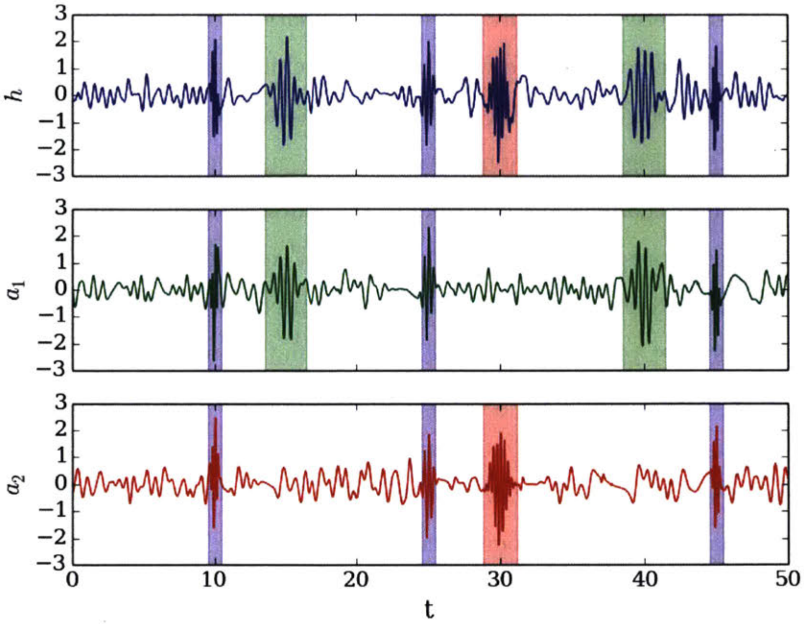


Figure 3-6: By analyzing multiple auxiliary channels simultaneously, Machine Learning Algorithms may detect higher-order correlations such as triple coincidences (blue). These can be removed separately from pair-wise correlations (green and red), thereby increasing classification accuracy and robustness.

$p(\text{data}|\text{signal})$ and the deadtime representing $p(\text{data}|\text{noise})$. This similarity often leads to better ROC curves for ϵ/f compared to other metrics (Fig. 3-7). $\%_{\text{use}}$ is the ratio of the number of $h(t)$ transients removed by this configuration to the number of auxiliary glitches used to construct the segments. High $\%_{\text{use}}$ favors one-to-one mappings between auxiliary and $h(t)$ glitches. p_{poisson} measures the probability of finding at least as many $h(t)$ transients within this configuration's segments given the observed rate of $h(t)$ transients, assuming they are Poisson distributed.

$$p_{\text{poisson}} = \sum_{k=n_c}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad \Bigg| \quad t : \text{total time contained within segments} \tag{3.18}$$

$$= \sum_{k=n_c}^{\infty} e^{-n_c f/\epsilon} \frac{(n_c f/\epsilon)^k}{k!} \tag{3.19}$$

where n_c is the number of coincident $h(t)$ glitches. We note that, for a fixed ϵ/f, p_{poisson} favors configurations with larger n_c because we order configurations by increasing p_{poisson}.

While OVL is able to determine many pair-wise correlations, it is fundamentally incapable of determining higher order correlations between multiple auxiliary channels and $h(t)$ glitches. An alternative approach is to vectorize the full auxiliary state of the detector and make use of Machine Learning Algorithms [27]. Fig. 3-6 demonstrates how we may be able to determine correlations between multiple auxiliary channels instead of only the pair-wise combinations shown in Fig. 3-5. We investigated Random Forests [30, 31], Analytic Neural Networks [75, 77], and Support Vector Machines [38, 45, 47]. Random Forests average over the predictions of many decision trees, each with access to only a subset of the total set of features and optimized over that subset. Neural Networks use collections of connected nodes (neurons) that are activated by weighted sums of their inputs, resulting in a single output node which combines the results passed through multiple layers into a single prediction. Support Vector Machines attempt to partition the feature space using hyper-planes separating regions dominated by a single type of sample. All these algorithms constitute supervised learning.

We vectorize the auxiliary state into 5 features per auxiliary channel: amplitude, time between auxiliary glitch and $h(t)$ glitch, central frequency of the auxiliary glitch, the duration of the auxiliary glitch, and the number of wavelet pixels included in the auxiliary glitch. These features are taken from the loudest auxiliary glitch within ± 100ms of the target time; if no auxiliary glitch exists, all values are set to zero. With historical data, we find that Random Forests typically perform best, although they performed only comparably to OVL [27].

Fig. 3-7 shows the performance of OVL's three metrics on a week of data surrounding GW151226 (§4.2). This week corresponds to excellent classification at LHO and above average performance at LLO, although not remarkably so. The ROC curves also show the expected ordering, with ϵ/f performing slightly better than $\%_{\mathrm{use}}$. For this data, p_{poisson} seems to perform comparably to the other metrics, although this is usually only true in the limit of large FAP or deadtime. The LHO panel of Fig. 3-7 suggests this somewhat in that the p_{poisson} ROC curve starts at higher deadtimes than the other metrics, a consequence of the preference for veto configurations which remove a large number of glitches and correspondingly large amounts of time.

Both OVL and all the MLAs were originally designed for offline analysis, where latency is not an issue. They have been adapted to run in an online, low-latency mode (§3.2.2). Before discussing this approach, there is another offline analysis worth mentioning.

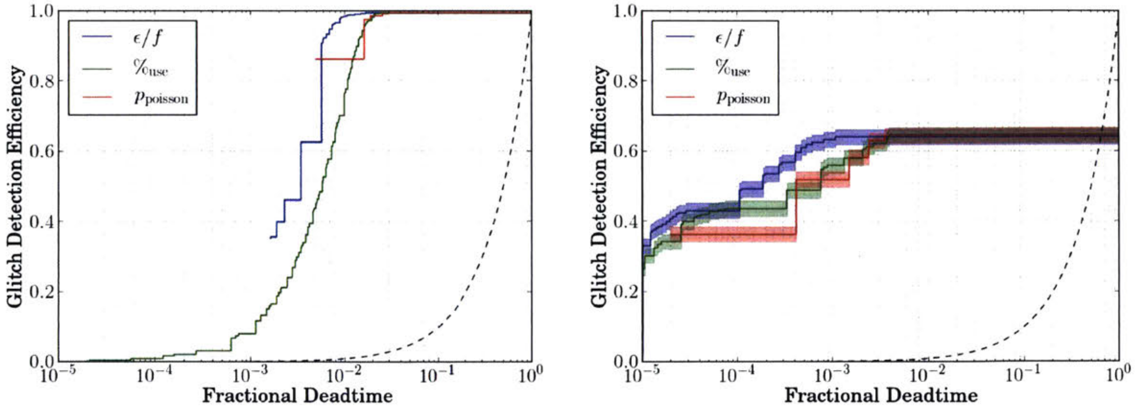


Figure 3-7: ROC curves from OVL for the week surrounding GW151226 (00:00:00 UTC on December 17, 2015 to 00:00:00 UTC on December 27, 2015) from (left) LHO and (right) LLO, analogous to Fig. 3-14. This corresponds to 168.61 and 129.17 hours of livetime and 65,323 and 505 glitches at LHO and LLO, respectively. We note that the error bars in the LHO data are too small to see (typically $\lesssim 10^{-3}$) given the large number of glitches.

Pointed Poisson Analysis

OVL is basically a counting experiment; we count how many $h(t)$ glitches fall within a set of segments defined by an auxiliary channel. However, to capture an important aspect of coincidences, namely how much time separates the participating glitches, we must include many different windows in the algorithm. Motivated by this, we can reformulate the problem of measuring the significance of a coincidence in terms of the proximity of the events directly. In this formulation, we assume the auxiliary glitches are Poisson distributed[5] and measure the significance of observing a coincidence at least as close to a particular time. This, in its essence, is a null test for the hypothesis that the time is not correlated with the triggers in the auxiliary channel.

We assume we can measure the rate of auxiliary glitches above some amplitude threshold (λ_ρ) and then compute the p-value for the nearest such trigger (Δt_ρ away from the time of interest) as

$$p = 1 - e^{-\lambda_\rho \Delta t_\rho} \quad \Big| \quad \lambda_\rho,\ \Delta t_\rho \text{ are functions of the amplitude threshold } (\rho) \tag{3.20}$$

Now, we expect loud triggers to be rarer than quiet ones. Therefore, a loud but distant glitch may

---

[5] We assume Poisson distributed triggers because it models most channels well. However, the process could be repeated with other distributions, even empirically measured distributions.

be more significant than a quieter but closer glitch. It is the product of $\lambda_\rho \Delta t_\rho$ that matters. For this reason, we take the minimum p-value[6] over all ρ

$$p_{\min} = \min_{\rho} \left\{ 1 - e^{-\lambda_\rho \Delta t_\rho} \right\} \tag{3.21}$$

For a given time, we compute $p_{\min}$ for each channel. We can then measure the expected *accidental coincidence* distribution by measuring $p_{\min}$ for randomly chosen times. Fig. 3-8 demonstrates just such a distribution. We note that the tail of the distribution is linear ($\mathrm{prob}(p_{\min} \leq q) \propto q$), meaning $p_{\min}$ is a legitimate p-value in this regime, as desired. The distribution's overall normalization reflects the number of channels used. We note that, by normalizing Δt by λ for each channel separately, we are able to put each channel on an equal footing. Accidental coincidences from any channel are drawn from the same background distribution.

We also note that, for times characteristic of a single type of event (e.g. a single $h(t)$ glitch family), we can repeat this experiment and stack the measurements. This is because each measurement is individually identically distributed and the joint probability is the product of the individual probabilities.

This pointed null test for uncorrelated, Poisson-distributed auxiliary glitches has been used to confirm the lack of terrestrial noise sources for all confirmed GW detections to date (§4, [145, 146, 151]). What's more, it has been used to study populations of troublesome noise within the detectors, a good example of which are *blip glitches*. These broad-band, short duration signals currently limit the background of all transient searches. Upon investigation of ~ 600 blips identified in the O1 Coherent WaveBurst ([86, 87]) background, we found several populations of blips using these techniques. Fig. 3-9 shows the behavior of LHO blips under the assumption of a single population and multiple populations. For a single population, we simply stack all measurements; for multiple populations, we stack all measurements for which a particular channel was found to be coincident ($\min_\rho\{\Delta t_\rho\} < 1$ sec). This should mitigate the washing-out effects of stacking unrelated measurements at the cost of a larger tail. Compared to Fig. 3-8, which shows the equivalent distributions for random times, there is a clear excess in the multi-population distributions. This excess is associated with very loud glitches that are witnessed in magnetometers and microphones

---

[6] A more careful analysis would include marginalization over the measurement uncertainty of λ_ρ and a more careful accounting of the trials factor incurred by the minimization over ρ. We find these to be unnecessary for our immediate goals.

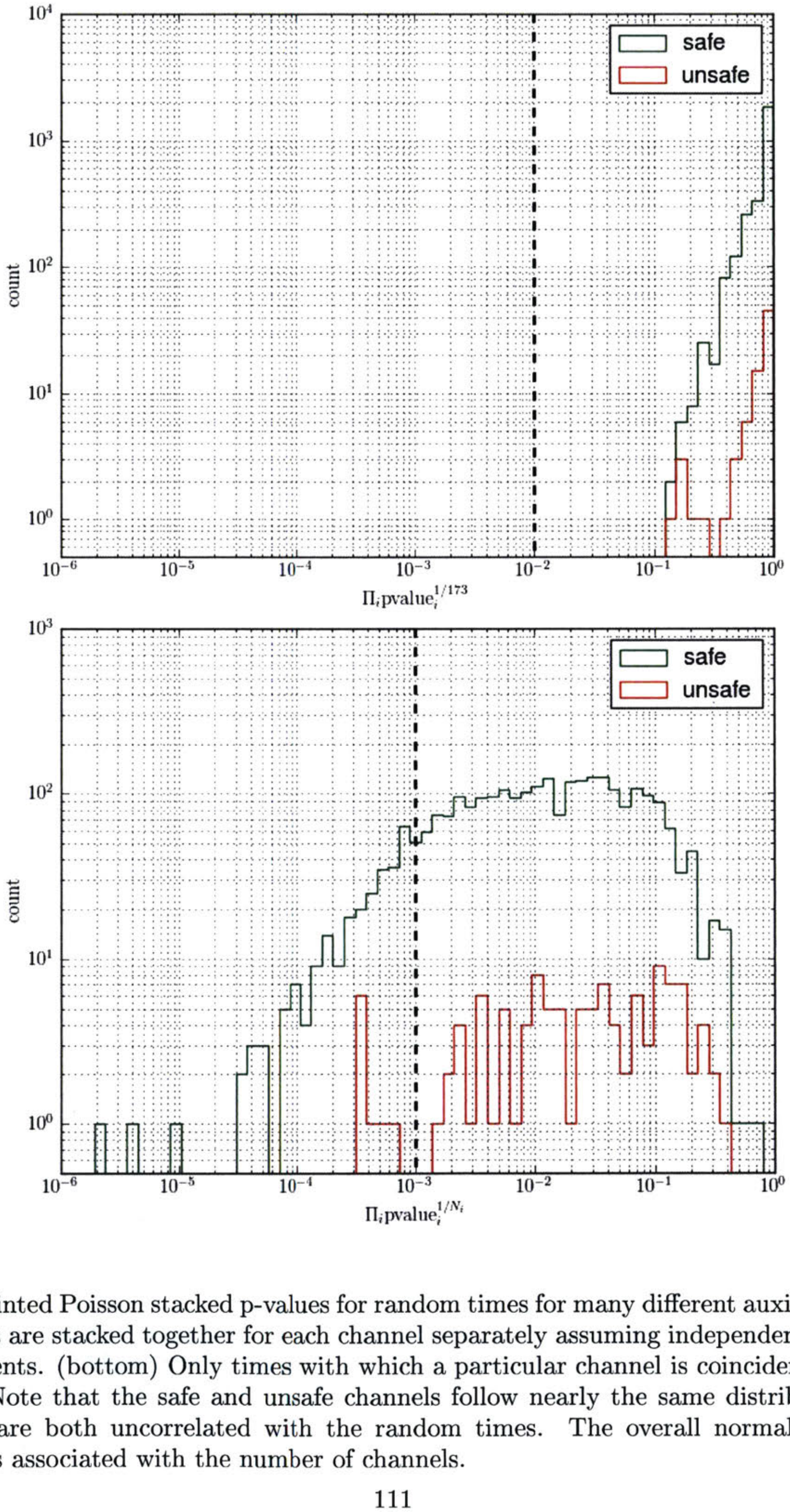


Figure 3-8: Pointed Poisson stacked p-values for random times for many different auxiliary channels. (top) All times are stacked together for each channel separately assuming independently identically distributed events. (bottom) Only times with which a particular channel is coincident ($\Delta t \leq 1\mathrm{sec}$) are stacked. Note that the safe and unsafe channels follow nearly the same distributional forms because they are both uncorrelated with the random times. The overall normalization of the distributions is associated with the number of channels.

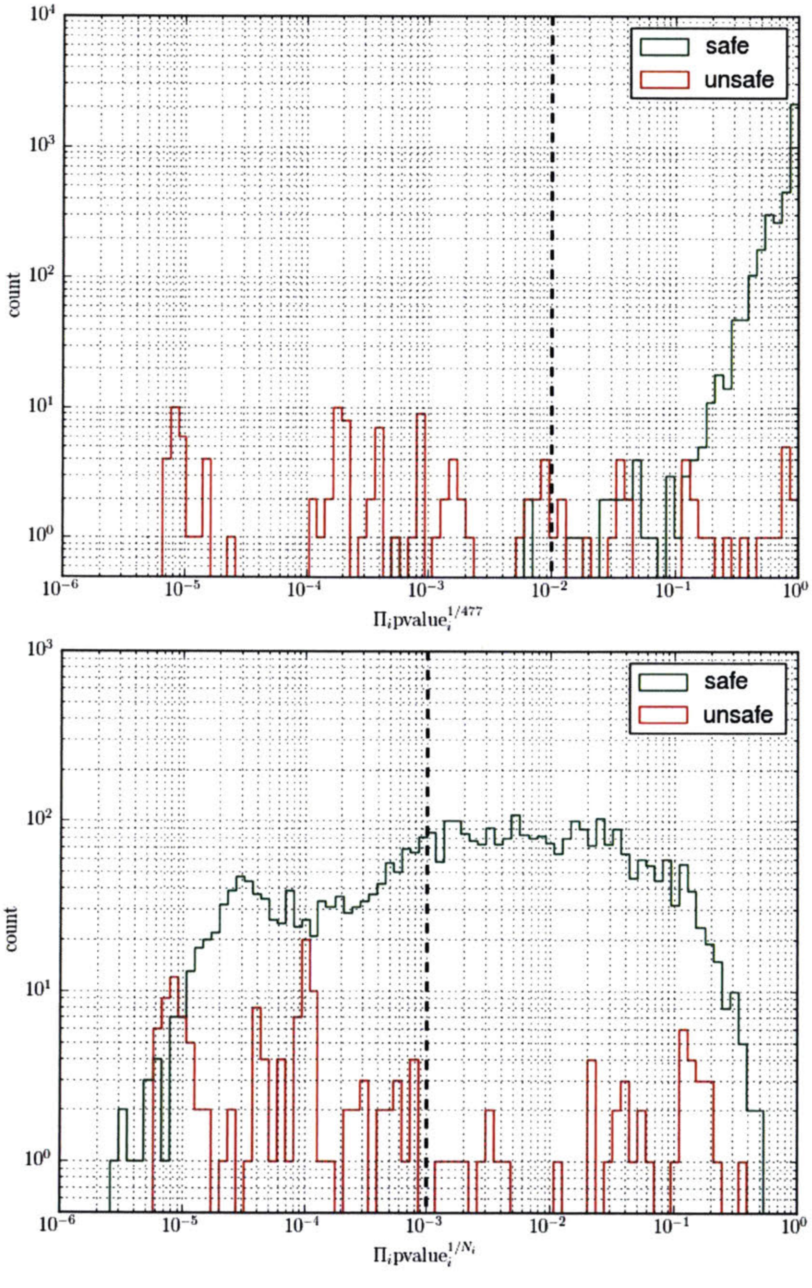


Figure 3-9: Pointed histograms for blippy times from LHO, analogous to Fig. 3-8. Although we see clear outliers in the unsafe distributions assuming a single population and multiple populations, there is a clear excess of safe outliers only in the multiple populations distribution.

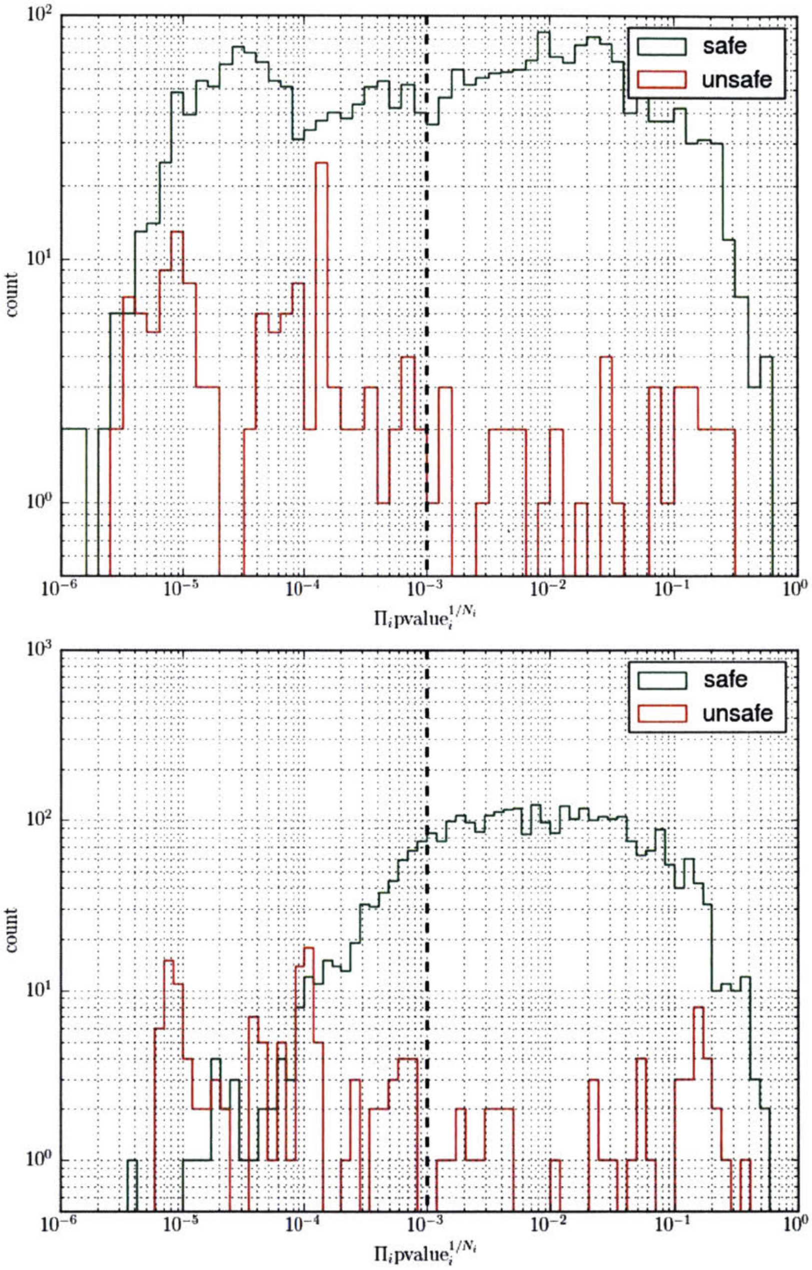


Figure 3-10: Pointed histograms for (top) blips that are witnessed in magnetometers (53 times) and (bottom) blips that have no auxiliary witness except for unsafe channels (554 times), which strongly resembles the bottom panel of Fig. 3-8.

(top panel of Fig. 3-10). However, the vast majority of blips ($\sim 91\%$) have no auxiliary witness (bottom panel of Fig. 3-10). Importantly, the only channels that correlate with blips are within the differential-arm motion control loop or otherwise *unsafe*; that is, they can be sensitive to real GW signals. This study conclusively shows that our current set of *safe* auxiliary witnesses is insufficient for blip glitches.

3.2.2 Online Infrastructure for Low-Latency Vetting

The online detection of glitches in $h(t)$ provides additional constraints. While computational cost and efficiency are important for offline analyses, they are vital for online predictions for which latency is a factor. Here I describe an online architecture used to robustly identify $h(t)$ glitches based on auxiliary channel information alone, often with lower latency than GW searches themselves: iDQ.

iDQ ingests triggers from a rapid dyadic wavelet transform based on the Haar wavelet decomposition (KleineWelle [39]). While the precise source of triggers is unimportant algorithmically, the speed of the Haar transform makes these triggers extremely useful. KleineWelle is able to process 32 seconds of several thousand channels, each sampled at several kHz, on a single core in approximately 30× real time. The main latency for iDQ is KleineWelle trigger generation and the main latency for KleineWelle is associated with waiting for enough data to whiten the wavelet transform.

The problem of glitch classification is compounded by non-stationarity. For this reason, we divide the classification problem into separate, asynchronous steps that mesh to provide low latency predictions while automatically updating the trained classifiers and calibrating their output. Specifically, a low latency evaluation loop searches for new auxiliary triggers and uses pre-trained classifiers to predict the existence of glitches in $h(t)$. The classifiers' output, which by itself has little physical meaning, is then automatically calibrated into probabilistic statements such as the FAP, glitch detection efficiency, and even the likelihood ratio (§3.1.3). These are reported as streaming time-series. Periodically, training and calibration jobs are separately launched to account for non-stationarity. These jobs update the trained classifiers and calibrated mappings from classifier output to probabilities based on historical data, usually up to a week's worth. The training jobs perform the offline OVL and/or MLA analyses while the calibration job performs a simple counting experiment using the results of the online predictions.

A crucial aspect of this process is the cross-validation performed by the calibration step. The low latency evaluation, in addition to generating time-series, reports predictions for a few discrete times.

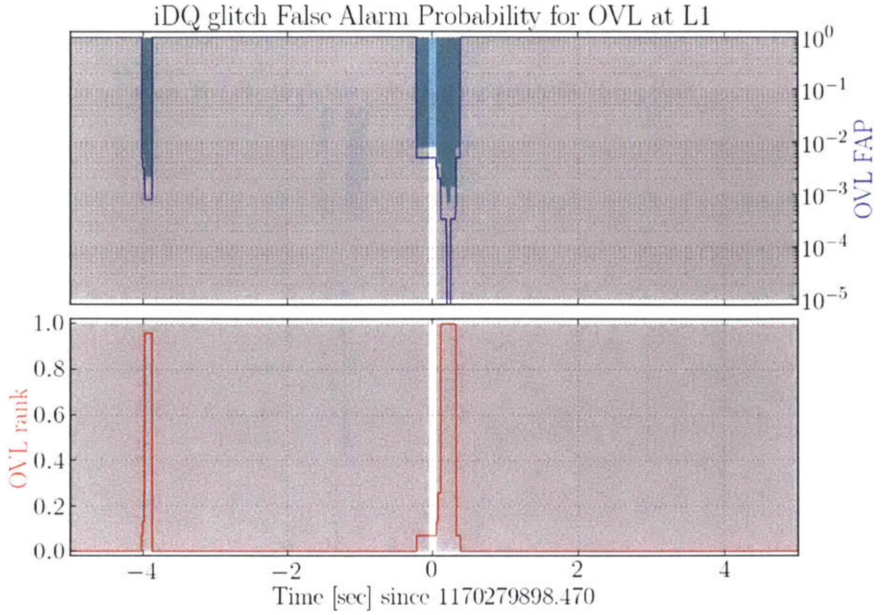


Figure 3-11: iDQ time-series from LLO for G272601 (a clear radio-frequency whistle from 21:44:40 UTC on February 4, 2017); blue shaded regions correspond to the 90% upper limit on the estimated FAP based on the counting statistics within iDQ's calibration job. Note that the classifier's rank increases when the FAP decreases. A spectrogram of $h(t)$ for this event is provided in Fig. 3-31.

These are chosen to represent both glitchy and clean times based on the actual $h(t)$ signal. Glitchy times are defined by the presence of a loud trigger in $h(t)$ and cleans are drawn from a Poisson process and are required to be at least 100ms away from glitchy times. Because the classifiers make their predictions *without* knowledge of the glitch/clean label and are causally trained (using only historical data), this procedure allows us to automatically measure both the FAP and detection efficiency in an unbiased way.

Another important aspect of iDQ is its automatic, low-latency response to candidate events reported by search pipelines (Fig. 3-11, 3-12, 3-13, and 3-1). Whenever a new candidate is reported, iDQ generates a full report of its predictions, the features relevant to those predictions, and measures of how well the classification as a whole is performing in the neighborhood of that event. Specifically, each instance of iDQ (one per detector) reports the time-series along with the raw classifier output (called *rank*). Over the same time-window, OVL reports which veto configurations were active and their associated ranks. This immediately allows analysts to identify the auxiliary channel

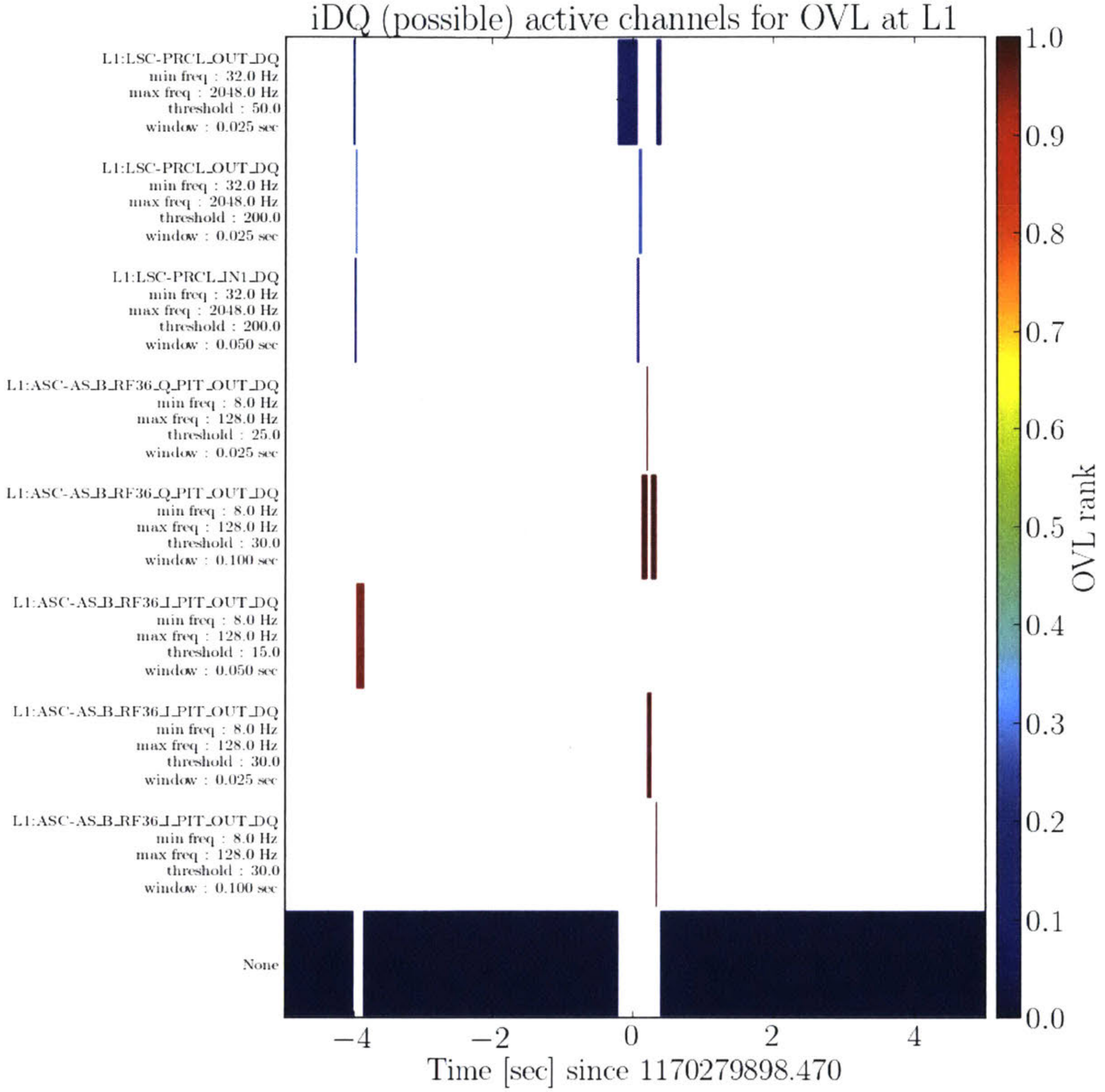


Figure 3-12: Auxiliary channels identified by iDQ to be correlated with $h(t)$ and when the corresponding veto segments are active for each configuration from LLO for G272601. Note, the rank reported here corresponds to the rank reported in Fig. 3-11. Fig. 3-31 shows spectrograms for the key witness channels along with $h(t)$ for this event.

