BINARY BLACK HOLE ASTROPHYSICS WITH GRAVITATIONAL WAVES

by

BRUCE EDELMAN

A DISSERTATION

Presented to the Department of Physics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2023

DISSERTATION APPROVAL PAGE

Student: Bruce Edelman

Title: Binary Black Hole Astrophysics with Gravitational Waves

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

Prof. Ray Frey	Chairperson
Prof. Ben Farr	Advisor
Prof. Tien-Tien Yu	Core Member
Prof. Carol Paty	Institutional Representative

and

Krista Chronister	Vice Provost of Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2023

DISSERTATION ABSTRACT

Bruce Edelman

Doctor of Philosophy

Department of Physics

June 2023

Title: Binary Black Hole Astrophysics with Gravitational Waves

Gravitational Waves (GWs) have quickly emerged as powerful, indispensable tools for studying gravity in the strong field regime and high-energy astrophysical phenomena since they were first directly detected by the Laser Interferometer Gravitational-Wave Observatory (LIGO) on September 14, 2015. Over the course of this dissertation work gravitational-wave astronomy has begun to mature, going from 11 GW observations when I began to 90 at the time of writing, just before the next observing run begins. As the network of GW observatories continues to grow and these observations become a regular occurrence, the entire population of merging compact objects observed with GWs will provide a unique probe of the astrophysics of their formation and evolution along with the cosmic expansion of the universe. In this dissertation I present four studies that I have led using GWs to better understand the astrophysics of the currently most detected GW source, binary black holes (BBHs). We first present a novel data-driven technique to look for deviations from modeled gravitational waveforms in the data, coherent across the network of observatories, along with an analysis of the first gravitational-wave transient catalog (GWTC-1). The following three studies present the three different approaches to modeling populations of BBHs, using parametric, semi-parametric and non-parametric models. The first of these studies uses a parametric model that imposes a gap in the mass distribution of black holes, looking for evidence of effects caused by pair-instabiliy supernovae. The second study introduces a semi-parametric model that aims to take advantage of the benefits of both parametric and non-parametric methods, by imposing a flexible perturbation to an underlying simpler parametric description. This study was among the first data-driven studies revealing possible structure in the mass distribution of BBHs using GWTC-2, namely an additional peak at $10M_{\odot}$. The final study introduces a novel non-parametric model for hierarchically inferring population properties of GW sources, and performs the most comprehensive data-driven study of the BBH population to date. This study is also the first that uses non-parametric models to simultaneously infer the distributions of BBH masses, spins and redshifts. This dissertation contains previously published and unpublished material.

CURRICULUM VITAE

NAME OF AUTHOR: Bruce Edelman

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
University of Oregon, Eugene, Oregon
Purdue University, West Lafayette, Indiana

DEGREES AWARDED:
Doctor of Philosophy, Physics, 2023, University of Oregon
Master of Science, Physics, 2020, University of Oregon
Bachelor of Science, Physics, 2017, Purdue University

AREAS OF SPECIAL INTEREST:
Gravitational Waves, Compact Object Astrophysics

PROFESSIONAL EXPERIENCE:

Research Assistant, University of Oregon 2018-Present

Teaching Assistant, University of Oregon 2017-2019

Undergraduate Teaching Assistant, Purdue University, 2015-2017

PUBLICATIONS:

J. Godfrey, **B. Edelman**, B. Farr (2023). Cosmic Cousins: Identification of a Subpopulation of Binary Black Holes Consistent with Isolated Binary Evolution. Submitted to *Astrophys. J.*. 2304.01288.

A. Farah, **B. Edelman**, M. Zevin, M. Fishbach, J. Ezquiga, D. Holz, B. Farr (2023). Things that might go bump in the night: Assessing structure in the binary black hole mass spectrum. Submitted to *Astrophys. J.*. 2301.00834.

R. Abbott, et al. incl. **B. Edelman** (2023). The population of merging compact binaries inferred using gravitational waves through GWTC-3. Submitted to *Phys. Rev. X*, **13**, 011048 (2023). *Contributions*: Served on the editorial and review teams and led analyses inferring the BBH mass distribution.

B. Edelmam, B. Farr, Z. Doctor (2023). Cover Your Basis: Comprehensive Data-driven Characterization of the Binary Black Hole Population. *Astrophys. J.*, **946** 16 (2023).

B. Edelman. Z. Doctor, J. Godfrey, B. Farr (2022). Ain't No Mountain High Enough: Semi-Parametric Modeling of LIGO/Virgo's Binary Black Hole Mass Distribution. *Astrophys. J.*, **924**, 101 (2022).

J.D. Merritt, B. Farr, R. Hur, **B. Edelman**, Z. Doctor (2021). Transient glitch mitigation in Advanced LIGO data with *glitschen*. *Phys. Rev. D*, **104**, 102004 (2021).

B. Edelman, Z. Doctor, B. Farr (2021). Poking Holes: Looking for Gaps in LIGO/Virgo's Black Hole Population. *Astrophys. J. Lett.*, **913**, L23 (2021).

R. Abbott, et al. incl. **B. Edelman** (2021). Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. *Astrophys. J Lett.*, **913**, L7 (2021). *Contributions*: Event characterization, and comments on the draft.

M. Fishbach, Z. Doctor, T. Callister, **B. Edelman**, et al (2021). When are LIGO/Virgo's Big Black Hole Mergers? *Astrophys. J.*, **912**, 98 (2021).

R. Abbott, et al. incl. **B. Edelman** (2021). GWTC-2: Compact Binary Coalescences Overserved by LIGO and Virgo during the First Half of the Third Observing Run. *Phys. Rev. X*, **11**, 021053 (2021). *Contributions*: Event characterization.

B. Edelman, F. Rivera-Paleo, J.D. Merritt, B. Farr, et al (2021). Constraining Unmodeled Physics with Compact Binary Mergers from GWTC-1. *Phys. Rev. D*, **103**, 042004 (2021).

I. Romero-Shaw, C. Talbot, S. Biscoveanu, et al. incl. **B. Edelman** (2020). Bayesian Inference for Compact Binary Coalescences with bilby: Validation and Application to the First LIGO–Virgo Gravitational-Wave Transient Catalogue. *MNRAS*, **499**, 3295–3319 (2020).

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the numerous pillars of support I had surrounding me throughout this journey. First and foremost I want to thank my best friend, partner, and wife, Lydia, for the endless support through this long challenging journey. I also want to thank my family, namely my parents and my brothers, who have always my biggest supporters.

The endless, unwavering support and confidence from my advisor, Professor Ben Farr, has consistently provided a beacon of light to follow, making a hard journey feel easy. I could not imagine a better advisor, person to work with or a scientist to look up to. I also want to thank my close mentor and collaborator Professor Zoheyr Doctor for encouraging me in my research goals and aspirations along with being genuinely one of the nicest people I have ever met. I would not have evolved into the scientist I am today with the skills and confidence to do independent research without these two monumentally helpful mentors there by my side.

Throughout the course of my PhD, I have been very lucky to work with and collaborate with many wonderful people. A few people I am especially grateful to have crossed paths with include (but not limited to): Amanda Farah, Jaxen Godfrey, JD Merritt, Gino Carrillo, Sangeet Paul, Tom Callister, Mike Zevin, Adrian Helmling-Cornell, Phillipe Nguyen, Matthew Ball, Rachel Hur and Lexi Vives. I also want to thank my thesis defense committee: Professor Ray Frey, Professor Tien-tien Yu, and Professor Carol Paty. I am also incredibly lucky to have developed great collaboration among the wonderful members of the LIGO scientific collaboration across the globe. The science that the LIGO collaboration puts out takes the effort of so many people it is impossible to thank everyone but this dissertation would not have been possible without the thousands of scientists in our collaboration.

Finally, I want to thank my dog, Cooper, for always keeping me happy and company during late nights. I wish you could have been here to see me finish this journey.

Dedicated to my father, who has constantly done everything possible to give me a path to success.

This dissertation is the product of unionized labor as part of the Graduate Teaching Fellows Federation, AFT Local 3544.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1. Theory of Gravitational Waves

1.1.1. General Relativity

Einstein's theory of general relativity has been immensely successful in explaining astrophysical phenomena throughout the century following the ground-breaking work, culminating with the verification of a robust prediction unique to this new view of gravity with the successful direct detection of gravitational waves (GWs) in 2015. This theory describes the nature of gravity not as a fundamental force as the previously held Newtonian view, but arising from the curved geometry of spacetime. Specifically it states that mass or energy in the universe distorts or warps spacetime such that the geometry of spacetime is not flat, but instead has a curvature. The result of this curvature looks like a gravitational force in flat space at certain scales but in this view there are only bodies freely moving in "straight lines" (or geodesics) through a curved 4-dimensional manifold we call spacetime. The entire theory can be described through Einstein's field equations, which are 10 coupled, non-linear, partial differential equations, shown below in its succinct form.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.1}$$

Where the right-hand side of this equation describes the distribution of mass and energy in the universe with the stress-energy tensor, $T_{\mu\nu}$. The left-hand side describes the geometry of spacetime with the metric tensor, $g_{\mu\nu}$. The first term on the left-hand side is the Einstein tensor, $G_{\mu\nu}$, which is a function of the metric tensor and its derivatives. The second term is the cosmological constant, Λ, which is a constant of proportionality that describes the energy density of the vacuum. To build up the nicely packaged Einstein tensor, let's start with the line element of 4-dimensional spacetime, $ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$. Through the principle of least action one can derive equations of motion for a particle in such a geometry governed by $g_{\mu\nu}$, which we call the geodesic equation, defined as:

$$\frac{d^2x^{\mu}}{ds^2} + \Gamma^{\mu}_{\alpha\beta}\frac{dx^{\alpha}}{ds}\frac{dx^{\beta}}{ds} = 0 \tag{1.2}$$

Above we have introduced the Christoffel symbols, $\Gamma^{\mu}_{\alpha\beta}$, which you can get from the metric tensor with:

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}g^{\mu\nu}\left(\frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} + \frac{\partial g_{\nu\alpha}}{\partial x^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\nu}}\right) \tag{1.3}$$

The next component we need is the Riemann curvature tensor, $R_{\rho\mu\sigma\nu}$, which encapsulates the curvature of spacetime, and is defined as:

$$R_{\rho\mu\sigma\nu} = g_{\rho\lambda}\left(\frac{\partial\Gamma^{\lambda}_{\nu\sigma}}{\partial x^{\mu}} - \frac{\partial\Gamma^{\lambda}_{\mu\sigma}}{\partial x^{\nu}} + \Gamma^{\lambda}_{\mu\eta}\Gamma^{\eta}_{\nu\sigma} - \Gamma^{\lambda}_{\nu\eta}\Gamma^{\eta}_{\mu\sigma}\right) \tag{1.4}$$

The Einstein tensor has two free indices which means we need to contract two indices of the Riemann tensor to get the Einstein tensor. This is done in two parts: first by constructing the Ricci tensor, $R_{\mu\nu} = g^{\rho\sigma}R_{\rho\mu\sigma\nu}$, followed with the Ricci scalar, $R = g^{\mu\nu}R_{\mu\nu}$. With these pieces in place we have the definition of the Einstein tensor, $G_{\mu\nu}$, or the left-hand side of the Einstein field equations (eq. 1.1.1).

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{1.5}$$

Now that we have built up the ingredients of Einstein's field equations, the heart of General Relativity, the hard part begins: actually solving them. This is an especially difficult task because they are non-linear, coupled, partial differential equations. There are only a few known exact solutions, all of which have to make simplifying assumptions and thus are not very realistic in most situations, especially those of two merging black holes which this thesis is focused on. This is one reason why we often need to resort to numerical solutions to solve the equations for most scenarios, which is what the field of numerical relativity focuses on. The next section focuses on one method of simplifying the equations to make them more tractable in the weak field limit, namely linearizing them.

1.1.2. Linearized Gravity

To illustrate the propagating wave solutions (gravitational waves) of Einstein's equations, we will linearize the equations in the weak field limit. We start by describing the metric as the flat Minkowski metric, $\eta_{\mu\nu}$ with some small perturbation $h_{\mu\nu}$ around it with $|h_{\mu\nu}| << 1$, which we can write as:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1.6}$$

Next we can derive the linearized Ricci tensor and Ricci scalar by substituting eq. 1.1.2 into the procedure laid out in the previous section, while neglecting terms that are higher than first order in $h_{\mu\nu}$, which yields:

$$R_{\mu\nu} = \frac{1}{2}\left(\frac{\partial\partial h^{\alpha}_{\nu}}{\partial x^{\alpha}\partial x^{\mu}} + \frac{\partial\partial h^{\alpha}_{\mu}}{\partial x^{\alpha}\partial x^{\nu}} - \frac{\partial\partial h}{\partial x^{\mu}\partial x^{\nu}} - \Box h_{\mu\nu} + \mathcal{O}(h^2)\right) \tag{1.7}$$

$$R = \eta_{\mu\nu}R^{\mu\nu} = \frac{\partial\partial h}{\partial x^{\mu}\partial x^{\nu}} - \Box h_{\mu\nu} + \mathcal{O}(h^2) \tag{1.8}$$

Where above $h := \eta^{\mu\nu}h_{\mu\nu}$ is the trace of the metric perturbation and $\Box := \eta^{\mu\nu}\partial_{\mu}\partial_{\nu}$ is the d'Alembertian operator in Minkowski spacetime. Next we make this transformation to the trace-reversed metric perturbation: $\bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$. We can now write the linearized Einstein tensor as:

$$G_{\mu\nu} = \frac{1}{2}\left(\frac{\partial\partial\bar{h}^{\rho}_{\nu}}{\partial x^{\mu}\partial x^{\sigma}} + \frac{\partial\partial\bar{h}^{\rho}_{\mu}}{\partial x^{\nu}\partial x^{\sigma} - \Box\bar{h}_{\sigma\rho} - \eta_{\mu\nu}\frac{\partial\partial\bar{h}^{\sigma\rho}}{\partial x^{\sigma}\partial x^{\rho}}}\right) + \mathcal{O}(h^2) \tag{1.9}$$

With the appropriate choice of gauge transformation, i.e. one that satisfies two conditions: $\partial^{\mu}\bar{h}_{\mu\nu} = 0$ and $\bar{h} := \eta^{\mu\nu}\bar{h}_{\mu\nu} = 0$, we arrive at the transverse-traceless gauge where all terms without a $\Box$ operator in eq. 1.1.2 vanish such that the linearized Einstein's field equations simplify to:

$$-\Box\bar{h}_{\mu\nu} + \mathcal{O}(h^2) = \frac{16\pi G}{c^4}T_{\mu\nu} \tag{1.10}$$

When considering vacuum solutions to the linearized Einstein field equations, i.e. $T_{\mu\nu} = 0$, we arrive at a familiar simple wave equation: $\Box \bar{h}_{\mu\nu} = 0$, with gravitational waves as solutions. Now consider a distribution of matter that is slowly varying. The gravitational radiation far away from the source described by eq. 1.1.2 is found to only depend on the second time derivative of the quadrupole moment of the source matter distribution, Q^{ij}, along with the luminosity distance to the source, d_{L}:

$$h^{ij} = \frac{2G}{c^4 d_{\mathrm{L}}} \frac{d^2}{dt^2} Q^{ij} \tag{1.11}$$

This equation is implying that any distribution of accelerating matter that is spherically asymmetric (i.e. $Q^{ij} \neq 0$) will radiate gravitational waves, although the second time derivatives of the quadrupole moments need to be very large to produce GWs with non-neglible magnitudes, considering $\frac{G}{c^4} \sim 10^{-20}\,\mathrm{Mpc}/(M_{\odot} c^2)$. In the next section we discuss the indirect and direct detections of GWs and their sources.

1.2. Gravitational-Wave Detections

Gravitational Wave solutions to Einstein's field equations have been known for many years before any physical evidence for their existence was found. In 1974 Hulse and Taylor discovered a neutron star - pulsar binary, PSR 1913 + 16, with a decreasing orbital period that aligned with general relativity's predicted energy loss due to emitted gravitational radiation. While this indirect detection was strikingly strong, later winning the Nobel Prize in physics in 1993, it would take more time to be able to construct instruments sensitive to directly detect these miniscule ripples in spacetime. As mentioned before in 2015 this was finally achieved by the Laser Interferometer Gravitational-Wave Observatory (LIGO) with the detection of GWs emitted in the last few seconds of the energetic merger of a binary black hole (BBH) system with each black hole around 30 times the mass of the sun. This gravitational wave detection (and all others) are named by the date they were detected, in this case GW150914 – September 14th, 2015. Less than two years later in LIGO's second observing run (now with Virgo added), the first binary neutron star (BNS) merger was detected, GW170817. This happened to be coincident with a gamma-ray burst (GRB) – and through both this and the GWs, an approximate location for the source was produced such that many telescopes were able to observe the kilonova that was predicted to follow BNS mergers, marking the beginning of GW astronomy and multi-messenger astronomy with GWs. While there are other types of GW sources that can produce a wide spectrum of GWs across many frequencies, ground-based detectors (e.g. LIGO, Virgo) are most sensitive to these compact binary mergers. To date only GWs from compact binary mergers have been detected, including binary neutron stars (BNSs), binary black holes (BBHs), and neutron star black hole binaries (NSBHs), with the most recent count being 90 GW detections (2 BNSs, 3-6 NSBHs, 82-85 BBHs) [1].

GW detectors like LIGO are in principle very large Michelson Interferometers, which sends laser light down two perpendicular arms (4km long for LIGO) and back in order to measure minute fluctuations in the distance of each as GWs pass by, from the interference signal from each arms' beams. We quantify the effect of GWs on the detector by calculating the fractional change in arm lengths of the interferometer, or strain, yielding the strain time-series

$h(t)$ sampled at a rate of 16kHz. Instruments this sensitive are bound to have lots of noise sources, but these can be mitigated with appropriate modeling and environmental monitoring. With multiple detectors spaced geographically far apart, we can impose coincidence across the network to reduce the false alarm rate of detections, while also gaining directional information from the time delay between detectors allowing for better sky localization of sources.

The first step of analyzing this swath of data is to use algorithms that search the data for possible GW signals. This is most commonly done with matched filtering procedures where we store a large bank of simulated gravitational waveforms from a variety of sources and parameters, which are then compared against the data to see if there is a match. In practice these methods or search pipelines are much more complex, but the primary focus of this thesis is on the next step in the data-analysis procedure after the pipelines have identified a stretch of data that likely contains a GW signal. Since solving Einstein's equation is challenging even in the simplest cases, constructing accurate models of GWs from compact binary mergers is another active field of study within GR. These models are constructed either by phenomenological approaches (i.e. [2, 3]), effective one body formalism (i.e. [4]), fits to expensive numerical relativity simulations (i.e. [5]), or combinations of such methods, and each come with their own simplifying assumptions and domains of applicability. The study presented in Chapter 2 explores novel methods that can be used to validate such waveform models while looking for signs in un-modeled physics within GW data. With waveform models from compact binary mergers, we can then use Bayesian inference and forward modeling to estimate the parameters of the source that produced the GWs, which is discussed in detail in the next section.

1.3. Bayesian Inference and Parameter Estimation

The goal of Bayesian inference or parameter estimation is to infer the probability distribution of a set of model parameters, θ, conditioned on the data of interest, d, and the assumed model or hypothesis, $\mathcal{H}$ (i.e. the waveform model in the GW context). This is called the posterior distribution and is given by Bayes' theorem:

$$p(\theta|d, \mathcal{H}) = \frac{\mathcal{L}(d|\theta, \mathcal{H})\pi(\theta|H)}{\mathcal{Z}(d|H\mathcal{H})} \tag{1.12}$$

Where $\mathcal{L}(d|\theta, \mathcal{H})$ is the likelihood of the data given the model and parameters, $\pi(\theta|H)$ is the prior distribution of the parameters, and $\mathcal{Z}(d|H\mathcal{H})$ is the evidence or marginal likelihood. This quantity is a factor that ensures the normalization of the posterior, found by integrating the numerator of eq. 1.3 over parameters θ:

$$\mathcal{Z}(d|H\mathcal{H}) := \int d\theta \mathcal{L}(d|\theta, \mathcal{H})\pi(\theta|H) \tag{1.13}$$

While the evidence is often ignored since it is just a normalizing constant of the posterior and does not depend on the parameters of interest, note that it does depend on the model $\mathcal{H}$, and is often used for Bayesian model selection, which is used later in chapters 3 and 4. Oftentimes the desired quantity, the posterior distribution, is highly dimensional, leaving one forced to sample the distribution using Markov Chain Monte Carlo (MCMC) methods. The works in this thesis explore the use of different MCMC methods that are each well suited in different scenarios including, but not limited to, parallel tempering, nested sampling, and Hamiltonian Monte Carlo.

For the likelihood in GW data, we make the assumption that the noise is stationary and Gaussian, which is a valid approximation for the short stretches of data that contain the signal we wish to analyze. Now let d^i_j be the frequency domain strain data in interferometer, i, at frequency bin j. Additionally, each interferometer's noise can be represented by its Power Spectral Density (PSD), which can be thought of as the power of the noise in each frequency bin, denoted as P^i_j. Finally, as before $\mathcal{H}$ represents our assumed physical waveform model, with $h^i_j(\theta)$ being waveform model evaluated with parameters, θ, at frequency bin j projected into interferometer i. With the assumption that the data at each frequency bin in each interferometer is independent of each other, the total log-likelihood is then calculated by summing over the N_{ifos} interferometers (indexed by i) and N_{bin} frequency bins (indexed by j):

$$\log \mathcal{L}(\{\mathbf{d}\}|\theta, \mathcal{H}) = -\sum_{i=0}^{N_{\text{ifos}}} \sum_{j=0}^{N_{\text{bin}}} \left[2\pi P^i_j + \Delta f \frac{|d^i_j - h^i_j(\theta)|^2}{P^i_j} \right] \tag{1.14}$$

With this formalism in place we are now able to use various MCMC methods to infer the posterior distribution over the set of GW source parameters, θ, given the observed strain data from our network of interferometers. This is usually a challenging task as the full parameter space for a BBH is 15-dimensional, with highly structured posteriors. While chapter 2 focuses on the validation of waveform models through parameter estimation, the following chapters focus on inferring the astrophysical distribution of BBHs using Hierarchical Bayesian

Inference. The method for extending this formalism to hierarchical models are discussed in the additional analysis details of chapters 3-5.

1.4. Polynomial Splines

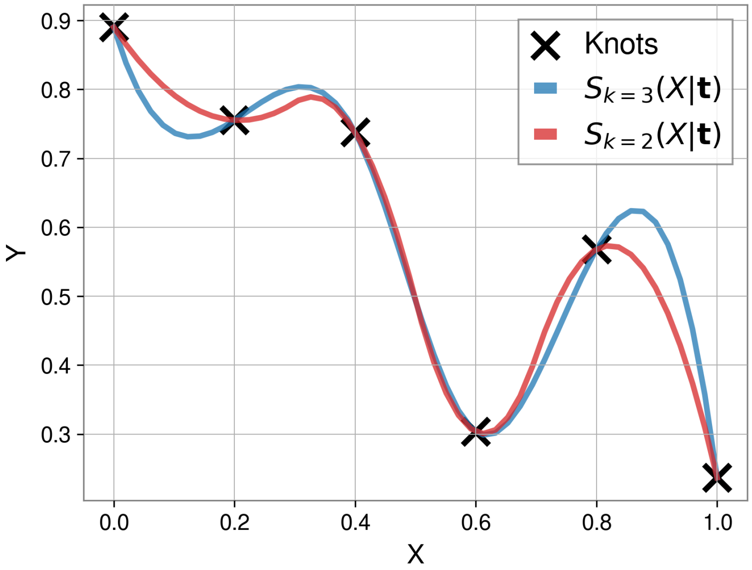

FIGURE 1.1. Spline interpolation from 5 knots or control points with a cubic (blue) and quadratic (orange) spline function interpolated from the same knots.

A powerful tool in statistical modeling that we make use of in 3 out of the next 4 chapters is polynomial splines. A polynomial spline is a function of piece wise polynomials of a given order that are stitched together at control points or knots, which are often used for interpolation [6]. To construct a spline function of order k with n degrees of freedom across the interval $[a, b]$, we first define a knot vector, $\mathbf{t}$, consisting of $n + k + 1$ knots in increasing order where $t_0 = a$ and $t_{n+k} = b$. Given these specifications the resulting spline function, $S_k(t|\mathbf{t})$ consists of $n - 1$ polynomial sections, $P_{n-1}(t)$, $t_{n-1} \leq t \leq t_n$. There is an additional condition that the derivatives of both polynomials around each knot must match, for all derivatives up to the $k-1$ order. Figure 1.1 shows a simple example with a cubic spline function ($k=3$) interpolated from 5 control points. It can further be shown that the set of spline functions of order k with n degrees of freedom and knot vector, $\mathbf{t}$ (i.e. $S_k(t|\mathbf{t}, n)$), forms a vector space, which can be decomposed into different basis sets that span this space, which we call basis splines [6, 7]. Basis splines are discussed further and put to use later in this dissertation in chapter 5.

1.5. Outline

The remainder of this thesis consists of selected published works that I have led studying different aspects of the astrophysics of binary black holes with gravitational waves. Each of the following chapters consists of these publications along with additional brief summaries of each publication to put the work into the broader context of the field. The included publications and their chapters are:

- **Chapter 2:** *Constraining Unmodeled Physics with Compact Binary Mergers from GWTC-1*
- **Chapter 3:** *Poking Holes: Looking for Gaps in LIGO/Virgo's Black Hole Population*
- **Chapter 4:** *Ain't No Mountain High Enough: Semi-Parametric Modeling of LIGO-Virgos Binary Black Hole Mass Distribution*
- **Chapter 5:** *Cover Your Basis: Comprehensive Data-Driven Characterization of the Binary Black Hole Population*

CHAPTER II

CONSTRAINING UNMODELED PHYSICS WITH COMPACT BINARY MERGERS FROM GWTC-1

2.1. Summary

All analyses of gravitational-wave data rely on the foundation of robust and accurate waveform models of the GWs emitted from merging compact binaries. Directly solving Einstein's equations to get this is a monumentally challenging task, leaving the field to either phenomenological or numerical approaches in building waveform models that each have their own simplifying assumptions. In order to validate these models and to look for any impact on the observed signals from possible physical effects neglected in the simplified models, we present a data-driven method to constrain deviations from modeled waveforms that are coherent across the network of detectors. In this chapter we present the study, published as Edelman et al. [8], introducing this novel non-parametric method to validate and constrain deviations from the models used to analyze data from the compact binary coalescences (CBCs) contained in the first gravitational-wave transient catalog, GWTC-1 [9].

2.2. Abstract

We present a flexible model to describe the effects of generic deviations of observed gravitational wave signals from modeled waveforms in the LIGO and Virgo gravitational wave detectors. With the detection of 11 gravitational wave events from the GWTC-1 catalog, we are able to constrain possible deviations from our modeled waveforms. In this paper we present our coherent spline model that describes the deviations, then choose to validate our model on an example phenomenological and astrophysically motivated departure in waveforms based on extreme spontaneous scalarization. We find that the model is capable of recovering the simulated deviations. By performing model comparisons we observe that the spline model effectively describes the simulated departures better than a normal compact binary coalescence (CBC) model. We analyze the entire GWTC-1 catalog of events with our model and compare it to a normal CBC model, finding that there are no significant departures from the modeled template gravitational waveforms used.

2.3. Introduction

General relativity (GR) has passed a multitude of tests over the past years [10], but until the detection of gravitational waves (GWs) from binary black holes[11, 12] it had not been widely tested for strong dynamical gravitational fields. Gravitational-wave astronomy and more specifically that of compact binary coalescences (CBC) gives us access to a genuinely strong gravitational field regime to both test GR [13, 14, 15, 16], and to provide constraints on new physics not predicted by GR modeled waveforms of these systems. Contemporary GW analyses employ matched filtering and forward modeling techniques, which both inherently rely on accurately modeled waveforms [12, 17, 18, 19]. We introduce here a model that can account for and measure in data deviations in phase and amplitude from a modeled waveform, either due to approximations inherent in the waveform calculation or a mismatch between theory (GR) and nature.