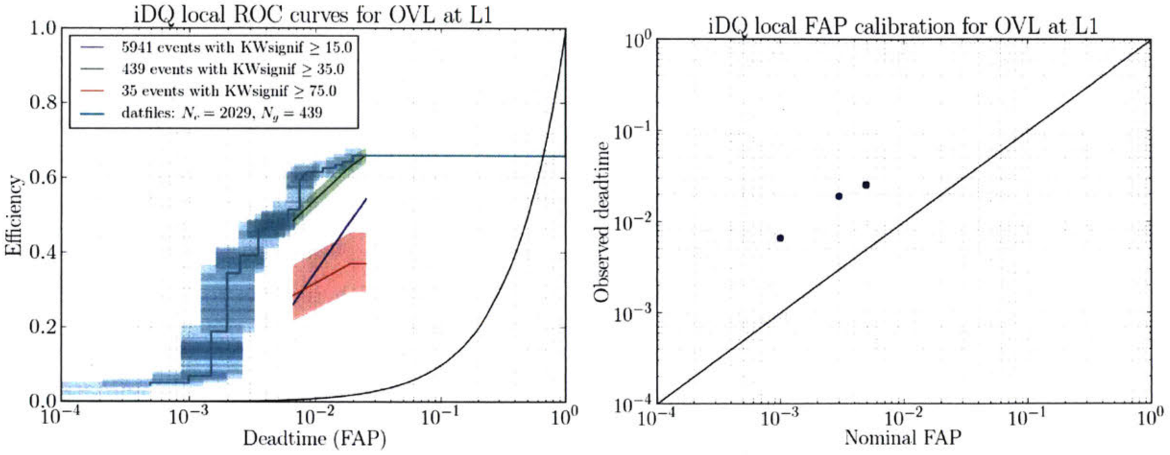


Figure 3-13: iDQ (left) ROC and (right) calibration coverage plots from LLO for G272601. The observed FAP is higher than the nominal FAP for this event, which is likely due to large numbers of whistles in the local neighborhood. iDQ assumes that glitches are relatively rare and therefore calibrates the FAP using segments generated from its time-series. However, if glitches are not rare, or suddenly become more common, they can make our calibration estimates misleading at times.

responsible for the veto should one occur. In addition to this, iDQ summarizes its local performance in both an ROC curve and a coverage plot, the latter of which shows how well the FAP time-series is calibrated[7]. Fig. 3-11, 3-12, and 3-13 demonstrates these annotations for a radio-frequency whistle at the Livingston detector during the second observing run. Although iDQ can run multiple classifiers simultaneously within a single process, we typically only upload information from OVL using ϵ/f.

iDQ can combine the output of multiple classifiers in an attempt to boost their joint performance; we call these *combiners*. Using basic assumptions about the form of the joint probability distributions over classifier output, we model the likelihood ratio to generate a new ranking. Typically, we take $p_{\text{joint}} = \max\limits_i \{p_i(\text{rank}_i)\}$ when we expect the classifier to be highly correlated (e.g. glitchy times) and $p_{\text{joint}} = \prod\limits_i p_i(\text{rank}_i)$ when they are independent (e.g. clean times). Furthermore, because we are limited by relatively small sample sizes for the glitchy distributions and cannot easily generate more glitchy times from historical data because of non-stationarity, we often find better performance by replacing the probability distributions in the above expressions with their

---

[7] Properly calibrated FAP time-series should produce deadtimes commensurate with the numerical value of any threshold applied to the time-series to create segments.

associated cumulative distributions so that

$$\Lambda_{\text{joint}} = \frac{\max_i \{p(\text{rank} \geq \text{rank}_i | \text{glitch})\}}{\prod_i p(\text{rank} \geq \text{rank}_i | \text{clean})} \tag{3.22}$$

While we do not see dramatic improvement over individual algorithms, we do find that the combiner often follows the maximum ROC curve from a selection of single algorithms. While this can complicate the interpretability of the algorithms' predictions with regards to which auxiliary channel is responsible, it does allow us to automatically select the best features from each OVL ranking metric. The combiner tends to follow ϵ/f and $\%_{\text{use}}$ at low FAP where they dominate and smoothly transitions to p_{poisson} at high FAP where ϵ/f and $\%_{\text{use}}$ can often stop reporting useful information.

Fig. 3-14 shows iDQ's performance over the same week surrounding GW151226 as Fig. 3-7. In general, we see comparable shapes for the ROC curves with excellent classification at LHO and above average at LLO. Importantly, Fig. 3-14 shows the FAP instead of the dead-ime, meaning it excludes times flagged that actually did contain a glitch. The difference is usually negligible (e.g. at LLO), but if there is a high glitch rate, it can be important. This is why the ordinates are different for the LHO curves in Fig. 3-7 and Fig. 3-14. We also note that the offline OVL runs typically reach higher detection efficiencies eventually, as expected, but the low FAP classification is often comparable. In particular, the general ordering of OVL's ranking statistics is mostly preserved in the online data.

We might ask how these ROC curves actually impact searches. At a basic level, searches perform coincidence measurements between detectors and, assuming uncorrelated Poisson distributed background in each detector, we expect the rate of accidentals to be $\lambda_{\text{acc}} = 2\Delta t \lambda_{\text{LHO}} \lambda_{\text{LLO}}$, where $\Delta t \sim 10\text{ms}$ corresponds to the time-of-flight between detectors. Although all classifiers give slightly different results, ϵ/f provides a decent benchmark. Within the week surrounding GW151226 (Fig. 3-14), iDQ's predictions using ϵ/f reduce the rate of glitches in LLO (LHO) to 47.6% (4.14%) of its original rate, requiring FAP $\lesssim 10^{-2}$. This means the rate of accidental coincidences is reduced to only 1.97% of its original value, over a 50-fold reduction in background.[8]

We also note that Λ_{joint} under-performs compared to the individual algorithms at times. This is associated primarily with detector non-stationarity and typically Λ_{joint} either follows the individual

---

[8] In reality, the impact will be somewhat less for searches because individual search algorithms can reject some of this background themselves.

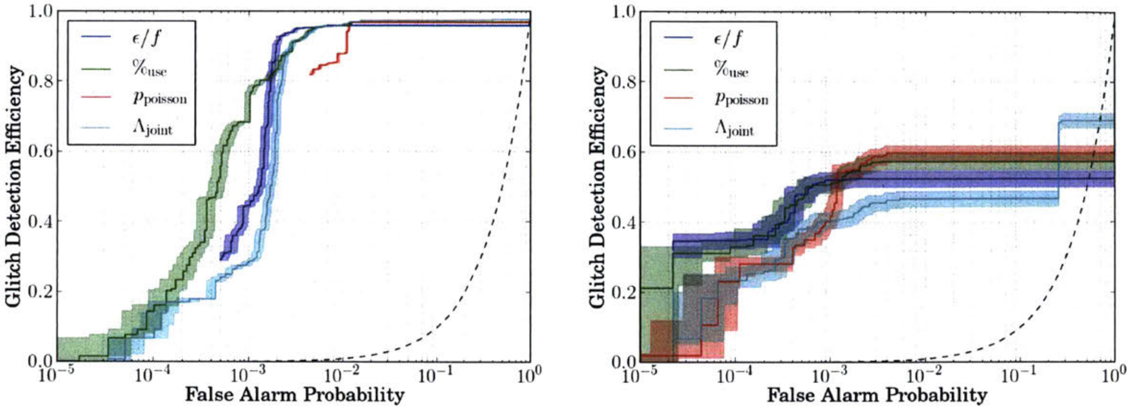


Figure 3-14: ROC curve from online iDQ runs in the week surrounding GW151226 (00:00:00 UTC on December 17, 2015 to 00:00:00 UTC on December 27, 2015) from (left) LHO and (right) LLO, analogous to Fig. 3-7. This corresponds to 168.61 and 129.17 hours of live-time at LHO and LLO, respectively, containing 65,320 and 496 glitches, respectively. We note that the FAP reported here are lower than the deadtimes reported in Fig. 3-7 for LHO because the deadtime includes glitchy times, while the FAP specifically excludes ±100ms surrounding each glitch. The difference is usually unimportant (e.g. the LLO results) unless there are *many* glitches in the data. For this training set, LHO has a glitch every ~ 9.3 seconds whereas LLO has a glitch every ~ 15 minutes. The error bars at LHO are also dominated by the estimation of FAP rather than glitch detection efficiency.

ROC curves much closer or outperforms them. Fig. 3-15 demonstrates this in a subset of Fig. 3-14's data at LHO.

We note that MLA algorithms are also supported under iDQ but due to computational limits associated with the I/O required for these particular implementations, we found they often struggled to keep up with real-time. OVL operates strictly in memory during the prediction step and is much faster.

3.3 Parameter Estimation and Localization

Now that we have discussed both trigger generation and data quality, we can consider parameter estimation assuming that there is certainly a signal within the data and that it is not polluted by noise artifacts. The basic goal of parameter estimation is to reconstruct the GW strain recorded by a network of detectors in the presence of noisy measurements. When there are well trusted source models, we can parametrize the GW strain with quantities like the component masses and spins for

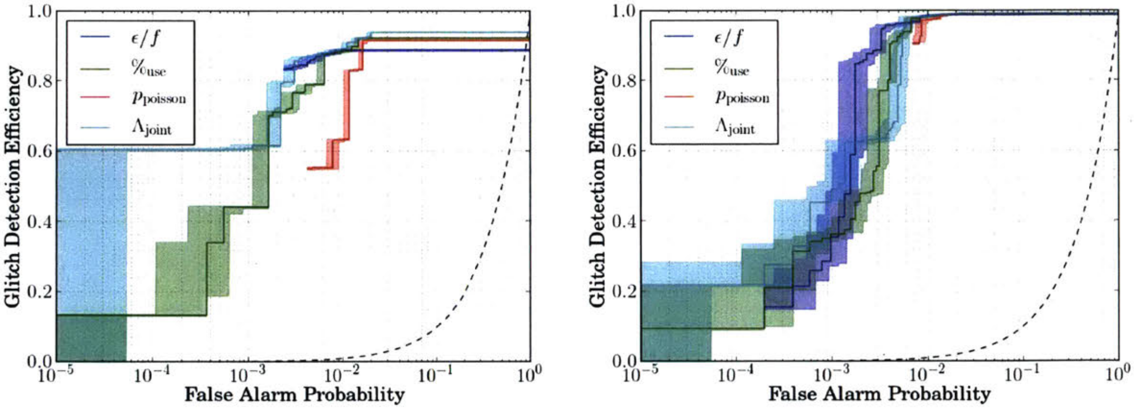


Figure 3-15: The general performance of $\Lambda_{\rm joint}$ relative to single ranking metrics using subsets of Fig. 3-14's LHO data from December (left) 22nd and (right) 23rd, 2015.

CBC sources. §2.2 gives an example of this for binary NS systems. Two important parameters that are present regardless of whether we have a confident signal model are the Equatorial coordinates of the GW point source; the Right Ascension (α) and Declination (δ).[9] Determining the sources' location on the sky is critical for multi-messenger follow-up observations, but it poses some interesting challenges. I will describe the basic physical mechanisms used to localize sources in §3.3.1 as well as comprehensive methods to characterize and compare localization estimates from different algorithmic approaches in §3.3.2.

3.3.1 Basics of Triangulation

GW detectors are nearly omni-directional instruments; their antenna responses are dominated by the $l = 2$ spherical harmonic and therefore vary slowly over the sky (Fig. 1-5). This means they are sensitive to many possible source locations but also means a single IFO cannot determine the source location accurately.[10] Instead, GW localizations rely on the time-of-flight between distant detectors via triangulation. According to GR, GWs travel at the speed of light and therefore time-delays correspond to extra distance traversed. As a toy example, let us consider localizing an un-modeled signal with a network of idealized all-sky instruments sensitive to a single GW

---

[9] We could also consider extended sources, but typically our expected sources are nearly point sources, e.g. binary stars.

[10] If a signal's spectral shape is known a priori, and the signal has large enough $\rho_{\rm net}$, single IFO localizations may be reduced. However, the necessary $\rho_{\rm net}$ is much larger than we expect from realistic source distributions.

polarization. In this case, the maximum likelihood estimate (Eqn. 3.14) yields

$$\ln p = 2\int df \sum_{a,b} d_a^* B_{ai}^* \left(A^{-1}\right)_{ij} B_{bj} d_b \tag{3.23}$$

$$= 2\int df \sum_{a,b} d_a \frac{Fe^{-2\pi i f \Delta t_a}}{S_a} \left(\sum_d \frac{F^2}{S_d}\right)^{-1} \frac{Fe^{2\pi i f \Delta t_b}}{S_b} d_b \tag{3.24}$$

$$= 2\sum_a \int df \, |d_a|^2 \frac{F^2}{S_a^2} \left(\sum_d \frac{F^2}{S_d}\right)^{-1} + 2\sum_{a\neq b} \int df \, d_a^* d_b \frac{F^2}{S_a S_b} \left(\sum_d \frac{F^2}{S_d}\right)^{-1} e^{2\pi i f (\Delta t_b - \Delta t_a)} \tag{3.25}$$

where I have explicitly included the time-delays at each IFO as part of the antenna response ($\Delta t_d = n_i r_d^i$ where the wave propagates in the direction defined by n_i and the detector is located at r_d^i). We recognize the second term in the final line as a measurement of the time-delay between detectors ($\Delta t_a - \Delta t_b$), which is the basis of triangulation. If we further consider this term and assume the data recorded by each detector is $d_a = Fe^{-2\pi i f (\Delta t_a)_o} h$, we obtain

$$\ln p \supset 2\sum_{a\neq b} \int df \, |h|^2 \frac{F^4}{S_a S_b} \left(\sum_d \frac{F^2}{S_d}\right)^{-1} e^{2\pi i f \delta t_{ab}} \tag{3.26}$$

$$= 2\mathcal{R}\left\{ \sum_{a\neq b} \int df \, |h|^2 \frac{F^4}{S_a S_b} \left(\sum_d \frac{F^2}{S_d}\right)^{-1} e^{2\pi i f \delta t_{ab}} \right\} \tag{3.27}$$

where $\delta t_{ab} = -\delta t_{ba} = (\Delta t_b - \Delta t_a) - ((\Delta t_b)_o - (\Delta t_a)_o)$. Therefore, we see that the posterior is the product of individual posteriors on the time-delay between pairs of detectors. What's more, the individual posteriors are the real parts of the Fourier transform of the noise weighted signal power in the frequency domain, with the time-delay as the Fourier conjugate to frequency. This means that broadband signals produce narrow posteriors on the time-delay and narrow band signals produce broad posteriors on the time-delay. Furthermore, high-frequency narrow band signals can produce oscillatory posteriors on the time-delay, which we call *fringe peaks.* When we have only two detectors, fringe peaks produce concentric triangulation rings. If there are fringe peaks from each pair of three detectors, we can obtain a checker-board pattern. Interestingly, because detectors can have different sensitivities at different frequencies, we could have some baselines with fringe peaks (at least one narrow band detector) and other baselines without fringe peaks (broadband sensitivity). This would produce a single triangulation ring between the broadband detectors modulated by the fringe peaks from baselines involving the narrow band detector, reminiscent of a string of beads. Fig. 3-16 demonstrates exactly this analytic beading behavior using our toy model of idealized all-sky instruments with two broadband detectors located at the LIGO sites and a narrow-band detector at Virgo, with an approximate binary black hole coalescence. The exact morphology depends on the relative sensitivity and number of detectors.

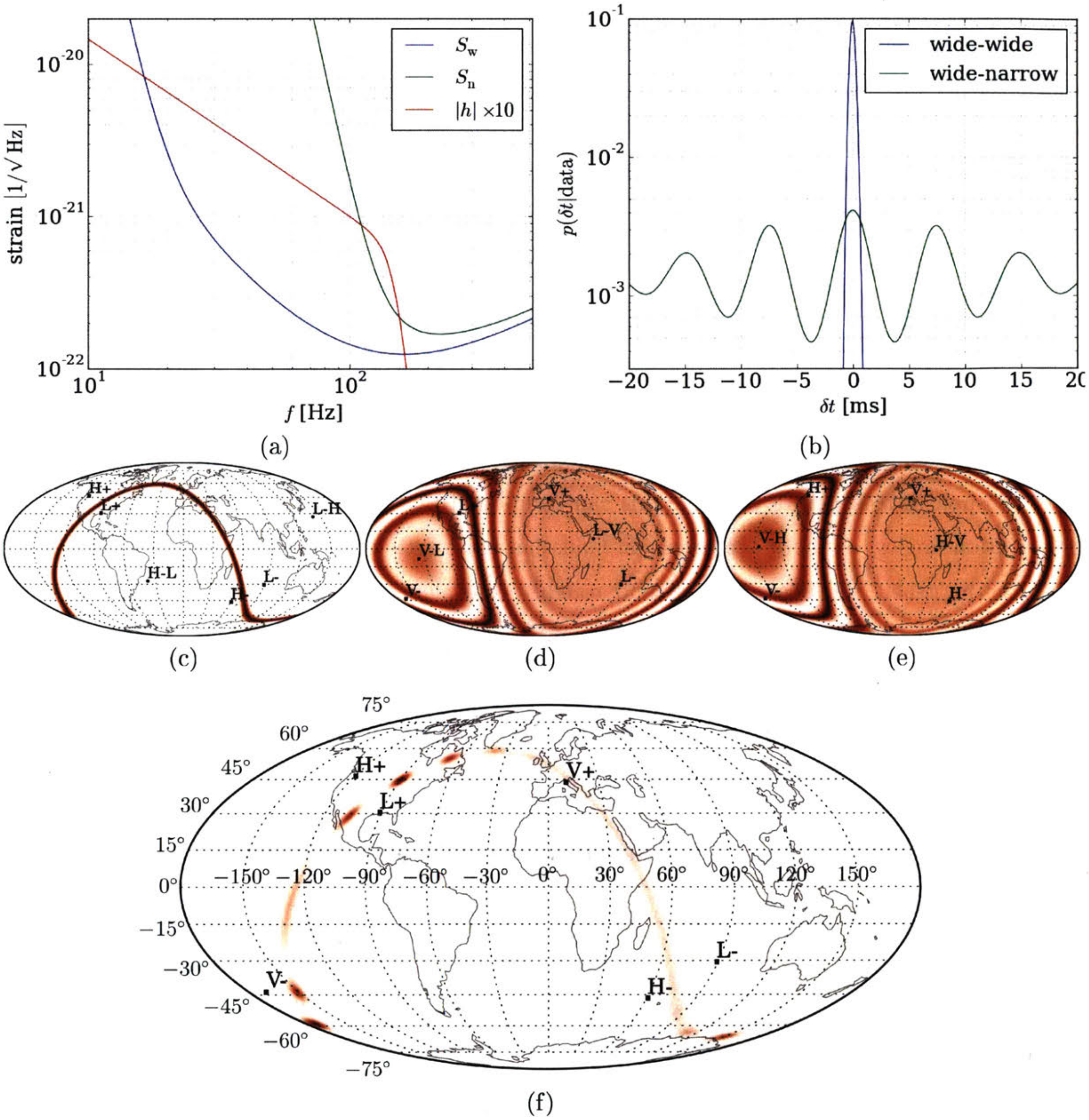
It is worthwhile to note that there can be strong degeneracies across the sky when we only have a few detectors. In particular, the likelihood vanishes unless there are more GW detectors than polarizations (in GR, this means $N_{\mathrm{IFO}} \leq 2$). That is because we are trying to reconstruct two functions with two measurements, and without prior knowledge of the waveform's shape we can infer strain signals from any direction that match the data equally well.[11] This degeneracy is typically broken by the application of prior knowledge, such as the signal's spectral shape [162], the distribution of the distance to sources, or by additional constraints such as reconstructing only one polarization [86].

It may not be immediately clear how knowledge of sources' distance distributions can improve localization, particularly if the detectors are nearly omni-directional. However, the argument is rather straightforward. For locations far from the true source location's triangulation ring, the reconstructed strain must contain a careful balance of multiple polarizations to reproduce the observed strain. This balance usually involves more signal energy than the true location because of cancellations between the projected strain from each polarization. Therefore, the inferred fluence from the GW signal is larger away from the correct triangulation ring because the GW polarizations add in quadrature within the fluence. If we assume all sources have similar energy distributions independent of their distance, then the observation of a higher fluence implies a closer distance. Thus, a priori knowledge that distant sources are more likely than close ones informs the posterior for source location, even when the likelihood is degenerate across the entire sky. Of course, if the IFOs also have directional sensitivity, the basic logic remains unchanged but the details become more complicated. For example, the *antenna response prior* derived therefrom modulates the posterior along triangulation rings [60].

Triangulation also provides sanity checks of localizations. For example, an ensemble of localiza-

---

[11] To be precise, whenever the rank of A_{ij} from Eqn. 3.14 is $\geq N_{\mathrm{IFO}}$, this problem exists. Perfectly aligned detectors produce rank-deficient A_{ij}, as do an arbitrary number of truly all-sky instruments ($F(\theta, \phi)$ = constant). These networks are effectively sensitive to fewer polarizations, allowing for non-trivial localizations even with only two detectors.



(a) (b)

(c) (d) (e)

(f)

Figure 3-16: A toy model of analytic triangulation from an un-modeled signal reconstruction using two broadband detectors and one narrow-band detector. (a) The PSDs and signal. (b) The associated time-delay likelihoods; note the oscillatory pattern for baselines including the narrow-band detector. This is because the signal is truncated at low frequencies in that detector. (c), (d), (e) Localization estimates using pairs of detectors; (c) shows a single ring while (d) and (e) demonstrate fringe peaks. (f) The combined likelihood, which resembles a string of beads.

tions for the same two IFOs *cannot* cross in Geographic coordinates because they are all determined by time-delays relative to the same line-of-sight. It also provides a natural coordinate system in which the localization posteriors are nearly separable: $p(\theta_{\rm los}, \phi_{\rm los}) \sim p(\theta_{\rm los})p(\phi_{\rm los})$. Therefore, sensible localizations in the line-of-sight frame should have small Mutual Information Distances (MIDs)[12]

$$\mathrm{MID} = \frac{I(\theta_{\rm los}, \phi_{\rm los})}{H(\theta_{\rm los}, \phi_{\rm los})} = \frac{\sum_i p(\theta_i, \phi_i) \ln \frac{p(\theta_i, \phi_i)}{p(\theta_i)p(\phi_i)}}{-\sum_i p(\theta_i, \phi_i) \ln p(\theta_i, \phi_i)} \in [0, 1] \tag{3.28}$$

This turns out to be an effective way to reject certain classes of noise. For the LHO-LLO network, large ρ imbalances in the detectors are often localized to dots that correspond to $\sim \pm 9$ms time-delays between the detectors (Fig. 3-17). These locations are associated with poles in the LLO antenna response,[13] meaning the ρ imbalance can be compensated by the antenna response alone. Differences in the responses to separate polarizations can compound this for certain orbital inclinations and polarization angles. If there is a loud non-Gaussian noise artifact in one IFO and pure Gaussian noise in the other, the posterior will find roughly equivalent support at both $+9$ms and -9ms time-delays because the Gaussian noise is roughly equivalent at both time shifts. This produces localizations with dots at nearly antipodal points in the LHO-LLO line-of-sight frame (Fig. 3-17), importantly with different $\phi_{\rm los}$. These localizations produce large MID and are distinguishable from true signals (Fig. 3-18).

3.3.2 Comparison and Characterization of Localizations

With a basic understanding of how localizations are produced from GW strain measurements, we can now characterize typical localizations from various algorithms and compare their estimates. Localization estimates are probability distributions defined on a 2-sphere. As such, information theoretic metrics are immediately applicable.

We focus on comparisons of localizations from algorithms targeting un-modeled transients (bursts), specifically Coherent WaveBurst [86, 87], LALInferenceBurst [98, 162], and BayesWave [44]. However, our techniques are applicable to maps from any algorithm. We use an ensemble of simulated signals injected into simulated Gaussian noise characteristic of projected noise budgets for the

---

[12] We define MID for discretized localizations in terms of a series of pixels centered at (θ_i, ϕ_i).

[13] The corresponding poles in the LHO response are less problematic because they occur along triangulation rings that span regions of higher network sensitivity, and are therefore rarely favored over the higher sensitivity regions.

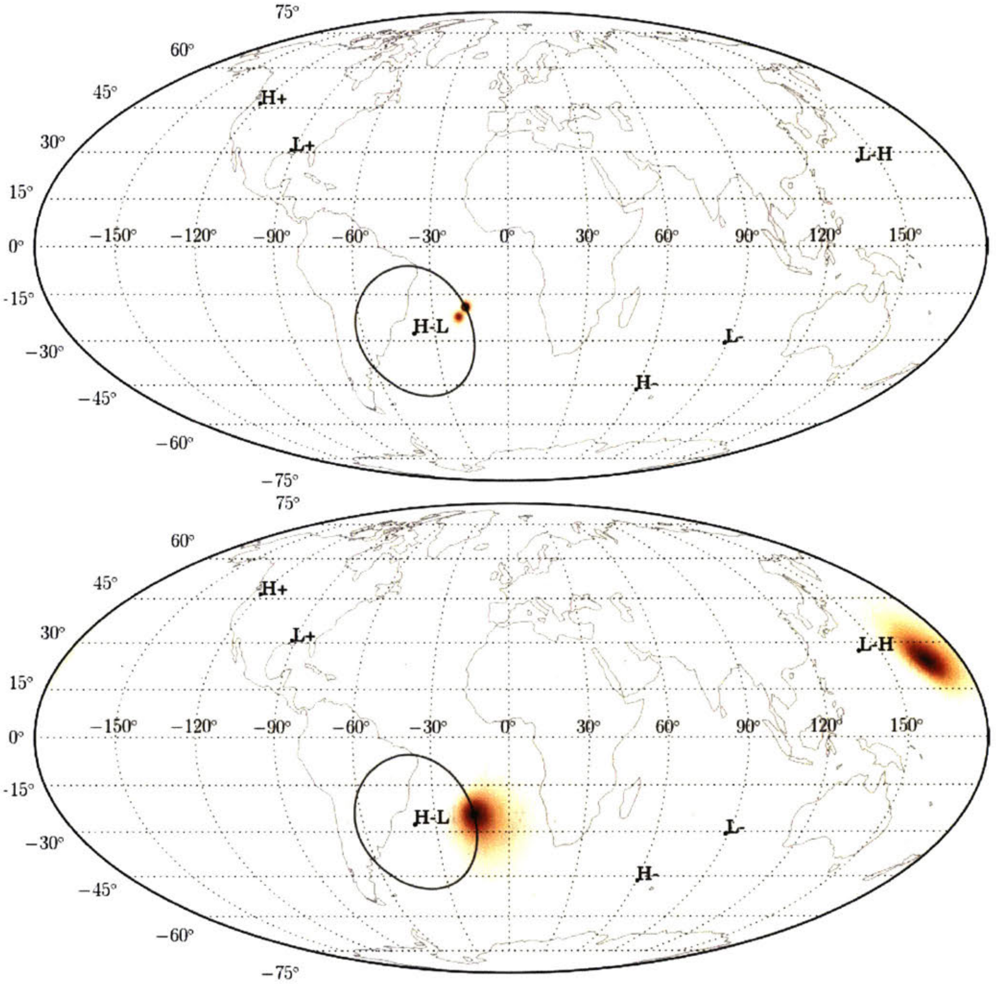


Figure 3-17: Two example localizations from (top) G273538 (22:43:34 UTC February 11, 2017) and (bottom) G272474 (23:40:34 UTC February 3, 2017), both of which are associated with glitches. G272474 shows support at both +9ms and −9ms whereas G273538 shows the characteristic dot behavior near only a single pole. Note that the dots for G273538 lie on different triangulation rings. For both events, the localization's shapes are driven by the poles of the antenna response rather than the data, which should produce a ring.

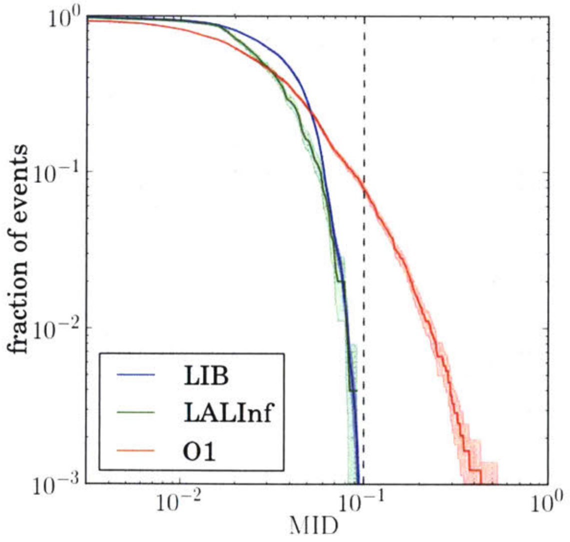


Figure 3-18: MID for two separate algorithms using campaigns of simulated injections and all localization estimates uploaded to GraceDb [140] during O1. We note that $\sim 8\%$ of all localization estimates can be vetoed based on the MID alone without losing a single injection.

early advanced detector era [22, 60, 61]. This includes a 2-detector network composed of the LIGO Livingston and Hanford sites as well as a 3-detector network containing the two LIGOs and Virgo. We investigate four signal morphologies: sine-Gaussians (SG), Gaussians (G, the limit of SG as $f \rightarrow 0$), White-Noise Bursts (WNB), and binary black hole (BBH) coalescences with (anti-)aligned spins.