We present a parameterization that quantifies the deviations between the observed waveform ($h_{\mathrm{intrinsic}}$) and the GR waveform models (h_{model}), with few assumptions about the deviations. This provides the ability to perform additional tests of GR and also presents a generic model for describing and possibly constraining additional effects in a binary merger like presence of higher-order modes [20] and tidal effects [21]. Quantifying such deviations is one of the major challenges in GW data analysis. The numerical method we use to parameterize deviations here is based on cubic spline interpolation in which the deviations (in phase and amplitude) are modeled as independent cubic spline functions interpolated from node points in the frequency domain. The cubic spline interpolant generates deviations that vary smoothly in frequency, but otherwise does not constrain the type or nature of deviation.

The splines employed to characterize the deviation provide a uniform way of describing GR departures rather than fitting separate parameters to the inspiral, merger, and ringdown (IMR) of the waveform separately as commonly done in GR IMR consistency tests [16, 22, 23]. In addition, IMR consistency tests have the same limitation as matched filtering, as they inherently assume perfect accuracy of the template waveform.

Another common test of GR is the parameterized test where one expands the waveform model in different regimes with post-Newtonian (PN) correction parameters added [24]. This test also builds in assumptions about how possible deviations may occur and has to fit different parts of the waveform with separate models as in the IMR consistency test. One commonly used class of model-agnostic tests of GR are residual [15] tests, where best-fit waveform models are subtracted from the data and normality tests (e.g., Anderson-Darling) are conducted on the residuals. Such tests would be sensitive to very large deviations from the signal model, but not to small but correlated deviations across many frequency bins. Our proposed model differs from these tests and constraints by allowing uncertainty in the template waveforms and letting it vary smoothly across the entire frequency range. Our model is able to describe and fit the inaccuracies in waveform models. With the assumption that the template GR waveform is completely accurate, it provides a clear way of describing and constraining unmodeled physics of our waveform models or departures from GR across the frequency range of the waveform.

In Section 2.4 we describe the model and methods for incorporating it in the `LALInference` Bayesian Analysis software [19]. Our implementation is similar to the calibration spline model described in [25]. We then present simulated deviations on which we validated the performance of our model in Section 2.5, followed by discussions and implications of the results. In Section 2.6 we present results of this model on the entire first LIGO and Virgo Gravitational Wave Transient Catalog (GWTC-1) [9] which includes the results of the first and second observing runs of the Advanced LIGO [26] and Virgo [27] detectors. This catalog of gravitational wave events includes ten binary black hole detections and one binary neutron star detection [9, 28]. Lastly the results and conclusions are summarized and discussed in Section 2.6.1.

2.4. Methods

2.4.1. Waveform Representation

When a gravitational wave enters a gravitational wave detector, the detector records a data stream which we can describe in the frequency domain as $d(f) =$

$h_{\text{observed}}(f) + n(f)$, which is an additive combination of a waveform $h_{\text{observed}}(f)$ and noise $n(f)$. The observed waveform in a detector can be represented as the sum of the intrinsic waveform polarizations, projected across that detector. This is done by multiplying the two (time-dependent) antenna pattern terms of that detector, F_+ and $F_\times$, to the plus and cross gravitational-wave polarizations as:

$$h_{\text{observed}}(f) = h_+(f)F_+(f) + h_\times(f)F_\times(f) \tag{2.1}$$

with

$$h_{+,\times}(f) = h_{\text{model},+,\times}(f)\left[1 + \delta A(f)\right]\exp\left[i\delta\phi(f)\right]. \tag{2.2}$$

Since we are searching for deviations from the coherent modeled waveforms, every detector observing a GW should see the same deviations. We model the uncertainties in the intrinsic waveform, $h_{\text{model},+,\times}(f)$, as frequency-dependent amplitude and phase departures in $h_{+,\times}$ with respect to $h_{\text{model},+,\times}$. This is the same technique that Farr et al. [25] use to model calibration errors in each detector independently. We also take the assumption that there are only two polarizations and that GWs travel at the speed of light. While our model is degenerate with the calibration model, coherent deviations observed across all detectors are modeled with a single set of parameters for this spline model, as opposed to modeling them independently in each detector using the calibration model. Thus we expect our prior distribution to result in deviations to the intrinsic waveform being preferably described by the coherent spline, and the calibration splines to measure only detector-dependent deviations.

We assume that phase deviations are small and under this assumption, we can approach the exponential term as

$$\exp\left(i\delta\phi(f)\right) = \frac{2 + i\delta\phi(f)}{2 - i\delta\phi(f)} + \mathcal{O}(\delta\phi^3) \tag{2.3}$$

which is more computationally efficient [25]. Then, we can rewrite the intrinsic waveform as:

$$h_{+,\times}(f) = h_{\text{model},+,\times}(f)\left[1 + \delta A(f)\right]\frac{2 + i\delta\phi(f)}{2 - i\delta\phi(f)} \tag{2.4}$$

This replacement agrees with the exponential term to third order for small phase deviations, and differs by 5% from the exponential term for the largest simulated deviation used in this paper of 60 degrees.

The intrinsic waveform, after being modified as Eq. 2.4, is then projected across the detectors to get h_{observed} in each detector as in Eq. 2.1. Despite the expectation that these departures are small, they have the potential to impact the measurement of all parameters of the source (masses, spins, distance, etc.) [18]. A consequence of modeling these deviations with a purely phenomenological model is that we can no longer trust the inference of astrophysically modeled parameters when deviations are present.

Under the assumption that $\delta A(f)$ and $\delta\phi(f)$ vary smoothly in frequency, they can be modeled by a spline function [25].

2.4.2. Spline Model

A spline function is a piece-wise polynomial interpolation that obeys smoothness conditions at the nodal points where the pieces connect. In the following, we use the case of cubic splines defined by 15 nodal points confined to a finite frequency interval. Formally these departures can be written as

$$\delta A(f) = \mathcal{I}_3(f; \{f_i, \delta A_i\}), \tag{2.5}$$

$$\delta \phi(f) = \mathcal{I}_3(f; \{f_i, \delta \phi_i\}), \tag{2.6}$$

where $\mathcal{I}_3$ is a cubic spline interpolant, the $\{f_i\}$ are the nodes of the spline interpolant in frequency, and $\{\delta A_i\}$ and $\{\delta \phi_i\}$ are the values of the spline at those nodal points. To better generalize the model to fit a larger variety of possible departures, we freely let the nodal points move around in frequency space during sampling (with the condition that they do not exchange orders or get too close to each other) after being initialized linearly in log-frequency space, as done in Farr et al. [25], Vitale et al. [29]. We choose the node locations to be the same for the amplitude and phase spline functions as we expect deviations to happen at similar frequencies and to reduce our model degrees of freedom. One could choose an independent set of node locations for phase and amplitude if expecting to have deviation effects that alter the amplitude and phase at different points in frequency. We prevent nodal points from getting arbitrarily close as this causes the spline to be too sensitive to very small changes in node positions, creating extreme deviations to satisfy the smoothness conditions of the spline function. The parameters added to our inference for this model are then the $\{\delta A_i\}$, $\{\delta \phi_i\}$ and $\{f_i\}$.

With our assignment of Gaussian priors (see § 2.4.3) to the δA and $\delta \phi$ parameters, our spline model implements a Gaussian process prior for the waveform deviations from the model (c.f. [30]).

2.4.3. Statistical Framework

We place a Gaussian prior on the departure parameters, $\{\delta A_i\}$ and $\{\delta\phi_i\}$, centered around zero, with σ_A and σ_ϕ characterizing our prior uncertainties about the magnitude of the departures in amplitude and phase from the modeled waveforms.

$$\begin{aligned} p(\delta A_i) &= N(0, \sigma_A) \qquad (2.7) \\ p(\delta\phi_i) &= N(0, \sigma_\phi) \qquad (2.8) \\ p(f_\mathrm{i}) &= U(f_\mathrm{low}, f_\mathrm{high}) \qquad (2.9) \end{aligned}$$

It is important however, not to think of the Gaussian priors on the node values of $\{\delta A_i\}$ and $\{\delta\phi_i\}$ as the broadband uncertainties of the interpolated spline function across the frequency range. In practice the prior widths on the nodes are wider than the broadband uncertainty resulting from sampling the prior. We constrain the spline nodes to be increasing in frequency:

$$f_\mathrm{i-1} < f_\mathrm{i} < f_\mathrm{i+1}. \qquad (2.10)$$

The spline model introduces some challenges that need to be accounted for, the first being the freedom in the frequency values of the node points. Since we also use the node positions in frequency space as sampling parameters and we place uniform priors on the spline node locations across the frequency band from f_low to the Nyquist frequency, the model is degenerate under exchange of node positions. To circumvent this degeneracy, we impose that the node positions stay ordered as in Eq. 2.10. These parameters can then be fit and the corresponding calibration errors marginalized over during inference.

Another challenge with the spline function is that if two nodes get too close together, obeying the conditions required of the cubic spline can lead to the interpolated deviations becoming very extreme. To account for this we prevent the frequency nodes from getting closer than 4 frequency bins away from one another as shown in equation 2.11, where the frequency bin width (in Hz) is determined from the segment length of data we are analyzing as $df = 1/T$ with T the segment length. That is, we reject any configurations with

$$f_{\mathrm{i}+1} - f_{\mathrm{i}} < 4 * \delta f = \frac{4}{T_{\mathrm{obs}}}. \tag{2.11}$$

Other spline based interpolation methods such as `BayesLine` combat this by keeping the spline nodes on a fixed frequency grid then turns them on/off during inference using trans-dimensional Markov-Chain Monte Carlo sampling [31]. We experimented with a few other fixes to this issue. We implemented a Gaussian prior on every frequency bin location; we also tried evaluating the Gaussian prior on some number of points between nodes of the spline interpolants to disfavor the large spline excursions as well. In practice both of these priors turned out to be too restrictive while attempting to recover the simulated deviations presented in Section 2.5; they may still be useful when exploring astrophysical events.

The last challenge is that our model of amplitude deviations is perfectly degenerate with the distance to the signal. The distance to the source can be increased while simultaneously producing a positive amplitude deviation across the entire frequency band of the signal to compensate. The default astrophysical prior used in `LALInference` [19] is $\propto\ D_L^2$ with D_L the luminosity distance.

When this prior distribution is coupled with the zero-mean Gaussian priors on amplitude deviations, the strong $\propto D_L^2$ prior on distance almost always results in systematically positive amplitude deviations in the spline component of the model. However, by allowing for broad-band changes to the amplitude of the signal in the first place we are no longer able to meaningfully infer the distance to the source. In other words, we use D_L (with a uniform prior) as a phenomenological parameter to fit the broad-band amplitude of the signal, and the spline model describes any frequency-dependent deviations that may be present.

With the prior assumptions and modified waveform we can now construct the posterior distribution according to Bayes' Theorem,

$$p(\boldsymbol{\theta}, \{\delta A_i\}, \{\delta\phi_i\}, \{f_i\}|d) \propto \mathcal{L}(d|\boldsymbol{\theta}, \{\delta A_i\}, \{\delta\phi_i\}, \{f_i\}) p(\boldsymbol{\theta}) p(\{\delta A_i\}) p(\{\delta\phi_i\}) p(\{f_i\}) \quad (2.12)$$

then use the `LALInference` Markov Chain Monte Carlo (MCMC) algorithm to draw samples from the posterior distribution in Eq. 2.12, with $\boldsymbol{\theta}$ the normal CBC parameters, d the gravitational wave strain data, $\mathcal{L}$ the standard GW likelihood with the modified intrinsic waveform as shown in Eq. 2.4.

2.5. Simulated Deviations

In order to validate our spline model, we run on data with astrophysically motivated deviations included. We choose a toy model that presents an extreme case of spontaneous scalarization [32] which modifies a simulated high-mass BBH waveform with similar astrophysical parameters as GW150914 [11, 12]. We generate two waveforms where one has simulated modifications according to the

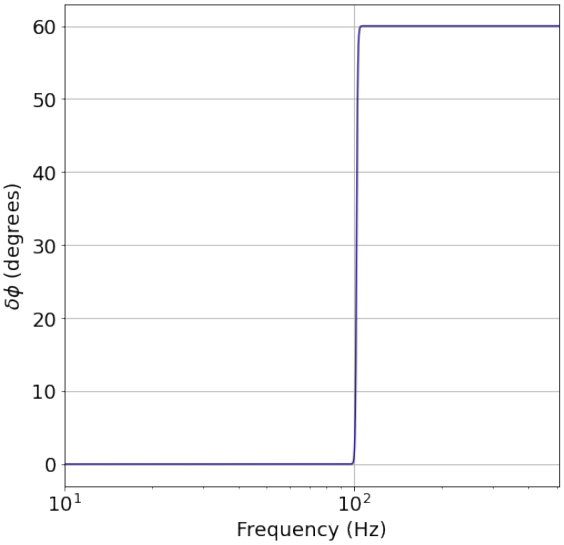

FIGURE 2.1. Simulated phase deviation $\delta\phi$ for extreme spontaneous scalarization toy model.

toy model and the other is exactly as the modeled GR waveform describes. We then simulate colored Gaussian noise according to the the Advanced LIGO design sensitivity noise curve or power spectral density (PSD) [26] and add it to the simulated waveforms.

m_1	36 $M_\odot$
m_2	29 $M_\odot$
D_L	450 Mpc
ϕ	2.76 rad
α	1.37 rad
δ	-1.26 rad
$S_{1,\mathrm{z}}$	0.0
$S_{2,\mathrm{z}}$	0.0

TABLE 2.1. Parameters for simulated validation signals

We first generate a GR-based frequency domain waveform using the IMRPhenomD [2] waveform approximant with the simulated parameters shown in Table 2.1, then modify the same waveform according to our toy model described above extreme scalarization case [32]. For the phase $\delta\phi$, the toy model includes an

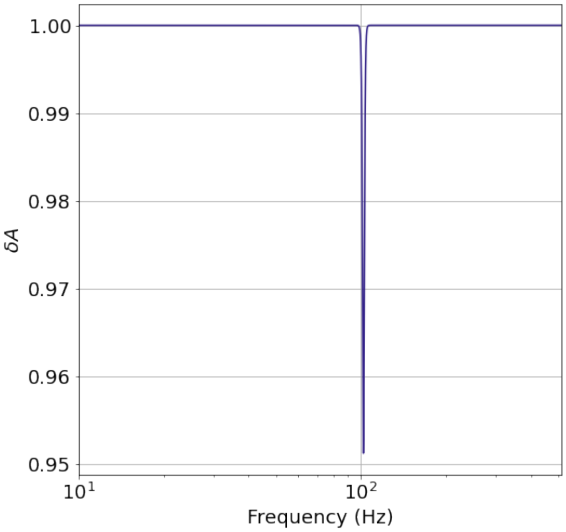

FIGURE 2.2. Simulated amplitude deviation δA for extreme spontaneous scalarization toy model.

abrupt increase of 60 degrees centered at $f_z = 102$Hz, with a width of $df = 1$Hz, see Figure 2.1. The amplitude temporarily drops by 5% in frequency, again centered at $f_z = 102$Hz and with a width of $df = 1$Hz, see Figure 2.2. To get these modifications we used these parameters in Eqs. 2.13 and 2.14 with $dA = 0.1$, $d\phi = 60°$.

$$\delta A(f) = e^{\frac{1}{2}dA\left(\tanh\left(\frac{f-f_z}{df}\right)^2 - \tanh\left(\frac{f_{\mathrm{ref}}-f_z}{df}\right)^2\right)} \tag{2.13}$$

$$\delta\phi(f) = \frac{1}{2}d\phi\Big[\tanh\left(\frac{f-f_z}{df}\right) - \tanh\left(\frac{f_{\mathrm{ref}}-f_z}{df}\right)\Big] \tag{2.14}$$

2.5.1. Results on simulated signals

To compare the effect of the waveform modification and the spline model, we ran the `LALInference` parameter estimation software on both the modified and unmodified signals with our spline deviation model turned on and off. Figures

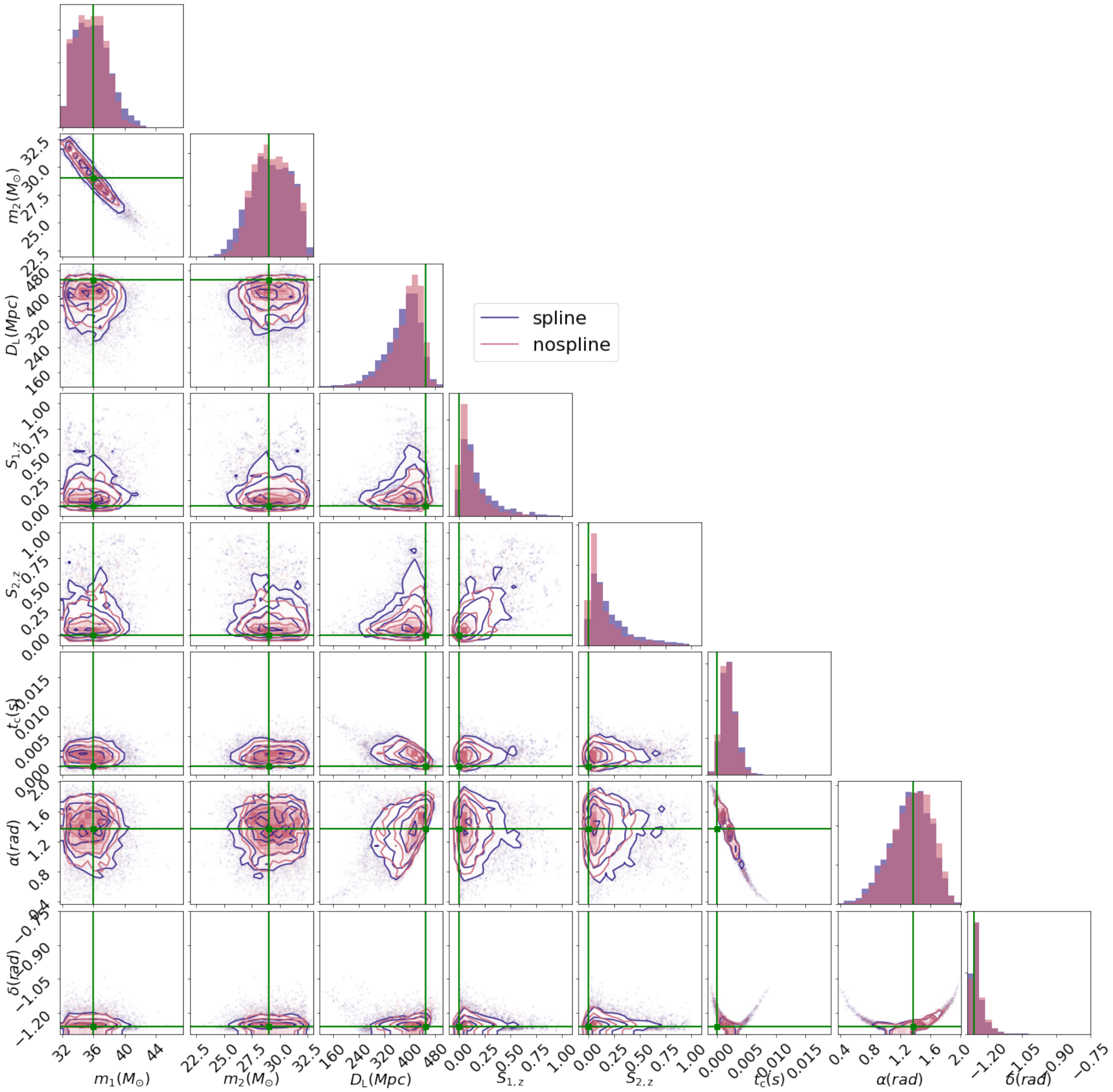

FIGURE 2.3. Corner plots showing the 1-D and 2-D marginalised posterior distributions for simulated parameters for PE runs on the Unmodified Signal with spline model on (purple) or off (pink). This demonstrates the impact of the model flexibility on astrophysical parameter uncertainties.

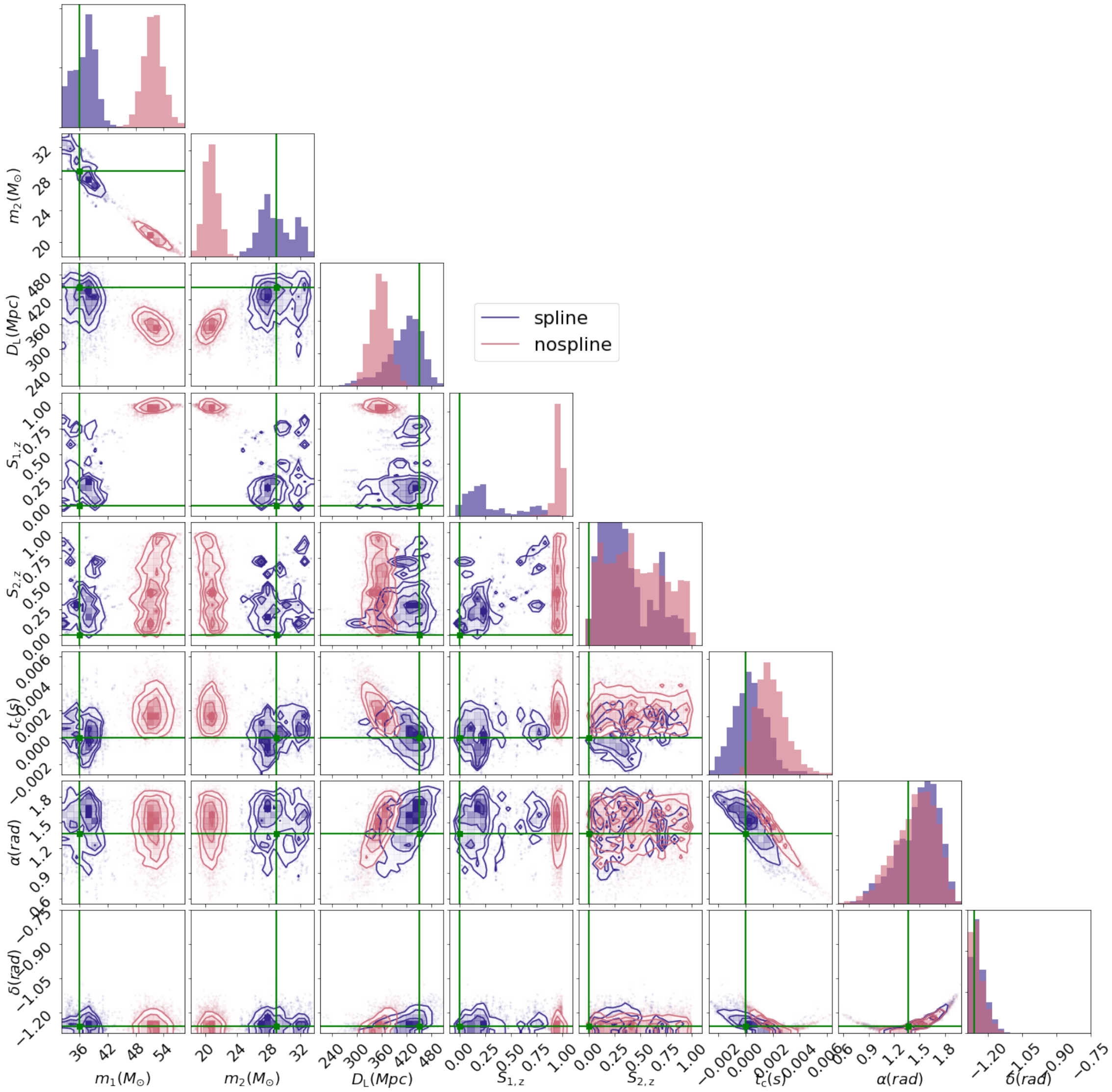

FIGURE 2.4. Corner plots showing the 1-D and 2-D marginalised posterior distributions for simulated parameters for PE runs on the Modified Signal with spline model on (purple) or off (pink). This demonstrates inaccuracies in parameter estimation performed on signals containing deviations from the physically modeled waveforms.

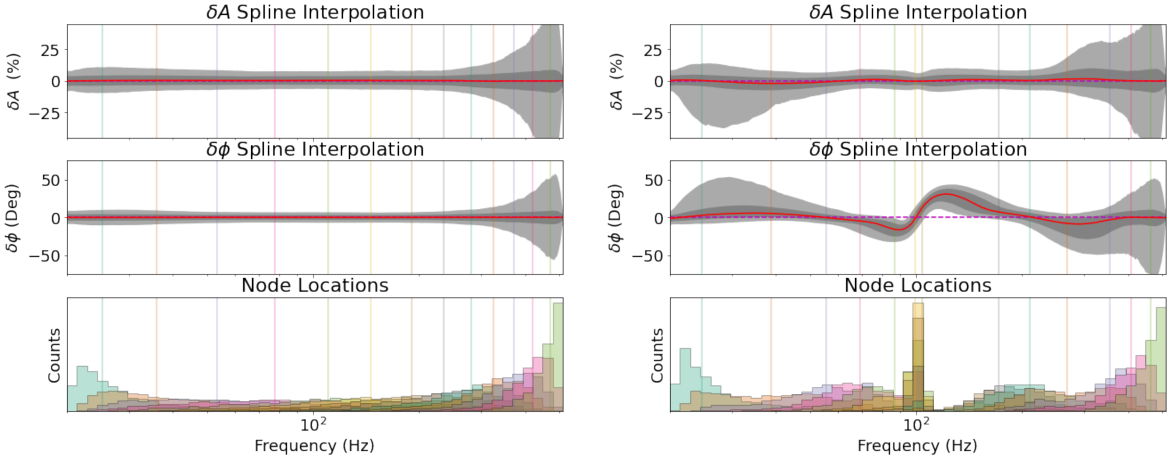

FIGURE 2.5. Spline Interpolation of unmodified (left) and modified (right) signal deviations. 1 and 2 σ credible intervals (grey) and the median spline (red) are shown with top panel the amplitude deviations and middle panel the phase deviations. In the bottom frame we plot the node position posterior distributions, which clump towards the frequency of deviations in the modified case in the bottom right most panel. For the unmodified case they are more uniformly distributed as they are exploring the prior.

2.3 and 2.4 show the corner plots for the simulated intrinsic (component masses, spins, etc) and extrinsic (luminosity distance and sky localisation) GW parameters comparing the 1-D and 2-D marginalised posterior distributions with the spline model on or off. In figure 2.3 we see that there is minimal difference in posteriors for the unmodified signal for most of the parameters which is what we would expect for the case of no deviations. We also expect to see greater uncertainties on certain parameters with the spline model as there are possible degeneracies with the spline and other parameters as seen in the different 1-D spin parameter posteriors in figure 2.3. This means the normal CBC model or template waveform is able to explain the entire coherent signal in the data.

For the Modified case (fig 2.4) we see that the simulated values are included in most of the spline model posterior distributions but not the posteriors with the spline model inactive. This is because the modified signal includes the abrupt

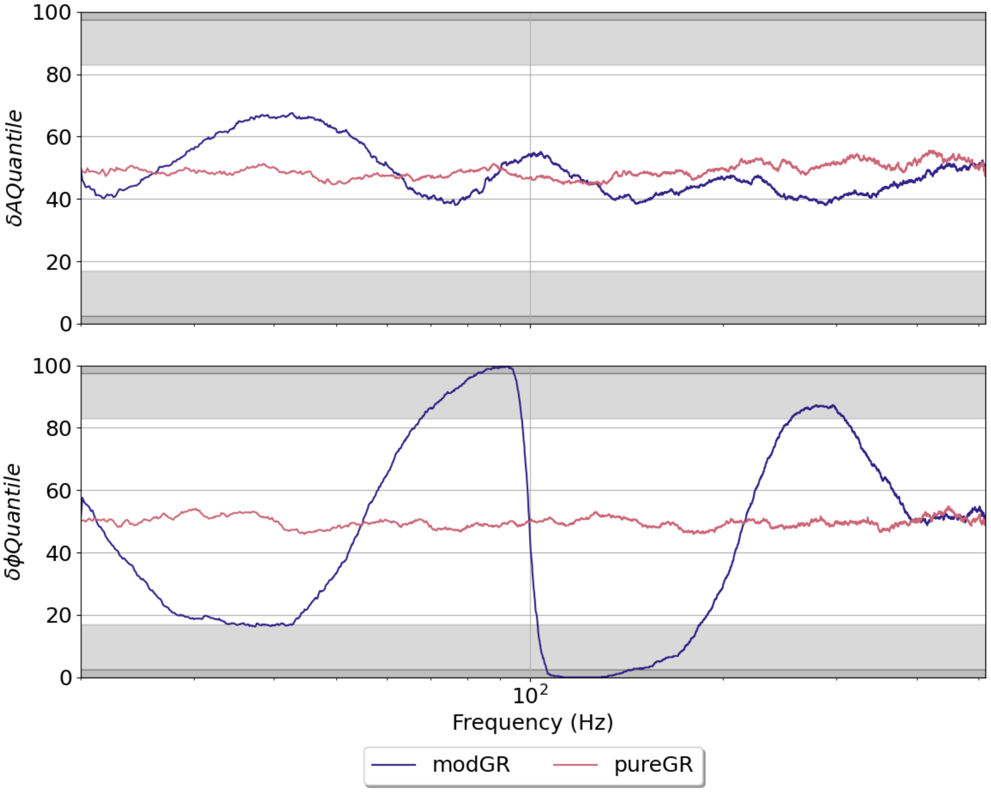

FIGURE 2.6. Posterior quantile of the spline interpolant that 0-deviation corresponds to for the simulated unmodified (pink) and modified (purple) signals.