Single-Algorithm Statistics

When analyzing a single localization estimate, it is natural to consider both the accuracy and the precision of the map. We consider several statistics that quantify both.

Accuracy can clearly be approximated by the angle separating the maximum a posteriori (MAP) location and the true location ($\delta\theta$). Fig. 3-19 shows this for 2- and 3-detector networks. We note that the general behavior is similar for both networks; this is because the LIGOs are more sensitive than Virgo and therefore dominate. Furthermore, we also see a symmetry in $\delta\theta$, with local maxima in $p(\delta\theta)$ near 0° and 180°. This is due to a degeneracy in the LLO-LHO network, which has maxima in the network sensitivity above both North America and the Indian Ocean. Therefore, occasionally the MAP estimate picks the wrong side of the sky.

Another measure of accuracy that is commonly used is the searched area, defined as the amount of sky assigned a probability greater than or equal to the probability assigned to the source's true location. If follow-up is conducted according to the posterior's ranking of locations, this approximates the amount of area observed before the true location is surveyed. Unlike $\delta\theta$, we see a marked improvement in the searched area between 2- and 3-detector networks, although the details depend on the algorithm and signal morphology. This is because, even though Virgo is less sensitive than either LIGO, it can rule out certain parts of the sky and greatly reduce the posterior's support along triangulation rings. We also note that the searched areas achievable for burst signals are only a factor of 2-3 larger than what is expected for broadband BNS signals recovered with perfect templates [24, 126]. This is consistent with what was observed for GW150914, which was first detected and localized by burst searches (§4.1). Tables 3.1 and 3.2 list some characteristics of the measured distributions of $\delta\theta$ and searched area for cWB and LIB, respectively.

Table 3.1: Summary of cWB localization. Statistical error is on the order a few percent.

year		2015				2016			
network		HL				HLV			
morphology		BBH	SG	G	WNB	BBH	SG	G	WNB
fraction (in %) with searched area less than	5 deg^2	3.1	3.4	3.1	6.2	6.7	9.3	12.0	17.2
	20 deg^2	11.5	12.7	10.9	17.8	18.9	22.2	28.8	32.5
	100 deg^2	35.3	37.2	37.1	51.8	47.3	52.3	54.9	61.3
	200 deg^2	51.6	52.2	49.2	69.7	62.3	66.9	69.1	75.8
	500 deg^2	75.9	69.2	73.8	86.5	82.2	85.8	84.9	91.2
	1000 deg^2	89.2	82.2	87.1	95.6	93.1	94.8	95.0	98.0
fraction (in %) with $\delta\theta$ less than	1°	1.3	1.4	0.8	2.7	3.6	2.8	9.8	10.0
	5°	12.9	8.5	11.7	13.0	22.6	13.0	29.3	19.0
	15°	37.2	27.1	33.2	34.1	37.6	26.4	45.8	32.6
	45°	73.1	61.4	66.0	70.9	61.3	57.3	67.9	59.2
	60°	79.5	68.4	71.1	74.9	66.7	62.2	71.7	64.7
	90°	83.1	74.0	75.8	77.9	71.4	67.1	74.3	70.0
median searched area		184.6 deg^2	181.6 deg^2	209.9 deg^2	93.0 deg^2	112.5 deg^2	91.7 deg^2	71.3 deg^2	61.3 deg^2
median $\delta\theta$		23.1°	31.6°	25.7°	23.9°	27.5°	36.7°	18.6°	33.9°

Table 3.2: Summary of LIB localization. Statistical error is on the order of a few percent.

year network morphology		2015 HL				2016 HLV			
		BBH	SG	G	WNB	BBH	SG	G	WNB
fraction (in %) with searched area less than	5 deg^2	1.8	4.0	2.3	4.7	8.6	25.5	18.5	10.4
	20 deg^2	9.4	14.3	7.0	14.2	23.1	47.4	43.0	23.9
	100 deg^2	31.8	39.0	34.8	35.2	53.8	75.8	73.6	46.8
	200 deg^2	46.8	52.9	49.2	51.2	65.7	84.6	84.4	59.2
	500 deg^2	70.2	71.7	72.7	65.1	82.9	92.6	93.3	69.8
	1000 deg^2	88.2	82.6	89.1	74.9	93.0	94.9	97.4	76.0
fraction (in %) with $\delta\theta$ less than	1°	1.0	2.1	1.2	2.1	6.2	11.4	12.0	5.5
	5°	8.6	8.8	11.7	9.5	34.5	31.9	51.4	17.2
	15°	32.2	25.8	30.1	28.6	54.6	53.8	66.8	31.3
	45°	66.8	63.7	63.7	61.9	77.1	78.3	83.7	63.6
	60°	72.8	71.0	68.8	67.2	81.3	81.8	85.8	70.1
	90°	77.4	75.9	74.2	70.4	83.7	84.6	86.5	76.2
median searched area		238.5 deg^2	171.0 deg^2	208.4 deg^2	180.9 deg^2	82.5 deg^2	22.2 deg^2	31.3 deg^2	121.3 deg^2
median $\delta\theta$		26.6°	29.4°	27.1°	30.4°	11.1°	13.3°	4.9°	27.5°

The searched area also estimates the precision of a localization estimate. More precise (unbiased) estimates will produce smaller searched areas. We often use the size of confidence regions as well. These are immediately interpretable as the uncertainty in the localization. However, instead of choosing a few somewhat arbitrary confidence levels, we often consider the Shannon entropy of the discretized map normalized to a fraction of the sky

$$\mathcal{H} = A_{\mathrm{pix}} e^{-\sum\limits_i p_i \ln p_i} \tag{3.29}$$

This correlates strongly with the size of individual confidence regions but includes information from all of them. Fig. 3-19 shows the entropy of typical maps. Again, we tend to see more precise localizations for 3-IFO networks.[14]

Comparisons of cWB, LIB, BW, and LALInference

In addition to comparing statistics from ensembles of localization estimates separately, we can also compare estimates on an event-by-event basis. This better captures the relevant information for practical situations when we have multiple estimates for a single event and must choose which to believe.

A natural metric is the Kullback–Leibler (KL) divergence [90]

$$D_{\mathrm{KL}}(p|q) \equiv \sum_i p_i \ln\left(\frac{p_i}{q_i}\right) \tag{3.30}$$

which is a measure of the information gained when refining an estimate of a probability distribution from an initial guess (q) to the true distribution (p). However, we often find this is unstable for our localization estimates because if there is even a single pixel that is assigned non-zero posterior support by one algorithm and zero support by another, the KL divergence diverges. This renders it extremely sensitive to numeric noise in the distributions' tails. Instead, we focus on the Fidelity (F) between two distributions p and q

$$F(p,q) = \sum_i \sqrt{p_i \cdot q_i} \in [0,1] \tag{3.31}$$

---

[14] Entropy is really only meaningful when the posteriors are correctly calibrated and produce diagonal coverage plots. Nonetheless, it provides a useful heuristic in all cases.

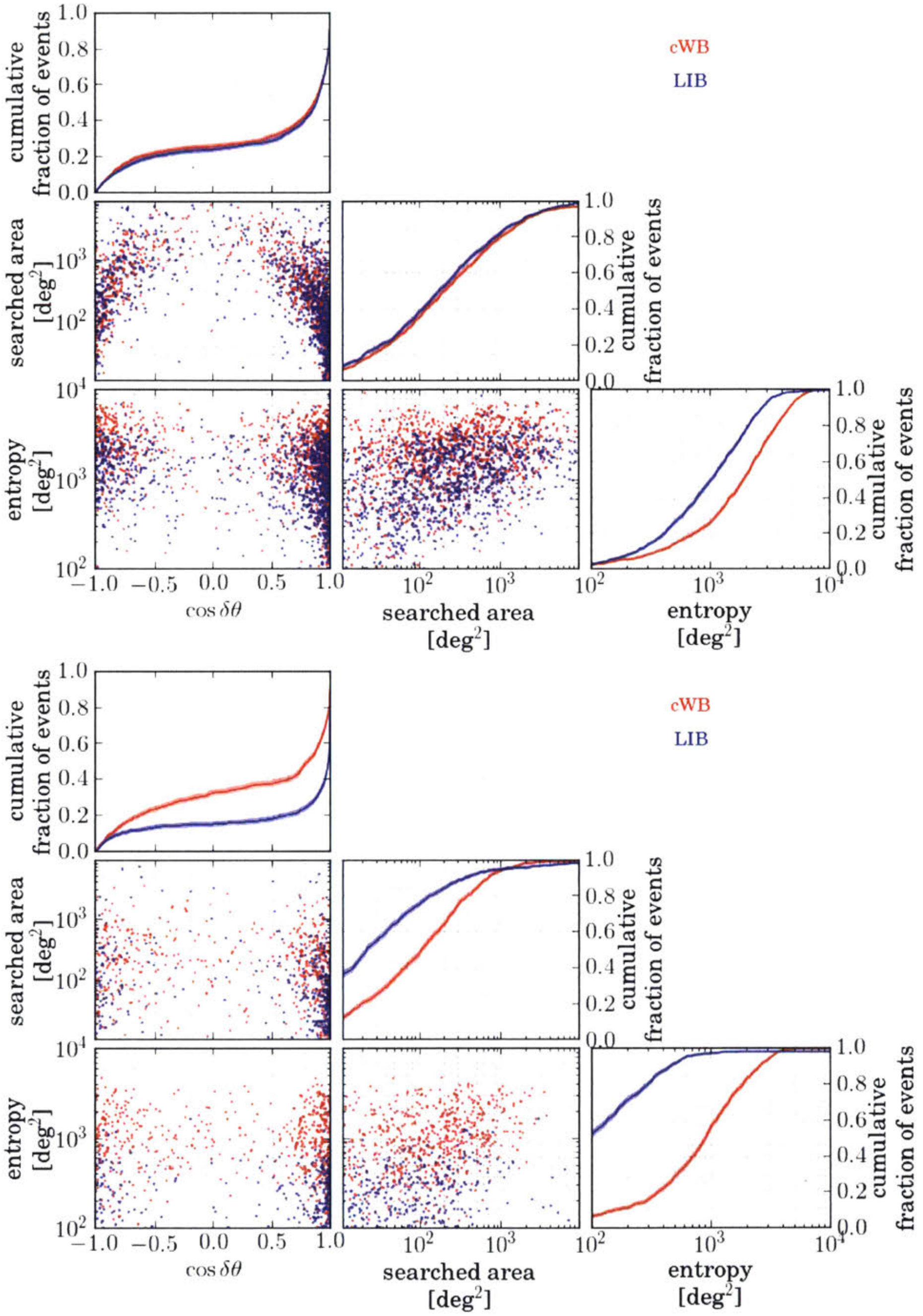


Figure 3-19: $\delta\theta$, searched area, and $\mathcal{H}$ for SG injections in the (top) 2- and (bottom) 3-detector networks. There is a clear correlation between $\delta\theta$ and searched area, but only a weak correlation between either and $\mathcal{H}$. This is because the overall precision of the map may not change much even though some injections lie near the maximum a posteriori and some do not.

where the square root guarantees the range for F because the individual distributions are L^1 normalized. This is much less sensitive to noise in our localization estimates.

For BBH signals, we have reliable signal models and can also reconstruct these signals with the full template using LALInference [162]. Fig. 3-20 shows the Fidelities for cWB, LIB, and LALInference for these simulations [165]. We find that the burst algorithms are much more similar to each other than they are to LALInference and that differences are dominated by the selection of different triangulation rings. The latter is apparent by comparing the Fidelity using the full localization estimate with the Fidelity computed using the time-delay marginals. We also note the strong correlation between the two, which is driven by the fact that large full-localization Fidelities are only possible when there are large Fidelities between the time-delay marginals. However, differences in the distribution around the ring could produce small all-sky Fidelities even with large time-delay Fidelities. For the most part, we do not see this, meaning different algorithms primarily select the same neighborhoods of nearby rings even if they do pick separate rings.

We also find that the full LALInference maps can be significantly more accurate than burst maps. Therefore, if the follow-up is not time critical, it is almost certainly better to wait for the few days required to produce a LALInference map.

The situation can be more complicated for 3-detector networks, but this is mostly due to LALInference's ability to localize the signal significantly better than burst algorithms by making use of the known waveform and Virgo's data. The general trends observed in the 2-detector data are simply exaggerated.

Model Averaging Localizations

We have seen that, while individual algorithms can differ in their localization estimates, they generally tend to broadly agree. Nonetheless, the differences can be non-trivial at times (e.g. triangulation rings separated by $\gtrsim 1^\circ$) and it can be difficult to determine which map to use. In particular, burst algorithms adopt different signal models when reconstructing the source location and it is not always clear which set of assumptions are more appropriate. An alternative to choosing a single localization is to average all estimates together. This is equivalent to marginalizing over the unknown signal model with equal priors for each model, in some sense. We investigate the effects of this procedure for our ensemble of maps.

Surprisingly, the searched area for averaged maps seems to select the best searched area of

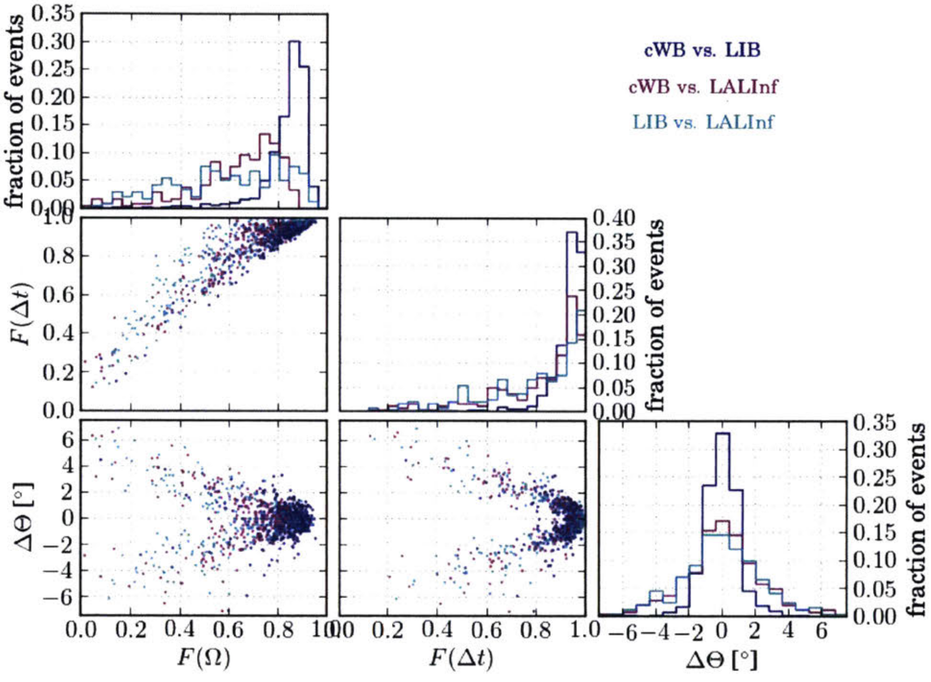


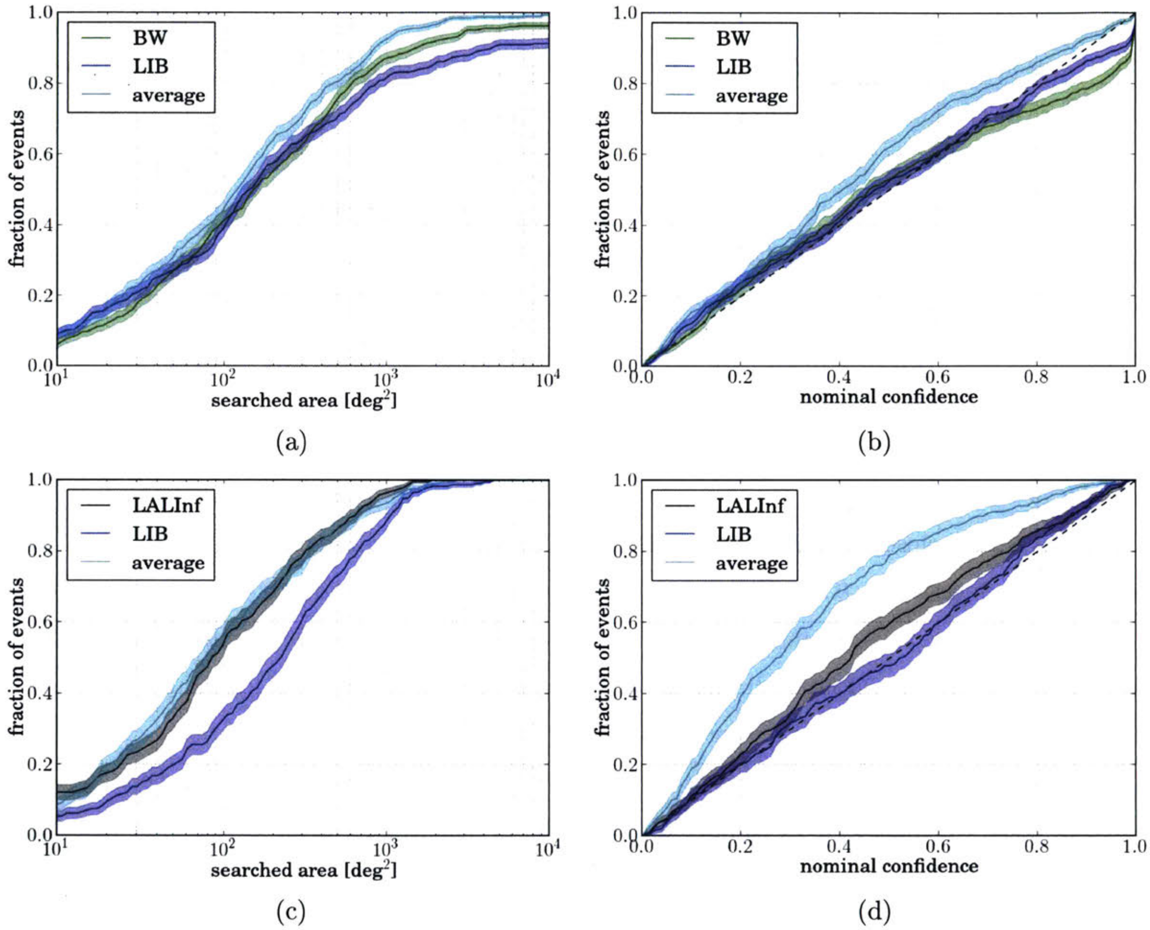
Figure 3-20: Fidelities of the full-localization and the time-delay marginals along with the angular offset between triangulation rings for massive BBH simulations with the LHO-LLO network. We see a clear correlation between both Fidelities and $\Delta\theta$, showing that localization estimates differ primarily in which triangulation rings they select.

any constituent map both when the performances of the individual maps are similar and when they are very different. This is likely because the smallest constituent map places high posterior weight near the correct location and, even when this is averaged with other maps, there is still significant support nearby. The intuition may also work the other way. A less accurate but more concentrated map may not hurt the better constituent map much when averaged because the inaccurate concentrated region will be exhausted quickly and the tails of the distribution will be dominated by the better constituent. We even see strict improvement over any individual algorithm for some sets of averaged maps (e.g. the LIB+BW averaged map for WNB with 2-detectors outperforms either individual algorithm, Fig. 3-21). Therefore, model averaging appears very robust when accounting for modeling uncertainty between different algorithms.

Beyond the searched area, we also expect model averaging to affect the coverage plots. Different algorithms produce different coverage plots and not all algorithms are always correctly calibrated. We find that averaged maps always produce over-covered maps, meaning a larger fraction of events are contained in confidence regions than is expected based on the nominal confidence. This behavior is also apparent in the entropy distributions; all averaged maps produce entropy distributions similar to the least well-calibrated individual algorithms included in the average. This appears to be true regardless of the individual algorithm's behavior, even if one of the algorithms systematically produces under-covered localizations. Interestingly, averaging perfectly calibrated algorithms also produces over-coverage (Fig. 3-21c, d). While this can complicate the interpretation of confidence regions for averaged maps, we can be confident that we err on the side of caution with this procedure. For this reason, and the robust nature of averaged maps' searched areas, model averaging has been adopted as *the* procedure for handling multiple localization estimates when there is no clear preference for one estimate over another.

3.4 Importance of Frequency Dependent Response for 3G Detectors

Determining where sources come from depends on knowledge of how the detectors respond. As described in §1.3.1, GW detectors' response to GW signals depends on the frequency of the signal; the round-trip light travel time corresponds to the key frequency. For aLIGO, this free spectral range is $f_{\mathrm{fsr}} = c/2L = 37.2\mathrm{kHz}$, which is much larger than GW frequencies of interest. This



(a) (b)

(c) (d)

Figure 3-21: Coverage and searched area histograms showing how model averaging performs in comparison to individual algorithms with the LHO-LLO network. (a) and (b): BW and LIB estimates for WNB showing strict improvement in searched area with the averaged maps. (c) and (d): LIB and LALInference estimates for BBH coalescences, showing over-coverage even when both algorithms are well calibrated separately.

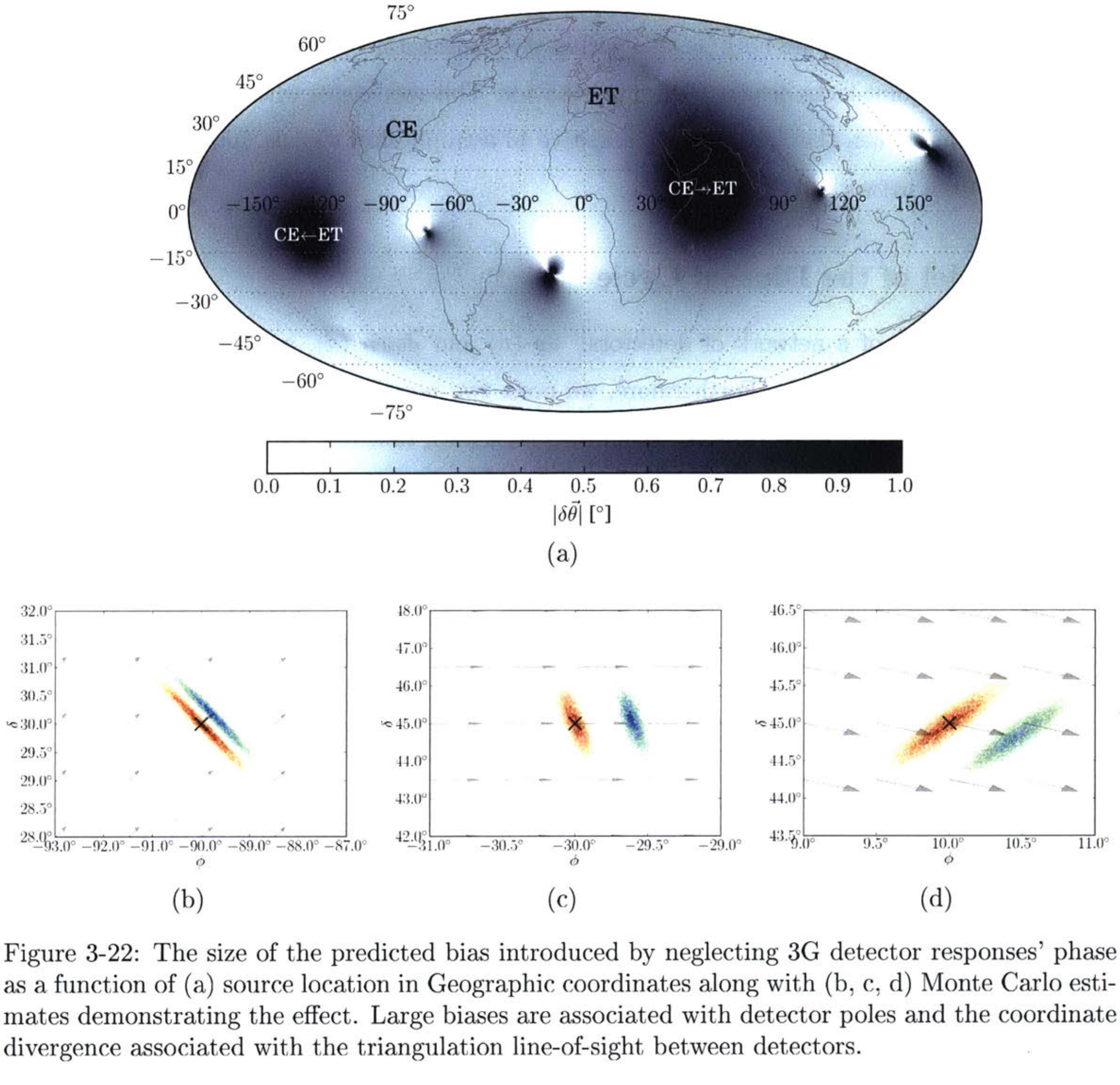
motivates the common approximation that the detector's response is independent of frequency and is a purely geometrical effect of projecting strain onto the detector arms. However, proposed 3^{rd} generation detectors may have much longer arms. Cosmic Explorer [139] may have arms in excess of 40-km, corresponding to $f_{\text{fsr}} \leq 3.7\text{kHz}$. Assuming a $1.4M_{\odot}$, 12km-radius neutron star, the dynamical frequency of internal modes will be approximately $f = \sqrt{GM/R^3}/2\pi \sim 1.65\text{kHz}$, which is an appreciable fraction of f_{fsr}. It is not unreasonable to question whether the frequency dependence of the antenna response is important for such signals.

While the change in the relative directional sensitivity can be dramatic, the main impact involves the additional phase introduced by the detector response. Indeed, at low frequencies this phase is nearly linear in frequency (Fig. 1-6), with a mild dependence on the source direction relative to the detector arms. This phase acts exactly like a time-of-arrival delay, and therefore differences in this phase between detectors can easily be confused with time-delays associated with triangulation, thereby biasing localization estimates. These biases should be on the order of $10''$ for current 4-km detectors, which is much smaller than even the most optimistic statistical uncertainties. However, they can be as large as $\sim 0.5^\circ$ for networks of 3G detectors with arms as long as 40-km. What's more, for systems close enough to make electromagnetic follow-up interesting ($D_L \leq 500$ Mpc), the statistical uncertainty (the triangulation ring's width) will be considerably smaller (Fig. 3-22). In fact, we find that these systematics remain comparable to or larger than statistical uncertainty even for sources at several Gpc. Importantly, the systematic bias does *not* depend on the signal and will apply equally to all detections if not taken into account.

I have focused on CBC sources, which are dominated by their low frequency inspiral components. Higher frequency signals may be impacted differently, although the bias in localization will apply equally as long as $f \ll f_{\text{fsr}}$.

3.5 Impact of Diurnal Cycles on the Exposure of GW Detector Networks

A vital aspect of determining where detected signals come from is determining where you looked. This is quantified by the detector responses, but those are only fixed in Geographic coordinates that rotate with the Earth. A more natural coordinate system is fixed to the distant stars. Naively, one would expect the Earth's rotation to average out any dependence on the Right Ascension (α),



(a)

(b) (c) (d)

Figure 3-22: The size of the predicted bias introduced by neglecting 3G detector responses' phase as a function of (a) source location in Geographic coordinates along with (b, c, d) Monte Carlo estimates demonstrating the effect. Large biases are associated with detector poles and the coordinate divergence associated with the triangulation line-of-sight between detectors.

leaving only a gentle modulation with Declination (δ). However, that is only the case if the IFOs operate uniformly through time. In reality, ground-based IFOs show a strong preference for locking during their local night [40]. Fig. 3-23 shows histograms for S5 [3, 6], S6 [4, 5], and O1 [7, 149, 150] wrapped with a 24-hour period. All of these recent runs, which represent very different detectors, show strong preferences toward locking at night.

Because the two LIGOs are both located within the continental United States, their local nights coincide and their joint-locks display the same behavior. This modulates when the detectors record data and therefore when they are actually sensitive to astrophysical signals, introducing a dependence on α to the network exposure [40].