(in frequency) modifications or deviations simulated and these types of extreme abrupt deviations are very poorly described by the template waveform models and especially the IMRPhenomPv2 waveform template used in this analysis. However, we see that with our spline model turned on, the posterior distributions do encompass most of the simulated parameter values. This illustrates that parameter estimation (PE) without including deviation parameters would not be reliable at estimating the true signal parameters in cases where there may be departures from the waveform template used. In particular we see from figure 2.4 that the masses and trigger time posterior distributions are more consistent with the simulated values with the spline model fitting the deviations.

Figure 2.5 shows the interpolated spline functions for the deviations in amplitude and phase, δA and $\delta\phi$, for both the modified and unmodified simulated signals (right, left). The plots show the 1σ and 2σ bounds in the spline interpolants along with the median. We see here that our model consistently recovers zero deviations for the unmodified signal across the frequency band. The ranges included in the 1σ and 2σ bands show the exploration of the prior bounds while sampling while also being symmetric around the median. Looking at the modified case, we see a presence of deviations away from zero in phase around 100 Hz. The phase recovery does not show the clear step function behavior that was simulated and shown in figure 2.1 as that extreme deviation is disfavored by the prior, because the priors used on the node positions are Gaussian distributions centered about zero; to recover the large flat step in frequencies greater than 100 Hz would require very low prior probability at those nodes. The spline posterior does however show that there is a transition at the modified frequency ($\sim$ 100 Hz) and shows that the phase increases roughly by 60 degrees as we simulated.

The model compensates for this by “ramping” down the phase modification so that it slopes back to zero after the merger frequency ($\sim$ 250 Hz for this signal) since there is no signal content to infer from after merger. The model may also be able to coherently ramp down as a result of possible degeneracies with other parameters. The amplitude recovery is less revealing since the simulated deviation is on order of the prior width along with the sharp resolution of the deviation in frequency, it is harder to clearly recover that feature in our model. The presence of deviations is corroborated by the fact that the posterior distribution on the deviations excludes zero deviation at some frequencies at the 95% level, which we do see in our modified case. To focus on this we can calculate the posterior quantile of phase and amplitude deviation that the x-axis (0-deviations) falls at for each frequency bin. This is shown in figure 2.6.

From this we see in the phase plot that there are significant portions of the frequency range where the phase interpolant ruled out zero deviations at greater than the 99% credible level. The amplitude recovery shows less excursion which tells us we would not be able to constrain amplitude deviations of this type and magnitude. We do see significantly more freedom in amplitude than the no-modifications case which considerably constrained the amplitude interpolants to zero, however zero deviations lies within the 1σ range across the frequency band meaning it could still be consistent with no deviations present in amplitude.

2.5.2. Model Comparisons

A useful way to see how well the spline model performs on a given segment of data is to perform a model comparison between the spline on and spline off

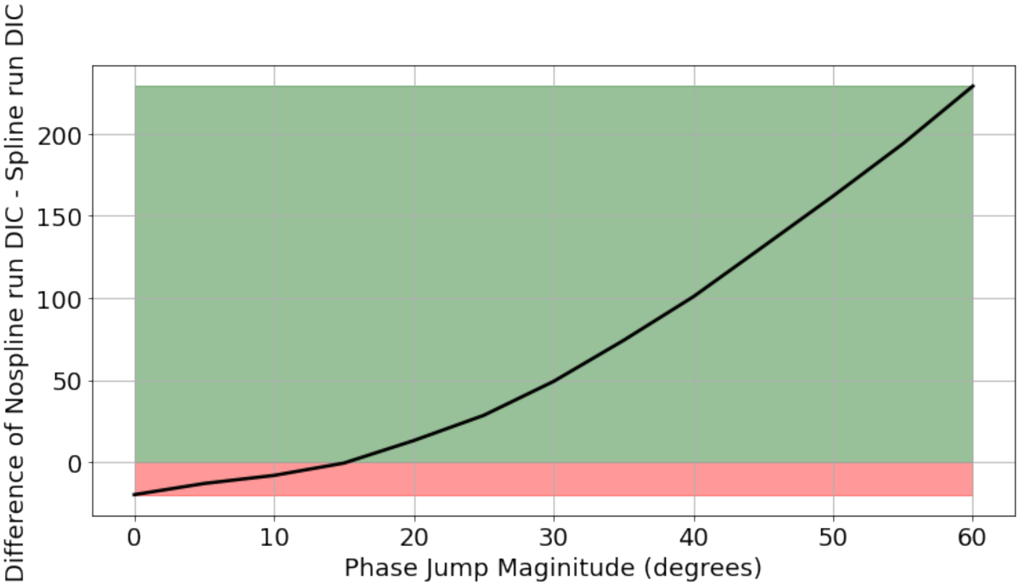

FIGURE 2.7. The DIC value from spline-model-off run minus spline-model-on run for different simulated phase deviations. Green shaded regions are where the spline model is preferred and red shaded region is where the spline model off is preferred.

models. To do this we take an Information Theory approach by computing the Deviance Information Criterion (DIC) [33, 34]. This measure of fitness has the feature that lower values correspond to a better fit and includes an "Occam's Factor" that penalizes models with greater numbers of parameters. To effectively compare the spline on to spline off models we take the DIC value from the spline off and subtract the DIC value from the spline on model giving us a single value that is positive if the spline model is preferred and negative if the spline model is disfavored. We now perform a study where we run on simulated signals with modifications as before but with increasing jumps in phase from 0 to 60 degrees incrementing by 5 degrees. Looking at figure 2.7, we see that as the phase jump increases the spline model becomes more favored. We can compare the DIC differences from figure 2.7 and to the differences for the un-modified signal which has DIC difference of -74.46, showing that for the DIC the spline model off is preferred at about the same level the spline model on is preferred for a phase jump of 35 degrees.

2.5.3. Model Limitations and Alternative Parameterizations

We have attempted a few other iterations of this model to increase its flexibility and performance on simulated signals which were not included in our final analyses. The first was to parameterize the deviations in the derivative of the phase/amplitude. This was tried for the same reason that the model was insensitive to the step function deviation that we have simulated in the phase. The step function was very incompatible with the normal model priors as each node after stepping up is penalized from our priors yet if we parameterized in the derivative the step function derivative looks like a delta function which is much more compatible with zero-centered Gaussian priors on the spline nodes. In practice this brought more challenges than it solved increasing other degeneracies with parameters and did not seem to qualitatively improve our efficiency or performance of correctly fitting the simulated deviations.

As a further test of our model, we attempted to recover modifications to a neutron star-black hole (NSBH) merger waveform. In the event of tidal disruption of the neutron star, there is an expected deviation to the waveform predicted by GR for non-deformable bodies. Specifically, we constructed a toy model featuring a roll-off in amplitude beyond a disruption frequency. Physically, this corresponds to a spreading and redistribution of mass after the moment of disruption, which would decrease the intensity of GWs emitted from then on [35, 36]. We made no change to the phase. Our model was consistently unable to recover these deviations. This is likely due to more degeneracies in our parameters,

GWTC-1 Events							
Event	χ_{eff}	d_L (Mpc)	m_1 ($M_\odot$)	m_2 ($M_\odot$)	$\mathcal{M}$ ($M_\odot$)	ρ_{gstlal}	Source
GW150914	$-0.01^{+0.12}_{-0.13}$	$430.0^{+150.0}_{-170.0}$	$35.6^{+4.8}_{-3.0}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.6}_{-1.5}$	24.4	BBH
GW151012	$0.04^{+0.28}_{-0.19}$	$1060.0^{+540.0}_{-480.0}$	$23.3^{+14.0}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.0}_{-1.1}$	10.0	BBH
GW151226	$0.18^{+0.2}_{-0.12}$	$440.0^{+180.0}_{-190.0}$	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.6}$	$8.9^{+0.3}_{-0.3}$	13.1	BBH
GW170104	$-0.04^{+0.17}_{-0.2}$	$960.0^{+430.0}_{-410.0}$	$31.0^{+7.2}_{-5.6}$	$20.1^{+4.9}_{-4.5}$	$21.5^{+2.1}_{-1.7}$	13.0	BBH
GW170608	$0.03^{+0.19}_{-0.07}$	$320.0^{+120.0}_{-110.0}$	$10.9^{+5.3}_{-1.7}$	$7.6^{+1.3}_{-2.1}$	$7.9^{+0.2}_{-0.2}$	14.9	BBH
GW170729	$0.36^{+0.21}_{-0.25}$	$2750.0^{+1350.0}_{-1320.0}$	$50.6^{+16.6}_{-10.2}$	$34.3^{+9.1}_{-10.1}$	$35.7^{+6.5}_{-4.7}$	10.8	BBH
GW170809	$0.07^{+0.16}_{-0.16}$	$990.0^{+320.0}_{-380.0}$	$35.2^{+8.3}_{-6.0}$	$23.8^{+5.2}_{-5.1}$	$25.0^{+2.1}_{-1.6}$	12.4	BBH
GW170814	$0.07^{+0.12}_{-0.11}$	$580.0^{+160.0}_{-210.0}$	$30.7^{+5.7}_{-3.0}$	$25.3^{+2.9}_{-4.1}$	$24.2^{+1.4}_{-1.1}$	15.9	BBH
GW170817	$0.0^{+0.02}_{-0.01}$	$40.0^{+10.0}_{-10.0}$	$1.46^{+0.12}_{-0.1}$	$1.27^{+0.09}_{-0.09}$	$1.186^{+0.001}_{-0.001}$	33.0	BNS
GW170818	$-0.09^{+0.18}_{-0.21}$	$1020.0^{+430.0}_{-360.0}$	$35.5^{+7.5}_{-4.7}$	$26.8^{+4.3}_{-5.2}$	$26.7^{+2.1}_{-1.7}$	11.3	BBH
GW170823	$0.08^{+0.2}_{-0.22}$	$1850.0^{+840.0}_{-840.0}$	$39.6^{+10.0}_{-6.6}$	$29.4^{+6.3}_{-7.1}$	$29.3^{+4.2}_{-3.2}$	11.5	BBH

TABLE 2.2. Posterior estimates for LIGO-Virgo's GWTC-1 catalog of events with +/- 1σ errors shown as well

and specifically in component mass. By moving to a higher mass, parameter estimation can push the amplitudes lower at high frequencies, and the additional flexibility of the spline model was unable to capture any of the deviation. Better results may be had with a more realistically modified NSBH waveform, or by changing the way we manage degeneracies in our parameters.

2.6. Results from LIGO-Virgo Public Data

The LIGO-Virgo GWTC-1 catalog [9] presents a population of GW events to study deviations of GR along with a suite of data with which to validate our gravitational waveform models. We first discuss the results of our model on the ten binary black hole merger events listed with a few selected median posterior parameters in table 2.2, and then we will discuss the one binary neutron star event, GW170817.

We performed parameter estimation on the ten binary black hole (BBH) events with the template waveform IMRPhenomPv2 [2], and the single binary neutron star (BNS) event, GW170817, with the comparable

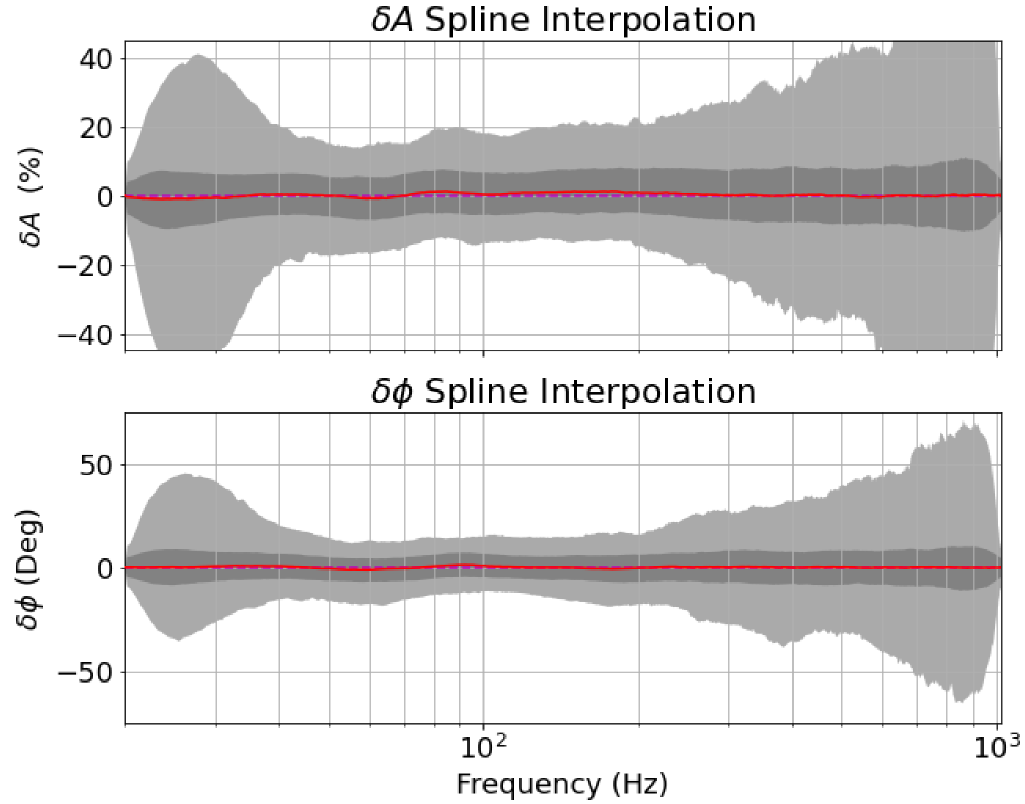

FIGURE 2.8. Spline interpolation of GW170823 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.

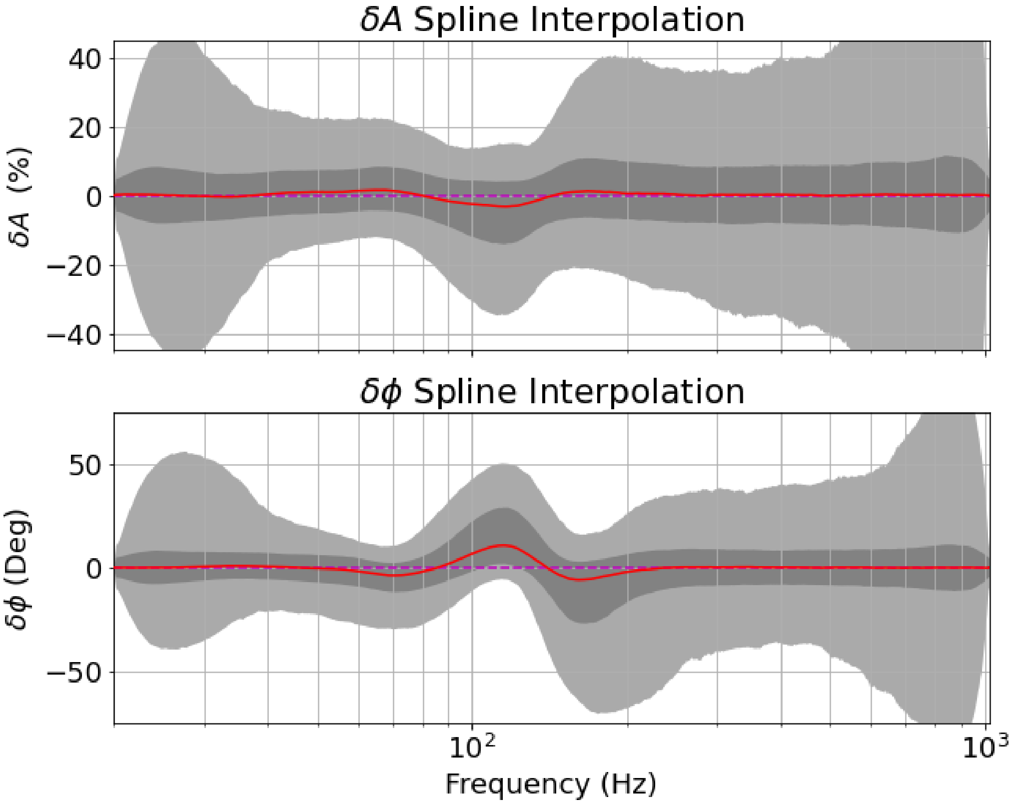

FIGURE 2.9. Spline interpolation of GW170729 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.

IMRPhenomPv2_NRTidal [37] template waveform that includes neutron star tidal effects, each with our spline model turned on and off. For GW170817, since there was an Electromagnetic Counterpart detected, and to help optimize the speed of sampling, we have fixed the sky location of the source in our analysis. These runs used the same prior settings detailed in Section 2.4 which is a 5% uncertainty of the amplitude and 5 degrees uncertainty on phase. First we look at the spline plot recoveries for GW170823 and GW170729 to show two different cases, both still consistent with zero deviations. In figure 2.8 we see the case where our model is consistent with zero deviations consistently across the frequency band with the credible intervals illustrating the prior exploration during sampling. Contrasting to this we can look at figure 2.9 to see a case in which our model has less posterior support for zero deviations at some frequencies. We see here that there is an area on the plot, most importantly the phase portion, around 100 Hz in which zero deviations falls nearly outside our 1σ interval. The first thing one might wonder seeing deviations here is whether this can be explained by the normal calibration uncertainties of each detector being similar. We check this in fig 2.10 along with a run using a waveform model including higher-order modes (HOM) or greater than $\ell = 2$ modes in the spherical harmonic expansion [38], IMRPhenomPv3HM [39]. As seen in table 2.2 GW170729 is the most massive event from GWTC-1 and HOM are more important for higher mass systems along with asymmetric mass systems [20, 40]. We can clearly see a very similar spline interpolant recovery leading us to believe this is fitting features in the data that are unexplained by either a higher-order mode waveform or similar calibration errors across the network of detectors. However for each of the GW170729 results we still have posterior support for zero deviations at the frequencies of largest excursions.

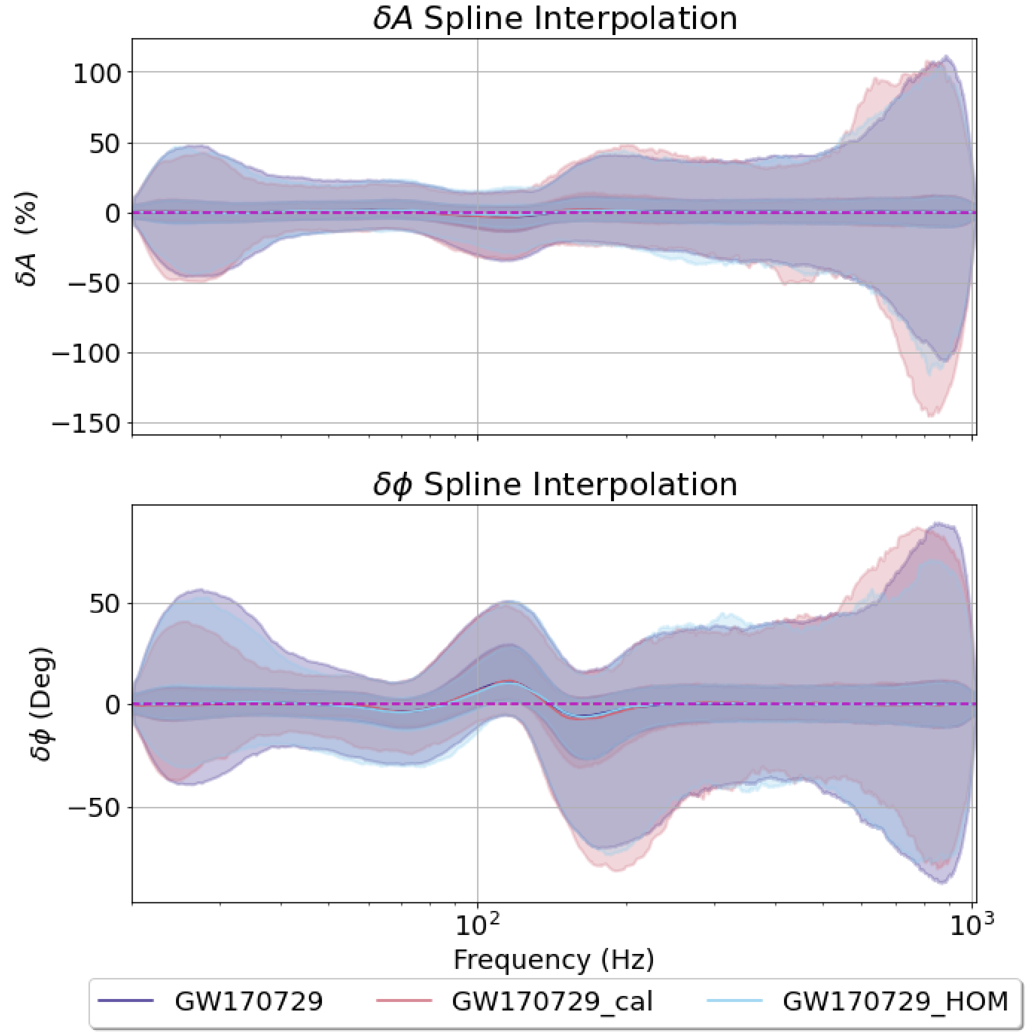

FIGURE 2.10. Spline interpolation of GW170729 with 1σ percent credible intervals shown, comparing runs on GW170729 with calibration turned on/off and using a HOM waveform also with calibration turned off.

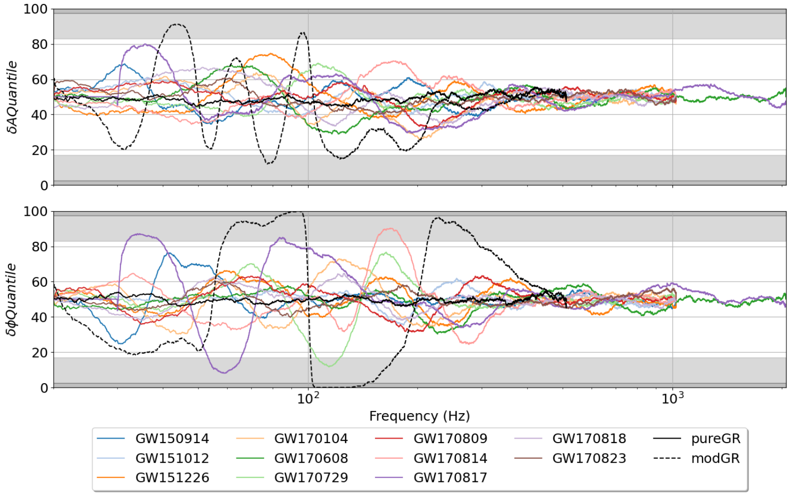

FIGURE 2.11. Quantile of spline interpolation that 0-deviation corresponds to for GWTC-1 events. (Different choices of $f_{\rm low}$ and sampling rates were chosen for some events.)

As in the previous section we can now look at which quantile of the spline interpolant posteriors will fall along the x-axis (corresponding to no deviation), at each frequency bin for each event. This is shown in figure 2.11 in which we highlight the $> 1\sigma$ and $> 2\sigma$ bands. We notice in this figure that for all ten binary black hole events, no deviations or modifications in amplitude fall outside the 1-sigma interval. We do see that for two events, namely GW170729 and GW170814, there are regions in which no deviation falls outside of the 1σ band. However there are significant portions where GW170817 also falls outside the 1σ band. Looking at the spline posterior in figure 2.12 for GW170817 shows the clear departures away from zero at some frequencies but overall outside of that small region looks behaved.

To further check our analysis we perform the same model comparisons described earlier in this paper on an event by event basis comparing the spline model on to the spline model off. This is seen in figure 2.13 with the same results

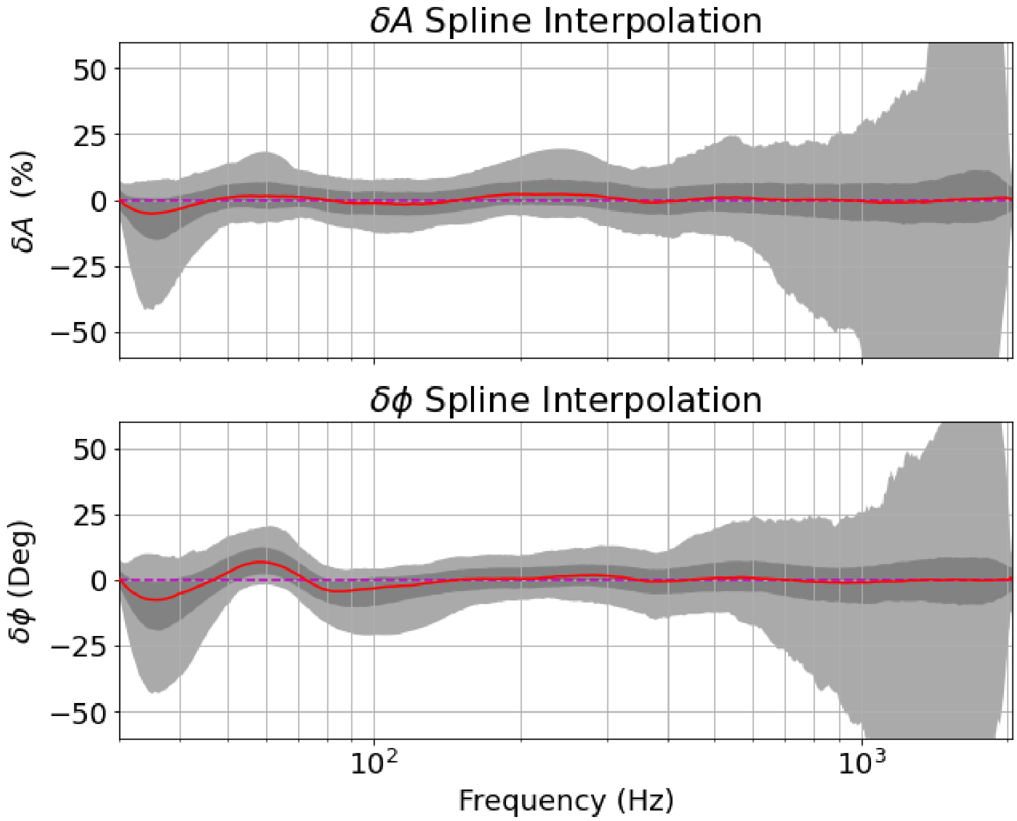

FIGURE 2.12. Spline interpolation of GW170817 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.

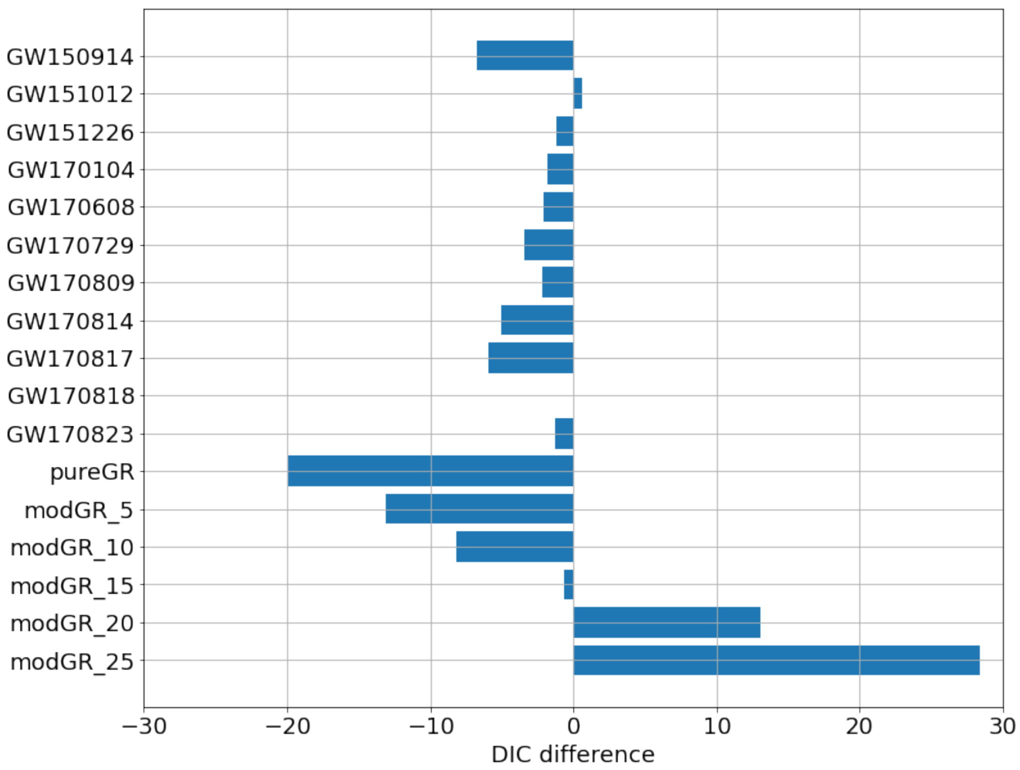

FIGURE 2.13. DIC of spline-model-off run minus DIC of spline-model-on run for GWTC-1 events and simulated signals. Simulated modified signals are denoted with the magnitude of phase deviation jump in degrees. Negative values correspond to spline model disfavored while positive values show the spline model favored.

from differing the phase jump as discussed in the previous section. This shows that for each event there is no preference to either model from their DIC values. Even for the events with more extreme spline interpolant posterior (i.e, GW170729 and GW170817) we still see that from a model comparison approach the spline model does not significantly describe the data better than without the spline model.

2.6.1. Conclusions

We have presented a useful model and parameterization to describe general departures or deviations from gravitational waveform models. Our model can be used to look for departures from any of modeled waveforms by generically fitting the entire frequency band at once with spline functions. We find that for the 11 gravitational wave events of both BBH and BNS origin in GWTC-1, the data are consistent with the IMRPhenomPv2 waveform. Shown in figure 2.11 there are two events that we consider "outliers" with some portion of the phase deviation outside of the 1σ range but still lying well within the 2σ bounds of no deviations, which for a sample size of 11 events would not be unexpected, even with no deviations present.

Currently, more investigation into possible degeneracies of our model would be necessary to vet any significant sign of deviation. Further studies also need to be done to evaluate effects of detector sensitivity on our model, expand the validation of our model on other physically motivated deviations that can be simulated, and possibly incorporating information from proposed alternatives to GR into the priors. However, the model presented in this paper can be used as a model agnostic test to look for first signs of departures in the modeled waveforms. With more events and increased detector sensitivity, we will be able to better constrain any general deviations with our model while at the same time giving us a better testing set to look for hidden degeneracies between our model and other parameters.

2.7. Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1807046 and work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. We would also like to thank the Niels Bohr Institute for its hospitality

while part of this work was completed, and acknowledge the Kavli Foundation and the DNRF for supporting the 2017 Kavli Summer Program. This work made use of the `PyCBC` [41], `GWPy` [42], and `LALInference` [19] software packages along with creating all plots and figures with `Matplotlib` [43] and `corner` [44].

2.8. Additional Analysis Details

2.8.1. Degrees of Freedom and Model Comparison

If deviations from a modeled waveform are primarily in phase but not the amplitude of the signal, our model's added flexibility could unnecessarily penalize our model comparisons in the effective dimensionality penalty of the DIC. The DIC test statistic is defined as:

$$\mathrm{DIC} = -2\overline{\log(\mathcal{L})} + p_D = -2\left(\overline{\log(\mathcal{L})} - \mathrm{var}(\log(\mathcal{L}))\right) \tag{2.15}$$

With $\overline{\log\mathcal{L}}$ the average log-likelihood, and p_D the effective number of dimensions, defined as $p_D = \frac{1}{2}\mathrm{var}(-2\log\mathcal{L})$ with var(...) denoting the variance. In this definition the effective dimension term penalizes models with more degrees of freedom. Here we explore the constraints of a phase-only modified model. We only do so in an approximate way, assuming that the goodness of fit (i.e., mean of log-likelihood) term obtained with our full phase and amplitude model is the same that a phase-only model would produce. We then halve the model's effective dimension, p_D, in the DIC calculation shown in 2.15, effectively removing the degrees of freedom due to allowing the amplitude to vary.

Figure 2.14 shows that when approximating a phase-only deviation model as described above, we still do not show a clear favoring of the spline model for

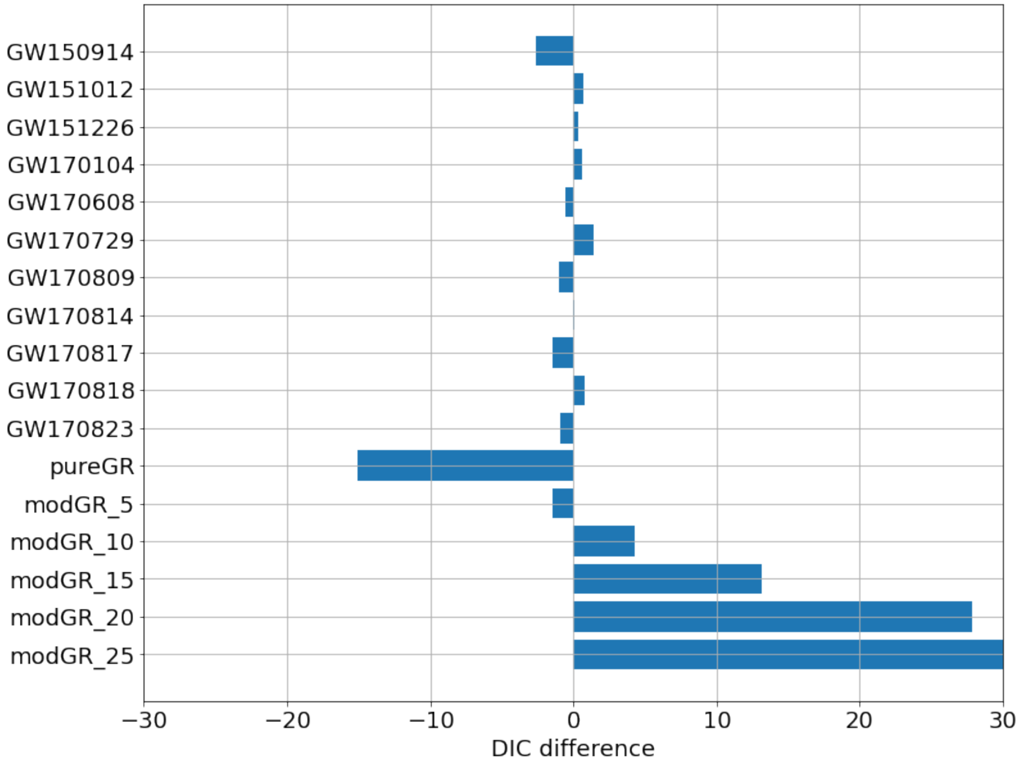

FIGURE 2.14. DIC of spline-model-off run minus DIC of spline-model-on run for GWTC-1 events and simulated signals. DICs in this plot are calculated with half of the variance degree of freedom penalty term. Simulated modified signals are denoted with the magnitude of phase deviation jump in degrees. Negative values correspond to spline model disfavored while positive values show the spline model favored.

any event in GWTC-1, but compared to figure 2.13 there is more ambiguity about which model is favored. This illustrates that with the current data, a phase-only spline deviation model as presented does not qualitatively alter our conclusions but may be useful in analyzing future catalogs of GW events.

CHAPTER III

POKING HOLES: LOOKING FOR GAPS IN LIGO/VIRGO'S BLACK HOLE POPULATION

3.1. Summary

Near the end of 2020 the LIGO-Virgo-KAGRA (LVK) collaboration released data and analyses from the first half of their third observing run (O3), with an additional 35 GW detections bringing the cumulative second catalog, GWTC-2, to 46 detections [45]. One of the especially interesting events in GWTC-2 is GW190521 [46, 47], which was the contained the most massive black holes of the catalog. The detection of these massive black holes challenged current theoretical understanding of black hole formation through stellar collapse, as the phenomena of Pair-Instability Supernovae (PISNe) is expected to preclude the creation of black holes with masses $> 55\mathrm{M}_\odot$ through stellar collapse, while the components of GW190521 were estimated to be about $\sim 85\mathrm{M}_\odot$ and $\sim 65\mathrm{M}_\odot$ [46, 47]. This led to many new interpretations of this peculiar merger, including one where Fishbach and Holz [48] showed that under a population-informed prior, the masses of GW190521 could be straddling the bounds of the expected PISNe mass gap. Additionally, Nitz and Capano [49] showed a re-analysis of the parameter estimation of GW190521 with newer waveforms models that also showed considerable support a high mass-ratio mode, letting the masses straddle the mass gap. In this chapter we present work published as Edelman et al. [50], that looks for evidence of such a gap in the mass distribution of black holes, with simple parametric population models containing a gap. We find mild evidence for the presence of such a feature, however this depends both on the choice of underlying mass distribution, and the parameter estimates for the most massive binary black hole (BBH) in GWTC-2, GW190521. This model presented would be later used in the LVK collaboration's population analysis of the next catalog, GWTC-3 [1, 51].

3.2. Abstract

Stellar evolution models predict the existence of a gap in the black hole mass spectrum from $\sim 55\mathrm{M}_{\odot} - 120\mathrm{M}_{\odot}$ due to pair-instability supernovae (PISNe). We investigate the possible existence of such an "upper" mass gap in the second gravitational wave transient catalog (GWTC-2) by hierarchically modeling the astrophysical distribution of black hole masses. We extend the TRUNCATED and POWERLAW+PEAK mass distribution families to allow for an explicit gap in the mass distribution, and apply the extended models to GWTC-2. We find that with the TRUNCATED model there is mild evidence favoring an upper mass gap with log Bayes Factor $\ln \mathcal{B} = 2.79$, inferring the lower and upper bounds at $56.12^{+7.54}_{-4.38}\mathrm{M}_{\odot}$, and $103.74^{+17.01}_{-6.32}\mathrm{M}_{\odot}$ respectively. When using the POWERLAW+PEAK model, we find no preference for the gap. When imposing tighter priors on the gap bounds centered on the expected PISNe gap bounds, the log Bayes factors in favor of a gap mildly increase. These results are however contingent on the parameter inference for the most massive binary, GW190521, for which follow up analyses showed the source may be an intermediate mass ratio merger that has component masses straddling the gap. Using the GW190521 posterior samples from the analysis in Nitz and Capano [49], we find an increase in Bayes factors in favor of the gap. However, the overall conclusions are unchanged: There is no preference

for a gap when using the POWERLAW+PEAK model. This work paves the way for constraining the physics of pair-instability and pulsational pair-instability supernovae and high-mass black hole formation.

3.3. Introduction

With the recent release of its second gravitational wave transient catalog (GWTC-2), the LIGO/Virgo collaboration (LVC) has now detected 50 gravitational wave (GW) events since the start of the advanced detector era, at least 46 of which came from binary black hole (BBH) systems [9, 26, 27, 45]. GWTC-2 therefore provides a rich data set to infer properties of the astrophysical population of stellar mass black holes [52, 53]. A robustly predicted feature that we can look for, specifically in the black hole (BH) mass distribution, is the theorized upper mass gap produced from effects due pair instability supernovae (PISNe) which precludes formation of BHs with masses $\sim 55\mathrm{M}_\odot - 120\mathrm{M}_\odot$ from stellar collapse.

Stellar evolution simulations show that stars with core masses from $\sim 40 - 135\mathrm{M}_\odot$ undergo PISNe in which the highly energetic gamma rays produced in the core can collide with atomic nuclei and produce electron-positron pairs [54]. This production process absorbs energy that was previously counteracting the gravitational pressure causing the core to contract. Heavier core stars in the $\sim 65 - 135\mathrm{M}_\odot$ range ignite oxygen leading to an unstable thermonuclear explosion which leaves behind no compact remnant [55, 56]. Lighter stars with core masses $\sim 40 - 65\mathrm{M}_\odot$ can temporarily stabilize themselves after the ignition and thus go through a series of pulsations (PPISN) that shed large amounts of mass with each pulse [57]. This continues until the mass of the star is too light

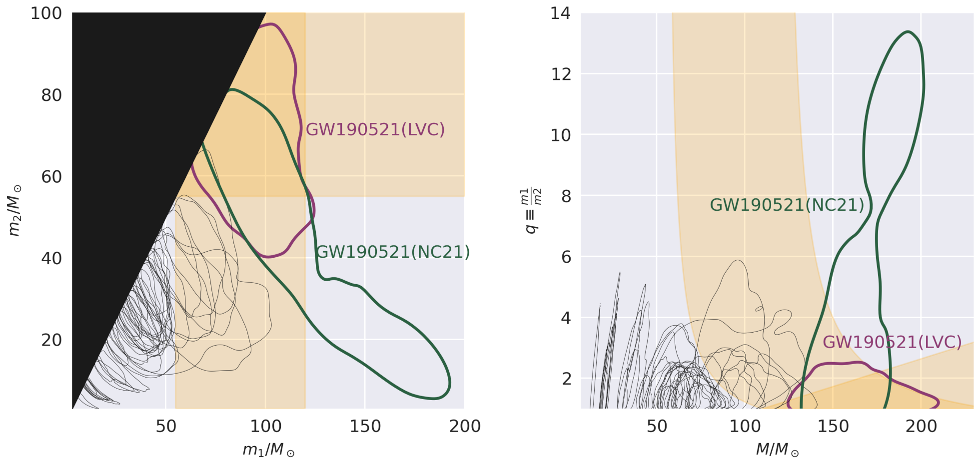

FIGURE 3.1. 90% credible level contour of the posterior samples for each of the 46 BBH mergers in GWTC-2. We show both sets of posterior samples for the highest mass event, GW190521, from [49] (green) and from the LVC analysis (purple). Posterior samples from Nitz and Capano [49] have been re-weighted to the same priors as the LVC analyses. The approximate expected region ($\sim 55\mathrm{M}_\odot - 120\mathrm{M}_\odot$) of the PISNe mass gap is highlighted in orange.

to pair-produce, leaving the star to undergo a normal core-collapse supernova (CCSN) that leaves behind remnant black holes of masses $\lesssim 55\mathrm{M}_\odot$ [58]. Even more massive low metallicity stars can bypass PISN completely and possibly form intermediate-mass black holes (IMBHs) with masses $> 120\mathrm{M}_\odot$ [59]. While simulations consistently predict the existence of this proposed mass gap in the distribution of binary black hole masses, the precise locations of the lower and upper boundaries remain uncertain [60, 61, 62, 63, 64, 65]. Analyses of GW data from the first and second Advanced LIGO-Virgo observing runs put constraints on the lower bound of this mass gap by modeling the black hole primary mass distribution as a powerlaw with a sharp high-mass cutoff, and found support for this lower edge at $\sim 45\mathrm{M}_\odot$ [52, 66, 67]. While the GWTC-2 catalog is consistent with 97% of observed BBH primary masses lying below $45\mathrm{M}_\odot$, population analyses using GWTC-2 and parameterized toy models find that there is less support for a sharp cutoff and instead preference for a shallow tail to larger masses [53]. This shallow tail extending into the supposedly forbidden range of masses is primarily driven by the few GW detections that have posterior support in the mass gap, most notably GW190521 [46, 47]. This leads to the possibility that there could be a separate population of black holes contributing in that mass range, formed through some other mechanism. One possibility is that remnants of previous black hole mergers undergo subsequent "hierarchical" mergers which, in dynamical environments such as globular clusters or active galactic nuclei, can contribute a significant fraction to the overall rate of mergers [68, 69, 70, 71, 72, 73, 74]. For example, Kimball et al. [75] finds evidence that GWTC-2 includes at least one merger with a second generation component under certain assumptions about the 1st generation black hole mass distribution.

The hierarchical merger scenario is not the only explanation for high mass events like GW190521, though. Other recent work has looked more closely at the inferred source parameters of GW190521. Using gravitational waveforms with quasi-circular black hole inspirals and a standard "agnostic" prior, the LVC found the source of GW190521 to have component source-frame masses within the theorized bounds of the mass gap with $m_1^{\mathrm{src}} = 85^{+21}_{-14}\mathrm{M}_\odot$ and $m_2^{\mathrm{src}} = 66^{+17}_{-18}\mathrm{M}_\odot$ [46, 47]. However, other waveform models and priors lead to other interpretations of this event. For example, the source of GW190521 could have had a highly eccentric orbit, been a head-on merger, or been subject to new physics allowing formation within the PISNe mass gap [68, 71, 76, 77, 78]. Alternatively, assuming the secondary component of GW190521 comes from the same prior distribution of secondaries as other events, there is more support for the components of GW190521 to straddle the lower and upper bounds of the mass gap [48]. There has also been work by Nitz and Capano [49] (hereafter NC21) reanalyzing the parameter estimates of GW190521 with the recently released IMRPhenomPXHM waveform [79] which supports mass ratios $q \equiv m_1/m_2 > 4$ that were not considered in the LVC analysis. NC21 found that GW190521 may be an intermediate mass ratio merger, reporting a multimodal posterior with an additional high mass ratio mode not identified in the LVC analysis. The reported source frame component masses for the high mass ratio mode squarely puts each outside of the theorized mass gap with $m_1^{\rm src} = 166^{+16}_{-35}\mathrm{M}_\odot$ and $m_2^{\rm src} = 16^{+14}_{-3}\mathrm{M}_\odot$ [49]. Figure 3.1 shows the 90% contours on the posterior samples from events in GWTC-2 with both the LVC GW190521 posterior samples in addition to samples from NC21 highlighted. This illustrates how differences in the analysis of GW190521 can considerably change the posterior support for its component masses to lie in the theorized PISNe mass gap. If GW190521 does "straddle" the gap, it would signal the existence of a high-mass population that could inform questions in both astrophysics and cosmology [80].

In this letter we present a simple phenomenological population model parameterizing the PISNe mass gap that enforces a zero rate of BBH mergers within the gap. Our model is a complementary approach to other physically motivated models that describe the impact of PISNe on the mass spectrum [e.g., 81]. Using this model we evaluate the evidence for the presence of a mass gap in LIGO/Virgo's second gravitational wave transient catalog, and constrain its properties. We conduct each analysis twice, first using posterior samples for GW190521 released by the LVC, and alternatively using samples produced in NC21. In Section 3.4 we introduce our parameterized mass gap model, and the methods used to infer population properties. In Section 3.5 we present the results

Primary Mass Model Parameters			
Model	**Parameter**	**Description**	**Prior**
Truncated	α	slope of the powerlaw	U(-4, 12)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
PowerLaw+Peak	α	slope of the powerlaw	U(-4, 12)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
	μ_p	mean of Gaussian peak ($M_\odot$)	U($20M_\odot$, $70M_\odot$)
	σ_p	width of the Gaussian peak ($M_\odot$)	U($0.4M_\odot$, $10M_\odot$)
	λ_p	fraction of BBH in the Gaussian component	U(0, 1)
Mass Ratio Model Parameters			
PowerLaw MassRatio	β_q	slope of the mass ratio powerlaw	U(-4, 12)
Redshift Evolution Model Parameters			
PowerLaw Redshift	γ	slope of redshift evolution powerlaw $(1+z)^\gamma$	U(-6, 6)
Mass Gap Parameters			
Agnostic MassGap	m_g	lower bound of PISNe mass gap ($M_\odot$)	U($40M_\odot$, $100M_\odot$)
	w_g	width of the PISNe mass gap ($M_\odot$)	U($0M_\odot$, $160M_\odot$)
Informed MassGap	$m_{\mathrm{g,min}}$	lower bound of the PISNe mass gap ($M_\odot$)	$\mathcal{N}(\mu = 55M_\odot, \sigma = 10M_\odot)$
	$m_{\mathrm{g,max}}$	upper bound of the PISNe mass gap ($M_\odot$)	$\mathcal{N}(\mu = 120M_\odot, \sigma = 20M_\odot)$

TABLE 3.1. Prior choices and description of hyperparameters for used population models.

of our inference with both sets of posterior samples and two underlying mass distributions. We then discuss our interpretation of the results and astrophysical implications in Section 3.6 and finish with our conclusions on the support for the presence of an upper mass gap in LIGO/Virgo's BH population in Section 3.7.

3.4. Methods

3.4.1. Hierarchical Inference

We use hierarchical Bayesian inference to simultaneously infer hyperparameters of the population distribution of the primary masses (m_1), mass ratios (q) and the redshifts (z) of observed BBHs. We assume the BBH merger rate $d\mathcal{R}$ over a given interval of masses and redshifts can be factored as:

$$\frac{d\mathcal{R}(m_1, q, z|\mathcal{R}_0, \Lambda)}{dm_1 dq} = \mathcal{R}_0 p(m_1|\Lambda) p(q|m_1, \Lambda) p(z|\Lambda) \tag{3.1}$$

with Λ the population hyperparameters and $\mathcal{R}_0$ the local ($z = 0$) merger rate. Under the condition that $p(m_1|\Lambda)$ and $p(q|m_1, \Lambda)$ are both normalized, and $p(z)$ chosen such that $p(z = 0) = 1$, integrating the merger rate density across all primary massses and mass ratios at a given z, returns the total BBH merger rate density at that redshift, $\mathcal{R}(z)$. The number density of BBH mergers can be related to the merger rate density by:

$$\frac{dN(m_1, q, z|\mathcal{R}_0, \Lambda)}{dm_1 dq dz} = \frac{dV_c}{dz}\left(\frac{T_{\text{obs}}}{1+z}\right)\frac{d\mathcal{R}(m_1, q, z|\mathcal{R}_0, \Lambda)}{dm_1 dq} \tag{3.2}$$

with V_c the comoving volume element and T_{obs} the total observing time with the factor of $1 + z$ converting source-frame time to detector-frame. Integrating the above number density across all primary masses, mass ratios and redshifts out to a maximum z_{max} returns the expected number of BBH mergers in the universe out to z_{max}. Given a set of data $\{d_i\}$ from N_{obs} gravitational wave events, we can calculate the posterior on Λ following e.g. Farr [82] and Mandel et al. [83]:

$$p\left(\mathcal{R}_0, \Lambda|\{d_i\}\right) \propto p(\Lambda)p(\mathcal{R}_0)e^{-\mathcal{R}_0\langle VT\rangle_\Lambda}\prod_{i=1}^{N_{\text{obs}}}\left[\int \mathcal{L}\left(d_i|m_1^i, q^i, z^i\right)\frac{d\mathcal{R}}{dm_1 dq dz}(\Lambda) dm_1 dq dz\right], \tag{3.3}$$

where $\mathcal{L}(d_i|m_1, q, z)$ is the single-event likelihood function used to infer each event's parameters, and $\langle VT\rangle_\Lambda$ is the average sensitive time-volume when assuming a population corresponding to hyper-parameters Λ. To estimate the $\langle VT\rangle_\Lambda$, we use the results of the LVC's injection campaign where the GWs from a fixed, broad population of sources were simulated, injected into real detector data, and searched for using the same analyses that produced GWTC-2 [1]. We use importance sampling over the detected simulated events to estimate $\langle VT \rangle_\Lambda$, marginalizing over the uncertainty in our estimate due to a finite number of simulated events, following [82]. We assume that repeated sampling of $\langle VT \rangle_\Lambda$ will follow a normal distribution (i.e. $\langle VT \rangle_\Lambda \sim \mathcal{N}(\mu(\Lambda), \sigma(\Lambda))$), with μ the raw importance sampled estimate and σ standard error. Now we define N_{eff}, the effective number of independent draws contributing to the importance sampled estimate, as $N_{\text{eff}} \equiv \frac{\mu^2}{\sigma^2}$, which we verify to be sufficiently high after re-weighting to a population (i.e. $N_{\text{eff}} > 4N_{\text{det}}$). After marginalizing over the uncertainty estimating the sensitive time-volume, we write the marginalized posterior as:

$$p\left(\mathcal{R}_0, \Lambda | \{d_i\}\right) \propto p(\Lambda) p(\mathcal{R}_0) \prod_{i=1}^{N_{\text{det}}} \left[\int \mathcal{L}\left(d_i | m_1^i, q^i, z^i\right) p(m_1^i, q^i, z^i | \Lambda) dm_1 dq dz \right] * \\ \mathcal{R}_0^{N_{\text{obs}}} e^{\frac{\mathcal{R}_0 \mu (\mathcal{R}_0 \mu - 2N_{\text{eff}})}{2N_{\text{eff}}}} \tag{3.4}$$

Finally, when using the commonly chosen log-Uniform prior on $\mathcal{R}_0$ [53], we can marginalize over the local merger rate, neglecting terms of $\mathcal{O}(N_{\text{eff}}^{-2})$ or greater: [82]

$$\log p\left(\Lambda | \{d_i\}\right) \propto \sum_{i=1}^{N_{\text{obs}}} \log \left[\frac{1}{K_i} \sum_{j=1}^{K_i} \frac{p(m_1^{i,j}, q^{i,j}, z^{i,j} | \Lambda)}{\pi(m_1^{i,j}, q^{i,j}, z^{i,j})} \right] - N_{\text{obs}} \log \mu + \\ \frac{3N_{\text{obs}} + N_{\text{obs}}^2}{2N_{\text{eff}}} + \mathcal{O}(N_{\text{eff}}^{-2}) \tag{3.5}$$

---

[1] For O3a we used the injection sets used by Abbott et al. [53], which can be found at https://dcc.ligo.org/LIGO-P2000217/public. For O1/O2 we used the mock injection sets used by Abbott et al. [52] which can be found at https://dcc.ligo.org/LIGO-P2000434/public

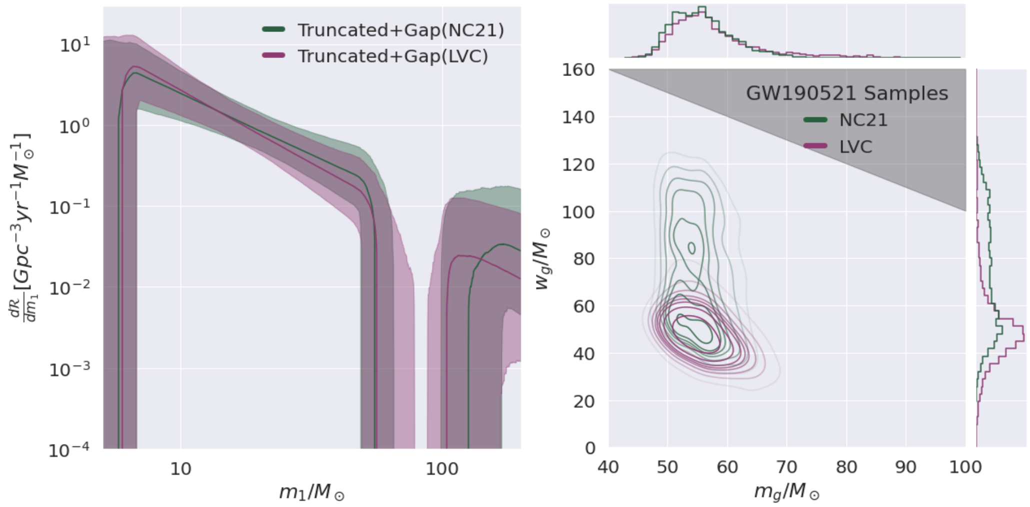

FIGURE 3.2. Posterior merger rate density (left) as function of primary mass inferred with the mass gap imposed on top of the Truncated model from Abbott et al. [53]. Solid curves shown the median posterior sample, while the shaded regions show the 90% credible level. 1-d and 2-d marginal posterior samples (right) of the two mass gap parameters, the lower edge, m_g, and the width, w_g, both with uniform agnostic priors over the range shown. The contour lines enclose 10-80% of the posterior. The grey region shows where our model reduces to the Truncated model with maximum mass at m_g. Results are shown using both the GW190521 posterior samples reported by the LVC (purple) and Nitz and Capano [49] (green).