3.5.1 Model of the Diurnal Cycle

We posit two states of a network of detectors: *up* (u) and *down* (d), with science-quality data available only in the *up* state. Furthermore, we posit two possible causes for detectors being in the *down* state: *random* causes (r) that are uncorrelated with time and *cyclic* causes (c) which are correlated with time, usually through a diurnal cycle. We note that the *random* and *cyclic* models are not necessarily mutually exclusive and the detector could be down for multiple reasons at the same time. Therefore, we have

$$p(d|t) = p(d|c,t)p(c|t) + p(d|r,t)p(r|t) - p(d|c \cap r, t)p(c \cap r|t) \tag{3.32}$$

Furthermore, because the *up* and *down* states are mutually exclusive and span the space of possible detector states at any single time, we have

$$p(u|t) = 1 - p(d|t) \Rightarrow p(u) = 1 - p(d) \tag{3.33}$$

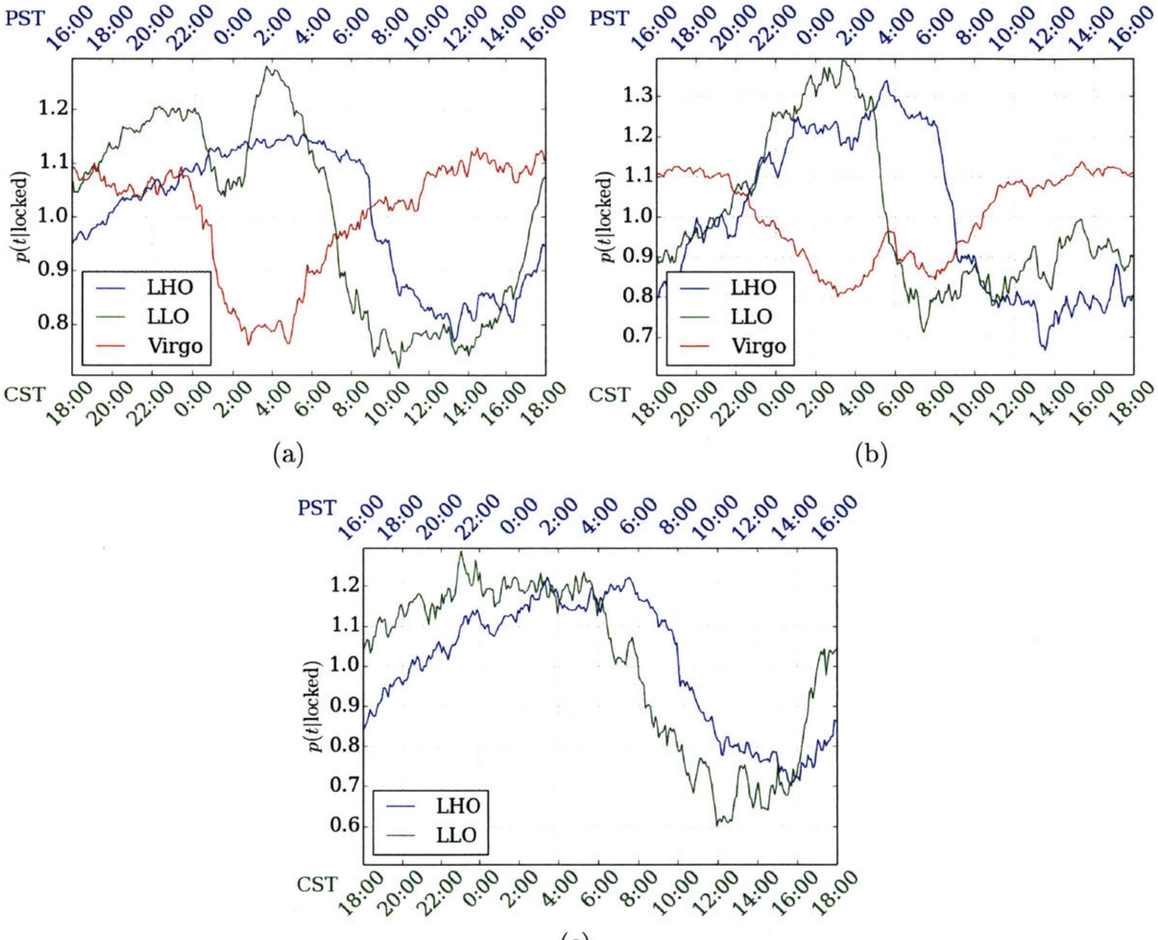
and Bayes theorem yields

$$p(t|u) = \frac{p(u|t)p(t)}{p(u)} = \frac{(1 - p(d|t))\,p(t)}{1 - p(d)} \tag{3.34}$$

We measure $p(u)$ and $p(t|u)$ in a straightforward manner from the data used in Fig. 3-23.[15]

---

[15] Typically, we assume some periodicity in $p(t|u)$ to generate a histogram with enough samples to be statistically meaningful, but we expect to be able to identify the periodic elements of $p(t|u)$ through Fourier analysis as well.



(a) (b) (c)

Figure 3-23: Diurnal cycles shown in local time zones for LLO (CST) and LHO (PST). Although a diurnal cycle is present in (a) S5 and (b) S6, there is an extremely clear two-hour shift during (c) O1 corresponding to the shift in local time between LHO and LLO. We also see similar cycles at Virgo in S5 and S6, which, somewhat interestingly, appear to be mostly anti-correlated with the LIGOs' cycle.

However, we are also interested in slightly different probabilities. To wit, we would like to know

$$p(c|d,t) = \frac{p(d|c,t)p(c|t)}{p(d|t)} \tag{3.35}$$

$$p(r|d,t) = \frac{p(d|r,t)p(r|t)}{p(d|t)} \tag{3.36}$$

$$p(c \cap r|d,t) = \frac{p(d|c \cap r,t)p(c \cap r|t)}{p(d|t)} \tag{3.37}$$

which express the probabilities that the detector is down due to a particular cause at a specific time. If we only care about long-term averages (over time-scales much longer than the *cyclic* model's periodicity), then we can marginalize away the time dependence:

$$p(c|d) = \int dt p(t) \frac{p(d|c,t)p(c|t)}{p(d|t)} \tag{3.38}$$

$$p(r|d) = \int dt p(t) \frac{p(d|r,t)p(r|t)}{p(d|t)} \tag{3.39}$$

$$p(c \cap r|d) = \int dt p(t) \frac{p(d|c \cap r,t)p(c \cap r|t)}{p(d|t)} \tag{3.40}$$

These equations should hold regardless of the specific form of the *cyclic* and *random* models.

By definition, we assume that the *random* causes are uncorrelated with time so that $p(d|r,t) = p(d|r)$. Furthermore, we assume some periodic function for the *cyclic* model so that $p(d|c,t) = p(d|c,t+\tau)$ for some τ. Specifically, we expand the periodic function in terms of the oscillating (AC) and constant (DC) components

$$p(d|c,t) = p_{DC}(d|c) + p_{AC}(d|c,t) \tag{3.41}$$

such that

$$p_{DC} = \int_0^\tau dt\, p(d|c,t) \tag{3.42}$$

Clearly, we require $p_{DC}(d|c) \geq -|p_{AC}(d|c,t)|\,\forall t$. We also assume the priors for the causes do not depend on time ($p(r|t) = p(r)$ and $p(c|t) = p(c)$) and that the priors are equal for the two causes ($p(r) = p(c) = p$).[16]

---

[16] By measuring $p(d)$, we only extract the combination of $p_{DC}(d|c)+p(d|r)$ and cannot separate these terms further. However, by measuring $p(t|u)$ as well, we are able to determine $p_{AC}(d|c,t)$ from which we can determine $p_{DC}(d|c)$ by requiring that $\min_t \{p(d|c,t)\} = 0$. Any other DC component to the *cyclic* model is indistinguishable from the

If we allow the cause models to overlap but require them to be independent, we obtain $p(d|c \cap r, t) = p(d|c,t)p(d|r,t)$. This implies

$$p(r) + p(c) - p(r)p(c) = 1 \Rightarrow p(r) = p(c) = p = 1 \tag{3.43}$$

and

$$\begin{aligned}
p(c|d,t) &= \frac{p_{AC}(d|c,t) + p_{DC}(d|c)}{(p_{AC}(d|c,t) + p_{DC}(d|c))\,(1 - p(d|r)p) + p(d|r)} && (3.44)\\
p(r|d,t) &= \frac{p(d|r)}{(p_{AC}(d|c,t) + p_{DC}(d|c))\,(1 - p(d|r)p) + p(d|r)} && (3.45)\\
p(c \cap r|d,t) &= \frac{(p_{AC}(d|c,t) + p_{DC}(d|c))\,p(d|r)p}{(p_{AC}(d|c,t) + p_{DC}(d|c))\,(1 - p(d|r)p) + p(d|r)} && (3.46)
\end{aligned}$$

A reasonable ansatz is $p(d|c,t) = B\left(1 + \sin\left(\frac{2\pi t}{\tau} - \phi\right)\right)$, in which case we obtain

$$\begin{aligned}
p(d) &= p(d|r)p + Bp(1 - p(d|r)p) && (3.47)\\
p(t|u) &= p(t)\left(1 + \frac{(1 - p(d|r)p)pB}{1 - p(d)}\sin\psi\right)\\
&= p(t)\,(1 + A\sin\psi) && (3.48)
\end{aligned}$$

We also see that a low duty cycle (small $p(u) = 1 - p(d)$) can amplify the amplitude of the day/night bias.

If we are interested in just the probability associated with the *cyclic* model regardless of the *random* model, we obtain

$$p(c|d) = \frac{1 - p(d)}{p(d)}\left(\frac{A}{1 - p(d) + (1 - p(d))A}\right) \approx 0.49$$

which implies that nearly half the time *cyclic* causes were at least partly responsible for bringing down the detector network during O1. If we restrict ourselves to times when the *cyclic* causes were the sole cause of the downtime, we obtain

$$p(c|d) - p(c \cap r|d) = \frac{(1 - p(d))A}{p(d)} \approx 0.29$$

---

random model.

which suggests that we could reduce the downtime by 30% if we completely removed *cyclic* causes of downtime.

3.5.2 Network Exposure Including the Diurnal Cycle

Fig. 3-27 demonstrates how the exposure swings through the celestial sphere as the Earth orbits the Sun. In the Northern hemisphere's summer, the exposure lines up nicely with the galactic plane, making it a good time to search for events within our galaxy. However, extinction can make electromagnetic follow-up of such events difficult, and the vast majority of detectable compact binaries are expected to be extra-galactic. This means that follow-up targeting binaries is better suited to the North's winter months, when the exposure is primarily directed away from the galactic plane.

The diurnal cycle also has implications for ground-based follow-up facilities; not all locations will have equal opportunity to follow-up GW candidates. While EM follow-up is a complex problem, we simplify it to three main features: the amount of posterior probability observable from a site, the minimum angle from an observatory's zenith through which the observations will be made, and the amount of time an observatory must wait before it can begin to observe. These represent the fraction of accessible sources, the quality of observations possible, and, if the sources' counterparts decay quickly, the apparent magnitude of the source when it becomes observable.

We simulate events by bootstrapping our library of localizations [61] to represent a collection of events drawn according to the network exposure. We focus on localizations using the two LIGO detectors only. For the foreseeable future, these IFOs will dominate the detection of GWs because of their higher sensitivity and therefore we expect the distribution of detected events to follow their antenna responses. Additional detectors may reduce the localization error areas, but this will just add more variance to our estimators, defined as

$$\hat{p}_{\mathrm{obs}}(\mathrm{lat}_{\mathrm{site}}) = \frac{1}{N_d}\sum_{i=1}^{N_d}\int d\Omega\, p_{\mathrm{sky},i}\Theta_{\mathrm{obs}}(\Omega, t_i, \mathrm{lat}_{\mathrm{site}}) \tag{3.49}$$

$$\hat{D}_{\mathrm{zen}}(\mathrm{lat}_{\mathrm{site}}) = \frac{1}{N_d\,\hat{p}_{\mathrm{obs}}}\sum_{i=1}^{N_d}\int d\Omega\, p_{\mathrm{sky},i}\Theta_{\mathrm{obs},i}D_{\mathrm{zen}}(\Omega, t_i, \mathrm{lat}_{\mathrm{site}}) \tag{3.50}$$

$$\hat{D}_{\mathrm{del}}(\Omega_{\mathrm{site}}) = \frac{1}{N_d\,\hat{p}_{\mathrm{obs}}}\sum_{i=1}^{N_d}\int d\Omega\, p_{\mathrm{sky},i}\Theta_{\mathrm{obs},i}D_{\mathrm{del}}(\Omega, t_i, \Omega_{\mathrm{site}}) \tag{3.51}$$

where $p_{\text{sky},i}$ is the localization probability for the i_{th} detection, $\Theta_{\text{obs},i}$ is a window function on what is observable from that site, and $\hat{p}_{\text{obs}}$ in Eqn. 3.50 and 3.51 is computed using the same set of localization estimates as the explicit sum. D_{zen} and D_{del} measure the minimum zenith distance of that location from that observatory and the time-delay before an observatory can first start observing a location, respectively.

We note that using the antenna response yields the mean of these statistics when simulating many detections. To wit, if we calculate χ using the antenna response and $\hat{\chi}$ using sets of simulated maps, we expect

$$\begin{aligned}\chi &= \lim_{N_d\to\infty} \hat{\chi} \\ &= \lim_{N_d\to\infty} \frac{1}{N_d} \sum_{i=1}^{N_d} \int d\Omega\, p_{\text{sky},i}\chi(t_i)\end{aligned} \tag{3.52}$$

We note that $\chi(t_i)$ only depends on the time the event occurs through the position of the Sun; the dependence will be the same for all detections that occur at the same time. If we break the sum into small segments of time, we can write

$$\chi = \lim_{N_d\to\infty} \frac{1}{N_d} \sum_{j=1}^{N_t} \sum_{\substack{i=1 \\ t_i\in[t_j,t_j+\Delta t)}}^{N_{d,j}} \int d\Omega\, p_{\text{sky},i}\chi(t_i) \tag{3.53}$$

where $\sum_j N_{d,j} = N_d$. Now, when $N_d \to \infty$, we can make the segments as small as we like while maintaining a large number of detections in each bin. We then obtain

$$\begin{aligned}\chi &= \lim_{\substack{N_t\to\infty \\ N_t\Delta t=T}} \sum_{j=1}^{N_t} \lim_{N_{d,j}\to\infty} \frac{1}{N_d} \sum_{\substack{i=1 \\ t_i\in[t_j,t_j+\Delta t)}}^{N_{d,j}} \int d\Omega\, p_{\text{sky,i}}\chi(t_i) \\ &= \lim_{\substack{N_t\to\infty \\ N_t\Delta t=T}} \sum_{j=1}^{N_t} \lim_{N_{d,j}\to\infty} \left(\frac{N_{d,j}}{N_d}\right) \int d\Omega \left(\frac{1}{N_{d,j}} \sum_{\substack{i=1 \\ t_i\in[t_j,t_j+\Delta t)}}^{N_{d,j}} p_{\text{sky,i}}\right) \chi(t_i) \\ &= \lim_{\substack{N_t\to\infty \\ N_t\Delta t=T}} \sum_{j=1}^{N_t} (p(t|\text{operating})\Delta t) \int d\Omega\, (p_{GW}(\Omega, t_j))\, \chi(t_j) \\ &= \int dt\, p(t|\text{operating}) \int d\Omega\, p_{GW}(\Omega, t)\chi(t)\end{aligned} \tag{3.54}$$

where T is the length of the observing season. We assumed $t_i \to t_j$ for all events within each bin and used the fact that many localization posteriors stacked on top of one another will average to the network antenna response in Equatorial coordinates ($p_{GW}(t_j)$), which depends on the bin's time. We have also used the fact that the fraction of events occurring in each bin is equal to the probability that the detectors are operating throughout that bin ($p(t|\text{operating})\Delta t$). By approximating these integrals, we obtain the limits of the means much more efficiently than through direct simulation. This procedure, or an equivalent, is used in Fig. 3-24, 3-25, and 3-26.[17]

We see that equatorial locations can access more probability on average (Fig. 3-25). This is because they can see both the Northern and Southern antenna response maxima. There is also a seasonal variation which favors whichever hemisphere is currently in winter, which can be comparable to the modulation due to observatory longitude. Fig. 3-25 shows the minimum angle from zenith through which observations can be made. The asymmetry between the North and South is associated with the Sun's statistical likelihood to be above the Southern hemisphere's antenna response maximum when detections occur. This means Southern observatories are more likely to observe near twilight and therefore must look through a higher airmass (larger D_{zen}). We also note that there is a moderately longer time-delay over Southern Africa (Fig. 3-26). This is because it is likely to be dawn there when detections are made, and observatories must wait nearly a full day before observing. We note that this distribution of $\hat{D}_{\text{del}}$ is very skew right for small numbers of detections.

It is also informative to examine these statistics when there are only a few events. This is particularly important when considering how an observatory could be impacted for any individual event, rather than for a collection of events. We also consider these distributions in the limit of extremely well localized sources: $p_{\text{sky},i} \to \delta(\Omega - \Omega_i)$. This point source limit describes the distributions obtained when only observing the location of the true source.[18]

Of all the statistics we consider, $\hat{p}_{\text{obs}}$ is the most Gaussian for small N_d for year-long averages. However, because it is bounded from above and below, the distribution does deviate at times. Fig. 3-28 shows the distributions for a few latitudes and a few values of N_d. We note that, in the limit $N_d \to \infty$, p_{obs} is the fraction of true counterparts that an observatory can observe. When

---

[17] We also note that we normalize by the total observable probability in Eqn. 3.50 and 3.51. This is because we restrict ourselves to only the fraction of the probability that is actually observable. The particular form of our normalization (dividing by $\hat{p}_{\text{obs}}$ for a set of events rather than each event separately) guarantees that we sample the antenna response in the limit.

[18] It also allows us to estimate the possible impact of more detectors and the associated improvements in localization.

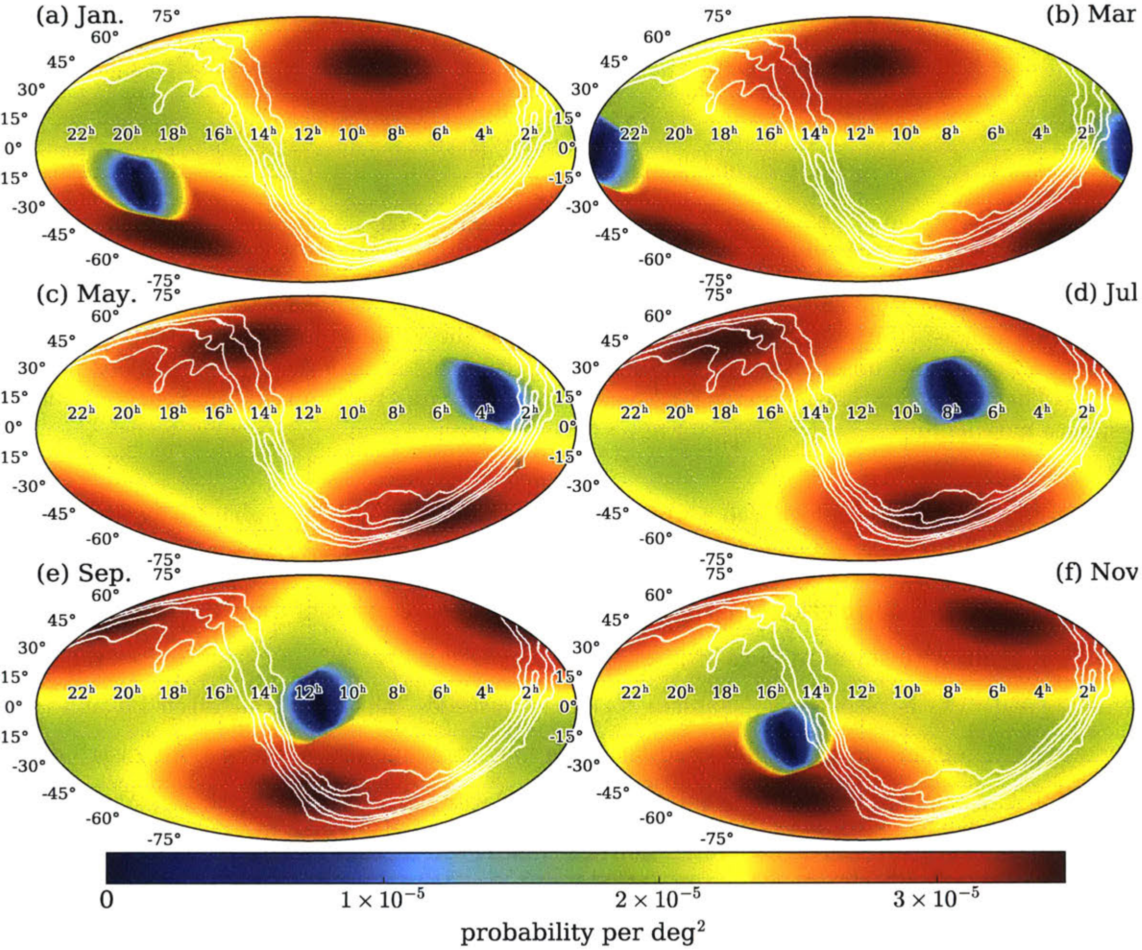


Figure 3-24: Mollweide projections the galaxy annotated showing seasonal variation of the network exposure. We include occultation by the Sun to demonstrate how EM observatories will be impacted.

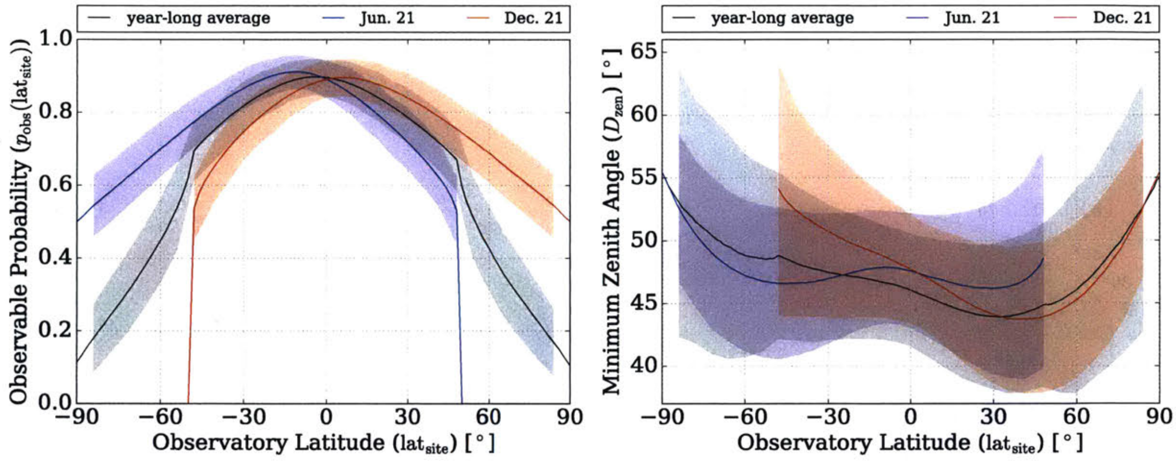


Figure 3-25: (left) p_{obs} as a function of observatory location. Shaded regions approximate the fluctuations for ~ 10 event by normalizing the limit of the variance for many detections. (right) D_{zen} as a function of observatory location. Shaded regions approximate the fluctuations for ~ 10 events. The North/South asymmetry is associated with the Sun's proximity to the Southern antenna response maximum.

considering extremely well localized sources, the observatory will either be able to observe the true source or not, and each trial will have nearly the same probability of success, modulo variations caused by the Sun's declination. Fig. 3-28 plots the point-source limit of these distributions as well ($\hat{p}_{\mathrm{obs}}^{(\mathrm{pt\,src})}$), again with several values of N_d.

$\hat{D}_{\mathrm{zen}}(N_d = 1)$ is also fairly Gaussian for some latitudes, but there can often be non-trivial deviations therefrom. In particular, mid-latitudes may show interesting skew right distributional shapes. Fig. 3-29 shows these distributions for a few latitudes as a function of N_d. We note that these distributions are much narrower than the point source limit in which all events are well localized ($\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}$), also shown in Fig. 3-29. There is more shot noise for $\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}$, which broadens the distributions, but the general distributional shapes are similar for both $\hat{D}_{\mathrm{zen}}$ and $\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}$.

Fig. 3-30 demonstrates the skew right nature of $\hat{D}_{\mathrm{del}}$. Typically, there is an extremely large lobe near zero, corresponding to events that are immediately observable, and a long tail comprised of events that require waiting. When we consider the point source limit of well localized events, this behavior is enhanced, also shown in Fig. 3-30. When we integrate over typical localizations, instead of point sources, we find that the peak is smeared out to longer delay times. This is because different parts of the localization estimate may become observable at different times, and that fuzz

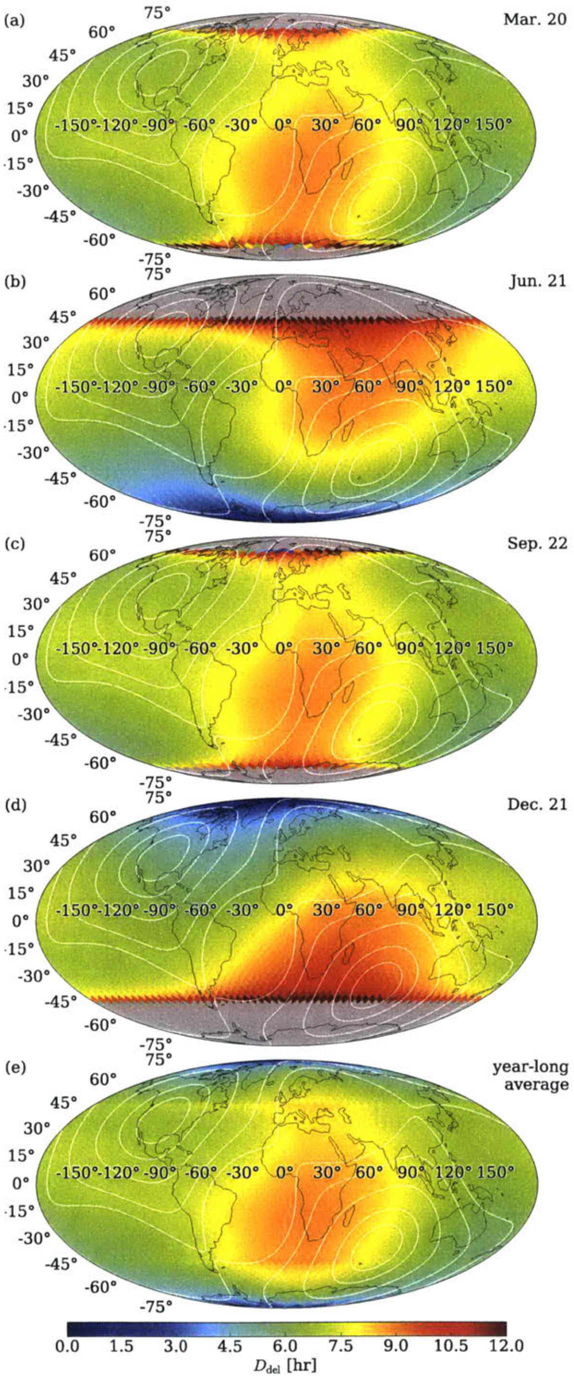
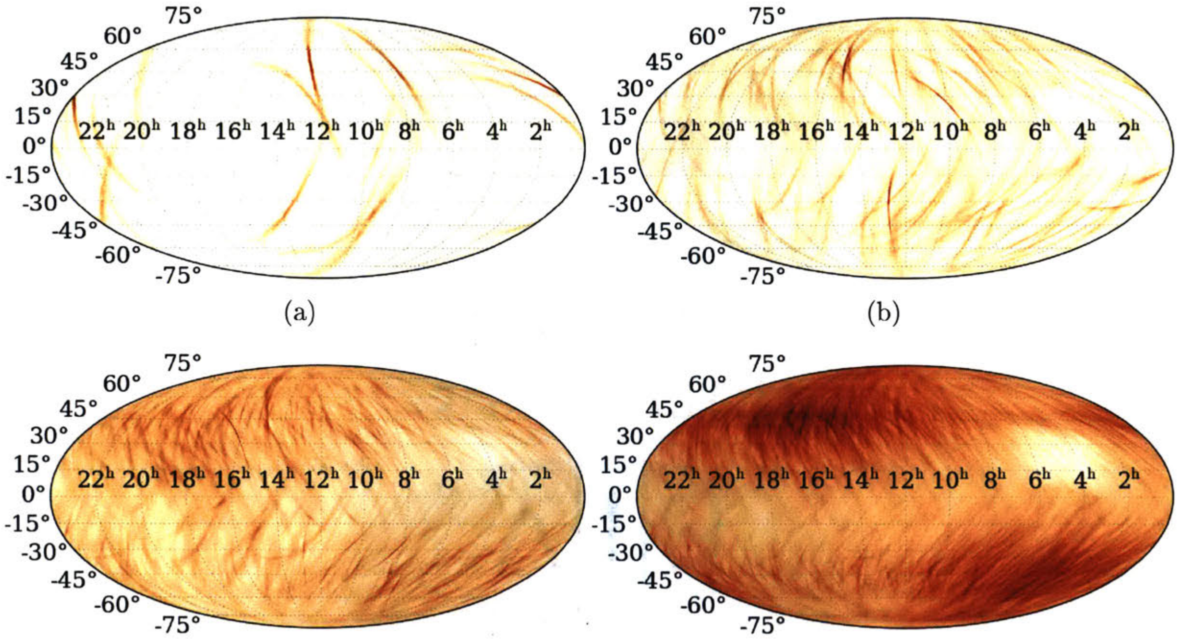


Figure 3-26: Average time-delay before observatories can first observe the localization estimate ($D_{\rm del}$) as a function of the observatory's position. We note there can be large seasonal variation, but the main effect is a moderately longer wait for Africa and the South Atlantic.



(a) (b)

(c) (d)

Figure 3-27: Stacked localization posteriors in Equatorial coordinates from bootstrapped samples of BBH localizations using LIB [54, 61] for (a) 10, (b) 100, (c) 1000, and (d) 10,000 events. This sampling corresponds to Fig. 3-24(c), and indeed the stacked posteriors converge toward this distribution.

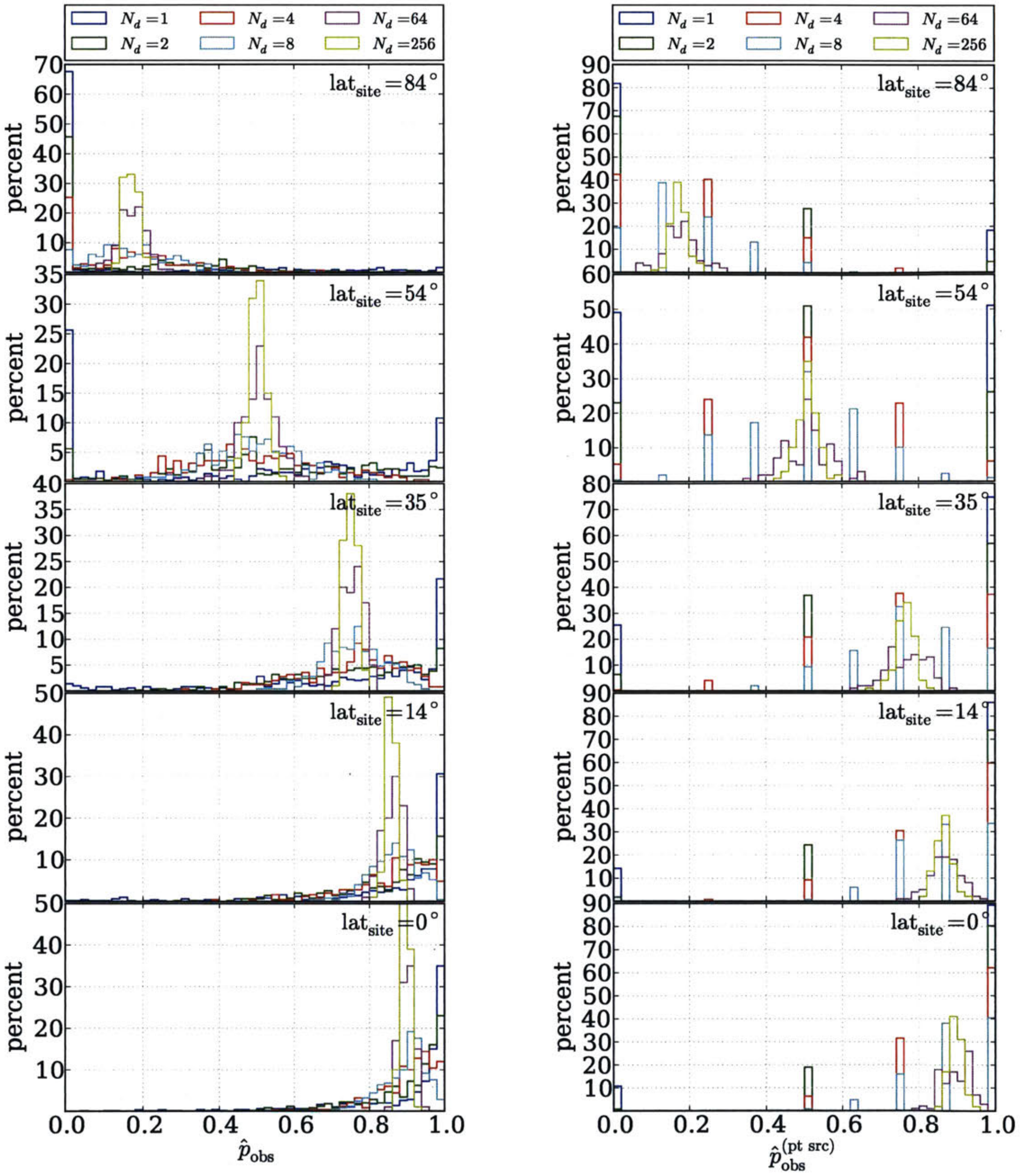


Figure 3-28: Distributions of (left) $\hat{p}_{\mathrm{obs}}$ and (right) $\hat{p}_{\mathrm{obs}}^{(\mathrm{pt\,src})}$ for a few latitudes and N_d for year-long averages. We note that the distributions with $N_d = 1$ may not be very Gaussian and display long tails. Nonetheless, as $N_d \to \infty$, the means of the distributions tend toward the values reported in Fig. 3-25.