In the last expression we further approximated the inner integral over the individual event parameters m_1^i, q^i, z^i with importance sampling over K_i single-event posterior samples generated from inference with prior $\pi(m_1^{i,j}, q^{i,j}, z^{i,j})$. To calculate marginal likelihoods and draw samples of the hyper parameters from the hierarchical posterior distribution shown in equation 3.5, we use the Bilby [84, 85] and GWPopulation [86] Bayesian inference software libraries with the Dynesty dynamic nested sampling algorithm [87].

3.4.2. Parameterized Mass Gap Model

We build on models used in Abbott et al. [53], Abbott et al. [52], and Fishbach et al. [88] for the mass ratio and redshift of our population, specifically

$p(z|\gamma) \propto (1+z)^{\gamma}$ and $p(q|m_1, m_{\min}, \beta_q) \propto q^{\beta_q}\Theta(qm_1 - m_{\min})\Theta(m_1 - qm_1)$ with Θ denoting the Heaviside step function ensuring m_2 is within the range $[m_{\min}, m_1]$. We choose to neglect a population model for the spins, assuming that their population follows the uniform and isotropic prior used in each event's initial analysis. For the primary mass distribution we use two different models presented in Abbott et al. [53], the Truncated and PowerLaw+Peak models. We choose to build upon the Truncated model for its simplicity and the Powerlaw+Peak model since Abbott et al. [53] found it to have the highest marginal likelihood of the models used. Additionally, it is important to include the peak (Powerlaw+Peak), as it was motivated to model the pileup of events due to PPISN mass loss which is expected from the same processes predicting the upper mass gap [67]. We then introduce a mass gap into both the primary and mass ratio distributions by enforcing that neither component mass can lie within the gap, which we parameterize with the location of the lower edge m_g and the width of the gap w_g:

$$p(m_1|m_g, w_g, \Lambda) \propto \begin{cases} 0 & m_g \leq m_1 \leq m_g + w_g \\ p(m_1|\Lambda) & \text{otherwise} \end{cases} \tag{3.6}$$

$$p(q|m_g, w_g, m_1, \Lambda) \propto \begin{cases} 0 & m_g \leq m_1 q \leq m_g + w_g \\ p(q|m_1, \Lambda) & \text{otherwise} \end{cases} \tag{3.7}$$

Our model prescribes a zero-rate within the mass gap, which might be expected if the entire population of sources is formed through stellar collapse. We enforce an overall maximum BH mass of $200\mathrm{M}_{\odot}$ so that if the upper edge of the gap is not constrained (i.e. $m_g + w_g \geq 200\mathrm{M}_{\odot}$), it is equivalent to the underlying primary mass model with a maximum mass at m_g. The hyperparameters' descriptions according to each of the models used along with their corresponding priors can be found in Table 3.1.

3.4.3. Injections and Sensitivity Estimates

The injection sets reported and used by the LVC in Abbott et al. [53] only include simulated signals with source frame masses up to $100\mathrm{M}_{\odot}$. However, since NC21 found GW190521 to have posterior support for its primary source frame mass to be up to $180\mathrm{M}_{\odot}$, we want to probe this region of parameter space. To prevent inferring an artificially high merger rate above the gap, our mass gap models are chosen to enforce the same powerlaw index in the region above the gap as below. This fixes the normalization above the gap based on the powerlaw fit below the gap that is influenced by the events which have $m_1 < 60\mathrm{M}_{\odot}$, which is the majority. This is also in-line with the expectation that very massive BHs can also be produced through stellar evolution, and thus come from the same stellar population as below the gap [89]. We additionally fit our population models with the same injection set truncated to only include injections with masses up to $80\mathrm{M}_{\odot}$, and found that it did not bias our results.

3.5. Results

We fit our population models to the 46 definitive BBH mergers in GWTC-2 (i.e., excluding GW170817, GW190425, GW190814, and GW190426_15215 [28, 90, 91]), and since we are focused on the details of the high-mass population we neglect the low-mass smoothing feature of the models used in Abbott et al. [53] for simplicity. Fig. 3.2 (left) shows the inferred merger rate density as a function of primary mass, when using the gap model on top of the underlying Truncated model, in which there is clear inference of a mass gap. When using the Truncated model, the choice of GW190521 posterior samples (either from the LVC or NC21) does not significantly affect the outcome, but the gap is inferred to be narrower and with an upper edge at lower masses when using the LVC samples. Fig 3.2 (right) shows the 1-d and 2-d marginal posterior distributions for the two gap parameters. Here we can see that the posterior distribution for the width of the gap is less constrained using NC21 samples, but both cases show little support for a zero-width gap. Using the LVC GW190521 samples we find support for the Truncated+Gap over Truncated model with Bayes factor $\ln\mathcal{B} = 2.79$, with lower and upper bounds at $55.12^{+7.54}_{-4.38}\mathrm{M}_{\odot}$, and $103.74^{+17.01}_{-6.32}\mathrm{M}_{\odot}$ respectively. While the gap model is clearly favored in that comparison, when using samples from NC21's for GW190521 the gap is more clearly favored with $\ln\mathcal{B} = 6.5$, with lower and upper bounds at $55.33^{+5.21}_{-4.21}\mathrm{M}_{\odot}$, and $126.03^{+30.25}_{-22.65}\mathrm{M}_{\odot}$.

Fig 3.3 (left) shows the inferred merger rate density as a function of primary mass when imposing a gap on the most favored mass model in Abbott et al. [53], Powerlaw+Peak. When including the Gaussian peak in our primary mass distribution, the support for the upper mass gap significantly reduces, regardless of which GW190521 samples are used. We find log-Bayes factors for inclusion of the gap to be $\ln\mathcal{B} = -0.5$ and $\ln\mathcal{B} = 0.5$ when using the LVC, and Nitz and Capano [49] GW190521 posterior samples respectively. Fig. 3.3 (right) shows the 1-d and 2-d marginal posterior distributions for the gap parameters, which show poorer constraints on the gap in the Powerlaw+Peak+Gap model relative to Powerlaw+Gap. In this case, both choices of posterior samples show support for a zero-width gap, as reflected in the Bayes factors.

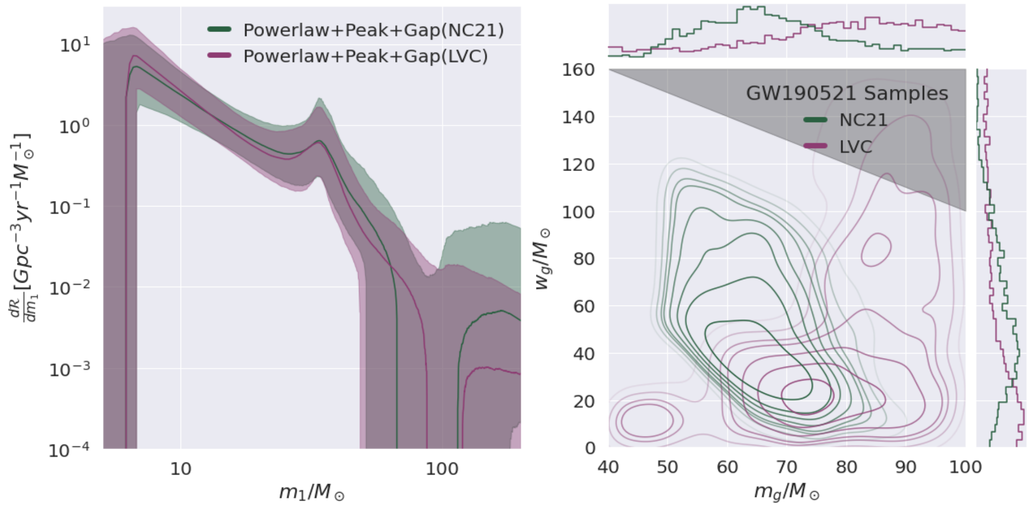

FIGURE 3.3. Posterior merger rate density (left) as function of primary mass inferred with the mass gap imposed on top of the Powerlaw+Peak model from Abbott et al. [53]. Solid curves shown the median posterior sample, while the shaded regions show the 90% credible level. 1-d and 2-d marginal posterior samples (right) of the two mass gap parameters, the lower edge, m_g, and the width, w_g, both with uniform agnostic priors over the range shown. The contour lines enclose 10-80% of the posterior. The grey region shows where our model reduces to the Powerlaw+Peak model with maximum mass at m_g. Results are shown using both the GW190521 posterior samples reported by the LVC (purple) and Nitz and Capano [49] (green).

Model	LVC	NC21
TRUNCATED	-4.98	-7.99
TRUNCATED+GAP	-2.20	-1.51
TRUNCATED+GAP (informed)	-0.87	0.0
POWERLAW+PEAK	0.0	-1.93
POWERLAW+PEAK+GAP	-0.57	-1.35
POWERLAW+PEAK+GAP (informed)	-1.05	-0.95

TABLE 3.2. Log Bayes factors for the models analyzed in this work, shown relative to the most favored model in each column. The two columns show results with the LVC reported GW190521 parameter estimation samples vs. those reported by NC21.

3.6. Discussion

Our results are inconclusive about the existence of a high-mass mass gap. While a gap is clearly inferred when using a pure power-law model of the population, adding a Gaussian peak to the mass distribution washes away the need for the gap. Furthermore, differences in parameter estimates with different priors and waveforms gives rise to different inferences on the gap parameters if a gap indeed exists. These results are summarized in Table 3.2 through Bayes factors comparing the marginal likelihood of each model to that of the model with highest marginal likelihood (which therefore has $\ln\mathcal{B} = 0$). The Bayes factors for LVC and NC21 parameter estimates are treated separately in the table. We also include Bayes factors for analyses with "informed" priors on the gap boundary, where we place Gaussian priors for $m_{\rm g,min}$ and $m_{\rm g,max}$ around centered on the approximate expected gap bounds (i.e. $p(m_{\rm g,min}) \sim \mathcal{N}(\mu = 55\mathrm{M}_\odot, \sigma = 10\mathrm{M}_\odot)$ and $p(m_{\rm g,max}) \sim \mathcal{N}(\mu = 120\mathrm{M}_\odot, \sigma = 20\mathrm{M}_\odot)$). With the smaller prior volume in these runs, the Bayes factors are higher than with the uninformed gap priors. Nevertheless, these Bayes factors do not increase enough to change the general finding of this work that the gap is favored with the Truncated model, but its existence is unclear when considering the Powerlaw+Peak model.

While the gap (if it exists) is difficult to resolve at present due to low number statistics, future detections will enable a finer look at the gap. Three extensions should be made when more high-mass detections are available, which we have eschewed for now due to the single event GW190521 driving the inference:

1. Allow a different mass ratio distribution for high-total-mass events than for low-total-mass.
2. Allow a non-zero rate in the gap, possibly with spins enforced to be near $\chi \sim 0.7$ to account for hierarchical mergers [69, 92, 93, 94].
3. Allow the merger rate normalization above the gap to be a free parameter.[2]

The location of the lower edge of the PISNe mass gap has been found to be insensitive to many variations in stellar physics, especially metallicities [61]. A metallicity independent feature in the BH mass spectrum can provide a "standard siren" that allows for independent measurements of redshift and luminosity distances to GW sources to directly measure the Hubble constant [95]. The lower edge of the PISNe mass gap has been found to be very sensitive to variations in the $^{12}\mathrm{C}(\alpha, \gamma)^{16}\mathrm{O}$ reaction rate, with some choices of rate pushing the lower bound up to $\sim 90\mathrm{M}_\odot$, illustrating that constraints on the lower bound can also be used to put constraints on nuclear physics going on inside stars' cores. [61, 63]. These astrophysical implications rely on the CO-BH mass relation from Farmer et al. [61] that predicts a pileup of BHs below the onset of the PISNe mass gap, implying

---

[2] This was not possible in this work due to LVC injections only reaching source frame component masses of $100\,M_\odot$, making the overall rate above that threshold unconstrained, as was discussed in Section 3.4.3

constraints on the gap lower bound that neglect a pileup could not be reliably used to constrain nuclear physics. The upper edge of the mass gap, is currently not well constrained, but Ezquiaga and Holz [80] argues LIGO/Virgo (at A+ sensitivity) will be sensitive to BBH's with component masses that could lie above the PISNe gap. Future constraints on the upper edge may also provide a novel probe of physics beyond the standard model [89].

3.7. Conclusions

Black holes formed through stellar collapse are expected to have a gap in their mass spectrum from $\sim 55\mathrm{M}_{\odot}-120\mathrm{M}_{\odot}$. We assess the support, or lack thereof, for the existence of such a gap in the GWTC-2 catalog, using two parameterized black hole binary merger population models. Our population models build on the TRUNCATED and POWERLAW+PEAK models previously fit to these catalogs, and explicitly allow for a zero-rate mass gap with a population of black holes above the gap. Our analyses also consider two separate inferences of GW190521 parameters, one from the LVC and the other from Nitz and Capano [49]. We find that the results of our inference regarding the existence of a gap are contingent in part on the choice of population model and GW190521 parameter estimation results.

If a pure power law is used to describe the distribution of primary masses, we infer a mass gap from $56.12^{+7.54}_{-4.38}\mathrm{M}_{\odot}$ to $103.74^{+17.01}_{-6.32}\mathrm{M}_{\odot}$, however if the data support more unequal masses for GW190521 as suggested in Nitz and Capano [49], we infer a mass gap from $55.33^{+5.21}_{-4.21}\mathrm{M}_{\odot}$ to $126.03^{+30.25}_{-22.65}\mathrm{M}_{\odot}$. When using a power law with an additional Gaussian component, we no longer find significant support for a zero-rate mass gap. This does not, however, imply the nonexistence of a mass gap due to PISNe but points towards there being a secondary population of BHs that

LIGO/Virgo is observing not formed through isolated stellar evolution. Future studies may be able to distinguish between these multiple formation channels in part by looking for a zero-rate gap in BH sub populations while additionally using informed constraints on expected properties that a hierarchically formed population of BHs would have.

3.8. Acknowledgements

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (`https://www.gw-openscience.org/`), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. This work benefited from access to the University of Oregon high performance computer, Talapas. This material is based upon work supported in part by the National Science Foundation under Grant PHY-1807046 and work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

CHAPTER IV

AIN'T NO MOUNTAIN HIGH ENOUGH: SEMI-PARAMETRIC MODELING OF LIGO-VIRGOS BINARY BLACK HOLE MASS DISTRIBUTION

4.1. Summary

After GWTC-2 the population of GW sources were beginning to be uncovered, revealing a possible peak or break in the mass distribution at $\sim 35 - 40\mathrm{M}_{\odot}$, found with simple parametric models, while more flexible non-parametric modeling was beginning to show more structure [51, 96]. In this chapter we present work published as Edelman et al. [97] which introduces a novel semi-parametric method of inferring the mass distribution of BBHs in GWTC-2. Out semi-parametric approach allows us to gain the advantages of both parametric and non-parametric modeling by using a flexible non-parametric modulation to an underlying simpler parametric description. We construct this model with cubic splines, similarly to the method used in Chapter 2, and find considerable support for the peak at $\sim 35\mathrm{M}_{\odot}$ along with signs of a peak at lower masses, $\sim 10\mathrm{M}_{\odot}$, validating hints of structure in the mass distribution. This model presented would be later used in the LVK collaboration's population analysis of the next catalog, GWTC-3 [1, 51], finding more significant signs of structure, especially at $10\mathrm{M}_{\odot}$. Since these papers came out there have been additional theoretical work to interpret this low mass peak in the mass distribution [98], simulation studies to calibrate significance of structure [99], as well as other groups borrowing this model to construct analyses of their own [100].

4.2. Abstract

We introduce a semi-parametric model for the primary mass distribution of binary black holes (BBHs) observed with gravitational waves (GWs) that applies a cubic-spline perturbation to a power law. We apply this model to the 46 BBHs included in the second gravitational wave transient catalog (GWTC-2). The spline perturbation model recovers a consistent primary mass distribution with previous results, corroborating the existence of a peak at $35\mathrm{M}_\odot(> 97\%$ credibility) found with the Powerlaw+Peak model. The peak could be the result pulsational pair-instability supernovae (PPISNe). The spline perturbation model finds potential signs of additional features in the primary mass distribution at lower masses similar to those previously reported by Tiwari and Fairhurst [96]. However, with fluctuations due to small number statistics, the simpler Powerlaw+Peak and BrokenPowerlaw models are both still perfectly consistent with observations. Our semi-parametric approach serves as a way to bridge the gap between parametric and non-parametric models to more accurately measure the BBH mass distribution. With larger catalogs we will be able to use this model to resolve possible additional features that could be used to perform cosmological measurements, and will build on our understanding of BBH formation, stellar evolution and nuclear astrophysics.

4.3. Introduction

The LIGO-Virgo Collaboration's second catalog of compact object mergers has shown that the universe is teeming with colliding compact objects with a variety of masses and spins [101]. In contrast to the 11 sources reported in the first LIGO-Virgo Collaboration (LVC) catalog [GWTC-1 9], the second catalog

[GWTC-2 45] contains 50 sources, enabling a deeper look into the formation environments of compact object binaries. The sources in GWTC-2 include two binary neutron stars (BNSs) [28, 90], 46 binary black holes (BBHs), and two neutron star black hole (NSBHs) candidates [91]. The 46 confirmed BBHs observed in GWTC-2 include the first clear evidence of an asymmetric mass binary, potentially the least massive black hole known, and the most massive stellar mass black hole to date [46, 47, 91, 102]. With this large catalog of BBH mergers, one can now begin to robustly infer the properties of the astrophysical BBH distribution in addition to each individual event properties [52, 53].

Prior to the release of GWTC-2, the inferred mass distribution for the more massive (primary) components in mergers was thought to be consistent with a declining power law that cuts off at $\sim 45 \mathrm{M}_\odot$ [52, 66]. When analyzing the BBH primary mass distribution including events in GWTC-2, Abbott et al. [53] found that a truncated power law is no longer consistent with the additional observations. The primary mass distribution was found to have some feature at $\sim$ 35–40 $\mathrm{M}_\odot$, which was best described by either a break to a steeper power law or a with the addition of a peak. The presence of a peak in the primary mass distribution in this mass range is not surprising: it would be expected if we are witnessing effects of pulsational pair-instability supernovae (PPISNe) [67]. Massive stars that are too light to be fully disrupted by a pair-instability supernova (PISN) can shed large amounts of mass in a series of explosive pulsations before collapsing to a black hole [57, 58, 61]. This process leads to a wide range of initial stellar masses that map onto remnant black holes with masses $30\,\mathrm{M}_\odot \leq m_{\mathrm{BH}} \leq 45\,\mathrm{M}_\odot$ [60, 62, 103]. GWTC-2 also includes more massive binaries than previously observed, most notably GW190521 [46, 47]. Both component black holes of GW190521 have masses that pose a challenge to the theoretical prediction that pair-instability (PI) would forbid isolated stellar evolution from creating remnant black holes with masses from ~ 50–$125\,\mathrm{M}_{\odot}$ [54, 55, 56, 59]. There is some evidence that GW190521 could be a mass gap straddling binary or the result of other physical processes that get around the conflict with PISN theory [48, 49, 50, 71, 72, 73, 77, 78, 104]. However, the presence of these high mass component black holes could also point towards there being a contribution to the observed population of BBHs detected by LIGO/Virgo, that formed in a way that avoids PI. These formation possibilities include hierarchical mergers in dense stellar environments, relativistic accretion onto heavy BHs in active galactic nuclei disks, isolated binary evolution of low-metallicity Population II stars, or even the presence of new physics beyond the standard model [64, 69, 70, 74, 75, 94, 105, 106, 107].

Incorrectly inferring the BBH mass distribution has been shown to significantly bias both estimates of merger rates and the stochastic gravitational wave background amplitude [67]. Additionally the effects of PI can imprint features onto the mass distribution such as a high mass cutoff in the mass distribution (PISN), or a possible a pileup of mergers at masses just below the cutoff (PPISN). Resolving either of these features can provide a mass scale, calibrated across cosmic time, that enables measurements of the redshift-luminosity-distance relation to infer cosmological parameters [95]. As catalogs of GWs from BBHs grow in size [108], we will be able to infer the BBH mass distribution with greater fidelity to determine if there is presence of additional structure beyond a power law. Such structure could yield insights about the nature of what environments BBHs form in and how they are connected to the rich fields of stellar evolution and nuclear astrophysics [61, 63, 109, 110].

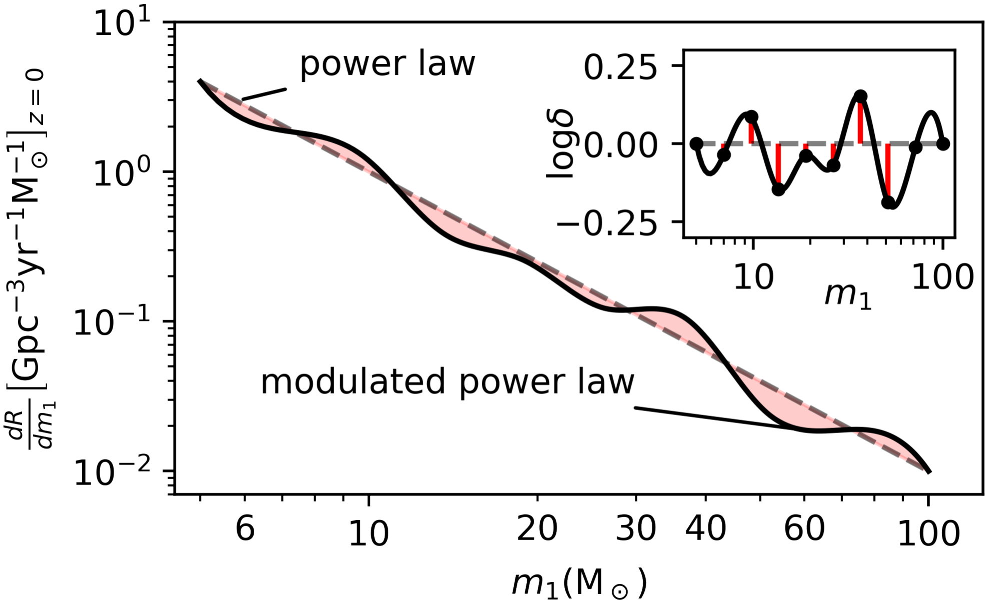

FIGURE 4.1. Sketch description of the spline perturbation primary mass model. The inset shows the interpolated cubic spline perturbation function for the plotted modulated power law.

Bayesian non-parametric models provide a useful data-oriented approach to modeling when one has little information or prior knowledge about the structure of a set of data. These approaches provide little to no constraints on the functional form imposed by the model and instead use very flexible functions that have large prior support for a wide variety of unknown densities. Non-parametric modeling has been widely applied across different areas of GW Astrophysics, including modeling deviations from GR waveform models, modeling the noise power spectrum of detectors, modeling the calibration of the detectors, and creating surrogate models for faster waveform execution [8, 31, 111, 112].

In this work, we approach the mass spectrum from a data-driven perspective, using a semi-parametric method rather than the low-dimensional parametric models used in Abbott et al. [52, 53]. Our semi-parametric method is complementary to both parametric and fully non-parametric approaches [113, 114] by incorporating a simple parametric description of the large-scale structure (i.e. a power law) with an additional non-parametrically modeled component on top. This approach can aid in searching for generic deviations to the underlying parametric descriptions that could be the result of astrophysical processes. Since non-parametric approaches make few assumptions on the form that the underlying distributions may take, our model minimizes biases to the structure such deviations could take. We expect a large fraction of stellar mass BHs to form at the end of life of massive stars, which motivates our choice of a power law form of the BBH primary mass distribution following a similar functional form to the stellar initial mass function [115]. We therefore reconstruct the primary mass distribution with the TRUNCATED power law model [53, 66], in which we modulate with a non-parametric perturbation. This method takes advantage of using a simple parametric form to capture the majority of the structure in the primary mass distribution while the perturbation function can find data-informed deviations from the power law. In Section 4.4 we describe our semi-parametric perturbation population model, and in Section 4.5 we present and discuss the inferred properties of the primary mass distribution when analyzing all 46 BBHs in GWTC-2. Finally, we explore possible interpretations of our results and conclude in Section 4.6.

4.4. Spline Perturbation Model

We use a hierarchical Bayesian inference framework to infer the properties of the astrophysical distribution of BBHs that incorporates software injections to estimate selection effects [82, 83, 116]. This procedure is described in detail in Appendix 4.8.1. In order to capture both the overall trends and any sharper features that may be in the primary mass distribution, we modulate a base parametric hyper-prior on primary mass, $p(m_1|\Lambda)$, by a highly flexible perturbation function – in this case, a cubic spline. We choose the simplest of previously used parametric models as our underlying mass distribution, $p(m_1|\Lambda)$, which is described by a power law in both primary mass and mass ratio with a sharp low and high mass cutoff [52, 53, 66]. This model was referred to as the TRUNCATED model in Abbott et al. [53]. While the TRUNCATED model alone does not describe GWTC-2 well [53], it captures the majority of the large-scale structure found in the primary mass distribution. For our underlying description, we extend the TRUNCATED model to allow for a tapering of the distribution at low masses following the same form used for the POWERLAW+PEAK model described in Talbot and Thrane [67] and Abbott et al. [53]. Figure 4.1 shows an illustration of our spline perturbation model on top of a power law without any mass cutoffs or tapering. We multiplicatively apply perturbations to the underlying distribution as:

$$p_{\mathrm{spline}}(m_1|\Lambda, \{m_i, f_i\}) = k * p(m_1|\Lambda) \exp(f(m_1; \{m_i, f_i\})) \tag{4.1}$$

In the above equation, k is a normalization factor found by numerically integrating $p_{\mathrm{spline}}(m_1|\Lambda, \{m_i, f_i\})$ over the entire range of primary masses, and $f(m_1; \{m_i, f_i\})$ is the perturbation function modeled as a cubic spline that is interpolated between n knots placed in m_1 space. These knots are denoted by their locations in mass space, $\{m_i\}_{i=1}^{n}$, and their heights at each knot, $\{f_i\}_{i=1}^{n}$. For readability, we hereafter drop explicit dependence of f on $\{m_i, f_i\}$ unless needed.

We fix the locations of each knot to be linear in log m_1 space from 5–100 $\mathrm{M}_\odot$ and restrict the perturbations to zero at the minimum and maximum knots. With these restrictions our spline model adds $n - 2$ extra hyper-parameters to the underlying primary mass model we are perturbing, one for each of the inner knots' heights. We log-space the knots and perturb our underlying model with the multiplicative factor, $\exp(f(m_1))$, to reflect the wide range in orders of magnitude of the underlying power law. We then impose Gaussian priors on the knot heights $\{f_i\}$ centered at 0 and with standard deviations, σ_{knot}. Our model then has two specifications which control the resolution (n) and the magnitude (σ_{knot}) of perturbations the model is sensitive to. We discuss the effect of changing these model settings on our prior assumptions and motivate the particular choices we made for this work in Appendix 4.8.2.

In addition to the primary mass distribution, we simultaneously fit for the mass ratio and redshift distributions, without any spline perturbations applied. We apply a power law distribution for the mass ratio as $p(q|m_1, m_{\mathrm{min}}, \beta_q) \propto q^{\beta_q}\Theta(qm_1 - m_{\mathrm{min}})\Theta(m_1 - qm_1)$, with Θ denoting the Heaviside step function that ensures m_2 is within the range $[m_{\mathrm{min}}, m_1]$. We then fit for the evolution of the merger rate with redshift also with a power law such that $p(z|\lambda) \propto \frac{dV_c}{dz}\frac{1}{1+z}(1+z)^{\lambda}$, where dV_c is the co-moving volume element [52, 53, 88, 117]. We do not fit for a population prior on the BBH spins, and assume the spin prior used for individual event parameter estimation in Abbott et al. [45], which is uniform in component spin magnitudes and isotropic in component spin orientations. We enumerate each of the model's hyper-parameters and corresponding hyper-prior distributions used in this work in tables 4.1 and 4.2 found in Section 4.8.1.

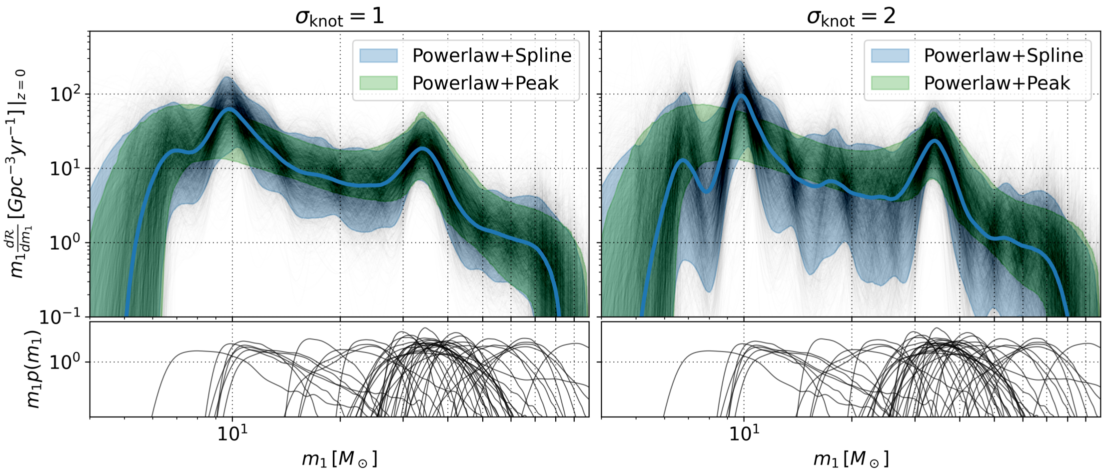

FIGURE 4.2. We plot the differential merger rate, $\frac{d\mathcal{R}}{d\log m_1} = m_1 \frac{d\mathcal{R}}{m_1}$ as a function of primary mass (top row) for the combined spline model marginalized over the 10, 15, and 20 knot models in blue. The solid line shows the median while the shaded region shows the 90% credible interval and the 90% credible interval found from the POWERLAW+PEAK model in green. The black traces show 1000 draws from the combined spline model posterior and we plot kernel density estimates (KDEs) of the posterior samples of primary source frame mass for each of the 46 BBHs in GWTC-2. We plot $m_1 p(m_1)$ on a log-scaled y-axis with a Gaussian KDE approximating $p(m_1)$ for each event. These posterior samples are not re-weighted to a population and come directly from the accompanying data release to Abbott et al. [45].

4.5. Results

4.5.1. Astrophysical BBH Primary Mass Distribution

We use a catalog containing each of the 46 BBH mergers reported in Abbott et al. [45] in which we perform a hierarchical Bayesian analysis to infer the astrophysical mass spectrum and merger rate evolution with redshift, as described in Appendix 4.8.1. We perform multiple iterations of our semi-parametric model with different numbers of knots and both a "conservative" ($\sigma_{\mathrm{knot}} = 1$) and "wide" ($\sigma_{\mathrm{knot}} = 2$) prior width on the knots. For both cases of prior width we additionally do a post hoc "marginalization" over the number of knots by combining posterior draws weighted according to the ratios in marginal likelihoods. Explicitly we take $\left\lfloor N_{\mathrm{min}} \frac{\mathcal{Z}_n}{\mathcal{Z}_{\mathrm{max}}} \right\rfloor$ samples from each inference where N_{min} is the minimum number of samples from each posterior, $\mathcal{Z}_n$ the marginal likelihood of inference with n knots, and $\mathcal{Z}_{\mathrm{max}}$, the maximum marginal likelihood of the combined posteriors. In Figure 4.2 we plot the posterior merger rate density as a function of primary mass (top row) for our combined spline model (combining 10, 15 and 20 knot models), compared to the Powerlaw+Peak model. The most prominent feature in the primary mass distribution is the apparent peak at $\sim 35 \mathrm{M}_{\odot}$, similar to the peak found at the same mass by the Powerlaw+Peak model [53, 67].

In addition to the peak at $\sim 35 \mathrm{M}_{\odot}$, there are signs of additional features — albeit less significant — at lower masses. We find signs of an inflated rate of mergers with primary masses $\sim 10 \mathrm{M}_{\odot}$ and reduced rate around $\sim 7.5 \mathrm{M}_{\odot}$, when compared with the power law structure. The model is less certain about the low-mass features as there are only a few events with support for $m_1 < 10\,\mathrm{M}_{\odot}$. The dearth of observed low mass BBHs, combined with their small sensitive volume,

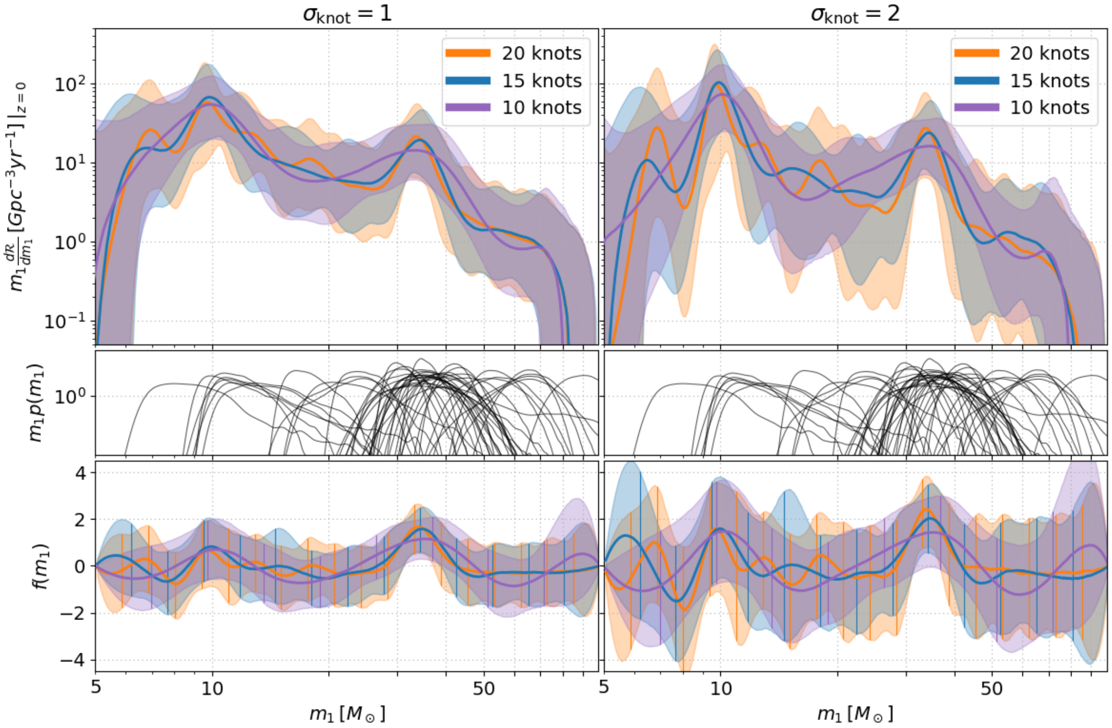

FIGURE 4.3. The median (solid) and 90% credible interval (shaded) of the inferred differential merger rate density as a function of primary mass (top row) with spline models with 10 (purple), 15 (blue), and 20 (orange) knots. We show the conservative prior case ($\sigma_{\rm knot} = 1$) in the left column and the wide prior case ($\sigma_{\rm knot} = 2$) on the right. The middle row shows kernel density estimates (KDEs) of the posterior samples of primary source frame mass for each of the 46 BBHs in GWTC-2. We plot $m_1 p(m_1)$ on a log-scaled y-axis with a Gaussian KDE approximating $p(m_1)$ for each event. These posterior samples are not re-weighted to a population and come directly from the accompanying data release to Abbott et al. [45]. The bottom row shows the median (solid) and 90% credible intervals (shaded) of the inferred perturbation function, $f(m_1)$, for each choice of n, with the vertical lines showing the locations of the spline knots.

significantly inflates our uncertainty at the low end of the mass distribution. In Figure 4.3, we show the results inferred by each of the 10, 15, and 20 knot spline models individually, showing that the spline model consistently finds common features regardless of the number and location of knots. The combined spline model based on the marginal likelihoods is mostly comprised of the 15 knot result because the ratios of marginal likelihoods favor 15 knots over 20 at 2:1 odds and 15 over 10 at 3:1. The spline models are best constrained in the regions of over/under densities discussed above and much less certain (prior dominated) in regions between the features where the parametric component (i.e. power law) can fully explain the trend. The bottom row of Figure 4.3 shows the perturbation function, $f(m_1)$, inferred from the different spline models. While there are some differences between knot choices due to the different length scales, they are all in agreement when taking into account the uncertainties and each consistently recovers similar merger rates and perturbations at both the $10\mathrm{M}_{\odot}$and $35\mathrm{M}_{\odot}$peaks. In Figure 4.4 we plot the posterior distribution of the perturbation function sliced at the approximate masses of the three, ($f(m_1 = 7.5\,\mathrm{M}_{\odot})$, $f(m_1 = 10\,\mathrm{M}_{\odot})$ and $f(m_1 = 35\,\mathrm{M}_{\odot})$). We find similar posteriors on the perturbation at these three mass regions, across the models varying the number of knots and the spline prior width. We calculate the percentile where $f = 0$ falls in the posterior distribution for each of these three cuts, which would be near 50% in the presence of no deviations to the power law or equivalently for draws from the spline model prior. The percentiles of zero perturbation for the combined model with the conservative (wide) priors are 70.8% (72.6%), 19.4% (13.1%), and 0.8% (2.5%) at $7.5\mathrm{M}_{\odot}$, $10\mathrm{M}_{\odot}$, and $35\mathrm{M}_{\odot}$, respectively.

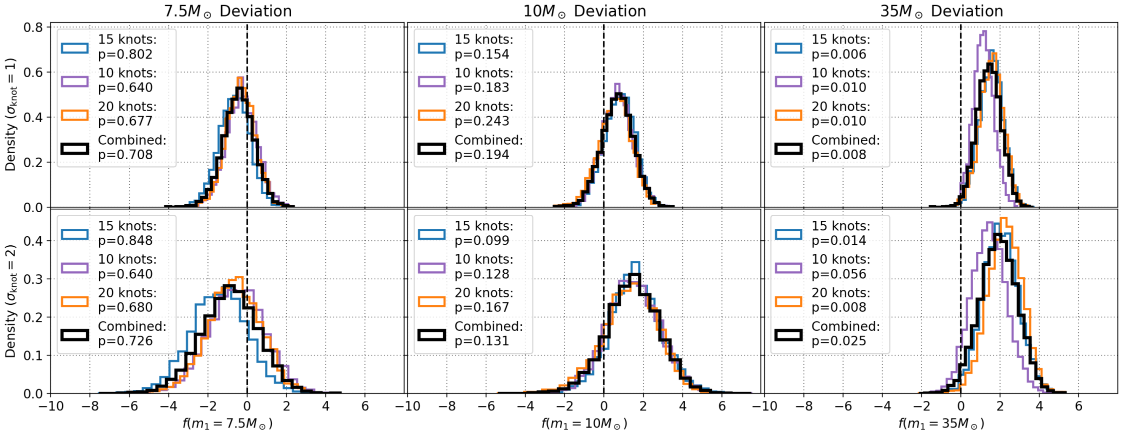

FIGURE 4.4. The posterior distribution of $f(m_1)$ at sliced at the three most apparent inferred perturbations in the posterior which roughly lie at $\sim$ 7.5M$_\odot$ (left column), $\sim$ 10M$_\odot$ (middle column), and $\sim$ 35M$_\odot$ (right column). We show the posteriors for 10 (purple), 15 (blue), and 20 (orange) knots and for both cases of prior width: $\sigma_{\rm knot} = 1$ (top row) and $\sigma_{\rm knot} = 2$ (bottom row). We additionally show the result when combining the models (weighted by their marginal likelihoods) across the three choices number of nodes in black. We report the quantile in which $f = 0$ falls for each of the models and perturbation regions knots in each legend.

The presence of the $\sim$ 35M$_\odot$ feature was previously found and reported in Abbott et al. [53] as being either a peak (likely due to the PPISN pileup [67]) or a break to a steeper power law. The POWERLAW+PEAK model returned the highest marginal likelihood of parametric mass models considered in Abbott et al. [53], but was only favored with roughly 3:1 odds. Due to the inherent nature of the spline perturbation model, we would be more likely to find features that look like peaks rather than a power law break in the distribution. We additionally fit for spline perturbations on top of the BROKENPOWERLAW model, which found little to no support for two different power law slopes and recovered a nearly identical primary mass distribution to what was found when modulating a single power law. The low mass feature recovered by our spline model was not identified in Abbott et al. [53] because the models considered there did not have the flexibility to fit such features at low mass: they only include a smooth rise to the low-mass end of the power-law. This could be explained by additional structure that cannot be described by a smooth rise to a power law, coming from the upper edge of the proposed neutron star black hole mass gap. The flexibility of our semi-parametric approach enables us to find additional structure in the astrophysical mass distribution beyond what can be found with simpler toy models. While the spline perturbation model clearly finds structure beyond the power law around $\sim 35\mathrm{M}_\odot$ in the analyzed catalog of 46 BBH mergers, we cannot say for certain that the low mass feature is inherent to the astrophysical mass distribution. There is still the possibility that our model could be latching onto fluctuations in our data due to small number statistics. This possibility is reflected in the percentiles in Figure 4.4. The perturbation function at $7.5\mathrm{M}_\odot$ and $10\mathrm{M}_\odot$ does not rule out $f = 0$ at high credibility regardless of prior choices in the spline model, while in contrast, the perturbation at $35\mathrm{M}_\odot$ rules out $f(35\mathrm{M}_\odot) = 0$ at 97-99.4% credibility across each variation of spline model used. We investigate the possibility that these subsequent deviations from a power law could appear due to our model's systematics in Appendix 4.8.3. We report no signs of correlations between the successive perturbations, which would be expected if the spline function was imposing biases onto our inferred perturbations. With future larger catalogs of gravitational-wave sources, we will be able to further resolve these low-mass features to determine if they are indeed present in the astrophysical mass distribution or a reflection of the current small catalog size.

4.5.2. Posterior Predictive Checks

With the large flexibility that comes from taking non-parametric approaches to modeling, one must be careful in validating inferences, especially in cases with

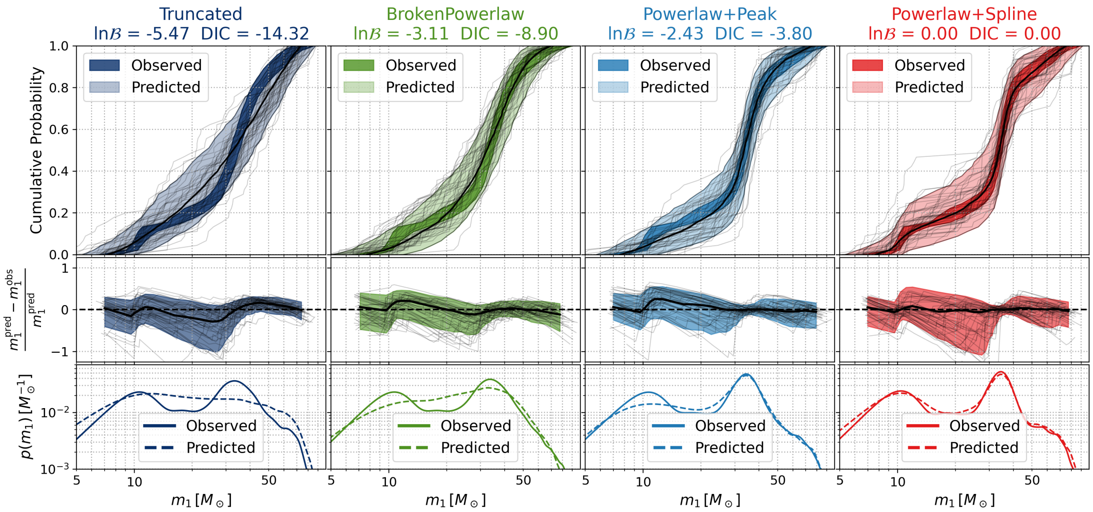

FIGURE 4.5. Posterior predictive checks for three of the parametric models used in Abbott et al. [53] and the spline perturbation model of this work. We show the spline model result with the highest marginal likelihood which was the 15 knot and $\sigma_{\rm knot} = 1$ model. The observed and predicted values for primary masses are generated by re-weighting either the injection set or the set of posterior samples for each BBH analyzed, for 500 draws from each models inferred posterior on the hyper-parameters, and then drawing 46 values from the re-weighted injections and a single fair draw from each of the 46 event re-weighted posterior samples. The top panel shows the CDF generated from these sets observed and predicted events for each of the 4 models, with the 90% credible levels enclosed by the bands, the median in dark black, and the thin black lines showing 50 of the 500 sets of 46 predicted events. The middle row uses the same set of predicted and observed events and the y-axis shows the relative error in predicted to observed mass ($(m_1^{\rm pred} - m_1^{\rm obs})/m_1^{\rm pred}$) as a function of $m_1^{\rm obs}$. The last row of plots shows the PDF of the top row averaged over the 500 draws from the posterior on the hyper-parameters for both sets of events.

small numbers of observations. One way we evaluate our semi-parametric model is through posterior predictive checks. We employ the same injection set used to estimate our search's selection effects mentioned in Appendix 4.8.1 to create mock detected populations under a given population described by a posterior on hyper-parameters. We then compare these mock populations to the observed population. To do this, we first re-weight the injections to our inferred population for $N_{\rm draw}$ draws of hyper-parameters from our population posterior, then take $N_{\rm obs}$ (46 for this BBH-only analysis) draws for each of the re-weightings. This generates $N_{\rm draw}$ sets of $N_{\rm obs}$ "Predicted" observations for a given population inference. Next, we re-weight the individual event posterior samples to the same inferred population for $N_{\rm draw}$ draws of hyper-parameters from the posterior. For each draw of hyper-parameters, we take a fair draw from each re-weighted event posterior to generate our corresponding $N_{\rm draw}$ sets of $N_{\rm obs}$ "Observed" observations for a given population inference. From this procedure we generated 500 sets of 46 "Observed" and "Predicted" catalogs, which we compare to each other in Figure 4.5 to confirm that our inferred population model predictions are consistent with the observations. We show the cumulative probability as a function of observed primary mass in the top row, the relative error in predicted primary masses to observed in the second row, and the last row shows the observed and predicted primary mass distributions averaged over all of the hyper-parameters inferred in our posterior. The colored bands in the top row of Figure 4.5 show that a model is inconsistent with observations when the dark "Observed" band extends outside of the lighter "Predicted" band. We see this behaviour at $\sim$ 40–50$\mathrm{M}_\odot$ for the TRUNCATED model which illustrates a conclusion from Abbott et al. [53]: the TRUNCATED model is inconsistent with the mergers in GTWC-2. The spline perturbation model is the only primary mass model that recovers both the low and high mass features seen in the observed distribution, but it does exhibit more uncertainty than other models in the regions between the $\sim 10\mathrm{M}_\odot$ and $\sim 35\mathrm{M}_\odot$ peak. When considering possible fluctuations due to small number statistics, the observations at low mass are still consistent with both the POWERLAW+PEAK and BROKENPOWERLAW models.

4.5.3. Astrophysical Implications

The BBH mass distribution is particularly well-suited to answering a wide range of astrophysical questions. In particular, the masses of detected events are relatively well-measured, and different channels of BBH formation result in different mass distributions [e.g. 109], implying that the formation history is encoded within these distributions. With the tens of detected events available now, disentangling the overlapping sub-populations in the full population (if they exist) is a challenge. Our perturbation model can be used to see if or where a distribution describing a single (dominant) sub-population or formation channel may fail to fully fit the data, which would provide evidence that there may be non-negligible contributions due to additional formation channels. The hints of structure we see at the low end of the mass distribution could point to such a superposition of multiple formation channels. Another factor that can affect the mass distribution is the physics of PISN or PPISN. Stellar evolution models describing mass loss and PPISN have uncertainties that can drastically change predictions on the masses at which the PPISN and PISN play a role. For example, choices in nuclear reaction rates within stellar cores can affect the BBH mass distribution [61, 63]. Our spline model enables us to measure these imprints of BBH formation in the observed distribution without enforcing the specific distribution shapes that are inherent to parametric models. Our findings corroborate the existence of a PISN "pile-up" in the 30 - 40 $M_\odot$ range, and we infer its shape without assuming a simple functional form. With more GW detections, further resolution of this peak with the spline model could offer insights into supernova physics.

4.6. Conclusions

Accurate estimation of the BBH mass distribution is paramount to getting accurate estimates of merger rates, the GW stochastic background, and false alarm rates for potential new triggers. Low-dimensional parametric models have the advantage of being easily interpreted but are limited in their flexibility and subject to a-priori expectations. We presented a semi-parametric approach to modeling the primary mass distribution of BBH mergers, using cubic splines that modulate a power law. We show that our flexible semi-parametric approach, when applied to the BBHs in GWTC-2, consistently recovers the previously reported excess of observed mergers near $\sim 35 M_\odot$, and shows potential signs of additional features in the low-mass end of the BBH distribution. These low-mass features beyond the power law structure, correspond with similar features found in the chirp mass distribution using a separate non-parametric approach based on a flexible Gaussian Mixture model [96]. We show through posterior predictive checks that the spline model is at least as good as the Powerlaw+Peak model at fitting the high mass structure in our catalog while having the flexibility beyond a smooth rise into a power law to capture the apparent excess at $10 M_\odot$. Structure in the mass distribution that could arise from many different astrophysical phenomena but if we are able to confidently identify either a high-mass cutoff or pileup in the mass distribution it is likely to be related to effects of PISN or PPISN. These two features can both be used as calibrated mass-scales to measure a redshift-luminosity-distance relation with which it is possible to infer cosmological parameters with [95].

Our semi-parametric approach has advantages compared to other fully non-parametric approaches modeling the BBH mass distribution [96, 113, 114]. The semi-parametric approach leverages the information learned from the parametric models to explain the majority of the structure, while reserving the flexibility to see where observations may start to diverge from previous inferences. This same method of applying cubic spline perturbations to simpler population models can be used on any of the other commonly modeled population distributions such as the mass ratio, spins or redshift evolution. With the relatively small catalog sizes currently available, structure, if present, would likely only appear in the best measured parameters. In future work, we plan to extend this method to incorporate multi-dimensional splines, that could uncover correlations between different parameters such as peaks in the mass distribution associated with a high spin magnitude. Such correlations would be a tell-tale sign of hierarchical mergers, for example [66, 69, 70, 92, 94]. Future work could also extend this method to more than just the BBH mass distribution, and allow for adaptive resolution splines that allow the knot locations to vary [8]. With additional observations of GWs associated with BNSs and NSBHs [118] our spline perturbation model is well suited to model the joint mass distribution of all GW-observed compact objects. This would complement the parametric model of [119], giving insights on the structure (or lack there of) of the "lower mass gap" that may exist between the heaviest neutron stars and lightest black holes.

4.7. Acknowledgements

We thank Will Farr for countless useful discussions related to this work and hierarchical modeling. This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (`https://www.gw-openscience.org/`), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. This work benefited from access to the University of Oregon high performance computer, Talapas. This material is based upon work supported in part by the National Science Foundation under Grant PHY-1807046 and work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

4.8. Additional Analysis Details

4.8.1. Hierarchical Inference

We use hierarchical Bayesian inference to simultaneously infer the population distributions of the primary masses (m_1), mass ratios (q) and the redshifts (z) of observed BBHs. For a set of hyper-parameters, Λ, and local ($z = 0$) merger rate density (units of mergers per co-moving volume per time), $\mathcal{R}_0$, we write the overall number density of BBH mergers in the universe as:

$$\frac{dN(m_1, q, z|\mathcal{R}_0, \Lambda)}{dm_1 dq dz} = \frac{dV_c}{dz}\left(\frac{T_{\mathrm{obs}}}{1+z}\right)\frac{d\mathcal{R}(m_1, q, z|\mathcal{R}_0, \Lambda)}{dm_1 dq} = \mathcal{R}_0 p(m_1|\Lambda)p(q|m_1, \Lambda)p(z|\Lambda) \tag{4.2}$$

Above, dV_c is the co-moving volume element [117] and T_{obs}, the observing time period that produced the catalog with the related factor of $1 + z$ converting this detector-frame time to source-frame. We assume a LambdaCDM cosmology using the cosmological parameters from Planck Collaboration et al. [120]. The redshift evolution of the merger rate follows $p(z|\lambda) \propto \frac{dV_c}{dz}\frac{1}{1+z}(1+z)^{\lambda}$. Integrating equation 4.2 across all masses, and up to some redshift, z_m, returns the total number of BBH mergers in the universe out to that redshift. Let $\{d_i\}$ represent a set of data from N_{obs} observed gravitational waves associated with BBH mergers. We model the merger rate as an inhomogenous point process and when imposing a log-uniform prior on the merger rate, we can marginalize over the merger rate to get the posterior distribution of our hyper-parameters, Λ [83, 111].

$$p\left(\Lambda|\{d_i\}\right) \propto \frac{p(\Lambda)}{\xi(\Lambda)^{N_{\text{obs}}}} \prod_{i=1}^{N_{\text{obs}}} \left[\int \mathcal{L}\left(d_i|m_1^i, q^i, z^i\right) p(m_1|\Lambda) p(q|m_1,\Lambda) p(z|\Lambda) dm_1 dq dz \right], \tag{4.3}$$

where, $\mathcal{L}(d_i|m_1, q, z)$, is the single-event likelihood function from each events original analysis, and $\xi(\Lambda)$ is the detection efficiency given a population distribution described by Λ. The procedure for calculating $\xi(\Lambda)$ is described in more detail below. The LVC reports out posterior samples for each observed event, with which we can use importance sampling to estimate the integrals above in equation 4.3. We replace the integrals with ensemble averages over K_i posterior samples associated with each event in the catalog:

$$p\left(\Lambda|\{d_i\}\right) \propto p(\Lambda) \prod_{i=1}^{N_{\text{obs}}} \left[\frac{1}{K_i} \sum_{j=1}^{K_i} \frac{p(m_1^{i,j}|\Lambda) p(q^{i,j}|\Lambda) p(z^{i,j}|\Lambda)}{\pi(m_1^{i,j}, q^{i,j}, z^{i,j})} \right] \tag{4.4}$$

Here j indexes the K_i posterior samples from each event and $\pi(m_1, q, z)$ is the default prior used by parameter estimations that produced the posterior samples for each event. In the analyses of GWTC-2 the default prior used is uniform in detector frame masses and Euclidean volume. The corresponding prior evaluated in source frame masses and redshift is $\pi(m_1, q, z) \propto m_1(1+z)^2 D_L^2(z)\frac{dD_L}{dz}$, where D_L is the luminosity distance.