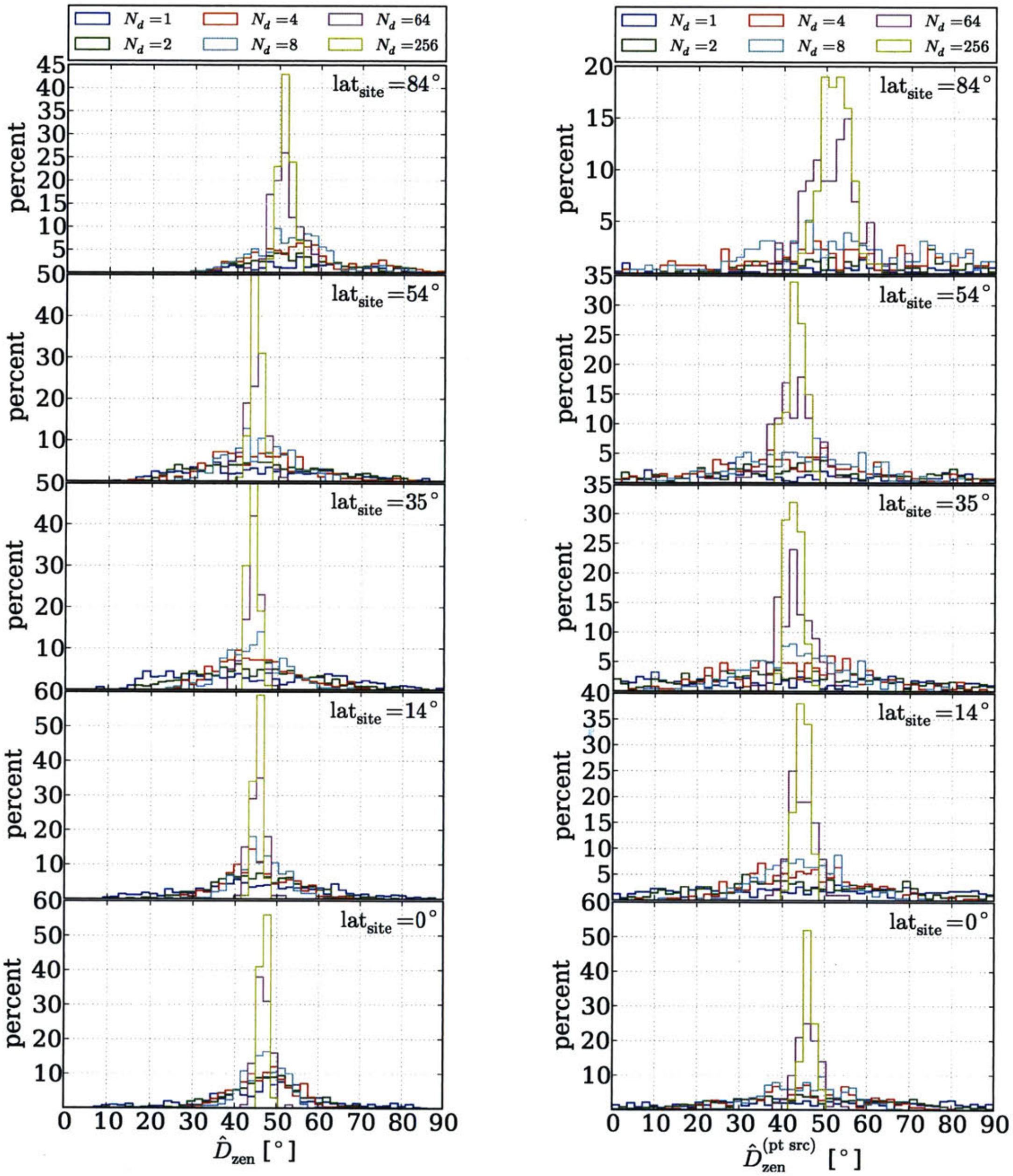


Figure 3-29: Distributions of (left) $\hat{D}_{\mathrm{zen}}$ and (right) $\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}$ for year-long averages. We note that the distributions are rather broad and all centered near similar values. Nonetheless, the mean of the distribution collapses to the same values as in $\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}$ as $N_d \to \infty$. $\hat{D}_{\mathrm{zen}}(N_d = 1)$ is often larger than $\hat{D}_{\mathrm{zen}}^{(\mathrm{pt\,src})}(N_d = 1)$ because the triangulation rings span most of the antenna response.

tends to smooth the distribution.

3.6 GW-EM Infrastructure

The previous work on both data quality and localization is immediately relevant to GW searches, particularly in low-latency when we must decide if a candidate is promising enough to be released to electromagnetic observers. To this end, several automated follow-up processes summarize relevant information to aid in the decision making process. Fig. 3-1 depicts a simplified workflow for low-latency searches; pipelines process $h(t)$ and report candidates to a remote server (GraceDb [140]). Through an automated messaging service (LVAlert [141]), actions within and annotations to GraceDb trigger a variety of follow-up processes, including automated data quality and localization. GraceDb also notifies analysts for human vetting. Fig. 3-1 sketches this process.

3.6.1 Data Quality Reports

In addition to iDQ's automated response (see §3.2.2), there are several other data quality flags that are automatically produced within minutes of real-time. These typically do *not* correspond to statistical inferences based on machine learning and instead represent well studied mechanisms by which noise is introduced into the detector. For example, when photo-diodes in the control loop saturate, the associated nonlinearity introduces very large glitches in $h(t)$. These saturations are relatively easy to track and processes at each site automatically identify these events. An automated query examines a set of data quality flags and reports the results to GraceDb [57]. Analysts then check these reports for obvious issues with the detectors, or, if it is allowed, the query itself can reject candidates automatically.

If automated inference about data quality are important for analysts, visualization of the actual data streams is even more so. Much confidence in data quality issues associated with artifacts in $h(t)$ comes from lining up spectrograms from auxiliary channels with similar features in $h(t)$. For this reason, automatically generated spectrograms are extremely useful. The GW community often uses the Q-transform [33, 69], specifically the transform implemented in the Ω-pipeline (Ω-scans, [39]). Fig. 3-31 shows an example Ω-scan for a radio-frequency whistle at the Livingston detector in $h(t)$ as well as the two auxiliary witnesses identified by iDQ (Fig. 3-12). With few exceptions,[19] all

---

[19] Some burst events are rejected based on iDQ information alone without human intervention, such as the event

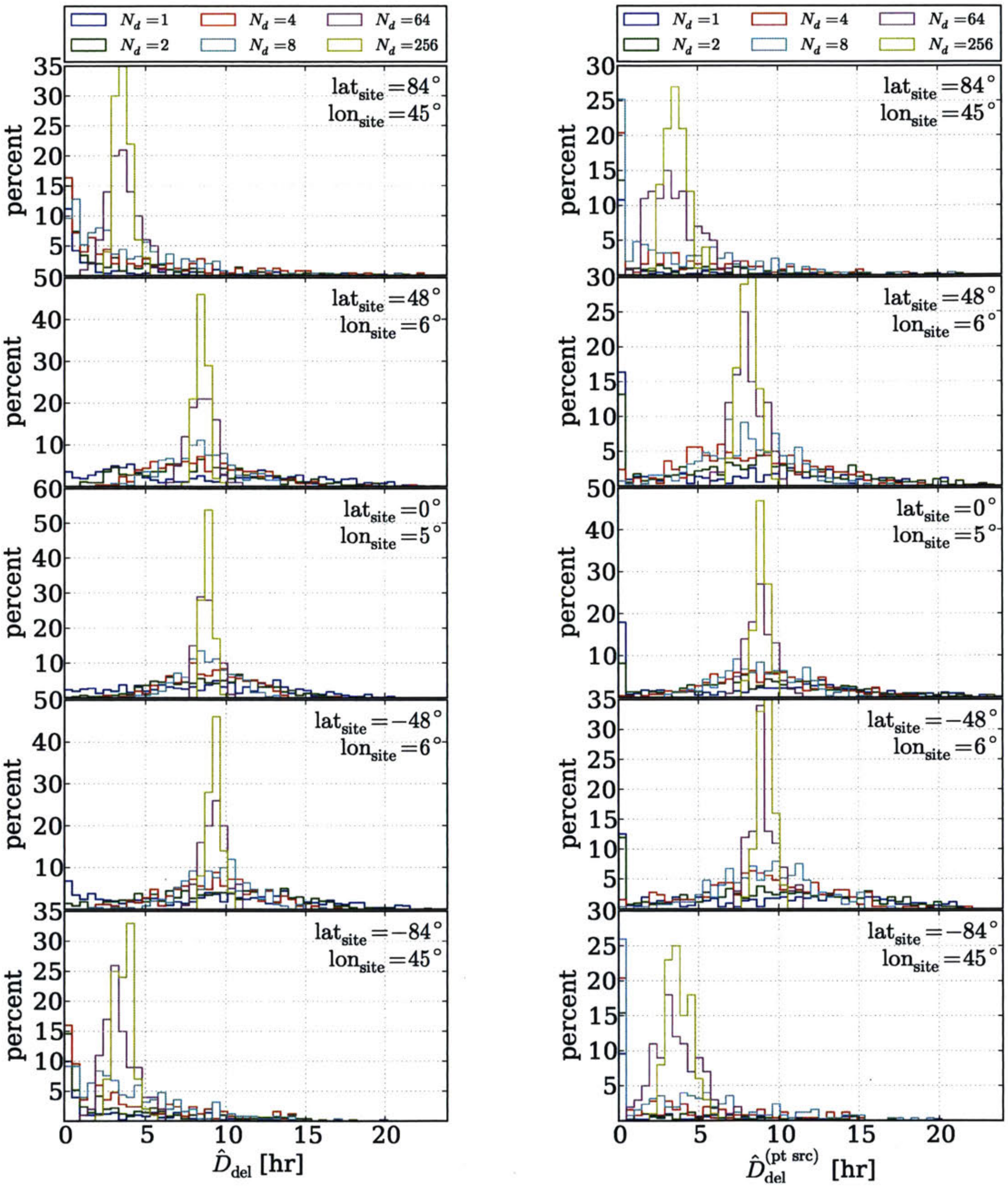


Figure 3-30: Distributions of (left) $\hat{D}_{\rm del}$ and (right) $\hat{D}_{\rm del}^{(\rm pt\,src)}$ for a few sites and N_d for year-long averages. We note that the $N_d = 1$ distributions have large modes near $\hat{D}_{\rm del} = 0$, corresponding to events that are immediately observable, as well as very broad support extending to long $\hat{D}_{\rm del}$. This is particularly evident near the poles. The modes near $\hat{D}_{\rm del}^{(\rm pt\,src)}(N_d = 1)$ are more pronounced than in Fig. 3-30 because there is more shot noise in the point source measurement. In fact, averaging over the localization estimate tends to broaden the distributions for $N_d \lesssim 5$.

vetting is still conducted by analysts. This currently dominates the latency for releasing candidates. Although the exact process depends somewhat on the data present, typically analysis first inspect the $h(t)$ Ω-scan to determine if it is consistent with any known glitch families [114, 115, 177]. This is then compared against the summaries from iDQ and other data quality flags for confirmation. If there is a clear auxiliary witness, the event is rejected. This is the case for G272601, depicted in Fig. 3-11, 3-12, and 3-31. Otherwise, it is released pending a reasonable localization. Future work will both automate the vetting process and tailor it for specific pipelines.

3.6.2 Localization Characterization and Comparison

Rapid characterization of localization estimates also aids in low-latency vetting. As such, an automated process [58] responds to any new localization estimates inserted into GraceDb. This process summarizes each estimate separately, calculating all the metrics discussed in §3.3.2 as well as several others (Fig. 3-32 and 3-33 show examples produced automatically). In particular, the MID is automatically computed for all IFO pairs which participated in the detection. The maps are also compared automatically, including overlays of confidence regions and Fidelities. This is particularly useful if multiple estimates are available before data quality vetting has been completed. For burst events, we use the Fidelities between maps to determine whether to model average before releasing localization estimates.

3.6.3 Monitoring

Monitoring is crucial to ensure any automated process behaves as expected. To this end, and given the scale and scope of automated responses triggered by annotations within GraceDb, Event Supervisor [55] proves a unified monitor of all expected responses and annotations within GraceDb. By listening to LVAlert messages automatically distributed by GraceDb, it is able to determine whether follow-up proceeded correctly, notifying analysts to possible problems in real-time. This task's complexity is compounded by the asynchronous, distributed nature of the collaboration's computing infrastructure. What's more, the number of processes that need to be monitored is dynamic. This forces our monitoring to validate follow-up results posted to GraceDb rather than ensuring each process itself is working separately. This *results-first* paradigm has proven to scale

---

in Fig. 3-31.

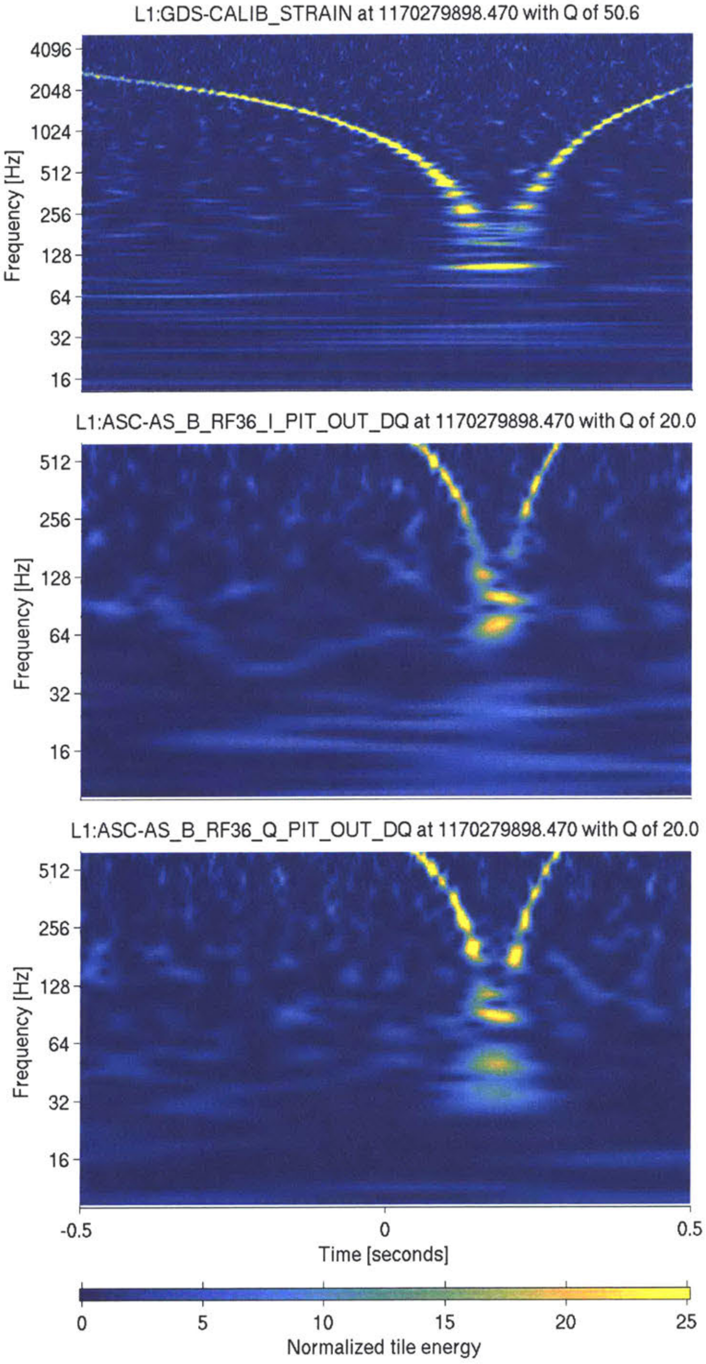


Figure 3-31: Ω-scans for $h(t)$ as well as L1:ASC-AS_B_RF36_{I,Q}_PIT_OUT_DQ for G272601 (a clear radio-frequency whistle at LLO at 21:44:40 February 4, 2017 UTC). The corresponding iDQ predictions are shown in Fig. 3-11 and 3-12; iDQ used these witness channels to reject this event within ~ 8 sec.

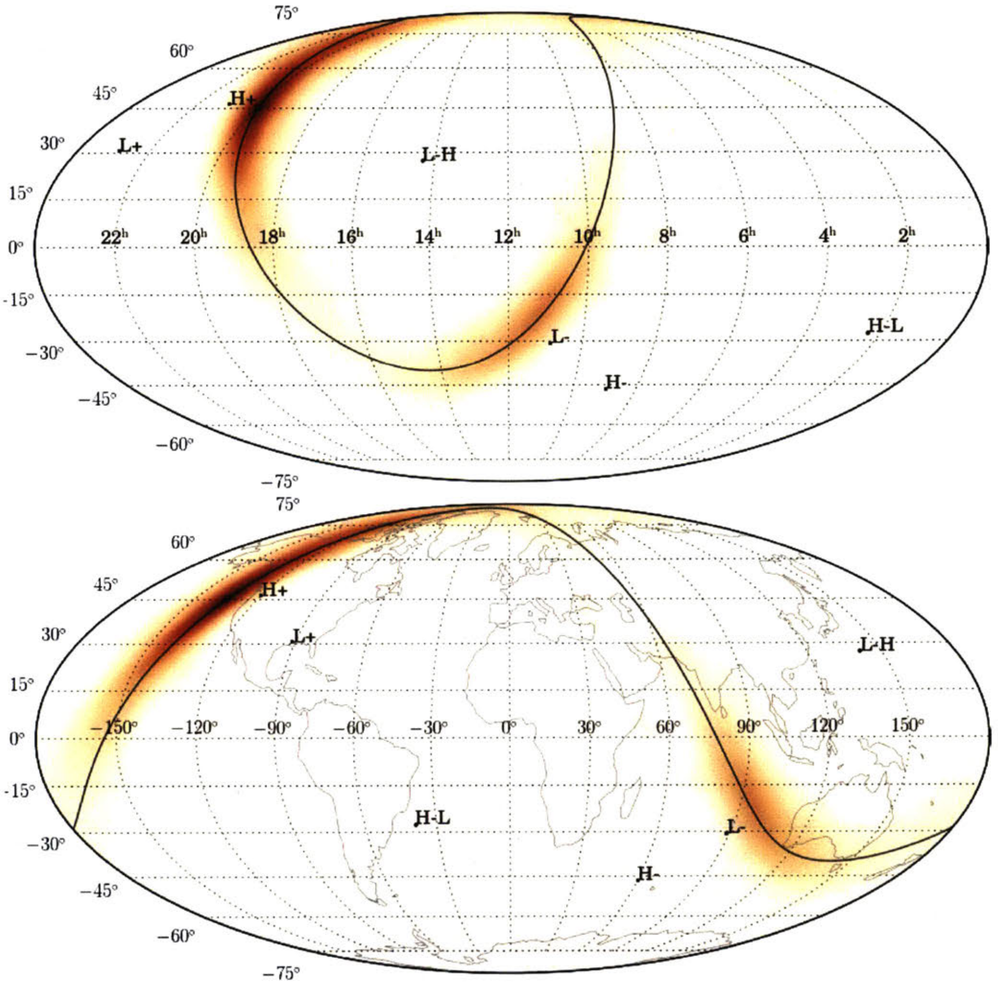
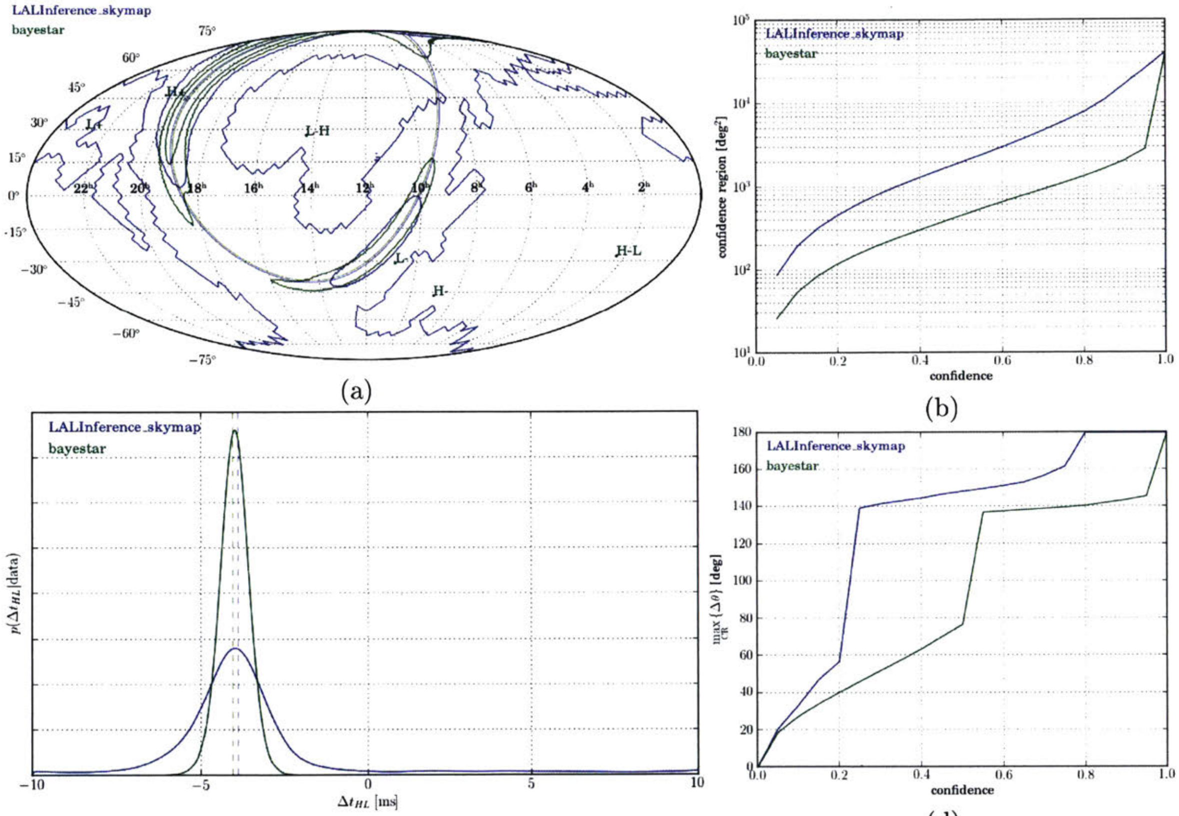


Figure 3-32: Example localization summary for G275404's LALInference estimate in both (top) Equatorial and (bottom) Geographic coordinates. For this map, MID = 0.00916.



(a) (b) (c) (d)

Figure 3-33: Example localization comparison for G275404 showing the (a) 50% and 90% confidence regions in Equatorial coordinates, (b) size of the confidence regions, (c) the time-delay marginals, and (d) the largest angular separation between any two points within each confidence region. For these maps, $F(\Omega) = 0.777$ and $F(\Delta t) = 0.794$.

well and adapt quickly to changes deployed during observation runs.

Monitoring follow-up processes themselves is also worthwhile when possible. This can be complicated due to the distributed nature of our network, but all processes interact with LVAlert in order to respond to GraceDb annotations. Even if the listener forks many processes (e.g. one for each of many candidates), there should be a static number of listening processes at any time. We can ensure they respond to LVAlert announcements by pinging them through the LVAlert infrastructure and listening for a response. This heartbeat [56] *process-first* monitoring has also been deployed throughout the network and is used to alert developers to issues before their processes fail to produce results, which would trigger warnings from Event Supervisor. In this way, we prevent failures before they have a chance to actually impact searches while still catching them in case they do.

Chapter 4

Confirmed Detections

To much fanfare, the advanced LIGO interferometers directly detected the first GWs in September 2015. At the time of writing, this is the first of three confirmed detections during the first and second observing runs (O1 and O2, respectively) [7, 145, 146, 151] spanning a total of more than 70 days of joint observation with both the LIGO Livingston and Hanford instruments [138]. While both instruments have undergone commissioning upgrades since the beginning of O1 to improve their sensitivity, their sky-averaged detection range for 1.4-1.4$M_\odot$ BNS sources has typically ranged between 60-80 Mpc. Of course, the LIGOs can (and do) observe BBH systems at much larger distances, depending on their masses, because they are intrinsically more luminous and because the loudest part of the BBH signal is in LIGO's most sensitive band [65, 66].

I will discuss the broad characteristics of each confident detection published at the time of writing, along with the less confident LVT151012. In addition, I will detail specific applications of my work described in §3 to each. I also provide a few general observations about the population of signals observed so far in §4.5.

4.1 GW150914

The first direct observation of GWs, GW150914 [146, 158], passed through the LIGO detectors at 09:05:45 UTC on September 14, 2015. The signal is consistent with the coalescence of two black holes, $36^{+5}_{-4}M_\odot$ and $29^{+4}_{-4}M_\odot$ in the source frame, respectively. This system was 410^{+160}_{-180}Mpc away from Earth ($z = 0.09^{+0.03}_{-0.04}$) and produced a single $62^{+4}_{-4} M_{\odot}$ Kerr black hole with dimensionless spin of $0.67^{+0.05}_{-0.07}$. The final black hole's spin is dominated by the angular momentum of the system before merger and there is no strong evidence for large spins for either component prior to merger. This signal was loud enough to be seen by eye in band-passed data (Fig. 4-1), with slightly larger amplitude in LHO than LLO. The difference in amplitudes is due to the IFOs antenna responses, corresponding to a time delay of $6.9^{+0.5}_{-0.4}$ms between LLO and LHO (Fig. 4-2).

GW150914 was first detected by online burst searches and later confirmed by matched filter techniques. The most sensitive search for signals like GW150914 bounds the FAP to above a 5.1σ deviation from background.

4.1.1 Data Quality

Offline OVL and pointed Poisson analyses (§3.2.1) showed no auxiliary couplings that could have been responsible for GW150914 [143]. This includes an analysis of 198,086 and 200,426 auxiliary channels at LLO and LHO, respectively, although we focus primarily on the 6,319 and 6,378 channels sampled faster than 16 Hz. While the OVL analysis did not find any couplings within 180 seconds of GW150914 at LLO or within 11 seconds at LHO, the pointed analysis found one clear outlier. A single auxiliary channel that was significant enough that we expected only 0.075 events with smaller p-values after accounting for the trials factor associated with the large number of channels. Indeed, upon repeated identical analyses of random times, it appears that this level of accidental coincidence does occur once in every $\sim$ 10 events. Upon further investigation, this accidental coincidence was confirmed to be causally unrelated to GW150914, to possess a different spectral structure than GW150914, and could not have been responsible for the signal in $h(t)$.

At the time of GW150914, iDQ was running and reported predictions to GraceDb within 18 seconds. However, it's output was corrupted by a miscommunication about the set of auxiliary channels used within the analysis. Because the offline analyses were more comprehensive, we did not re-analyze GW150914 with the corrected online iDQ configuration, which was deployed less than 24 hours after the event.