To carefully incorporate selection effects to our model we need to quantify the detection efficiency, $\xi(\Lambda)$, of the search pipelines that were used to create GWTC-2, at a given population distribution described by Λ.

$$\xi(\Lambda) = \int dm_1 dq dz P_{\rm det}(m_1, q, z) p(m_1|\Lambda) p(q|m_1, \Lambda) p(z|\Lambda) \tag{4.5}$$

The above integral is not tractable since there is no analytic prescription for $P_{\rm det}(m_1, q, z)$, the detection probability of an individual event. To estimate this integral we use a software injection campaign where GWs from a fixed, broad population of sources are simulated, put into real detector data, and then run through the same search pipelines that were used to produce the catalog we are analyzing [1]. With these search results in hand, we use importance sampling to evaluate the integral in equation 4.5:

$$\xi(\Lambda) = \frac{1}{N_{\rm inj}} \sum_{i=1}^{N_{\rm found}} \frac{p(m_1^i|\Lambda) p(q^i|m_1, \Lambda) p(z^i|\Lambda)}{p_{\rm inj}(m_1^i, q^i, z^i)} \tag{4.6}$$

Where the sum indexes only over the $N_{\rm found}$ injections that were successfully detected out of $N_{\rm inj}$ total injections, and $p_{\rm inj}(m_1, q, z)$ is the reference distribution

---

[1] For O3a we used the injection sets used by Abbott et al. [53], which can be found at https://dcc.ligo.org/LIGO-P2000217/public. For O1/O2 we used the mock injection sets used by Abbott et al. [52] which can be found at https://dcc.ligo.org/LIGO-P2000434/public

from which the injections were drawn. Additionally, we follow the procedure outlined in Farr [82] to marginalize the uncertainty in our estimate of $\xi(\Lambda)$ due to a finite number of simulated events. We make the assumption that repeated sampling of $\xi(\Lambda)$ will follow a normal distribution with $\xi(\Lambda) \sim \mathcal{N}(\mu(\Lambda), \sigma(\Lambda))$, where the mean, μ, is the estimate from equation 4.6, while the variance, σ^2, is defined as:

$$\sigma^2(\Lambda) \equiv \frac{\mu^2(\Lambda)}{N_{\text{eff}}} \simeq \frac{1}{N_{\text{inj}}^2} \sum_{i=1}^{N_{\text{found}}} \left[\frac{p(m_1|\Lambda)p(q|m_1,\Lambda)p(z|\Lambda)}{p_{\text{inj}}(m_1,q,z)} \right]^2 - \frac{\mu^2(\Lambda)}{N_{\text{inj}}} \tag{4.7}$$

Above we define N_{eff} as the effective number of independent draws contributing to the importance sampled estimate, in which we verify to be sufficiently high after re-weighting the injections to a given population (i.e. $N_{\text{eff}} > 4N_{\text{obs}}$). We write the hyper-posterior marginalized over the merger rate and uncertainty in estimation of $\xi(\Lambda)$, neglecting terms of $\mathcal{O}(N_{\text{eff}}^{-2})$ or greater [82], as:

$$\log p\left(\Lambda|\{d_i\}\right) \propto \sum_{i=1}^{N_{\text{obs}}} \log \left[\frac{1}{K_i} \sum_{j=1}^{K_i} \frac{p(m_1^{i,j}|\Lambda)p(q^{i,j}|\Lambda)p(z^{i,j}|\Lambda)}{\pi(m_1^{i,j}, q^{i,j}, z^{i,j})} \right] - N_{\text{obs}} \log \mu + \frac{3N_{\text{obs}} + N_{\text{obs}}^2}{2N_{\text{eff}}} + \mathcal{O}(N_{\text{eff}}^{-2}) \tag{4.8}$$

We explicitly enumerate each of the models used in this work for $p(m_1|\Lambda)$, $p(q|m_1,\Lambda)$, and $p(z|\Lambda)$ along with their respective hyper-parameters and prior distributions in tables 4.1 and 4.2. To calculate marginal likelihoods and draw samples of the hyper parameters from the hierarchical posterior distribution shown in equatio n 4.8, we use the Bilby [84, 85] and GWPopulation [86] Bayesian inference software libraries with the Dynesty dynamic nested sampling algorithm [87].

Primary Mass Model Parameters			
Model	**Parameter**	**Description**	**Prior**
Truncated	α	slope of the powerlaw	U(-4, 12)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
	δ_m	low-mass smoothing scale (smoothes from $m_{\min}$ to $m_{\min} + \delta_m$)	U(0, 10)
BrokenPowerlaw	α_1	slope of the first powerlaw	U(-4, 12)
	α_2	slope of the second powerlaw	U(-4, 12)
	b	fraction between $m_{\min}$ and $m_{\max}$ where the power law break lies	U(0, 1)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
	δ_m	low-mass smoothing scale (smoothes from $m_{\min}$ to $m_{\min} + \delta_m$)	U(0, 10)
PowerLaw+Peak	α	slope of the powerlaw	U(-4, 12)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
	μ_p	mean of Gaussian peak ($M_\odot$)	U($20M_\odot$, $70M_\odot$)
	σ_p	width of the Gaussian peak ($M_\odot$)	U($0.4M_\odot$, $10M_\odot$)
	λ_p	fraction of BBH in the Gaussian component	U(0, 1)
	δ_m	low-mass smoothing scale (smoothes from $m_{\min}$ to $m_{\min} + \delta_m$)	U(0, 10)
PowerLaw+MutliPeak	α	slope of the powerlaw	U(-4, 12)
	$m_{\min}$	minimum mass cutoff ($M_\odot$)	U($2M_\odot$, $10M_\odot$)
	$m_{\max}$	maximum mass cutoff ($M_\odot$)	200 $M_\odot$
	$\mu_p, 1$	mean of the first Gaussian peak ($M_\odot$)	U($5M_\odot$, $40M_\odot$)
	$\sigma_p, 1$	width of the first Gaussian peak ($M_\odot$)	U($0.4M_\odot$, $10M_\odot$)
	$\lambda_p, 1$	fraction of BBH in the first Gaussian component	U(0, 1)
	$\mu_p, 2$	mean of the second Gaussian peak ($M_\odot$)	U($5M_\odot$, $40M_\odot$)
	$\sigma_p, 2$	width of the second Gaussian peak ($M_\odot$)	U($0.4M_\odot$, $10M_\odot$)
	$\lambda_p, 2$	fraction of BBH in the second Gaussian component	U(0, 1)
	δ_m	low-mass smoothing scale (smoothes from $m_{\min}$ to $m_{\min} + \delta_m$)	U(0, 10)

TABLE 4.1. This table enumerates all the hyper-parameters for the parameteric mass distributions, their descriptions, and chosen priors for this work for each respective population model we use. The Truncated model is extended from the version used in Abbott et al. [53] to have the option of a low-mass taper of the same form as the Powerlaw+Peak model. Note that we do not describe a spin population model in this table since in this work we are not inferring a hyper-prior on the spins and instead assume they are described by the default (uniform in component magnitudes, isotropic in orientations) parameter estimation prior used to produce Abbott et al. [45].

4.8.2. Model Comparisons and Prior Specifications

To compare competing population models in the aforementioned Bayesian framework we calculate two different measures of model goodness-of-fit, namely the marginal likelihoods ($\mathcal{Z}$) and deviance information criterion (DIC) [121]. The marginal likelihood for a given model is a constant that enforces that the posterior distribution is normalized (i.e. $\mathcal{Z} = \int d\Lambda p(\Lambda|\{d_i\})$), which has the property that it is higher for models that fit the data better or find higher likelihoods,

Mass Ratio Model Parameters			
PowerLaw MassRatio	β_q	slope of the mass ratio powerlaw	U(-4, 12)
Redshift Evolution Model Parameters			
PowerLaw Redshift	γ	slope of redshift evolution powerlaw $(1+z)^{\gamma}$	U(-6, 6)
Spline Perturbation Model Parameters			
Cubic Spline	$\{m_n\}$	location in primary mass of the n spline iterpolant knots	FIXED
	$\{f_n\}$	y-value of the n spline iterpolant knots	$N(\mu=0, \sigma=1)$

TABLE 4.2. This table enumerates the rest of the hyper-paramters not included in fwfwf, their descriptions, and chosen priors for this work for each respective population model we use. Note that we do not describe a spin population model in this table since in this work we are not inferring a hyper-prior on the spins and instead assume they are described by the default (uniform in component magnitudes, isotropic in orientations) parameter estimation prior used to produce Abbott et al. [45].

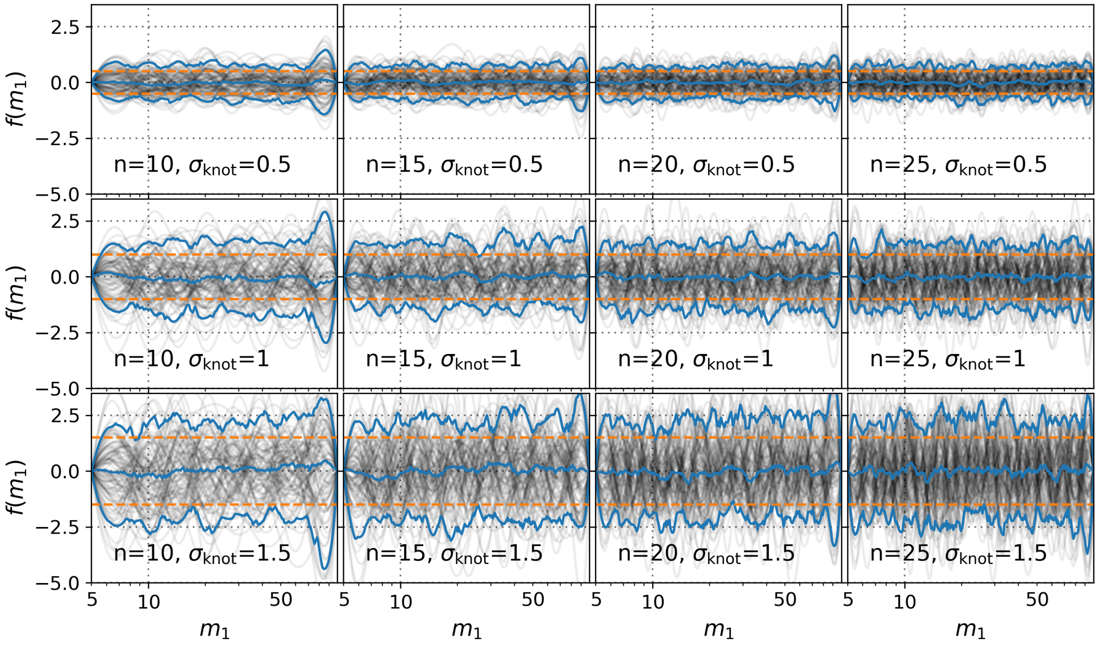

FIGURE 4.6. 100 draws from the prior predictive distribution of the cubic spline function, $f(m_1)$, for different choices for the number of knots, n, and the width of the Gaussian priors on the knots, $\sigma_{\rm knot}$. The orange dashed line shows $\pm\sigma_{\rm knot}$, while the blue solid lines show the median and 1σ credible bounds of the draws from the prior.

while penalizing more complicated models by their prior volumes. As our semi-parametric approach has arbitrary prior choices one needs to make, this can

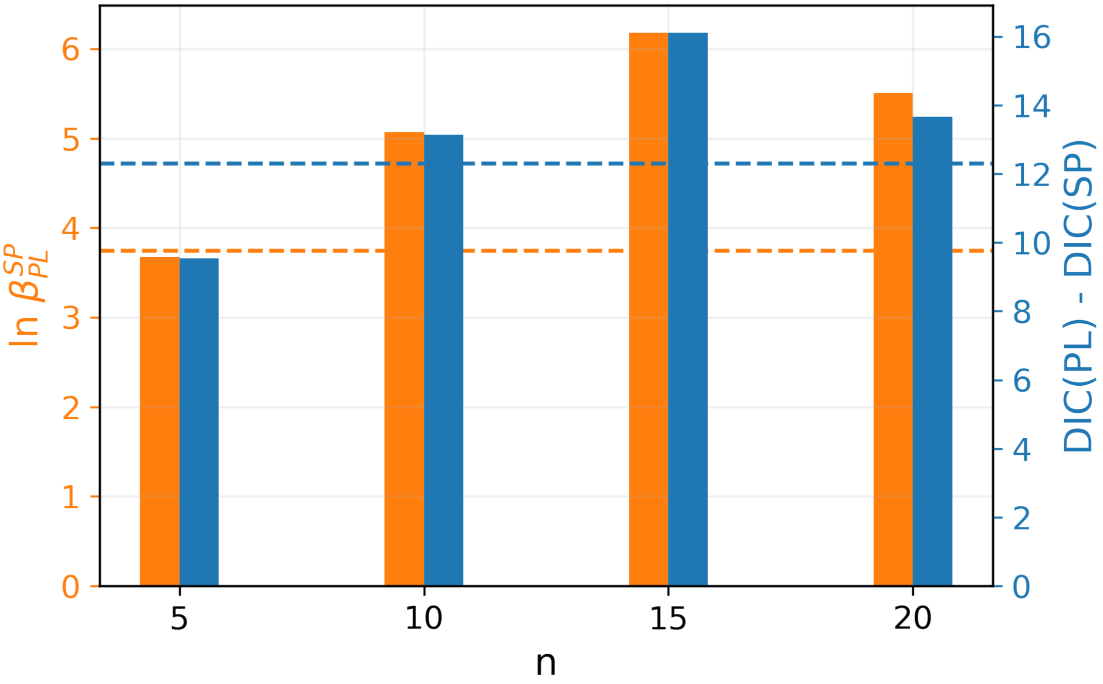

FIGURE 4.7. We show the two model comparison methods, Bayes factors (orange, left y-axis) and the DIC difference (blue, right y-axis), each comparing our spline perturbation model (denoted as SP) to the TRUNCATED model (denoted as PL). The comparisons are calculated such that positive values of either metric denote the spline perturbation model being favored over the TRUNCATED. Both values are shown for varying specifications for the spline prior. Along the x-axis we show different discrete choices (5, 10, 15, and 20) for the number of nodes, n. Each of these spline model analyses shown was performed with $\sigma_{\text{knot}} = 1$. The horizontal dashed lines show the Bayes factor (orange) and DIC difference (blue) found when comparing Abbott et al. [53]'s favorite mass model, POWERLAW+PEAK, to the TRUNCATED model.

significantly affect the marginal likelihoods inferred.We also calculate the DIC, a metric developed specifically for Bayesian hierarchical models [121], which is less sensitive to the arbitrary prior choices for our semi-parametric model. While there are some limitations to the DIC [33], it provides a secondary metric to validate our model choices. The DIC is defined as:

$$DIC = -2\overline{\log(\mathcal{L})} + p_D = -2\left(\overline{\log(\mathcal{L})} - \mathrm{var}(\log(\mathcal{L}))\right) \tag{4.9}$$

With $\overline{\log \mathcal{L}}$ the mean log-likelihood, and p_D the effective number of dimensions, defined as $p_D = \frac{1}{2}\mathrm{var}(-2\log \mathcal{L})$ with var(...) denoting the variance. Lower DICs indicate better models which, similarly to the marginal likelihood, favors models that find higher likelihoods while penalizing the more complicated models through the effective dimension term. We compare two models by calculating the ratio of their marginal likelihoods (i.e. Bayes Factors[2]), defined as the ratio of each models marginal likelihoods. To compare DICs, we take the difference of two models values (DIC dif $= DIC_A - DIC_B$) where positive differences indicate preference for model B, and negative differences indicate preference for model A.

We use these model comparisons to motivate a sensible choice for our spline models prior flexibility, namely the number of knots (n). Figure 4.6 shows changing prior widths on our knots only effects magnitude of perturbations the spline is sensitive to. Additionally, we see that as we add more knots, the model is free to fit sharper fluctuations. This flexibility comes with a penalty in our comparison metrics due to increased model complexity. Therefore, we would

---

[2] The true "Bayesian" way to compare models is using odds, which are Bayes factors multiplied by the ratio of prior odds of each model. Because we don't a priori have expectations of which population model would be more likely, we use Bayes factors which are odds ratios with equal prior weights for each model.

Model Name	$\ln \mathcal{Z}$	$\ln \mathcal{B}$	DIC dif
Powerlaw+Spline ($n = 15$, $\sigma_{\rm knot} = 1$)	-347.10	0.00	0.00
Powerlaw+Spline ($n = 20$, $\sigma_{\rm knot} = 1$)	-347.77	-0.67	-2.43
Powerlaw+Spline ($n = 15$, $\sigma_{\rm knot} = 1$, $\delta_m = 0$)	-347.83	-0.74	-3.27
Powerlaw+Spline ($n = 10$, $\sigma_{\rm knot} = 1$)	-348.21	-1.11	-2.96
Powerlaw+Spline ($n = 10$, $\sigma_{\rm knot} = 1$, $\delta_m = 0$)	-348.24	-1.14	-3.25
Powerlaw+Spline ($n = 20$, $\sigma_{\rm knot} = 1$, $\delta_m = 0$)	-348.26	-1.16	-3.57
Powerlaw+Spline ($n = 15$, $\sigma_{\rm knot} = 2$)	-348.40	-1.30	-1.00
Powerlaw+Spline ($n = 15$, $\sigma_{\rm knot} = 2$, $\delta_m = 0$)	-348.53	-1.43	-2.85
Powerlaw+Spline ($n = 20$, $\sigma_{\rm knot} = 2$)	-348.62	-1.52	-3.00
Powerlaw+Spline ($n = 20$, $\sigma_{\rm knot} = 2$, $\delta_m = 0$)	-349.01	-1.91	-3.39
Powerlaw+MultiPeak	-349.27	-2.18	-6.87
Powerlaw+Spline ($n = 10$, $\sigma_{\rm knot} = 2$, $\delta_m = 0$)	-349.43	-2.34	-3.12
Powerlaw+MultiPeak ($\delta_m = 0$)	-349.44	-2.35	-6.27
Powerlaw+Peak	-349.53	-2.43	-3.80
Powerlaw+Spline ($n = 10$, $\sigma_{\rm knot} = 2$)	-349.66	-2.57	-3.22
Powerlaw+Peak ($\delta_m = 0$)	-349.70	-2.60	-4.96
Broken Powerlaw ($\delta_m = 0$)	-349.92	-2.83	-9.32
Broken Powerlaw	-350.21	-3.11	-8.90
Truncated	-352.57	-5.47	-14.32
Truncated ($\delta_m = 0$)	-353.27	-6.18	-16.10

TABLE 4.3. Model comparison results, listing each model tested (semi-parametric spline model or parametric mass model from Abbott et al. [53]) and their respective marginal likelihoods ($\mathcal{Z}$) along with $\ln \mathcal{B}$ and DIC dif. Both comparison metrics for each of the listed models are relative to the "best performing" model or the one with the highest (lowest) marginal likelihood (DIC), which, in both cases, was the Powerlaw+Spline ($n = 15$, $\sigma_n = 1$) model. We note that the POWERLAW+MULTIPEAK finds higher marginal likelihoods than the POWERLAW+PEAK model which was not the case in Abbott et al. [53]. This is because we used different priors for the POWERLAW+MULTIPEAK model that allowed for a peak at lower masses than the $\sim 35\,\mathrm{M}_\odot$ peak instead of higher.

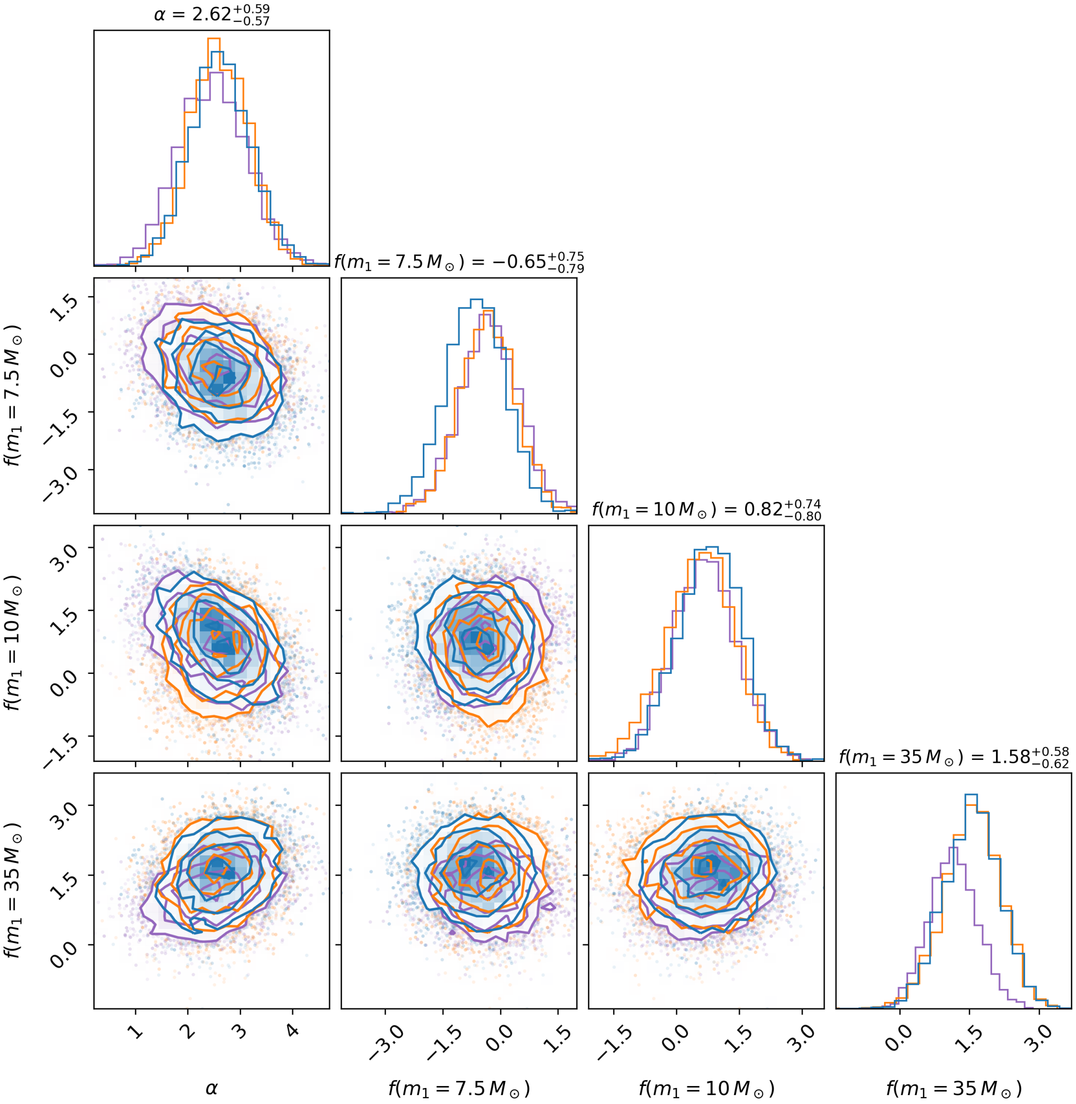

FIGURE 4.8. Corner plot that shows the posterior distribution on the power law slope, α, and the height of the perturbation function, $f(m_1)$, sliced at the three masses of most significant deviation: $7.5\,\mathrm{M}_\odot$, $10\,\mathrm{M}_\odot$, and $35\,\mathrm{M}_\odot$. We show the results for spline models with $\sigma_{\mathrm{knot}} = 1$ and 10 (purple), 15 (blue) and 20 (orange) nodes. The median and 90% credible intervals quoted are for the 15 knot model.

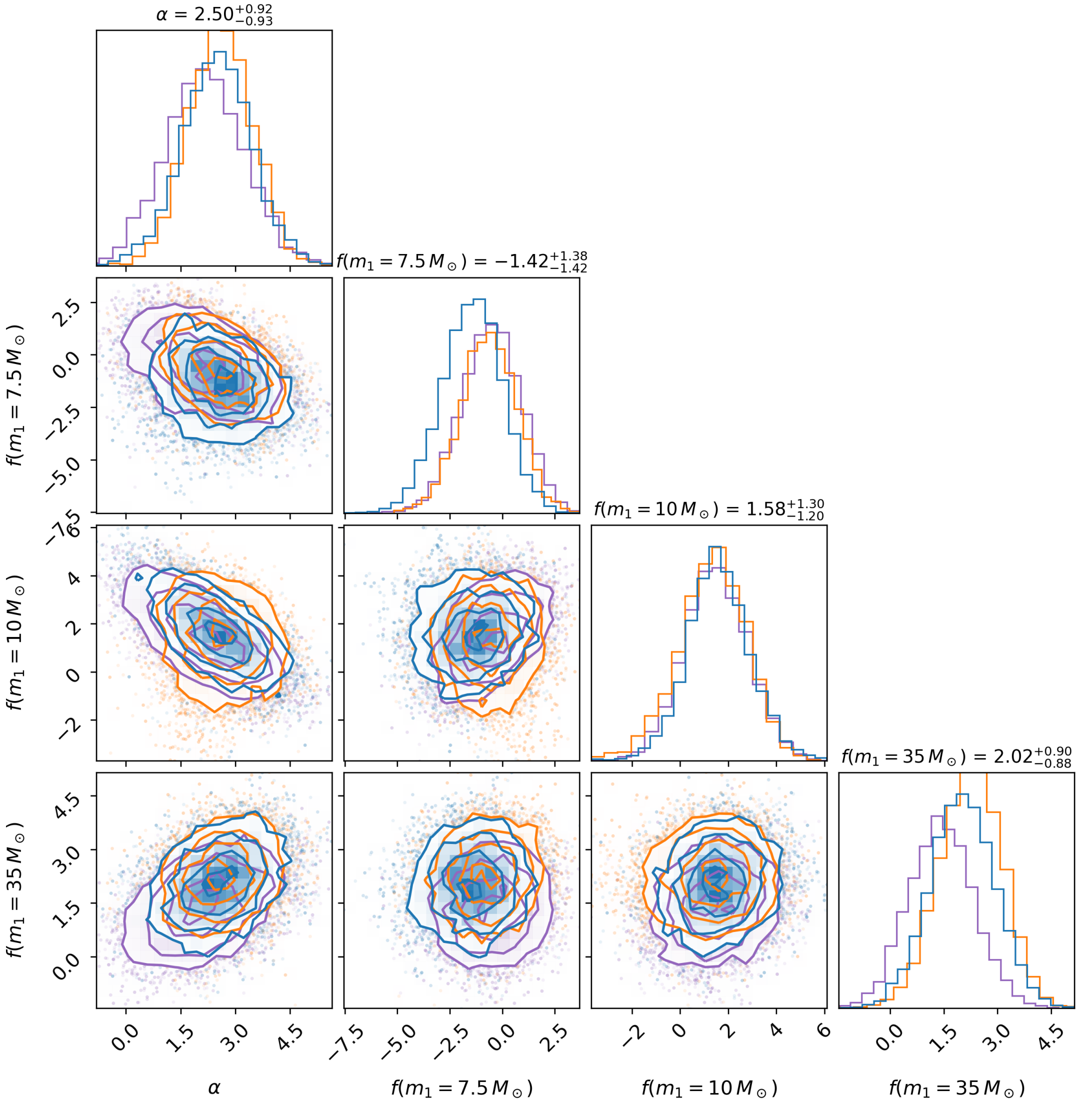

FIGURE 4.9. Corner plot that shows the posterior distribution on the power law slope, α, and the height of the perturbation function, $f(m_1)$, sliced at the three masses of most significant deviation: $7.5\,\mathrm{M}_\odot$, $10\,\mathrm{M}_\odot$, and $35\,\mathrm{M}_\odot$. We show the results for spline models with $\sigma_{\mathrm{knot}} = 2$ and 10 (purple), 15 (blue) and 20 (orange) nodes. The median and 90% credible intervals quoted are for the 15 knot model.

expect to see the model comparisons increasingly favor our spline perturbation model as we increase our model's complexity/flexibility up to a point where the penalty will start to overpower the higher likelihoods found with more flexibility. Figure 4.7 shows how the DIC differences and log Bayes factors ($\ln \mathcal{B}$) change when comparing the spline perturbation model to the TRUNCATED model with different choices for n. We see that the comparisons favoring our spline model peaks around 15 knots, indicating that 15 knots is a good trade-off between our models flexibility and goodness of fit. We also report the marginal likelihoods and model comparisons (relative to the most preferred model) for each of the parametric primary mass models from Abbott et al. [53] and various specifications of the spline model in table 4.3. From this table we see that the spline model is consistently favored despite our arbitrary model specifications, giving credence to the hypothesis that there are features in the data our semi-parametric method is capable of finding that previously used parametric mass models are not sensitive to. We do not use the comparisons in table 4.3 to determine the validity of the POWERLAW+SPLINE models over others, and further studies on simulated populations and the effect of small number statistics are needed to fully assess the significance and robustness of these features. However, as catalogs of BBH mergers increase in size, the impact of small number statistics will diminish.