To be clear, there is no evidence for any terrestrial source for GW150914. This includes exhaustive checks of physical environmental monitors at each site (magnetometers, seismometers, etc), even including lightning strikes and cosmic ray events. For example, coincident electromagnetic disturbances could not have been responsible for more than 1 part in 10^8 of GW150914's

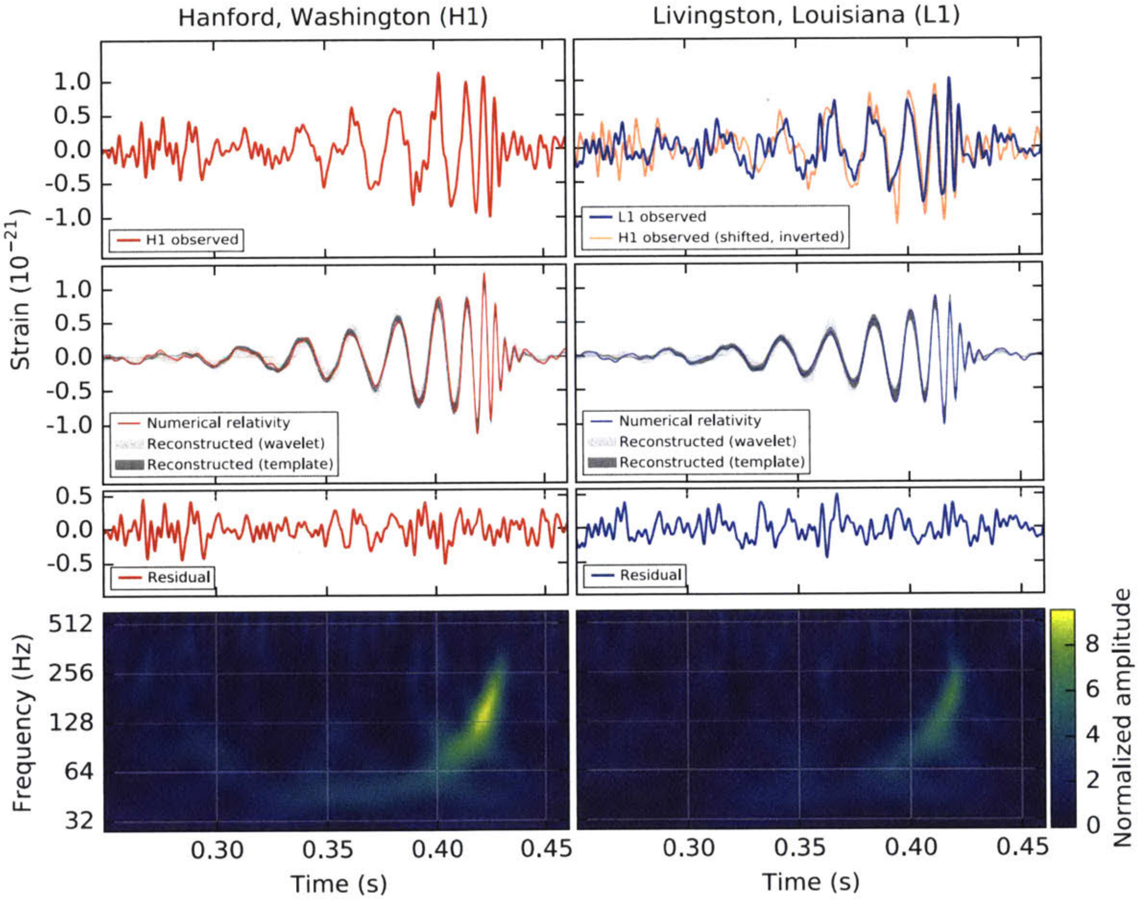


Figure 4-1: Band-passed data for GW150914 along with theoretical predictions, residuals, and Ω-scans [146].

amplitude [143].

4.1.2 Localization Characterization

GW150914 arrived at LLO $\sim$ 7ms before arriving at LHO. Both cWB and LIB produced localization estimates which were released to EM observing partners within a few days [152]. These estimates had rather low Fidelity (0.53), cause by their selection of slightly different triangulation rings, corresponding to the $\lesssim 10\%$ quantile based on simulations of similar BBH systems [165]. Two other localization estimates were released several months later, including the definitive LALInference analysis which included the effects of calibration uncertainty, the main effect of which is to widen the ring due to uncertainty in the phase within each IFO. Fig. 4-2 shows several representations of these estimates.

Notably, cWB's map shows a large Northern island, which is near a pole in the network antenna response. Because of the degeneracy in coherent likelihood localizations when there are two detectors and two polarizations (§3.3.1), cWB reconstructs a single polarization throughout most of the sky. However, near the pole, the analysis believes it can distinguish between multiple polarizations because of the comparable size of the (small) antenna response for each polarization, and therefore relaxes its constraint. Effectively, this is an artifact of cWB's rather complicated priors and assumptions, which is why the shape of the Northern island is dominated by the antenna response's pole (a blob) instead of the data (a triangulation ring).

Both LIB and LALInference show some support for Northern arcs as well. These correspond to slightly different triangulation rings compared to the more Southern parts of their localizations, and the difference appears to be associated with the handedness of the circularly polarized signal.

4.2 GW151226

GW151226, commonly referred to as the Boxing Day event [145, 158], was recorded on December 26, 2015 at 03:38:53 UTC, still Christmas day in the United States. This system was both lower mass and slightly more distant than GW150914, although there is significant overlap in their distance posteriors. Because of this, GW151226 had lower ρ_{net} than GW150914 and is not clearly distinguishable by eye. However, its lower mass means it spent more time within the LIGOs' sensitive band and matched filter analyses were able to identify it above the 5.3σ level. Burst searches were

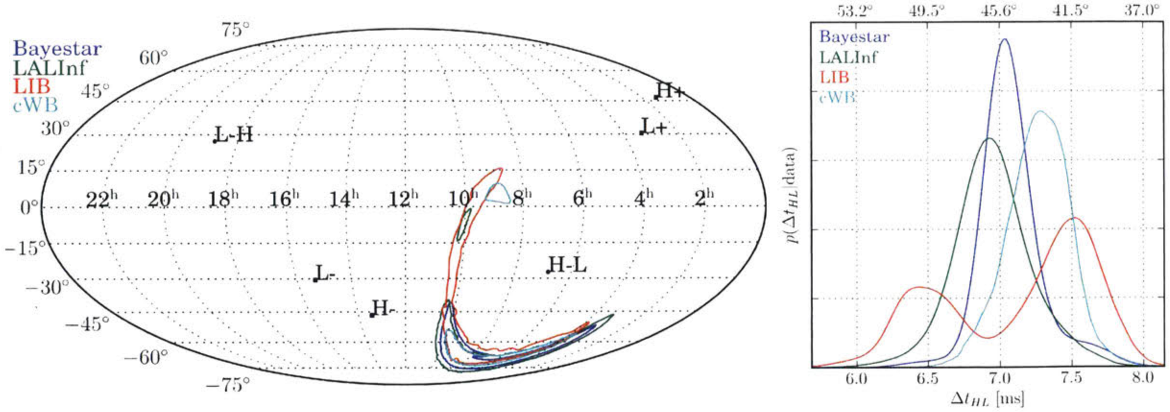


Figure 4-2: Overlay of 90% confidence regions for all published localization estimates for GW150914 in Equatorial coordinates along with time-delay marginals [158].

not able to confidently detect GW151226 due to its relatively long duration and lower amplitude.

This system was composed of a $14.2^{+8.3}_{-3.7}M_\odot$ BH and a $7.5^{+2.3}_{-2.3}M_\odot$ BH, and there is evidence for small (but non-zero) spin aligned with the orbital angular momentum. The final BH was $20.8^{+6.1}_{-1.7}M_\odot$ with a spin of $0.74^{+0.06}_{-0.06}$, again dominated by the orbital angular momentum. GW151226 was located 440^{+180}_{-190}Mpc from Earth ($z = 0.09^{+0.03}_{-0.04}$) and came from nearly directly overhead the LIGO detectors, near the maximum of their network sensitivity, arriving at LLO $1.1^{+0.3}_{-0.3}$ms before LHO [7].

4.2.1 Data Quality

Similar to GW150914, offline OVL and pointed Poisson analyses found no evidence for auxiliary couplings to $h(t)$ in the immediate neighborhood of GW151226. There were no significant outliers in the pointed analysis, and OVL's offline behavior was consistent with the low latency predictions produced by iDQ (see Fig. 3-7 and 3-14).

iDQ found no conclusive evidence for auxiliary couplings coincident with GW151226, and reported FAPs for misclassifying clean data as glitchy $\gtrsim 4 \cdot 10^{-3}$ at LHO and 1.0 at LLO within ~ 82 sec. This glitch FAP is completely consistent with expectations from background given the amount of time searched surrounding the event (searching through 45 seconds with a 25 ms veto window corresponds to an approximate trials factor of 1800). In contrast, Fig. 3-11 shows a confident identification of a glitch with FAP$\lesssim 10^{-5}$ within a smaller time window.

The data from several hours preceding GW151226 did contain a large amount of radio-frequency

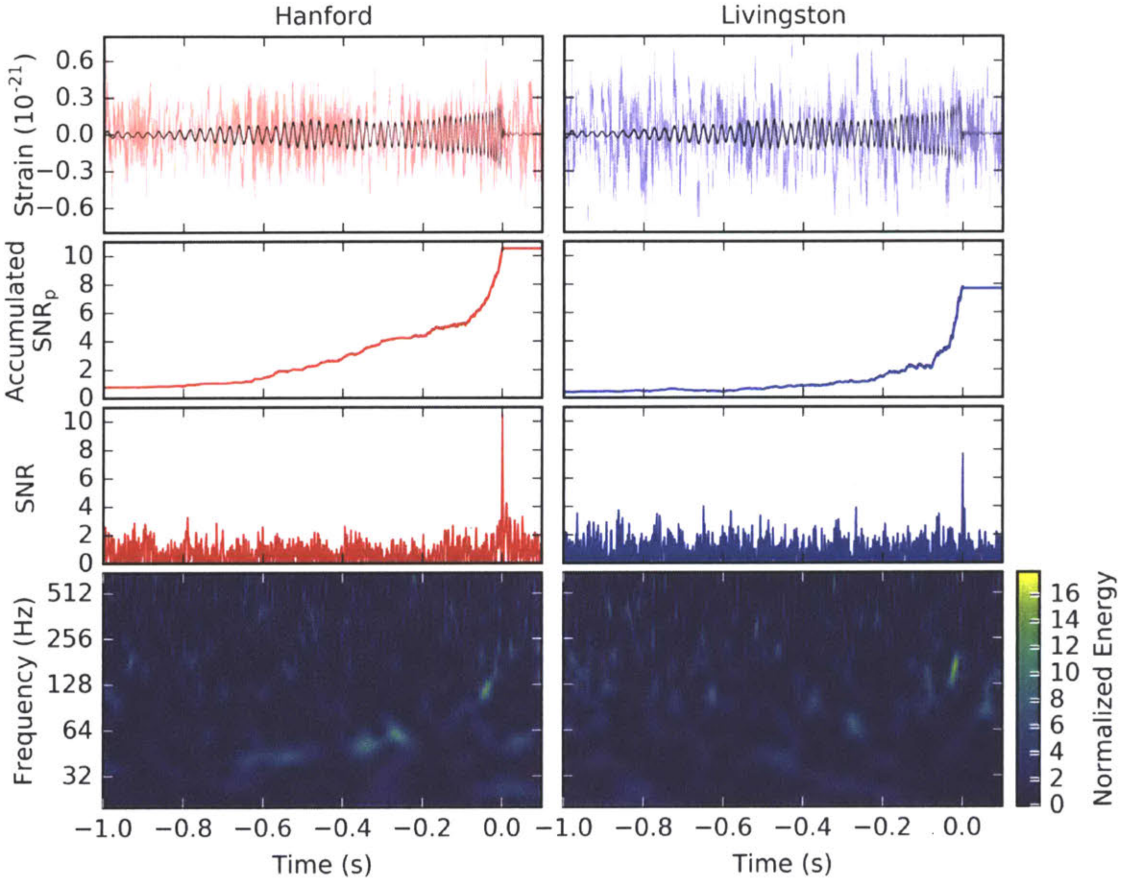


Figure 4-3: Whitened time-series with superimposed signals, signal-to-noise ratio time-series, and Ω-scans for GW151226 [145]

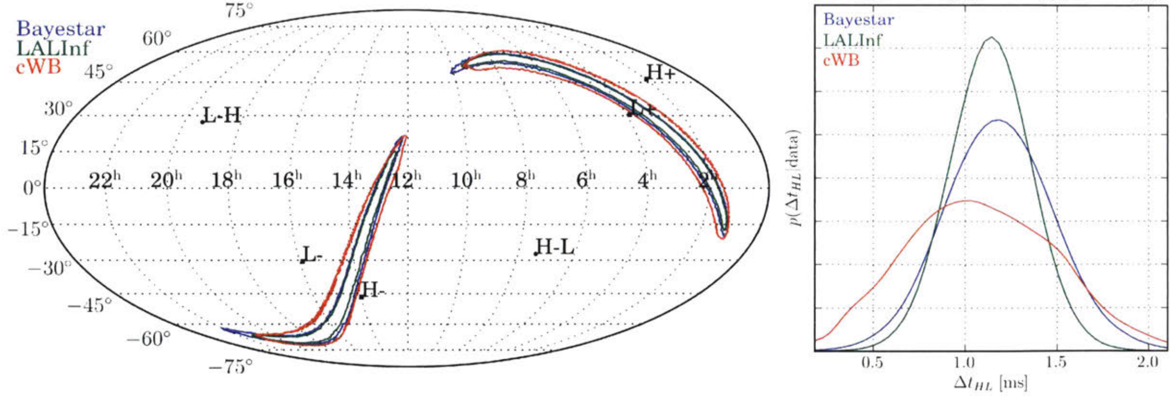


Figure 4-4: Equatorial Mollweide projections of 90% confidence regions from published maps for GW151226, along with their time-delay marginals [158].

noise (see Appendix A of [143]). However, these artifacts were witnessed extremely well by auxiliary channels and flagged by iDQ, often reaching glitch detection efficiencies in excess of 90% with FAP $\lesssim 10^{-3}$ (see §3.2, in particular Fig. 3-7 and 3-14). These artifacts are not present in the data immediately surrounding or overlapping GW151226, supported by iDQ's null result.

4.2.2 Localization Characterization

Several localization estimates were released for GW151226 (Fig. 4-4). They are broader than the estimates for GW150914 because of the lower ρ_{net} and also the location on the sky; GW151226 came from nearly directly overhead the LIGO's, meaning its triangulation ring nearly follows a great circle with a much larger diameter than the ring for GW150914. All estimates are consistent with one another, to a much greater degree than GW150914 (Fidelities $\gtrsim 0.9$ for low-latency and definitive estimates). Generally, the localization estimate for GW151226 is characteristic of what we would expect from a typical event detected by the LLO-LHO network.

4.3 GW170104

LIGO's third confirmed detection occurred on January 4, 2017 at 10:11:58 UTC and was inconsistent with background at higher than the 5σ level [151, 158]. GW170104 came from the coalescence of a $31.2^{+8.4}_{-6.0}M_\odot$ BH with a $19.4^{+5.3}_{-5.9}M_\odot$ BH approximately 880^{+450}_{-390}Mpc from Earth ($z = 0.18^{+0.08}_{-0.07}$), resulting in a $48.7^{+5.7}_{-4.6}M_{\odot}$ BH with dimensionless spin $0.64^{+0.09}_{-0.20}$. The data is consistent with non-spinning constituents, although there is a small preference for at least one of the spins to be anti-aligned with the orbital angular momentum. Similar to GW150914 and GW151226, the posteriors for the in-plane spin components are consistent with the prior, implying the data does not constrain them [7, 151].

4.3.1 Data Quality

Pointed Poisson analyses shows a single outlier at LLO, corresponding to an extremely high frequency trigger ($\sim 5,335$Hz compared to GW170104's peak frequency of ~ 150Hz) that is near the minimum threshold for events to be recorded (low amplitude). Upon further investigation, it was determined that this single auxiliary trigger was not actually causally related to GW170104 and is merely an accidental coincidence. This was motivated by the extreme frequency mismatch, the lack of any morphological similarity in the time-frequency structure, and the fact that the coincidence was not robust against different whitening procedures (i.e. it was not recovered when longer stretches of data were used to whiten the auxiliary channel). Offline OVL analyses similarly found no auxiliary couplings within ~ 9 seconds of the event, and even the nearest couplings were associated with FAP$\gtrsim 10^{-3}$. Nonetheless, OVL was able to identify $\sim 20\%$ of glitches at FAP$\sim 10^{-3}$ and $\sim 40\%$ at FAP$\sim 10^{-2}$. This is roughly 400-2000 times the expected efficiency from chance alone.

Due to peculiarities with the online analyses surrounding GW170104, iDQ's predictions are uninformative. While iDQ found no evidence for auxiliary couplings to $h(t)$ in either detector, it had not received any new data for training in several weeks. GW170104 occurred shortly after a holiday break during which there was scant new data. Detector non-stationarity would suggest that offline studies are more reliable than the online iDQ inference in this case, mostly because offline analyses can examine data acausally. Regardless, both the online iDQ and offline OVL analyses agreed that there are no auxiliary couplings that could be responsible for GW170104.

4.3.2 Localization Characterization

Two localizations were made available for GW170104, and they are extremely consistent with one another (Fidelity $\gtrsim 0.85$). Like GW151226, GW170104 came from nearly overhead the LIGOs, although GW170104 was located more directly overhead LHO and arrived at LLO $3.0^{+0.4}_{-0.5}$ms af-

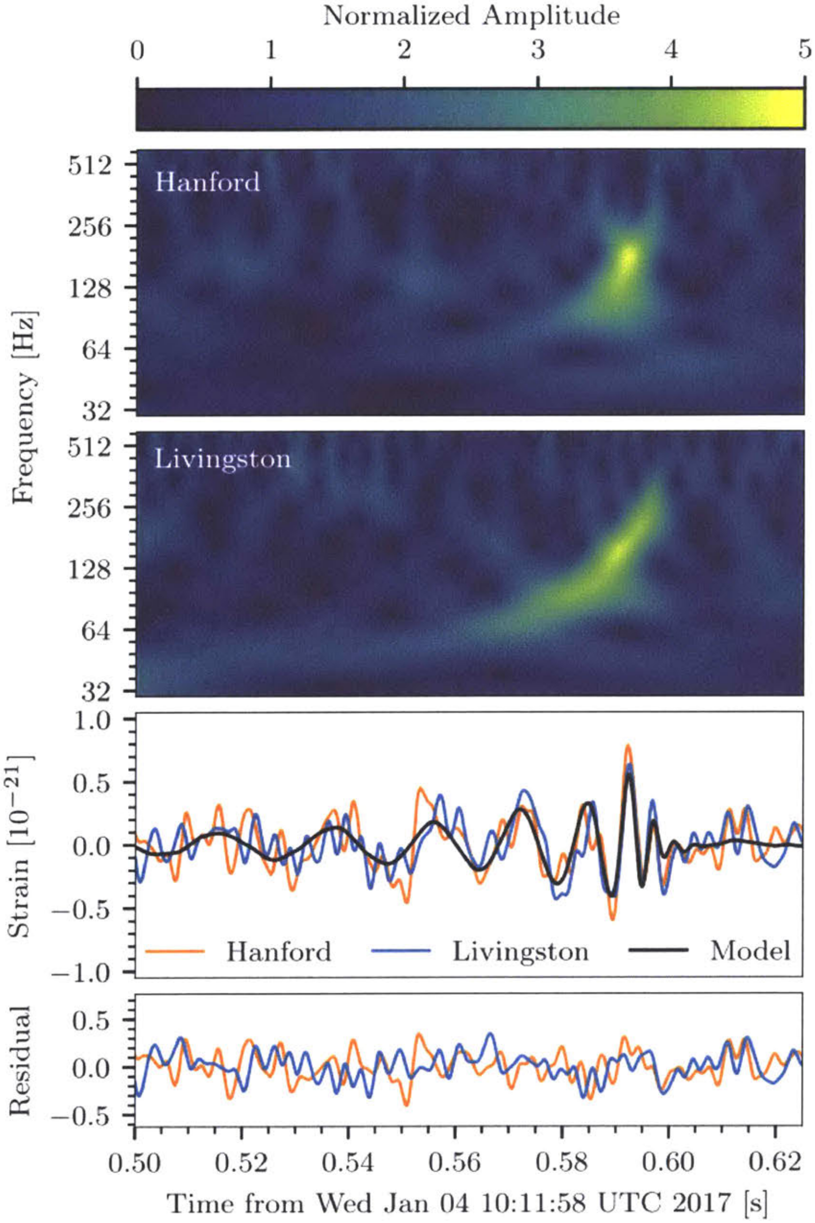


Figure 4-5: Ω-scans and whitened $h(t)$ with superimposed models for GW170104 [151].

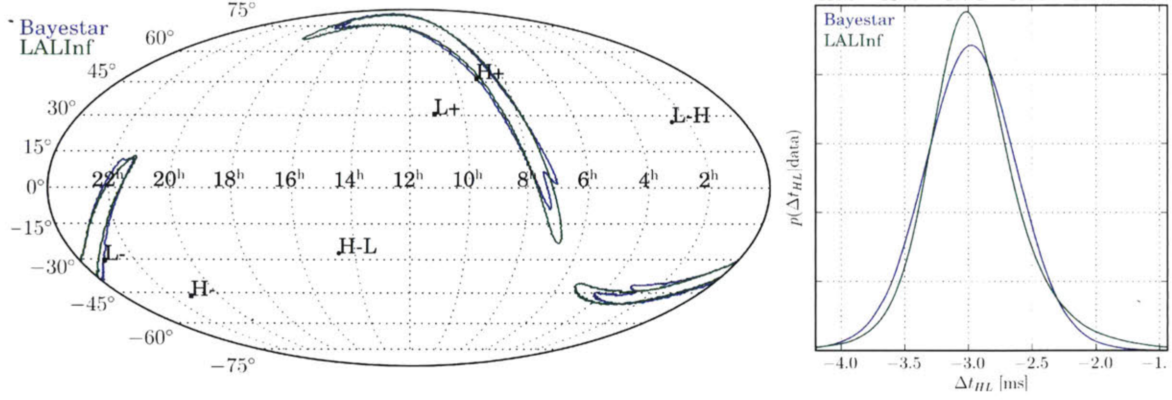


Figure 4-6: Equatorial Mollweide projection of 90% confidence regions and time-delay marginals for GW170104 [158].

ter reaching LHO. Also like GW151226, the localization is broader than GW150914's due to the lower ρ_{net} and larger diameter of the triangulation ring. GW170104 is another typical 2-detector localization.

4.4 LVT151012

Although not significant enough to be claimed as another detection, LVT151012 is nonetheless interesting. It corresponds to a $\lesssim 2\sigma$ deviation from background and, including estimates of the rate of signals in our detectors based on confirmed detections [147], it is more likely to be a signal than noise [7]. LVT151012 is consistent with the merger of a $23^{+18}_{-6} M_\odot$ BH with a $13^{+4}_{-5} M_\odot$ BH at a distance of $1.0^{+0.5}_{-0.5}$ Gpc, and arrived at the LIGO detectors on October 12, 2015 at 09:54:43 UTC. Marginal candidates like LVT151012 are expected from astrophysical distributions, particularly given the observed rate of more confident detections.

Similar data quality studies to those performed for GW150914, GW151226, and GW170104 showed no evidence for auxiliary couplings to $h(t)$ surrounding LVT151012. The localization estimate is again very broad due to the low ρ_{net} and location, corresponding to arrival at LHO $0.6^{+0.6}_{-0.6}$ms before LLO.

4.5 A Population of Binary Black Hole Systems

One detection is vindication. A second is confirmation. The third detection begs us to consider the population.[1] Indeed, even with this small number of sources we can already infer some properties of the source population. An obvious first question is the rate of coalescences. Although the precise bounds are somewhat model dependent, these three events constrain the rate of BBH coalescences to between 12–213 $\mathrm{Gpc}^{-3}\mathrm{year}^{-1}$ [151], consistent with the broad uncertainties predicted before the first detection [8]. Additionally, the distribution of ρ_{net} is roughly consistent with expectations from the predicted distance distribution, with ρ_{net}=23.7, 13.0, 13.0, and 9.7 for GW150914, GW151226, GW170104, and LVT151012, respectively. Because sources should be distributed uniformly within co-moving volumes on large scales, we expect $p(D_L) = dV/dD_L \propto D_L^2$, which implies $p(\rho_{\mathrm{net}}) \propto \rho_{\mathrm{net}}^{-4}$ because $\rho_{\mathrm{net}} \propto D_L^{-1}$. This implies that signals with $\rho_{\mathrm{net}} \geq 10$ should be 8 times more plentiful than sources with $\rho_{\mathrm{net}} \geq 20$. Although we have too few signals to conclusively demonstrate this (yet), the general trend agrees. In fact, if we assume 3 detections with $\rho_{\mathrm{net}} \geq 13$, we should expect 0.5 detections with $\rho_{\mathrm{net}} \geq 23.7$, corresponding to a $\geq 30\%$ chance of observing a *single* detection at least as loud as GW150914. Furthermore, the distribution of source locations appears roughly consistent with expectations. Fig. 4-7 shows the 50% and 90% confidence regions for all our events in Equatorial coordinates, and this distribution qualitatively resembles Fig. 3-27a. What's more, the localization estimates in Geographic coordinates clearly favor the antenna responses' maxima (Fig. 4-8). In fact, all triangulation rings span both maxima for both LHO and LLO, with the exception of GW150914, which falls slightly West of LLO's Southern maximum. This is, again, exactly what we expect from a population of BBH coalescences.

The observed events also span an interesting mass range, from $\sim 5M_{\odot}$ to as much as $\sim 40M_{\odot}$ (Fig. 4-9). These massive BH are approximately what is predicted to form within the cores of Globular Clusters [124] and typically are expected to form in low-metallicity environments [142], although the data is still consistent with formation through common envelop evolution. Additionally, the spins observed for each component, while not tightly constrained, suggest that large, aligned spins are not common. One might expect large, aligned spins to develop during the common-envelop phase of a BBH systems that forms from the main-sequence evolution of a pair of massive stars. However, dynamical capture within Globular Clusters would suggest a nearly isotropic spin

---

[1] Nonetheless, astronomers often begin to consider populations of sources with only a single (or fewer) detections (see, e.g. [142, 147]).

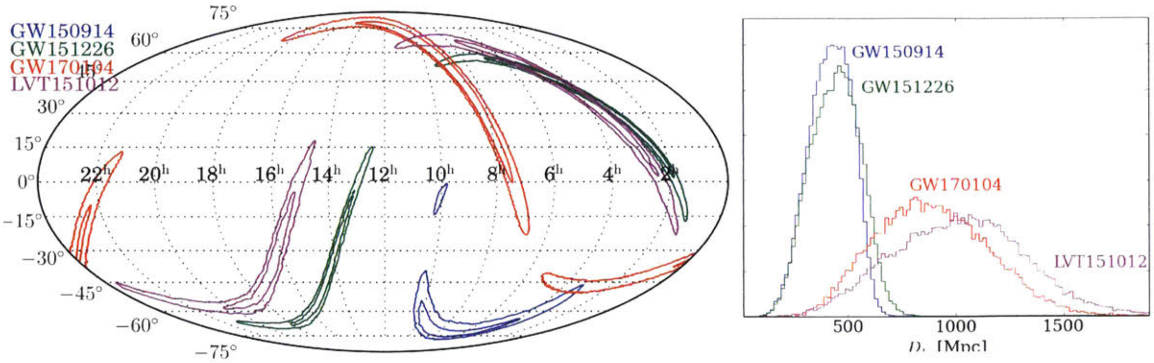


Figure 4-7: Celestial Mollweide projections of LALInference 50% and 90% confidence regions from localization estimates for GW150914, GW151226, GW170104, and LVT151012 along with luminosity distance marginals. We note that the distribution of localization estimates appears qualitatively similar to Fig. 3-27, as expected.

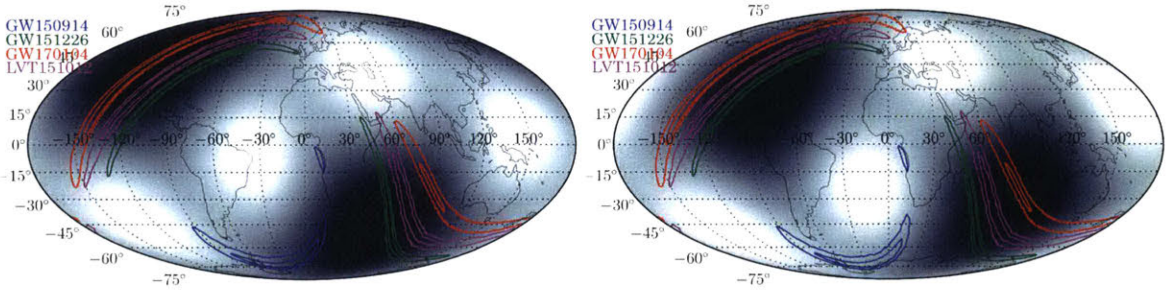


Figure 4-8: 50% and 90% confidence regions from LALInference localization estimates in Geographic coordinates with the antenna responses ($|F_+|^2 + |F_\times|^2$) for (left) LHO and (right) LLO superimposed. We note that all events come from near the maxima of each IFO's response, except GW150914, which is somewhat displaced from LLO's maximum.

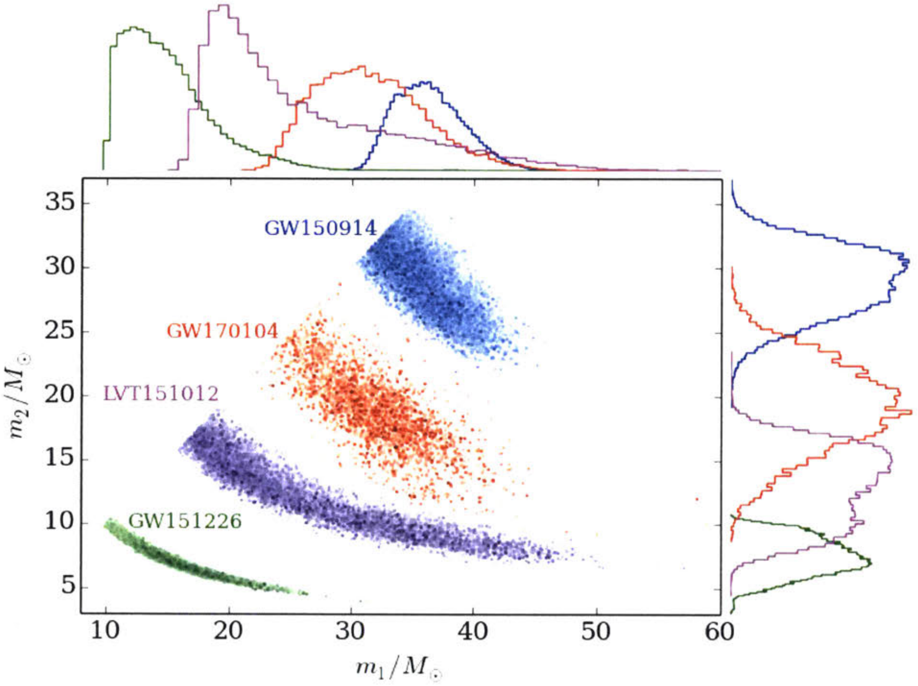


Figure 4-9: Joint posteriors for component masses for GW150914, GW151226, GW170104, and LVT151012.

distribution. Thus, the evidence seems to suggest that the BBH coalescences observed so far are likely to have formed through dynamical processes in distant Globular Clusters [142] roughly evenly distributed through co-moving volume.

We note that the non-detection of BNS and NSBH systems is currently expected from astrophysical models, in conflict with only the most optimistic rate estimates [149]. To wit, estimates made before O1 predicted between 0.0005–4 BNS detections [8], and the observed data are entirely consistent with these predictions [149]. What's more, detector sensitivity during O2 is comparable to O1, meaning the expected number of detections is only increased by the additional live-time (a factor of $O(1)$). These expected rates are low enough that we have no reason to suspect we are missing signals due to template mismatch caused by neglected nonlinear tidal interactions (§2.2.4), although this may change over the next few years [8, 149].

Chapter 5

Conclusions

Binary systems have always provided a testing ground for theories of gravity and orbital evolution, and they will continue to do so, from tests of strong-field General Relativity [148, 151] to nonlinear tidal interactions [62, 63, 168, 169, 170]. Exploiting the opportunities thereby provided will require firm theoretical understanding of the underlying physics along with robust statistical techniques to extract as much information as possible from noisy measurements.

In this thesis, I have shown how weakly nonlinear interactions between g-modes in the cores of sun-like main-sequence stars can be excited by Jupiter-mass companions. These interactions excite a cascade of modes, from long wavelengths that couple linearly to the tidal potential to short wavelengths that dissipate energy near the stellar core. We find that these networks of nonlinearly coupled modes can dissipate significant amounts of energy, producing orbital decay time-scales much shorter than their hosts' main-sequence lifetimes. In this sense, hot Jupiters really do just fall into their hosts. Work is ongoing to investigate similar behavior in other stellar hosts, including a variety of stellar masses and ages as well as evolved stars entering the horizontal and Red Giant branches [111].

Nonlinear interactions can also be important in GW sources. These systems primarily decay due to the emission of GWs, and resonant interactions do not have enough time to grow appreciably as the orbit spirals to ever increasing frequencies. Nonetheless, a nonresonant coupling between p-modes, g-modes, and the linear tide can grow throughout the inspiral and introduce a large phase shift in the GW waveform relative to a point-particle model. Although theoretical constraints are still loose, this nonlinear tidal interaction could bias inferred binary parameters as well as reduce detection rates by $\gtrsim 90\%$ if ignored. These effects should be well measured when taken into account, introducing yet another possible mechanism for measuring both redshift and luminosity distance directly from GWs alone. Furthermore, studies are ongoing to determine the impact of nonlinear tidal effects on signal consistency checks within actual detection pipelines as well as our ability to further constrain or detect these parameters using a population of marginal events.

Regardless of the presence of nonlinear tidal interactions, GWs from compact binaries offer a unique opportunity to study dynamics within strong-field GR as well as the formation mechanisms and distributions of compact objects that may not be observable in any other way. The three confirmed detections (GW150914, GW151226, GW170104) along with the weaker candidate (LVT151012) already show this to be true. Observations of more systems will constitute far more than butterfly collection; they will allow us to constrain cosmological models, stellar evolution, and even the basic dynamics of gravitating systems with ever increasing precision. However, to do so we must carefully analyze our observations to extract as much information as possible. Naturally, synthesizing information from all aspects of multi-messenger astronomy can improve our understanding of these sources and better place them in astrophysical context.

This includes rapid and robust identification of non-Gaussian noise artifacts within our detectors as well as comprehensive offline analyses to firmly rule out any terrestrial sources of noise. I have developed and described several such tools, demonstrating their importance in vetting all GW detections to date. We continue to study ways to improve our classification of noise artifacts, including novel machine learning techniques, feature space representations, and parametrizations of the classification problem to explicitly include multiple populations of $h(t)$ artifacts. This also extends to a thorough understanding of our detectors' responses and exposure. I have shown how inaccurate assumptions about the detector responses can bias triangulation and therefore localization estimates for 3^{rd} generation detectors. I have also shown how diurnal cycles within detectors can modulate the parts of the sky to which we are sensitive. In particular, future study will enable more stringent distributional tests for the sources' locations based on the network exposure. These facts, coupled with intensive characterization and comparisons of localization estimates, will help disentangle the complex problem of multi-messenger astronomy in the coming years.

There remain open questions in GW detection and tidal interactions within close binaries. However, this thesis has addressed several and provides the tools necessary to close others. Nature will certainly produce many surprises in the coming years, and we are now better prepared to understand and exploit them as much as possible.

Appendix A

Collective Instabilties

§2.1 provided a heuristic motivation for the scaling of collective parametric instabilities. Here we provide more rigorous derivations for a variety of coupling scenarios. While these examples are still somewhat artificial, they show that our intuition from the heuristic derivation holds in many situations.

We begin by analyzing the stability of a set of identical modes which are completely intercoupled (each mode is coupled to every other mode). Because we are interested in the stability of linear solutions, we assume the parent behaves as

$$q_\alpha = A_\alpha e^{-i(m_\alpha \Omega t - \delta_\alpha)}. \tag{A.1}$$

Each daughter then obeys

$$\begin{aligned} \dot{q}_\beta + (i\omega_\beta + \gamma_\beta) q_\beta =& i\omega_\beta \sum_\epsilon \kappa_{\alpha\beta\epsilon} A_\alpha e^{+i(m_\alpha \Omega t - \delta_\alpha)} q_\epsilon^* & \text{(A.2)} \\ =& i\omega_\beta \kappa_{\alpha\beta\beta} A_\alpha e^{+i(m_\alpha \Omega t - \delta_\alpha)} q_\beta^* + 2i\omega_\beta \sum_{\epsilon \neq \beta} \kappa_{\alpha\beta\epsilon} A_\alpha e^{+i(m_\alpha \Omega t - \delta_\alpha)} q_\epsilon^*. & \text{(A.3)} \end{aligned}$$

Defining a new set of variables $q = xe^{-i(\omega-\Delta)t}$, we re-write these as

$$\begin{aligned}
\dot{x}_\beta + (i\Delta_\beta + \gamma_\beta)x_\beta &= i\omega_\beta \kappa_{\alpha\beta\beta} A_\alpha x^*_\beta e^{+i(m_\alpha\Omega + 2\omega_\beta - 2\Delta_\beta)t - i\delta_\alpha} \qquad \text{(A.4)}\\
&\quad + 2i\omega_\beta \sum_{\epsilon\neq\beta} \kappa_{\alpha\beta\epsilon} A_\alpha x^*_\epsilon e^{+i(m_\alpha\Omega + \omega_\beta + \omega_\epsilon - \Delta_\beta - \Delta_\epsilon)t - i\delta_\alpha} \qquad \text{(A.5)}\\
&= i\omega_\beta \kappa_{\alpha\beta\beta} A_\alpha x^*_\beta e^{-i\delta_\alpha} + 2i\omega_\beta \sum_{\epsilon\neq\beta} \kappa_{\alpha\beta\epsilon} A_\alpha x^*_\epsilon e^{-i\delta_\alpha}, \qquad \text{(A.6)}
\end{aligned}$$

where in the last step we demanded that the time dependence cancels

$$m_\alpha\Omega + \omega_\beta + \omega_\epsilon - \Delta_\beta - \Delta_\epsilon = 0 \ \forall \ \{\beta, \epsilon\}. \tag{A.7}$$

We then separate x into real and imaginary parts $x = R + iI$,

$$\begin{aligned}
\begin{bmatrix} \dot{R}_\beta \\ \dot{I}_\beta \end{bmatrix} &= \begin{bmatrix} -\gamma_\beta + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \sin\delta_\alpha & \Delta_\beta + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \cos\delta_\alpha \\ -\Delta_\beta + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \cos\delta_\alpha & -\gamma_\beta - \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \sin\delta_\alpha \end{bmatrix} \begin{bmatrix} R_\beta \\ I_\beta \end{bmatrix} \\
&\quad + \sum_{\epsilon\neq\beta} \begin{bmatrix} 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \sin\delta_\alpha & 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \cos\delta_\alpha \\ 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \cos\delta_\alpha & -2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \sin\delta_\alpha \end{bmatrix} \begin{bmatrix} R_\epsilon \\ I_\epsilon \end{bmatrix}.
\end{aligned} \tag{A.8}$$

If we assume $[R_\beta, I_\beta] \propto e^{st} \ \forall \ \beta$, then this equation becomes

$$\begin{aligned}
0 &= \begin{bmatrix} -(\gamma_\beta + s) + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \sin\delta_\alpha & \Delta_\beta + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \cos\delta_\alpha \\ -\Delta_\beta + \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \cos\delta_\alpha & -(\gamma_\beta + s) - \omega_\beta A_\alpha \kappa_{\alpha\beta\beta} \sin\delta_\alpha \end{bmatrix} \begin{bmatrix} R_\beta \\ I_\beta \end{bmatrix} \\
&\quad + \sum_{\epsilon\neq\beta} \begin{bmatrix} 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \sin\delta_\alpha & 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \cos\delta_\alpha \\ 2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \cos\delta_\alpha & -2\omega_\beta A_\alpha \kappa_{\alpha\beta\epsilon} \sin\delta_\alpha \end{bmatrix} \begin{bmatrix} R_\epsilon \\ I_\epsilon \end{bmatrix}.
\end{aligned} \tag{A.9}$$

The general solution of this large eigenvalue problem typically must be obtained numerically. However, if we make several approximations, the problem becomes analytically tractable. Specifically, if we assume

$$\left.\begin{array}{r} \omega_\beta = \omega \\ \gamma_\beta = \gamma \\ \kappa_{\alpha\beta\beta} = \kappa_s \end{array}\right| \ \forall \ \beta, \qquad \kappa_{\alpha\beta\epsilon} = \kappa \ \forall \ \epsilon \neq \beta, \tag{A.10}$$

then we can define

$$M_S \equiv \begin{bmatrix} -(\gamma+s)+\omega A_\alpha \kappa_s \sin\delta_\alpha & \Delta+\omega A_\alpha \kappa_s \cos\delta_\alpha \\ -\Delta+\omega A_\alpha \kappa_s \cos\delta_\alpha & -(\gamma+s)-\omega A_\alpha \kappa_s \sin\delta_\alpha \end{bmatrix}, \tag{A.11}$$

$$M_I \equiv \begin{bmatrix} 2\omega A_\alpha \kappa \sin\delta_\alpha & 2\omega A_\alpha \kappa \cos\delta_\alpha \\ 2\omega A_\alpha \kappa \cos\delta_\alpha & -2\omega A_\alpha \kappa \sin\delta_\alpha \end{bmatrix}, \tag{A.12}$$

where $\Delta_\beta = \Delta \ \forall \ \beta$ since $\omega_\beta = \omega \ \forall \ \beta$. Writing this as a single matrix and requiring non-trivial mode amplitudes, we obtain

$$0 = \det \begin{vmatrix} M_S & M_I & M_I & M_I & \cdots & M_I & M_I \\ M_I & M_S & M_I & M_I & \cdots & M_I & M_I \\ M_I & M_I & M_S & M_I & \cdots & M_I & M_I \\ M_I & M_I & M_I & M_S & \cdots & M_I & M_I \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ M_I & M_I & M_I & M_I & \cdots & M_S & M_I \\ M_I & M_I & M_I & M_I & \cdots & M_I & M_S \end{vmatrix} = \det|M_S - M_I|^{N-1} \det|M_S + (N-1)M_I|. \tag{A.13}$$

by adding and subtracting columns and rows to make this lower-triangular and the applying the identity

$$\det \begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = \det \begin{vmatrix} A & 0 \\ C & D \end{vmatrix} = \det|A|\det|D| \tag{A.14}$$

repeatedly. We obtain $N-1$ repeated pairs of roots and one additional pair. The corresponding eigenvalues are

$$\begin{aligned} \det|M_S - M_I| =& (\gamma+s)^2 + \Delta^2 - \omega^2 A_\alpha^2 (\kappa_s - 2\kappa)^2 = 0 \\ \Rightarrow s =& -\gamma \pm \sqrt{\omega^2 A_\alpha^2 (2\kappa - \kappa_s)^2 - \Delta^2} \end{aligned} \tag{A.15}$$

and

$$\begin{aligned} \det|M_S + (N-1)M_I| =& (\gamma+s)^2 + \Delta^2 - \omega^2 A_\alpha^2 (2(N-1)\kappa + \kappa_s)^2 = 0 \\ \Rightarrow s =& -\gamma \pm \sqrt{\omega^2 A_\alpha^2 (2(N-1)\kappa + \kappa_s)^2 - \Delta^2}. \end{aligned} \tag{A.16}$$

In particular, we are interested in the values of A_α for which $\mathbb{R}\{s\} \rightarrow 0$. These are

$$A^2_{\mathrm{thr}} = \frac{\gamma^2 + \Delta^2}{4\omega^2(\kappa - \frac{1}{2}\kappa_s)^2} \tag{A.17}$$

and

$$A^2_{\mathrm{thr}} = \frac{\gamma^2 + \Delta^2}{4\omega^2((N-1)\kappa + \frac{1}{2}\kappa_s)^2}, \tag{A.18}$$

respectively. We see that there are $N-1$ modes that resemble standard 3-mode instabilities and one collective mode, with a threshold amplitude threshold suppressed by a factor of $\sim (N-1)$.

Because of the assumptions in Eqn. A.10, the actual value of A_{thr} will differ somewhat from this expression. Nonetheless, we expect it to generalize to the requirement that

$$(N-1)^2 A^2_\alpha \gtrsim \frac{\gamma_\beta \gamma_\epsilon}{4\omega_\beta \omega_\epsilon \kappa^2_{\alpha\beta\epsilon}} \left[1 + \frac{(\Delta_\beta + \Delta_\epsilon)^2}{(\gamma_\beta + \gamma_\epsilon)^2}\right] \ \forall \text{ modes } \alpha, \beta \in \text{collective set of } N \text{ modes}, \tag{A.19}$$

where $\Delta_\beta + \Delta_\epsilon = m_\alpha \Omega + \omega_\beta + \omega_\epsilon$. Similarly, we expect the growth rate to be enhanced by a factor of $\sim N$ over the 3-mode result. This is because the collective mode effectively sees $\sim N$ copies of the same coupling and κ appears $\sim N$ times larger ($s \propto \kappa$, $A_{\mathrm{thr}} \propto 1/\kappa$).

Now, let us consider two overlapping sets of collectively unstable modes, a coupling topology where there are three types of modes. A modes couple to other A modes and to C modes. B modes couple to other B modes and to C modes. C modes couple to all other modes. Furthermore, we assume that all A, B, and C modes couple to the same parent modes. The associated eigenvalue problem yields the following characteristic equation

$$0 = \det \begin{vmatrix}
M_S^A & M_I^A & \cdots & M_I^A & M_I^A & M_I^A & M_I^A & \cdots & M_I^A & M_I^A & 0 & 0 & \cdots & 0 & 0 \\
M_I^A & M_S^A & \cdots & M_I^A & M_I^A & M_I^A & M_I^A & \cdots & M_I^A & M_I^A & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_I^A & M_I^A & \cdots & M_S^A & M_I^A & M_I^A & M_I^A & \cdots & M_I^A & M_I^A & 0 & 0 & \cdots & 0 & 0 \\
M_I^A & M_I^A & \cdots & M_I^A & M_S^A & M_I^A & M_I^A & \cdots & M_I^A & M_I^A & 0 & 0 & \cdots & 0 & 0 \\
M_I^A & M_I^A & \cdots & M_I^A & M_I^A & M_S^C & M_I^C & \cdots & M_I^C & M_I^C & M_I^B & M_I^B & \cdots & M_I^B & M_I^B \\
M_I^A & M_I^A & \cdots & M_I^A & M_I^A & M_I^C & M_S^C & \cdots & M_I^C & M_I^C & M_I^B & M_I^B & \cdots & M_I^B & M_I^B \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
M_I^A & M_I^A & \cdots & M_I^A & M_I^A & M_I^C & M_I^C & \cdots & M_S^C & M_I^C & M_I^B & M_I^B & \cdots & M_I^B & M_I^B \\
M_I^A & M_I^A & \cdots & M_I^A & M_I^A & M_I^C & M_I^C & \cdots & M_I^C & M_S^C & M_I^B & M_I^B & \cdots & M_I^B & M_I^B \\
0 & 0 & \cdots & 0 & 0 & M_I^B & M_I^B & \cdots & M_I^B & M_I^B & M_S^B & M_I^B & \cdots & M_I^B & M_I^B \\
0 & 0 & \cdots & 0 & 0 & M_I^B & M_I^B & \cdots & M_I^B & M_I^B & M_I^B & M_S^B & \cdots & M_I^B & M_I^B \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & M_I^B & M_I^B & \cdots & M_I^B & M_I^B & M_I^B & M_I^B & \cdots & M_S^B & M_I^B \\
0 & 0 & \cdots & 0 & 0 & M_I^B & M_I^B & \cdots & M_I^B & M_I^B & M_I^B & M_I^B & \cdots & M_I^B & M_S^B
\end{vmatrix} . \tag{A.20}$$

We again note the high degree of symmetry and reduce the determinant to

$$\begin{aligned}
0 = \left(\det\left|M_S^A - M_I^A\right|\right)^{N_A-1} & \left(\det\left|M_S^B - M_I^B\right|\right)^{N_N-1} \left(\det\left|M_S^C - M_I^C\right|\right)^{N_C-1} \\
& \times \det\left|M_S^A + (N_A-1)M_I^A\right| \det\left|M_S^B + (N_B-1)M_I^B\right| \\
& \qquad \times \det\left|M_S^C + (N_C-1)M_I^C - N_C Z\right| ,
\end{aligned} \tag{A.21}$$

where

$$Z = N_A M_I^A \left(M_S^A + (N_A-1)M_I^A\right)^{-1} M_I^A + N_B M_I^b \left(M_S^B + (N_B-1)M_I^B\right)^{-1} M_I^B . \tag{A.22}$$

We recognize this as $N_A - 1$ independent A modes, $N_B - 1$ independent B modes, $N_C - 1$ independent C modes, one mode corresponding to the collective motion without the coupling to C modes for each of the A and B sets, and a collective set for the C modes with a modification due to the couplings to A and B (through Z). We further note that when $N_c \to 0$, the instabilities reduce to two separate collective sets, as expected.

The interesting eigenvalue is associated with the interaction between the C modes' collective set and the couplings to A and B modes. If we assume that all mode parameters are the same for all sets of modes, and further assume $N_a = N_b = N_c$, we obtain

$$(\gamma + s)^2 + \Delta^2 - \omega^2 A_\alpha^2 \left((k_s + 2(N-1)k)^2 + 8N^2k^2 \right) = 0 \tag{A.23}$$

and the threshold amplitude

$$\begin{aligned} A_{\mathrm{thr}}^2 &= \frac{\gamma^2 + \Delta^2}{4\omega^2 \left(3k^2N^2 + k(k_s - 2k)N + k(k - k_s) + k_s^2/4\right)} && \text{(A.24a)} \\ &\approx \frac{1}{3N^2}\left(\frac{\gamma^2 + \Delta^2}{4\omega^2k^2}\right) = \frac{1}{N_a^2 + N_b^2 + N_c^2}\left(\frac{\gamma^2 + \Delta^2}{4\omega^2k^2}\right), && \text{(A.24b)} \end{aligned}$$

where we assumed $N \gg 1$. We note that this is very similar to the case of a single collective set, except $N^2 \to N_a^2 + N_b^2 + N_c^2$. If we stitch together many separate collective sets by overlapping them, we only expect the effective number of modes to sum in quadrature. This was tested numerically without assuming equal numbers of modes, and found to be in reasonable agreement with this scaling.

Furthermore, let us consider sets of modes which are *not* self-coupled. This means that all modes in set A couple to all modes in set B but do not couple to other modes in set A. This is expected to be the most common case in which, e.g., all modes in set A have natural frequencies near 40% of the parent's and all modes in set B have frequencies near 60% of the parent's. Alternatively, if the parent azimuthal order (m) is odd, then the daughter modes must have different m even if they have identical ω and γ. If we consider two sets of modes, one with N daughters and one with n daughters, we find collective sets with characteristic equations like (capital letters correspond to the N-mode set and lower case letters correspond to the n-mode set)

$$0 = \det \begin{vmatrix} M_S & 0 & \cdots & 0 & 0 & M_I & M_I & \cdots & M_I & M_I \\ 0 & M_S & \cdots & 0 & 0 & M_I & M_I & \cdots & M_I & M_I \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & M_S & 0 & M_I & M_I & \cdots & M_I & M_I \\ 0 & 0 & \cdots & 0 & M_S & M_I & M_I & \cdots & M_I & M_I \\ M_i & M_i & \cdots & M_i & M_i & M_s & 0 & \cdots & 0 & 0 \\ M_i & M_i & \cdots & M_i & M_i & 0 & M_s & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ M_i & M_i & \cdots & M_i & M_i & 0 & 0 & \cdots & M_s & 0 \\ M_i & M_i & \cdots & M_i & M_i & 0 & 0 & \cdots & 0 & M_s \end{vmatrix} \tag{A.25}$$

We can simplify this to

$$0 = \left(\det|M_S|\right)^{N-1} \left(\det|M_s|\right)^{n-1} \det \begin{vmatrix} M_S & nM_I \\ NM_i & M_s \end{vmatrix}, \tag{A.26}$$

which, again, resembles two sets of independent modes and a collective mode. In general, the determinant must be solved numerically. However, if we again assume identical mode parameters and that $N = n$, we see that this reduces to

$$\det \begin{vmatrix} M_S & NM_I \\ NM_I & M_S \end{vmatrix}, \tag{A.27}$$

which looks just like the three-mode instability equations with $k \to Nk$ even though the coupling topology is completely different from previous cases.

An immediate question is how the presence of small differences in modes affects these collective instabilities; they would not be interesting physically if they required all modes to be precisely identical. We therefore consider the stability of a collective set of identical modes upon the addition a new mode with different parameters (ω', γ', etc.). Consider the following characteristic equation, with N identical modes and one slightly different mode indicated by δM

$$0 = \det \begin{vmatrix} M_S & M_I & \cdots & M_I & M_I \\ M_I & M_S & \cdots & M_I & M_I \\ \vdots & \vdots & & \vdots & \vdots \\ M_I & M_I & \cdots & M_S & M_I \\ M_I & M_I & \cdots & M_I & M_S + \delta M \end{vmatrix} \tag{A.28}$$

We can reduce this to

$$\begin{aligned} 0 &= \left(\det |M_S - M_I|\right)^{N-1} \det \begin{vmatrix} M_S - M_I & -\delta M \\ NM_I & M_S + \delta M \end{vmatrix} \\ &= \left(\det |M_S - M_I|\right)^{N-1} \det |M_S + (N-1)M_I| \det \left| M_S - NM_I \left(M_S + (N-1)M_I\right)^{-1} M_I + \delta M \right| . \end{aligned} \tag{A.29}$$

As $\delta M \to 0$, this reduces to a single collective set with $N \to N+1$. We also note that this appears to be a normal collective set with N modes and a new mode associated with δM. Furthermore, if δM dominates the new mode, then we see that it will decouple from the other modes. Clearly, there will be some threshold for how large δM needs to be before the different mode decouples, and that threshold will depend on the parent's amplitude in a non-trivial way. We expect that larger parent amplitudes will support larger δM before the mode decouples. This is equivalent to larger parent amplitudes exciting modes with larger detunings, for example.

Appendix B

Energy Scales within Sun-like Stars

There are several important energy scales for g-modes within sun-like stars. Specifically, we focus on the linear energy ($E_{\rm lin}$), the 3-mode stability threshold ($E_{\rm thr}$), and the breaking energy when the modes begin to overturn the background stellar stratification ($E_{\rm break}$). $E_{\rm lin}$ and $E_{\rm break}$ are properties of individual modes, while each mode may have several $E_{\rm thr}$ corresponding to different 3-mode couplings. We note that when $E_{\rm lin} \sim \min\{E_{\rm thr}\}$, we are in a weakly nonlinear regime with only a few modes excited. When $E_{\rm lin} \sim E_{\rm break}$, we are in a strongly nonlinear regime. What's more, we will show that, in some sense, the average 3-mode coupling corresponds to $E_{\rm thr} \sim E_{\rm break}$ and therefore most 3-mode triples become unstable as we enter the strongly nonlinear regime. Because we intentionally examine the weakly nonlinear regime, we expect $E_{\rm thr} \leq E_{\rm lin} \leq E_{\rm break}$.

We begin with $E_{\rm lin}$, which we write as

$$E_{\rm lin}/E_0 = \frac{(\omega U)^2}{\gamma^2 + (\omega - m\Omega)^2} \tag{B.1}$$

$$\sim \frac{\omega^2 U^2}{(\Delta\omega/2)^2} \tag{B.2}$$

$$\sim \left(\frac{2^5 \cdot 3 \cdot \pi^5}{5} \hat{I}^2_{a2\pm2} \alpha^2 l^2 \omega_0^{-11/3}\right) \left(\frac{M_p}{M + M_p}\right)^2 (\omega_0 P)^{-4} \omega^{5/3} \tag{B.3}$$

where we have assumed that the detuning is approximately half the mode spacing ($\Delta\omega \sim \omega/n = \omega^2/\alpha l$) and inserted Eqn. 1.55. If we assume a linearly resonant $l = 2$ mode ($\omega \sim 2\Omega_{\rm orb} = 4\pi/P$)

and a sun-like host ($\alpha \sim 4 \cdot 10^{-3}$, $\hat{I}_{\alpha 2 \pm 2} \sim 2.5 \cdot 10^{-3}$), we obtain[1]

$$E_{\mathrm{lin}} \sim \left(2.31 \cdot 10^{35}\mathrm{ergs}\right) \left(\frac{M_p}{M_J}\right)^2 \left(\frac{P}{\mathrm{day}}\right)^{-17/3} \tag{B.4}$$

Similarly, if we assume the detuning dominates the stability criterion (Eqn. 2.21), we find

$$E_{\mathrm{thr}}/E_0 = \frac{\gamma_\beta \gamma_\epsilon}{4\kappa^2 \omega_\beta \omega_\epsilon} \left(1 + \left(\frac{m_\alpha \Omega + \omega_\beta + \omega_\epsilon}{\gamma_\beta + \gamma_\epsilon}\right)^2\right) \tag{B.5}$$

$$\sim \frac{1}{4\kappa^2 \omega_\beta^2} \left(\frac{\varepsilon(\omega/2)^2/\alpha l_d}{2}\right)^2 \tag{B.6}$$

where l_d is the angular degree of the daughter mode, ε is a measure of how much better the actual detuning is than the maximum expected, and ω still refers to the natural frequency of the linearly resonant parent. Remembering that $\kappa \sim \hat{\kappa}(\omega \cdot \mathrm{day}/2\pi)^{-2}$, we obtain

$$E_{\mathrm{thr}} \sim \left(1.96 \cdot 10^{33}\mathrm{ergs}\right) \left(\frac{\varepsilon}{0.01}\right)^2 l_d^{-2} \left(\frac{P}{\mathrm{day}}\right)^{-6} \tag{B.7}$$

This is the approximate scaling of $\min\{E_{\mathrm{thr}}\}$, often corresponding to $\varepsilon \sim 10^{-2}$. We note that the scaling with P for both E_{thr} and E_{lin} is very similar, whereas the scaling with M_p is quite different. $M_p \sim 0.2 M_J$ implies $E_{\mathrm{thr}} \sim E_{\mathrm{lin}}$, which corresponds to the sharp turn-off we observe numerically (Fig. 2-9).

E_{break} is more complicated to express analytically, but using the WKB approximations in Appendix A of [170] we find

$$k_r \xi_r = \frac{\Lambda N}{\omega r} A \sin\phi \tag{B.8}$$

$$\frac{N}{r} = W \frac{\omega_0}{R} \tag{B.9}$$

$$A^2 = E \frac{\Delta P}{2\pi^2 \rho N r^3} = E \frac{P_0/\omega_0}{2\pi^2 \rho_0 (M/R^3)(N/r) r^4} \tag{B.10}$$

in the stellar core. Furthermore, the supremum of this should occur near the inner turning point

---

[1] This is slightly different than the estimate obtained in [63] because of slightly different assumed values of $\hat{I}_{\alpha 2 \pm 2}$.

when $\omega \sim N$, which yields

$$\sup_r\{k_r\xi_r\} = \sqrt{\frac{P_0 W^5}{2\pi^2\rho_0}}\left(\frac{\omega_0}{\omega}\right)^3 \Lambda\sqrt{\frac{E}{E_0}}. \tag{B.11}$$

Equating this with one and assuming stellar values yields

$$E_{\rm break} = \left(4.37\cdot 10^{36}{\rm ergs}\right)\left(\frac{6}{\Lambda^2}\right)\left(\frac{P}{\rm day}\right)^{-6} \tag{B.12}$$

We immediately note that both $E_{\rm break}$ and $E_{\rm thr}$ scale with the same power of P and similar powers of l, at least for large l, suggesting that $E_{\rm thr}$ and $E_{\rm break}$ are related in some way. Indeed, we expect breaking to occur through energy transfer into shorter wavelength modes, which must involve the 3-mode couplings in a fundamental way. When $E_{\rm break} \sim E_{\rm thr}$, our scaling relation suggests $\varepsilon \sim 0.47$. To say this another way, breaking occurs when 3-mode couplings which are not particularly resonant become unstable.

It is also worth comparing the 3-mode equilibrium energy for daughters in a unstable triple with the daughter's breaking energy. Assuming the parent obeys $E_{\rm lin} \gg E_{\rm thr}$, we expect the daughters to be driven toward equilibrium energies $(E_{\rm daughter})_{\rm 3md} \sim U_{\rm parent}/2\kappa$. This corresponds to

$$\frac{(E_{\rm daughter})_{\rm 3md}}{(E_{\rm daughter})_{\rm break}} \sim 82.9\left(\frac{\Lambda^2}{2}\right)\left(\frac{M_p}{M_J}\right)\left(\frac{P}{\rm day}\right)^{1/6} \tag{B.13}$$

and we see that the daughters are excited to large enough energies to break whenever they are unstable ($M_p \gtrsim 0.2M_J$), with a very weak dependence on the orbital period. Nonetheless, granddaughter modes become unstable before the daughters actually reach their breaking energy.

Our prediction for the relative size of all three energy scales is also validated; for linearly resonant modes, $E_{\rm thr} < E_{\rm lin} < E_{\rm break}$. We also recover the predicted threshold of $M_p \gtrsim 3.3M_J$ for linearly resonant g-modes to break in a $P \sim 3$ day orbit by comparing $E_{\rm break}$ and $E_{\rm lin}$ [20, 21]. Furthermore, using the limit of Eqn. 2.20 with parent energy E and $E \gg E_{\rm thr}$, we expect the mode amplitude growth rate to be $s \sim 2(\omega/2)\kappa\sqrt{E/E_0} \sim 0.145{\rm Hz}(P/{\rm day})\sqrt{E/E_0}$. If we assume the parent sits near $E_{\rm lin}$, a simple estimate of the energy dissipation is

$$2\left(\omega\kappa\sqrt{E_{\rm lin}/E_0}\right)E_{\rm lin} \sim \left(1.66\cdot 10^{28}{\rm erg/s}\right)\left(\frac{M_p}{M_J}\right)^3\left(\frac{P}{\rm day}\right)^{-7.5} \tag{B.14}$$

which is similar to, but smaller than, our numerical result (Eqn. 2.45)

$$\dot{E}_* \sim (3.5 \cdot 10^{29}\mathrm{erg/s}) \left(\frac{M_p}{M_J}\right)^{1.5} \left(\frac{P}{\mathrm{day}}\right)^{-7.4} \tag{B.15}$$

While the scaling with P is similar, we note the difference in the dependence on M_p along with the overall difference in energy scale. This is likely because the our simple estimate assumes only a single daughter pair is unstable, while in reality there are several that contribute; it is perhaps unsurprising that the complex dynamics observed in our numerical simulations are not perfectly captured by this back-of-the-envelop calculation.

Bibliography

[1] Exoplanets data explorer. http://exoplanets.org/, 2014. Accessed: Oct. 2014.

[2] J. Aasi et al. Constraints on cosmic strings from the ligo-virgo gravitational-wave detectors. *Phys. Rev. Lett.*, 112:131101, Apr. 2014.

[3] J. Abadie et al. Search for gravitational waves from compact binary coalescence in ligo and virgo data from s5 and vsr1. *Phys. Rev. D*, 82:102001, Nov. 2010.

[4] J. Abadie et al. All-sky search for gravitational-wave bursts in the second joint ligo-virgo run. *Phys. Rev. D*, 85:122007, June 2012.

[5] J. Abadie et al. Search for gravitational waves from low mass compact binary coalescence in ligo's sixth science run and virgo's science runs 2 and 3. *Phys. Rev. D*, 85:082002, Apr. 2012.

[6] B. P. Abbott et al. Search for gravitational-wave bursts in the first year of the fifth ligo science run. *Phys. Rev. D*, 80:102001, Nov. 2009.

[7] B. P. Abbott et al. Binary black hole mergers in the first advanced ligo observing run. *Phys. Rev. X*, 6:041015, Oct. 2016.

[8] B. P. Abbott et al. Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO and Advanced Virgo. *Living Reviews in Relativity*, 19, Feb. 2016.

[9] C. Aerts, J. Christensen-Dalsgaard, and D. Kurtz. *Asteroseismology.* Springer, 2010. ISBN: 978-1-4020-5178-4.

[10] M. Agathos et al. Tiger: A data analysis pipeline for testing the strong-field dynamics of general relativity with gravitational wave signals from coalescing compact binaries. *Phys. Rev. D*, 89:082001, Apr. 2014.

[11] M. Agathos et al. Constraining the neutron star equation of state with gravitational wave signals from coalescing binary neutron stars. *Phys. Rev. D*, 92:023012, July 2015.

[12] M. Agathos, W. D. Pozzo, T. Li, C. V. D. Broeck, J. Veitch, and S. Vitale. *Testing General Relativity using Gravitational Waves from Binary Neutron Stars; Effect of Spins*, pages 1710–1712. World Scientific, 2015.

[13] P. Ajith et al. Robust vetoes for gravitational-wave burst triggers using known instrumental couplings. *Classical and Quantum Gravity*, 23(20):5825, 2006.

[14] P. Ajith et al. Physical instrumental vetoes for gravitational-wave burst triggers. *Phys. Rev. D*, 76:042004, Aug. 2007.

[15] R. Alonso et al. Transiting exoplanets from the CoRoT space mission. II. CoRoT-Exo-2b: a transiting planet around an active G star. *Astronomy and Astrophysics*, 482:L21–L24, May 2008.

[16] E. Baird et al. Degeneracy between mass and spin in black-hole-binary waveforms. *Phys. Rev. D*, 87:024035, Jan. 2013.

[17] G. Á. Bakos et al. HAT-P-34b-HAT-P-37b: Four Transiting Planets More Massive than Jupiter Orbiting Moderately Bright Stars. *Astronomical Journal*, 144:19, July 2012.

[18] P. Balachandran and É. É. Flanagan. Detectability of Mode Resonances in Coalescing Neutron Star Binaries. *ArXiv e-prints*, Jan. 2007.

[19] A. J. Barker. Three-dimensional simulations of internal wave breaking and the fate of planets around solar-type stars. *Monthly Notices of the Royal Astronomical Society*, 414(2):1365–1378, 2011.

[20] A. J. Barker and G. I. Ogilvie. On internal wave breaking and tidal dissipation near the centre of a solar-type star. *Monthly Notices of the Royal Astronomical Society*, 404:1849–1868, June 2010.

[21] A. J. Barker and G. I. Ogilvie. Stability analysis of a tidally excited internal gravity wave near the centre of a solar-type star. *Monthly Notices of the Royal Astronomical Society*, 417:745–761, Oct. 2011.

[22] B. Bécsy et al. Parameter estimation for gravitational-wave bursts with the bayeswave pipeline. *The Astrophysical Journal*, 839(1):15, 2017.

[23] C. Bender and S. Orszag. *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*. Advanced Mathematical Methods for Scientists and Engineers. Springer, 1999.

[24] C. P. L. Berry et al. Parameter estimation for binary neutron-star coalescences with realistic noise during the advanced ligo era. *The Astrophysical Journal*, 804(2):114, 2015.

[25] J. L. Birkby et al. Wts-2 b: a hot jupiter orbiting near its tidal destruction radius around a k dwarf. *Monthly Notices of the Royal Astronomical Society*, 440(2):1470–1489, 2014.

[26] R. Biswas et al. Likelihood-ratio ranking of gravitational-wave candidates in a non-Gaussian background. *Phys. Rev. D*, 85(12):122008, June 2012.

[27] R. Biswas et al. Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. *Phys. Rev. D*, 88:062003, Sept. 2013.

[28] L. Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity*, 17(1):2, 2014.

[29] R. Bondarescu, S. A. Teukolsky, and I. Wasserman. Spinning down newborn neutron stars: Nonlinear development of the r-mode instability. *Phys. Rev. D*, 79(10):104003, May 2009.

[30] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.

[31] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[32] J. Brink, S. A. Teukolsky, and I. Wasserman. Nonlinear coupling network to simulate the development of the r mode instability in neutron stars. II. Dynamics. *Phys. Rev. D*, 71(6):064029, Mar. 2005.

[33] J. C. Brown. An efficient algorithm for the calculation of a constant Q transform. *J. Acoust. Soc. Am*, 92(5):2698–2701, 1992.

[34] A. Buonanno et al. Comparison of post-newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Phys. Rev. D*, 80:084043, Oct. 2009.

[35] K. Cannon et al. Toward early-warning detection of gravitational waves from compact binary coalescence. *The Astrophysical Journal*, 748(2):136, 2012.

[36] K. Cannon, C. Hanna, and J. Peoples. Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation. *ArXiv e-prints*, Apr. 2015.

[37] S. M. Carroll. *Spacetime and geometry: An introduction to general relativity*. Addison Wesley, 2004.

[38] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011.

[39] S. Chatterji et al. Multiresolution techniques for the detection of gravitational-wave bursts. *Classical and Quantum Gravity*, 21(20):S1809, 2004.

[40] H.-Y. Chen et al. Observational selection effects with ground-based gravitational wave detectors. *The Astrophysical Journal*, 835(1):31, 2017.

[41] N. Christensen et al. Vetoes for inspiral triggers in ligo data. *Classical and Quantum Gravity*, 21(20):S1747, 2004.

[42] N. Christensen et al. Veto studies for ligo inspiral triggers. *Classical and Quantum Gravity*, 22(18):S1059, 2005.

[43] N. Cornish et al. Gravitational wave tests of general relativity with the parameterized post-einsteinian framework. *Phys. Rev. D*, 84:062003, Sept. 2011.

[44] N. J. Cornish and T. B. Littenberg. Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Classical and Quantum Gravity*, 32(13):135012, 2015.

[45] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[46] A. D. Credico et al. Gravitational wave burst vetoes in the ligo s2 and s3 data analyses. *Classical and Quantum Gravity*, 22(18):S1051, 2005.

[47] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods.* Cambridge University Press, 2000.

[48] C. Cutler and É. É. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658–2697, Mar. 1994.

[49] C. Cutler and M. Vallisneri. Lisa detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms. *Phys. Rev. D*, 76:104018, Nov. 2007.

[50] T. Damour, A. Nagar, and L. Villain. Measurability of the tidal polarizability of neutron stars in late-inspiral gravitational-wave signals. *Phys. Rev. D*, 85(12):123007, June 2012.

[51] W. Del Pozzo. Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers. *Phys. Rev. D*, 86:043011, Aug. 2012.

[52] W. Del Pozzo et al. Demonstrating the feasibility of probing the neutron-star equation of state with second-generation gravitational-wave detectors. *Phys. Rev. Lett.*, 111:071101, Aug. 2013.

[53] W. Del Pozzo, T. G. F. Li, and C. Messenger. Cosmological inference using only gravitational wave observations of binary neutron stars. *Phys. Rev. D*, 95:043502, Feb. 2017.

[54] R. Essick. Selection effects. https://github.com/reedessick/selectionEffects, 2016.

[55] R. Essick. Event supervisor. https://github.com/reedessick/eventSupervisor, 2017.

[56] R. Essick. Lvalert heartbeat. https://git.ligo.org/lscsoft/lvalert-heartbeat, 2017.

[57] R. Essick. Segdb2gracedb. https://github.com/reedessick/SegDB2GraceDB, 2017.

[58] R. Essick. Skymap statistics. https://github.com/reedessick/skymap_statistics, 2017.

[59] R. Essick, L. Blackburn, and E. Katsavounidis. Optimizing vetoes for gravitational-wave transient searches. *Classical and Quantum Gravity*, 30(15):155010, 2013.

[60] R. Essick et al. Localization of short duration gravitational-wave transients with the early advanced ligo and virgo detectors. *The Astrophysical Journal*, 800(2):81, 2015.

[61] R. Essick et al. Burst first two years open data release. http://www.ligo.org/scientists/burst-first2years, 2017.

[62] R. Essick, S. Vitale, and N. N. Weinberg. Impact of the tidal $p-g$ instability on the gravitational wave signal from coalescing binary neutron stars. *Phys. Rev. D*, 94:103012, Nov. 2016.

[63] R. Essick and N. N. Weinberg. Orbital decay of hot jupiters due to nonlinear tidal dissipation within solar-type hosts. *The Astrophysical Journal*, 816(1):18, 2016.

[64] M. Favata. Systematic Parameter Errors in Inspiraling Neutron Star Binaries. *Physical Review Letters*, 112(10):101101, Mar. 2014.

[65] É. É. Flanagan and S. A. Hughes. Measuring gravitational waves from binary black hole coalescences. i. signal to noise for inspiral, merger, and ringdown. *Phys. Rev. D*, 57:4535–4565, Apr. 1998.

[66] É. É. Flanagan and S. A. Hughes. Measuring gravitational waves from binary black hole coalescences. ii. the waves' information and its extraction, with and without templates. *Phys. Rev. D*, 57:4566–4587, Apr. 1998.

[67] É. É. Flanagan and S. A. Hughes. The basics of gravitational wave theory. *New Journal of Physics*, 7(1):204, 2005.

[68] É. É. Flanagan and É. Racine. Gravitomagnetic resonant excitation of Rossby modes in coalescing neutron star binaries. *Phys. Rev. D*, 75(4):044001–+, Feb. 2007.

[69] D. Gabor. Theory of Communication. *J. IEE*, 93(26):429–457, Nov. 1946.

[70] D. Gandolfi et al. Kepler-423b: a half-jupiter mass planet transiting a very old solar-like star. *Astronomy and Astrophysics*, 576:A11, 2015.

[71] M. Gillon et al. Improved parameters for the transiting hot jupiters wasp-4b and wasp-5b. *Astronomy and Astrophysics*, 496(1):259–267, 2009.

[72] M. Gillon et al. The thermal emission of the young and massive planet CoRoT-2b at 4.5 and 8 μm. *Astronomy and Astrophysics*, 511:A3, Feb. 2010.

[73] P. Goldreich and S. Soter. Q in the solar system. *Icarus*, 5:375–389, 1966.

[74] J. Goodman and E. S. Dickson. Dynamical tide in solar-type binaries. *The Astrophysical Journal*, 507(2):938, 1998.

[75] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer series in statistics. Springer, 2001.

[76] L. Hebb et al. WASP-19b: The Shortest Period Transiting Exoplanet Yet Discovered. *The Astrophysical Journal*, 708:224–231, Jan. 2010.

[77] R. Hecht-Nielsen. Neural networks for perception (vol. 2). Harcourt Brace & Co., Orlando, FL, USA, 1992.

[78] C. Hellier et al. On the Orbit of the Short-period Exoplanet WASP-19b. *The Astrophysical Journal Lett.*, 730:L31, Apr. 2011.

[79] T. Hinderer et al. Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatures in binary inspiral. *Phys. Rev. D*, 81:123016, June 2010.

[80] T. Hinderer et al. Effects of neutron-star dynamic tides on gravitational waveforms within the effective-one-body approach. *Phys. Rev. Lett.*, 116:181101, May 2016.

[81] W. C. G. Ho and D. Lai. Resonant tidal excitations of rotating neutron stars in coalescing binaries. *Monthly Notices of the Royal Astronomical Society*, 308(1):153–166, 1999.

[82] T. Isogai et al. Used percentage veto for ligo and virgo binary inspiral searches. *Journal of Physics: Conference Series*, 243(1):012005, 2010.

[83] B. Jackson, R. Barnes, and R. Greenberg. Observational evidence for tidal destruction of exoplanets. *The Astrophysical Journal*, 698(2):1357, 2009.

[84] B. Jackson, R. Greenberg, and R. Barnes. Tidal evolution of close-in extrasolar planets. *The Astrophysical Journal*, 678(2):1396, 2008.

[85] E. Katsavounidis and P. Shawhan. Veto selection for gravitational wave event searches. LIGO Document LIGO-G060639-00-Z, 2006.

[86] S. Klimenko et al. Constraint likelihood analysis for a network of gravitational wave detectors. *Phys. Rev. D*, 72:122002, Dec. 2005.

[87] S. Klimenko et al. A coherent method for detection of gravitational wave bursts. *Classical and Quantum Gravity*, 25(11):114029, 2008.

[88] M. Kramer and I. Stairs. The double pulsar. *Annual Review of Astronomy and Astrophysics*, 46:541 – 572, 2008.

[89] M. Kramer and N. Wex. The double pulsar system: a unique laboratory for gravity. *Classical and Quantum Gravity*, 26(7):073001, 2009.

[90] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, Mar. 1951.

[91] B. D. Lackey and L. Wade. Reconstructing the neutron-star equation of state with gravitational-wave detectors from a realistic population of inspiralling binary neutron stars. *Phys. Rev. D*, 91:043002, Feb. 2015.

[92] D. Lai. Resonant oscillations and tidal heating in coalescing binary neutron stars. *Monthly Notices of the Royal Astronomical Society*, 270(3):611–629, 1994.

[93] D. Lai and Y. Wu. Resonant tidal excitations of inertial modes in coalescing neutron star binaries. *Phys. Rev. D*, 74(2):024007–+, July 2006.

[94] T. G. F. Li et al. Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence. *Phys. Rev. D*, 85:082003, Apr. 2012.

[95] T. G. F. Li et al. Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence: Further investigations. *Journal of Physics: Conference Series*, 363(1):012028, 2012.

[96] X.-D. Li. Formation of black hole low-mass x-ray binaries. *New Astronomy Reviews*, 64:1 – 6, 2015.

[97] L. Lindblom, B. J. Owen, and D. A. Brown. Model waveform accuracy standards for gravitational wave data analysis. *Phys. Rev. D*, 78:124020, Dec. 2008.

[98] R. Lynch et al. Information-theoretic approach to the gravitational-wave burst detection problem. *Phys. Rev. D*, 95:104046, May 2017.

[99] I. Mandel et al. Parameter estimation on compact binary coalescences with abruptly terminating gravitational waveforms. *Classical and Quantum Gravity*, 31(15):155005, 2014.

[100] E. Mathieu. Mémoire sur le mouvement vibratoire d'une membrane de forme elliptique. *Journal de Mathématiques Pures et Appliquées*, 13:137–203, 1868.

[101] A. McQuillan, T. Mazeh, and S. Aigrain. Stellar Rotation Periods of the Kepler Objects of Interest: A Dearth of Close-in Planets around Fast Rotators. *The Astrophysical Journal Lett.*, 775:L11, Sept. 2013.

[102] S. Meibom and R. D. Mathieu. A Robust Measure of Tidal Circularization in Coeval Binary Populations: The Solar-Type Spectroscopic Binary Population in the Open Cluster M35. *The Astrophysical Journal*, 620:970–983, Feb. 2005.

[103] C. Messenger et al. Source redshifts from gravitational-wave observations of binary neutron star mergers. *Phys. Rev. X*, 4:041004, Oct. 2014.

[104] C. Messenger and J. Read. Measuring a cosmological distance-redshift relationship using only gravitational wave observations of binary neutron star coalescences. *Phys. Rev. Lett.*, 108:091101, Feb. 2012.

[105] A. Mortier et al. New and updated stellar parameters for 90 transit hosts. The effect of the surface gravity. *Astronomy and Astrophysics*, 558:A106, Oct. 2013.

[106] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 231(694-706):289–337, 1933.

[107] G. I. Ogilvie. Tidal Dissipation in Stars and Giant Planets. *Annual Review of Astronomy and Astrophysics*, 52:171–210, Aug. 2014.

[108] G. I. Ogilvie and D. N. C. Lin. Tidal Dissipation in Rotating Solar-Type Stars. *The Astrophysical Journal*, 661:1180–1191, June 2007.

[109] R. O'Shaughnessy. Data-driven methods to explore a large space of computationally costly compact binary progenitor models. *Phys. Rev. D*, 88:084061, Oct. 2013.

[110] R. S. Park et al. Precession of mercury's perihelion from ranging to the messenger spacecraft. *The Astronomical Journal*, 153(3):121, 2017.

[111] K. C. Patra et al. The apparently decaying orbit of wasp-12b. *The Astronomical Journal*, 154(1):4, 2017.

[112] K. Penev and D. Sasselov. Tidal evolution of close-in extrasolar planets: High stellar q from new theoretical models. *The Astrophysical Journal*, 731(1):67, 2011.

[113] P. C. Peters and J. Mathews. Gravitational radiation from point masses in a keplerian orbit. *Phys. Rev.*, 131:435–440, July 1963.

[114] J. Powell et al. Classification methods for noise transients in advanced gravitational-wave detectors. *Classical and Quantum Gravity*, 32(21):215012, 2015.

[115] J. Powell et al. Classification methods for noise transients in advanced gravitational-wave detectors ii: performance tests on advanced ligo data. *Classical and Quantum Gravity*, 34(3):034002, 2017.

[116] S. Privitera et al. Improving the sensitivity of a search for coalescing binary black holes with nonprecessing spins in gravitational wave data. *Phys. Rev. D*, 89:024003, Jan. 2014.

[117] M. Rakhmanov. Response of ligo to gravitational waves at high frequencies and in the vicinity of the fsr (37.5 khz). Apr. 2005.

[118] M. Rakhmanov. On the round-trip time for a photon propagating in the field of a plane gravitational wave. *Classical and Quantum Gravity*, 26(15):155010, 2009.

[119] M. Rakhmanov, J. D. Romano, and J. T. Whelan. High-frequency corrections to the detector response and their effect on searches for gravitational waves. *Classical and Quantum Gravity*, 25(18):184017, 2008.

[120] J. S. Read et al. Measuring the neutron star equation of state with gravitational wave observations. *Phys. Rev. D*, 79(12):124033, June 2009.

[121] J. S. Read et al. Matter effects on binary neutron star waveforms. *Phys. Rev. D*, 88(4):044042, Aug. 2013.

[122] A. Reisenegger and P. Goldreich. Excitation of neutron star normal modes during binary inspiral. *The Astrophysical Journal*, 426:688–691, May 1994.

[123] R. A. Remillard and J. E. McClintock. X-ray properties of black-hole binaries. *Annual Review of Astronomy and Astrophysics*, 44:49 – 92, 2006.

[124] C. L. Rodriguez et al. Binary black hole mergers from globular clusters: Implications for advanced ligo. *Phys. Rev. Lett.*, 115:051101, July 2015.

[125] K. Schwarzschild. On the gravitational field of a mass point according to Einstein's theory. *ArXiv Physics e-prints*, May 1999.

[126] L. P. Singer et al. The first two years of electromagnetic follow-up with advanced ligo and virgo. *The Astrophysical Journal*, 795(2):105, 2014.

[127] J. Skilling. Nested sampling for general bayesian computation. *Bayesian Anal.*, 1(4):833–859, Dec. 2006.

[128] J. Slutsky et al. Methods for reducing false alarms in searches for compact binary coalescences in ligo data. *Classical and Quantum Gravity*, 27(16):165023, 2010.

[129] A. M. S. Smith et al. WASP-36b: A New Transiting Planet around a Metal-poor G-dwarf, and an Investigation into Analyses Based on a Single Transit Light Curve. *Astronomical Journal*, 143:81, Apr. 2012.

[130] J. R. Smith et al. A hierarchical method for vetoing noise transients in gravitational-wave detectors. *Classical and Quantum Gravity*, 28(23):235005, 2011.

[131] C. Staquet and J. Sommeria. Internal gravity waves: from instabilities to turbulence. *Annual Review of Fluid Mechanics*, 34:559–593, 2002.

[132] J. Steinhoff, T. Hinderer, A. Buonanno, and A. Taracchini. Dynamical tides in general relativity: Effective action and effective-one-body hamiltonian. *Phys. Rev. D*, 94:104028, Nov. 2016.

[133] S. Strogatz. *Nonlinear Dynamics And Chaos.* Sarat Book House, 2007.

[134] S. R. Taylor and J. R. Gair. Cosmology with the lights off: Standard sirens in the einstein telescope era. *Phys. Rev. D*, 86:023502, July 2012.

[135] S. R. Taylor, J. R. Gair, and I. Mandel. Cosmology using advanced gravitational-wave detectors alone. *Phys. Rev. D*, 85:023535, Jan. 2012.

[136] S. Teitler and A. Königl. Why is there a Dearth of Close-in Planets around Fast-rotating Stars? *The Astrophysical Journal*, 786:139, May 2014.

[137] C. Terquem et al. On the Tidal Interaction of a Solar-Type Star with an Orbiting Companion: Excitation of g-Mode Oscillation and Orbital Evolution. *The Astrophysical Journal*, 502:788–801, Aug. 1998.

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

[139] The LIGO Scientific Collaboration. Exploring the sensitivity of next generation gravitational wave detectors. *Classical and Quantum Gravity*, 34(4):044001, 2017.

[140] The LIGO Scientific Collaboration. Gracedb. https://gracedb.ligo.org/, 2017.

[141] The LIGO Scientific Collaboration. Lvalert. https://git.ligo.org/lscsoft/lvalert, 2017.

[142] The LIGO Scientific Collaboration and The Virgo Collaboration. Astrophysical implications of the binary black-hole merger gw150914. *The Astrophysical Journal Lett.*, 818(2):L22, 2016.

[143] The LIGO Scientific Collaboration and The Virgo Collaboration. Characterization of transient noise in advanced ligo relevant to gravitational wave signal gw150914. *Classical and Quantum Gravity*, 33(13):134001, 2016.

[144] The LIGO Scientific Collaboration and The Virgo Collaboration. Gw150914: First results from the search for binary black hole coalescence with advanced ligo. *Phys. Rev. D*, 93:122003, June 2016.

[145] The LIGO Scientific Collaboration and The Virgo Collaboration. Gw151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence. *Phys. Rev. Lett.*, 116:241103, June 2016.

[146] The LIGO Scientific Collaboration and The Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb. 2016.

[147] The LIGO Scientific Collaboration and The Virgo Collaboration. The rate of binary black hole mergers inferred from advanced ligo observations surrounding gw150914. *The Astrophysical Journal Lett.*, 833(1):L1, 2016.

[148] The LIGO Scientific Collaboration and The Virgo Collaboration. Tests of general relativity with gw150914. *Phys. Rev. Lett.*, 116:221101, May 2016.

[149] The LIGO Scientific Collaboration and The Virgo Collaboration. Upper limits on the rates of binary neutron star and neutron star–black hole mergers from advanced ligoâĂŹs first observing run. *The Astrophysical Journal Lett.*, 832(2):L21, 2016.

[150] The LIGO Scientific Collaboration and The Virgo Collaboration. All-sky search for short gravitational-wave bursts in the first advanced ligo run. *Phys. Rev. D*, 95:042003, Feb. 2017.

[151] The LIGO Scientific Collaboration and The Virgo Collaboration. Gw170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2. *Phys. Rev. Lett.*, 118:221101, June 2017.

[152] The LIGO Scientific Collaboration, The Virgo Collaboration, The Australian Square Kilometer Array Pathfinder (ASKAP) Collaboration, The BOOTES Collaboration, The Dark Energy Survey and the Dark Energy Camera GW-EM Collaborations, The Fermi GBM Collaboration, The Fermi LAT Collaboration, The INTEGRAL Collaboration, The J-GEM Collaboration, The Liverpool Telescope Collaboration, The MAXI Collaboration, The Pan-STARRS Collaboration, T. The Pi of the Sky Collaboration, The SkyMapper Collaboration, The Swift Collaboration, C. The TAROT, Zadko, Algerian National Observatory, The TOROS Collaboration, and The VISTA Collaboration. Localization and broadband follow-up of the gravitational-wave transient gw150914. *The Astrophysical Journal Lett.*, 826(1):L13, 2016.

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

[154] S. Udry and N. C. Santos. Statistical properties of exoplanets. *Annual Review of Astronomy and Astrophysics*, 45(1):397–439, 2007.

[155] S. A. Usman et al. The pycbc search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33(21):215004, 2016.

[156] L. Vaínshteín and V. Zúbakov. *Extraction of Signals from Noise*. Dover books on physics and mathematical physics. Dover Publ., Incorporated, 1970.

[157] M. Vallisneri. Use and abuse of the fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Phys. Rev. D*, 77:042001, Feb. 2008.

[158] M. Vallisneri et al. The ligo open science center. *Journal of Physics: Conference Series*, 610(1):012021, 2015.

[159] M. Vallisneri and N. Yunes. Stealth bias in gravitational-wave parameter estimation. *Phys. Rev. D*, 87:102002, May 2013.

[160] F. Valsecchi and F. A. Rasio. Planets on the Edge. *The Astrophysical Journal Lett.*, 787:L9, May 2014.

[161] T. Van Hoolst. Nonlinear, nonradial, isentropic oscillations of stars: Hamiltonian formalism. *Astronomy and Astrophysics*, 286:879–889, June 1994.

[162] J. Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library. *Phys. Rev. D*, 91:042003, Feb. 2015.

[163] J. Veitch and A. Vecchio. Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. *Phys. Rev. D*, 81:062003, Mar. 2010.

[164] T. Venumadhav, A. Zimmerman, and C. M. Hirata. The stability of tidally deformed neutron stars to three- and four-mode coupling. *The Astrophysical Journal*, 781(1):23, 2014.

[165] S. Vitale et al. On similarity of binary black hole gravitational-wave skymaps: to observe or to wait? *Monthly Notices of the Royal Astronomical Society Lett.*, 466(1):L78, 2017.

[166] R. Wald. *General Relativity.* University of Chicago Press, 1984.

[167] C. A. Watson and T. R. Marsh. Orbital period variations of hot Jupiters caused by the Applegate effect. *Monthly Notices of the Royal Astronomical Society*, 405:2037–2043, July 2010.

[168] N. N. Weinberg. Growth rate of the tidal p-mode g-mode instability in coalescing binary neutron stars. *The Astrophysical Journal*, 819(2):109, 2016.

[169] N. N. Weinberg, P. Arras, and J. Burkart. An instability due to the nonlinear coupling of p-modes to g-modes: Implications for coalescing neutron star binaries. *The Astrophysical Journal*, 769(2):121, 2013.

[170] N. N. Weinberg et al. Nonlinear tides in close binary systems. *The Astrophysical Journal*, 751(2):136, 2012.

[171] J. M. Weisberg and J. H. Taylor. The Relativistic Binary Pulsar B1913+16: Thirty Years of Observations and Analysis. In F. A. Rasio and I. H. Stairs, editors, *Binary Radio Pulsars*, volume 328 of *Astronomical Society of the Pacific Conference Series*, page 25, July 2005.

[172] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications.* Principles of Electrical Engineering Series. Massachusetts institute of technology, 1964.

[173] C. M. Will. The confrontation between general relativity and experiment. *Living Reviews in Relativity*, 9(1):3, 2006.

[174] K. Yagi and N. Yunes. Love number can be hard to measure. *Phys. Rev. D*, 89(2):021303, Jan. 2014.

[175] H. Yu and N. N. Weinberg. Resonant tidal excitation of superfluid neutron stars in coalescing binaries. *Monthly Notices of the Royal Astronomical Society*, Oct. 2016.

[176] H. Yu and N. N. Weinberg. Dynamical tides in coalescing superfluid neutron star binaries with hyperon cores and their detectability with third generation gravitational-wave detectors. *Monthly Notices of the Royal Astronomical Society*, May 2017.

[177] M. Zevin et al. Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science. *Classical and Quantum Gravity*, 34(6):064003, 2017.