4.8.3. Correlations of Peaks

We look for the effect of our spline function biasing the inferred perturbation function by plotting a corner plot of the value of $f(m_1)$ sliced through the masses that show the largest deviations. This is shown in Figure 4.8 for the conservative knot prior and Figure 4.9 for the wide knot prior. If the dip followed by the peak feature at 7.5 and 10 solar masses was found due to the nature of cubic splines we would expect to see correlations between the heights at these values which do not appear in either Figure 4.8 or 4.9. Since we are fitting for the underlying power law model simultaneously with the perturbations we also might expect to see some correlations with the power law slope and the peaks. There are slight signs of an expected anti-correlation of the $10\mathrm{M}_{\odot}$peak height and the power law slope, and corresponding correlation of the $35\mathrm{M}_{\odot}$peak height with the slope. This happens due to the degeneracy between the parametric and non-parametric portions of our model. If the $10\mathrm{M}_{\odot}$peak is small, the power law slope becomes steeper so that the "turnover" power law at low mass can contribute to fitting this over-density. With a steeper power law slope, the power law portion of the model under fits the excess at $35\mathrm{M}_{\odot}$, leading to a larger peak found in the perturbation. We also note these correlations with the power law slope are more apparent in the lower resolution spline models and under the wider prior on the knots.

CHAPTER V

COVER YOUR BASIS: COMPREHENSIVE DATA-DRIVEN CHARACTERIZATION OF THE BINARY BLACK HOLE POPULATION

5.1. Summary

Following the LVK collaboration's analyses of GWTC-3, the population of BBHs was found to exhibit significant structure in the mass distribution, which could be signs of competing formation channels. Non-parametric models can be useful to fully validate features in the population found with more restrictive models since they should be more flexible and data-driven. We construct a novel non-parametric model with basis splines that allow for simultaneous inference of non-parametric mass, mass ratio, spin magnitude, spin tilt, and redshift distributions of the BBH population. With basis splines the cost of computation is drastically reduced, along with moving to use Hamiltonian Monte Carlo (HMC) with GPU accelerators allows this massive inference problem with hundreds of parameters to be completed in less than an hour. In this chapter we present work published as Edelman et al. [122] which introduces this first completely non-parametric population study of the binary black holes (BBHs). To date this is the most comprehensive data-driven population study of BBHs and is the first to non-parametrically infer mass ratio and redshift distributions. This model will be more powerful as the catalog grows enabling us to probe additional subtle features of the population, and already has been shown to be useful with mixture modeling looking for signs of subpopulations of BBHs [123].

5.2. Abstract

We introduce the first complete non-parametric model for the astrophysical distribution of the binary black hole (BBH) population. Constructed from basis splines, we use these models to conduct the most comprehensive data-driven investigation of the BBH population to date, simultaneously fitting non-parametric models for the BBH mass ratio, spin magnitude and misalignment, and redshift distributions. With GWTC-3, we report the same features previously recovered with similarly flexible models of the mass distribution, most notably the peaks in merger rates at primary masses of $\sim 10\mathrm{M}_{\odot}$ and $\sim 35\mathrm{M}_{\odot}$. Our model reports a suppressed merger rate at low primary masses and a mass ratio distribution consistent with a power law. We infer a distribution for primary spin misalignments that peaks away from alignment, supporting conclusions of recent work. We find broad agreement with the previous inferences of the spin magnitude distribution: the majority of BBH spins are small ($a < 0.5$), the distribution peaks at $a \sim 0.2$, and there is mild support for a non-spinning subpopulation, which may be resolved with larger catalogs. With a modulated power law describing the BBH merger rate's evolution in redshift, we see hints of the rate evolution either flattening or decreasing at $z \sim 0.2 - 0.5$, but the full distribution remains entirely consistent with a monotonically increasing power law. We conclude with a discussion of the astrophysical context of our new findings and how non-parametric methods in gravitational-wave population inference are uniquely poised to complement to the parametric approach as we enter the data-rich era of gravitational-wave astronomy.

5.3. Introduction

Observations of gravitational waves (GWs) from compact binary mergers are becoming a regular occurrence, producing a catalog of events that recently surpassed 90 such detections [1, 9, 45]. As the catalog continues to grow, so does our understanding of the underlying astrophysical population of compact binaries [51, 52, 53]. Following numerous improvements to the detectors since the last observing run, the anticipated sensitivities for the upcoming fourth observing run of the LIGO-Virgo-KAGRA (LVK) collaboration suggest detection rates as high as once per *day* [26, 27, 108, 124]. With the formation history of these dense objects encoded in the details of their distribution [109, 125, 126, 127, 128], the likely doubling in size of the catalog with the next observing run could provide another leap in our understanding of compact binary astrophysics. Beyond formation physics, population-level inference of the compact binary catalog has also been shown to provide novel measurements of cosmological parameters [95, 129, 130], constrain modified gravitational wave propagation [131, 132, 133], constrain a running Planck mass [134], search for evidence of ultralight bosons through superradiance [135, 136], constrain stellar nuclear reaction rates [61, 63], look for primordial black holes [110, 137], and to constrain physics of neutron stars [138, 139]. Through a better understanding of the mass, spin, and redshift distributions of compact binaries that will come with the increased catalog size, one can probe a wide range of different physical phenomena with even greater fidelity.

The binary black hole (BBH) mass distribution was first found to have structure beyond a smooth power law with simpler parametric models, exhibiting a possible high mass truncation and either a break or a peak at $m_1 \sim 35 -$

$40\,\mathrm{M}_\odot$ [52, 53, 66, 67]. Starting with the moderately sized catalog, GWTC-2, more flexible models found signs of additional structure [96, 97]. The evidence supporting these features, such as the peak at $m_1 \sim 10\,\mathrm{M}_\odot$, has only grown after analyzing the latest catalog, GWTC-3, with the same models [51, 140]. While this shows the usefulness of data-driven methods with the current relatively small catalog size, they will become more powerful with more observations. The canonical approach to constructing population models has been to use simple parametric descriptions (e.g., power laws, beta distributions) that aim to describe the data in the simplest way, employ astrophysically motivated priors where appropriate, then sequentially add complexities (e.g., Gaussian peaks) as the data demands. This simple approach was necessary when data was scarce, but as we move into the data-rich catalog era, this approach is already failing to scale. More flexible and scalable methods, such as the non-parametric modeling techniques presented in this manuscript, will be necessary to continue to extract the full information contained in the compact binary catalog. In contrast to parametric models, flexible and non-parametric models are data-driven and contribute little bias to functional form. They additionally are particularly useful to search for unexpected features in the data, providing meaningful insight into features that parametric models may fail to capture.

While we eventually hope to uncover hints of binary formation mechanisms in the mass spectrum of BBHs, the distribution of spin properties have been of particular interest. The measurement of spin properties of individual binaries often have large uncertainties, but the theorized formation channels are expected to produce distinctly different spin distributions [109, 126, 127, 128, 141]. Isolated (or field) formation scenarios predict component spins that are preferentially aligned with the binary's orbital angular momentum, although some small misalignment can occur depending on the nature of the supernova kicks as each star collapses to a compact object [142, 143, 144]. Alternatively, dynamical formation in dense environments where many-body interactions between compact objects can result in binary formation and hardening (shrinking of binary orbits) should produce binaries with components' spins distributed isotropically [105, 126]. BBH spins have also been of controversial interest recently, with different parametric approaches to modeling the spin distribution coming to different conclusions. Studies have disagreed on the possible existence of a significant zero-spinning subpopulation, as well as the presence of significant spin misalignment (i.e. $\cos\theta_i < 0.0$) [51, 145, 146, 147, 148]. Another study recently showed that inferences of spin misalignment (or tilts) are sensitive to modeling choices and may not peak at perfectly aligned spins, as is often assumed [149]. While enlightening, these recent efforts to improve BBH spin models continue to build sequentially on previous parametric descriptions [146, 148, 149]. To ensure we are extracting the full detail the catalog has to offer, we extend our previous non-parametric modeling techniques to include spin magnitudes and tilts, as well as the binary mass ratio and redshift. Golomb and Talbot [150] was released concurrently with this work (based on our previous work Edelman et al. [97]), and find similar conclusions on the spin distribution when applying similar flexible models constructed with cubic splines. The work presented in this manuscript however, does not need to analyze a suite of different model configurations and includes flexible non-parametric models for each of the mass, spin and redshift distributions rather than spin alone.

Polynomial splines have been applied to success across different areas of gravitational-wave astronomy. They have been used to model the gravitational-wave data noise spectrum, detector calibration uncertainties, coherent gravitational waveform deviations, and modulations to a power law mass distribution [8, 25, 31, 97, 151] In this paper we highlight how the use of basis-splines can provide a powerful non-parametric modeling approach to the astrophysical distributions of compact binaries. We illustrate how one can efficiently model both the mass and spin distributions of merging compact binaries in GWTC-3 with basis splines to infer compact binary population properties using hierarchical Bayesian inference. We discuss our results in the context of current literature on compact object populations and how this method complements the simpler lower dimensional parametric models in the short run, and will become necessary with future catalogs. Should they appear with more observations, this data-driven approach will provide checks of our understanding by uncovering more subtle – potentially unexpected – features. The rest of this manuscript is structured as follows: a description of the background of basis splines in section 5.9.1, followed by a presentation of the results of our extensive, data-driven study of the mass and spin distributions of BBHs in GWTC-3 in section 5.5. We then discuss these results and their astrophysical implications in section 5.6 and finish with our conclusions in section 5.7.

5.4. Building the Model

We construct our data-driven model with the application of basis splines, or B-Splines [6]. B-Splines of order k are a set of order k polynomials that span the space of possible spline functions interpolated from a given set of knot locations.

For all B-Splines used in our model we use a third order basis which consists of individual cubic polynomials. The basis representation of the splines allows for the computationally expensive interpolation to be done in a single preprocessing step – amortizing the cost of model evaluation during inference. To mitigate the unwanted side effects of adding extra knots and to avoid running a model grid of differing numbers of knots (as in Edelman et al. [97]), we use the smoothing prior for Bayesian P-Splines [7, 152, 153], allowing the data to pick the optimal scale needed to fit the present features. We discuss basis splines, the smoothing prior, and our specific prior choices on hyperparameters in Appendix 5.9.1, 5.9.2 and 5.9.4.

We parameterize each binaries' masses with the primary (more massive component) mass (m_1) and the mass ratio ($q = m_2/m_1$) with support from 0 to 1. Furthermore, we model 4 of the 6 total spin degrees of freedom of a binary merger: component spin magnitudes a_1 and a_2, and (cosine of) the tilt angles of each component, $\cos\theta_1$ and $\cos\theta_2$. The tilt angle is defined as the angle between each components' spin vector and the binary's orbital angular momentum vector. We assume the polar spin angles are uniformly distributed in the orbital plane. For the primary mass distribution, we model the log probability with a B-Spline interpolated over knots linearly spaced in $\log(m_1)$ from a minimum black hole mass, which we fix to $5\mathrm{M}_\odot$, and a maximum mass that we set to $100\mathrm{M}_\odot$. We then have the hyper-prior on primary mass with log probability density $\log(p(m_1|\boldsymbol{c})) \propto B_{k=3}(\log(m_1)|\boldsymbol{c})$, where $B_{k=3}$ is the cubic B-Spline function with a vector of basis coefficients $\boldsymbol{c}$. We follow the same procedure for the models in mass ratio and spin distributions with knots spaced linearly across each domain so that we have $\log(p(\theta|\boldsymbol{c}_\theta)) \propto B_{k=3}(\theta|\boldsymbol{c}_\theta)$, where θ can be q, a_1, a_2, $\cos\theta_1$ or $\cos\theta_2$. For the spin magnitude and tilt distributions we construct two versions of the model: first, we model each component's distribution as independently and identically distribution (IID), where we have a single B-Spline model and parameters (coefficients) for each binary spin. Secondly, we model each component's distribution to be unique, fitting separate sets of coefficients for the B-Spline models of the primary and secondary spin distributions. Lastly, we fit a population model on the redshift or luminosity distance distribution of BBHs, assuming a ΛCDM cosmology defined by the parameters from the Planck 2015 results [120]. This defines an analytical mapping between each event's inferred luminosity distance, and its redshift, which we now use interchangeably. We take a semi-parametric approach to model the evolution of the merger rate with redshift, following Edelman et al. [97], that parameterizes modulations to an underlying model with splines (in our case basis splines). We use the PowerlawRedshift model as the underlying distribution to modulate, which has a single hyperparameter, λ_z, and probability density defined as: $p(z|\lambda_z) \propto \frac{dV_c}{dz}(1+z)^{\lambda_z - 1}$ [88]. For more detailed descriptions of each model and specific prior choices used for the hyperparameters see Appendix 5.9.4. Now that we have our comprehensive data-driven population model built, we simultaneously fit the basis spline models on the BBH masses, spins and redshift. We use the usual hierarchical Bayesian inference framework (see appendix 5.9.3 for a review; The LIGO Scientific Collaboration et al. [51], Abbott et al. [52, 53]), to perform the most extensive characterization of the population of BBHs to date using the most recent catalog of GW observations, GWTC-3 [1].

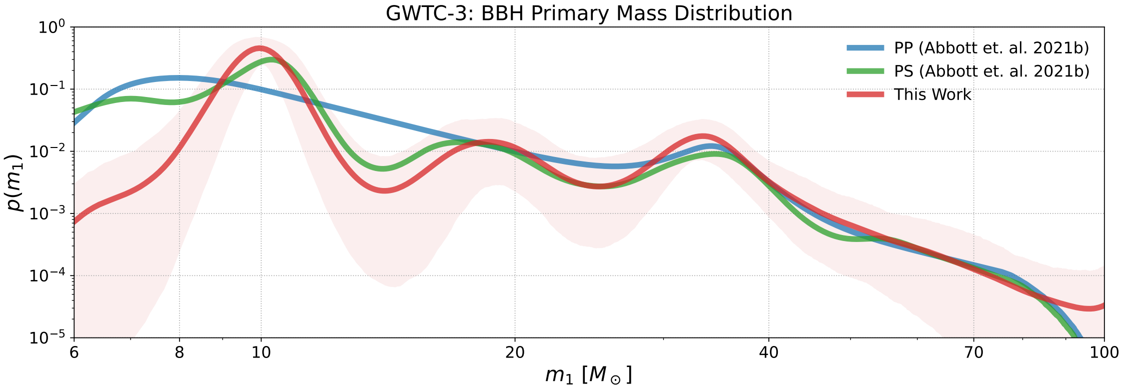

FIGURE 5.1. The marginal primary mass distribution inferred with the B-Spline model (red), with 64 knots spaced linearly in $\log m_1$, from $5\mathrm{M}_\odot$to $100\mathrm{M}_\odot$. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. We show the inferred PPD from the PowerlawPeak (blue) and PowerlawSpline (green) models from the LVK's GWTC-3 population analyses [51].

5.5. Binary Black Hole Population Inference with GWTC-3

We use hierarchical Bayesian inference (see Appendix 5.9.3) to simultaneously infer the astrophysical mass, spin, and redshift distributions of binary black holes (BBHs) given a catalog of gravitational wave observations. Following the same cut on the recent GWTC-3 catalog done in the LVK's accompanying BBH population analysis, we have 70 possible BBH mergers with false alarm rates less than 1 per year [1, 51, 154]. Since it was concluded to be an outlier of the rest of the BBH population in both GWTC-2 and GWTC-3, we choose to omit the poorly understood event, GW190814 [51, 53, 91, 155]. This leaves us with a catalog of 69 confident BBH mergers, observed over a period of about 2 years, from which we want to infer population properties. Following what was done in The LIGO Scientific Collaboration et al. [51], for events included in GWTC-1 [9], we use the published samples that equally weight samples from analyses with the IMRPhenomPv2 [156] and SEOBNRv3 [157, 158] waveforms.

For the events from GWTC-2 [45], we use samples that equally weight all available analyses using higher-order mode waveforms (PRECESSINGIMRPHM). Finally, for new events reported in GWTC-2.1 and GWTC-3 [1, 159], we use combined samples, equally weighted, from analyses with the IMRPHENOMXPHM [79] and the SEOBNRV4PHM [4] waveform models. Our study provides the first comprehensive data-driven investigation, simultaneously inferring all the BBH population distributions (i.e. mass, spin, and redshift), uncovering new insights and corroborating those found with other methods. We start with our inference of the mass distribution.

5.5.1. Binary Black Hole Masses

Figure 5.1 shows the primary mass distribution inferred with our B-Spline model (red), where we see features consistent with those inferred by the POWERLAWPEAK and POWERLAWSPLINE mass models [51, 53, 67, 97, 160]. In particular our B-Spline model finds peaks in merger rate density as a function of primary mass at both $\sim 10\mathrm{M}_\odot$ and $\sim 35\mathrm{M}_\odot$, agreeing with those reported using the same dataset in The LIGO Scientific Collaboration et al. [51]. The B-Spline model finds the same feature at $\sim 18\mathrm{M}_\odot$ as the POWERLAWSPLINE model, but remains consistent with the POWERLAWPEAK model; the mass distribution is more uncertain in this region. For each of these features we find the local maximums occurring at primary masses of $9.9^{+0.67}_{-0.48}\,\mathrm{M}_\odot$, $19^{+3.2}_{-2.3}\,\mathrm{M}_\odot$, and $33^{+2.1}_{-3.0}\,\mathrm{M}_\odot$ all at 90% credibility. We find the largest disagreement at low masses, where the power-law-based models show a higher rate below $\sim 8 - 9\mathrm{M}_\odot$. This is partly due to the minimum mass hyperparameter (where the power law "begins") serving as the minimum allowable primary and secondary masses of the catalog. This leads

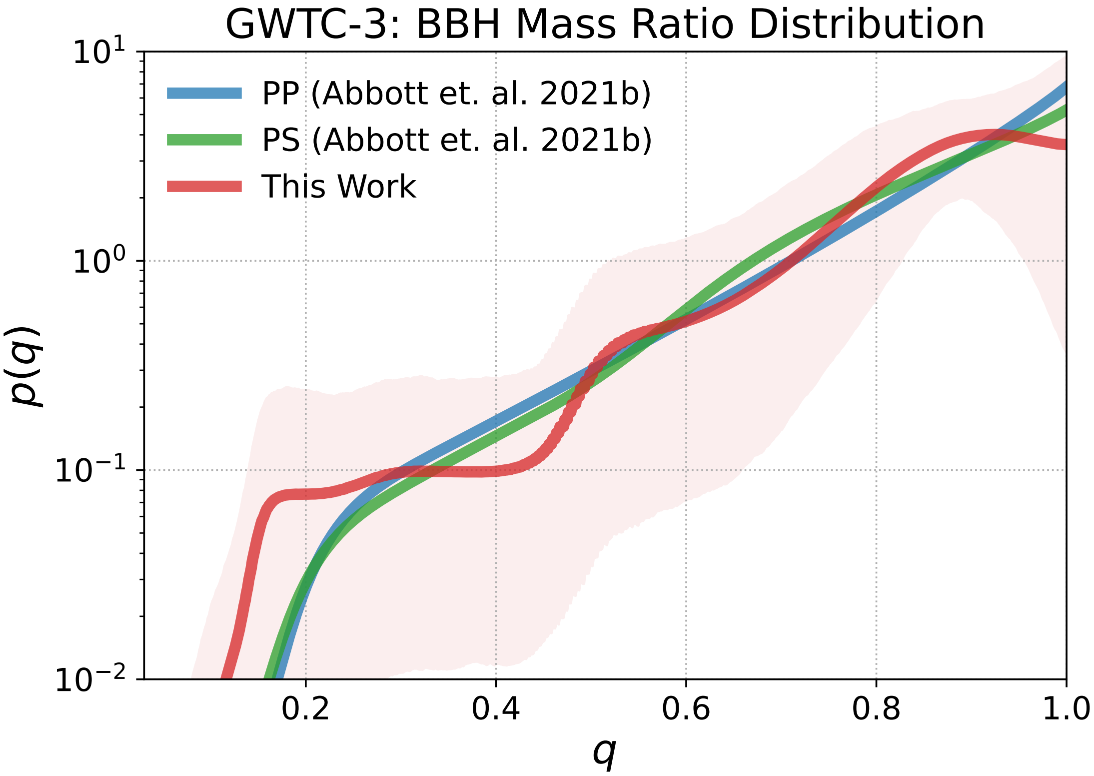

FIGURE 5.2. The marginal mass ratio distribution inferred with the B-Spline model (red), with 18 knots spaced linearly in q, from 0.05 to 1. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. We show the inferred PPD from the PowerlawPeak (blue) and PowerlawSpline (green) models from the LVK's GWTC-3 population analyses [51].

to inferences of m_{min} below the minimum observed primary mass in the catalog, which is $\sim$ 6.4$\mathrm{M}_\odot$, to account for secondary BBH masses lower than that. We choose to fix the minimum black hole mass for both primary and secondaries to 5$\mathrm{M}_\odot$, similar to the inferred minimum mass in The LIGO Scientific Collaboration et al. [51]. The lack of observations of binaries with low primary mass make rate estimates in this region strongly model dependent, while our flexible model provides an informed upper limit on the rate in this region and given the selection effects and that there are no observations. We could be seeing signs of a decrease in merger rate from a “lower mass gap” between neutron star and BH masses, or we could be seeing fluctuations due to low-number statistics [98]. Either way we expect this to be resolved with future catalog updates. We also find no evidence for a sharp fall off in merger rate either following the pileup at $\sim$35 $\mathrm{M}_\odot$ – expected if such a pileup was due to pulsational pair instability supernovae (PPISNe) – or where the maximum mass truncation of the power law models are inferred. The lack of any high mass truncation, along with the peak at $\sim$ 35$\mathrm{M}_\odot$ (significantly lower than expected from PPISNe) may pose challenges for conventional stellar evolution theory. This could be hinting at the presence of subpopulations that avoid pair instability supernovae during binary formation, but the confirmation of the existence of such subpopulations cannot be determined with the current catalog.

The marginal mass ratio distribution inferred by the B-spline model is shown in figure 5.2. These results suggest we may be seeing the first signs of departure from a simple power law behavior. We find a potential signs of a plateau or decrease in the merger rate near equal mass ratios, as well as a broader tail towards unequal mass ratios than the power law based models find, although a smooth power law is still consistent with these results given the large uncertainties. Our results also suggest a shallower slope from $q \sim 0.3$ to $q \sim 0.7$, though uncertainty is larger in this region. The sharp decrease in rate just below $q \sim 0.5$ is due to the minimum mass ratio truncation defined by $q_{\mathrm{min}} = \frac{m_{\mathrm{min}}}{m_1}$. When marginalizing over the primary mass distribution with a strong peak at 10$\mathrm{M}_\odot$, the mass ratio distribution truncates at $q \sim 0.5$: the minimum mass, 5$\mathrm{M}_\odot$, divided by the most common primary mass, $\sim$10$\mathrm{M}_\odot$.

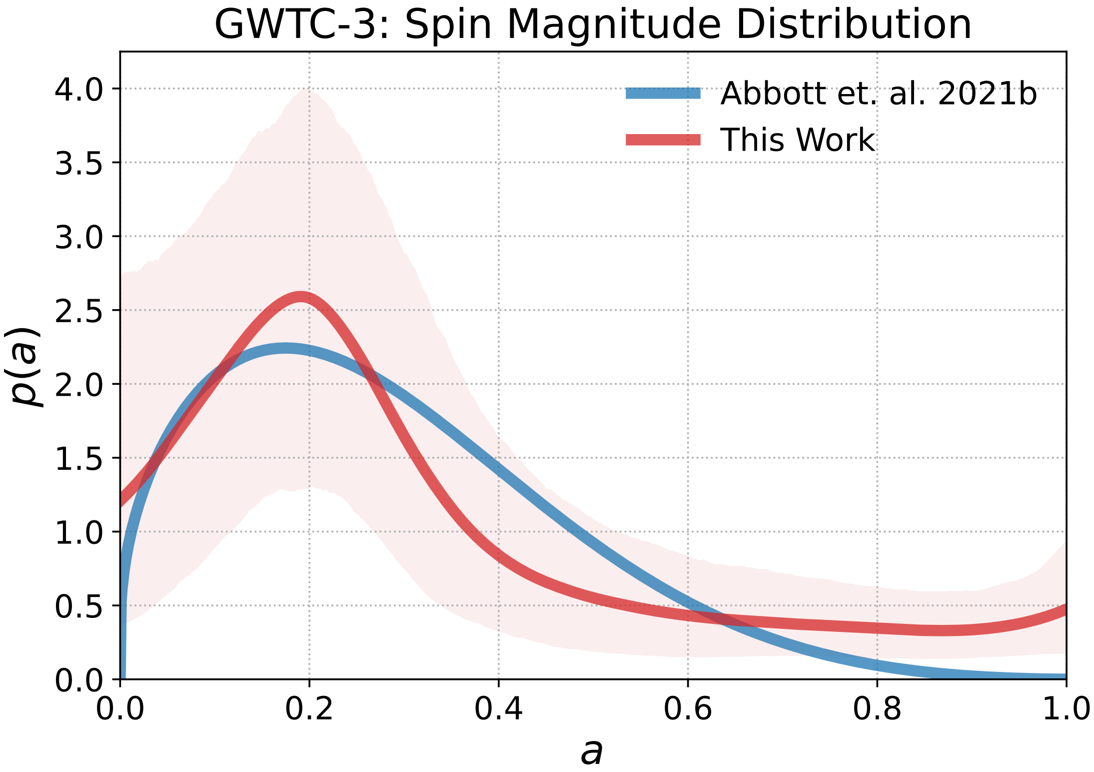

FIGURE 5.3. The spin magnitude distribution inferred with the B-Spline model (red) with 16 knots spaced linearly from 0 to 1, assuming the components are IID. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. For comparison, we show the inferred PPD from the DEFAULT (blue) model from The LIGO Scientific Collaboration et al. [51], the LVK's GWTC-3 population analyses.

5.5.2. Binary Black Hole Spins

Model	a_{peak}	$a_{90\%}$	$\cos\theta_{\mathrm{peak}}$	$f_{\cos\theta<0}$	$\log_{10} Y$
B-Spline IID	$0.19^{+0.12}_{-0.16}$	$0.71^{+0.13}_{-0.14}$	$0.44^{+0.56}_{-0.53}$	$0.35^{+0.11}_{-0.11}$	$0.24^{+0.46}_{-0.46}$
B-Spline Ind(primary)	$0.2^{+0.24}_{-0.2}$	$0.77^{+0.11}_{-0.13}$	$0.16^{+0.84}_{-0.84}$	$0.43^{+0.19}_{-0.16}$	$0.12^{+0.5}_{-0.53}$
B-Spline Ind(secondary)	$0.17^{+0.29}_{-0.17}$	$0.8^{+0.1}_{-0.14}$	$0.38^{+0.62}_{-1.0}$	$0.38^{+0.18}_{-0.15}$	$0.18^{+0.53}_{-0.54}$
DEFAULT [51]	$0.16^{+0.11}_{-0.13}$	$0.53^{+0.098}_{-0.073}$	$1.0^{+0.0}_{-0.0}$	$0.44^{+0.052}_{-0.12}$	$0.19^{+0.4}_{-0.17}$

TABLE 5.1. Summary of Component Spin distributions inferred both the independent and IID component spin B-Spline models and the DEFAULT spin model from The LIGO Scientific Collaboration et al. [51].

5.5.2.1. Spin Magnitude

The Default spin model (used by The LIGO Scientific Collaboration et al. [51]) describes the spin magnitude of both components as identical and independently distributed (IID) non-singular Beta distributions [161, 162]. The Beta distribution provides a simple 2-parameter model that can produce a wide range of functional forms on the unit interval. However, the constraint that keeps the Beta distribution non-singular (i.e. $\alpha > 1$ and $\beta > 1$) enforces a spin magnitude that always has $p(a_i = 0) = 0$. Recent studies have proposed the possible existence of a distinct subpopulation of non-spinning or negligibly spinning black holes that can elude discovery with such a model [145, 146, 147, 148, 163].

We model the spin magnitude distributions as IID B-Spline distributions. Figure 5.3 shows the inferred spin magnitude distribution with the B-Spline model, compared with the Default model from The LIGO Scientific Collaboration et al. [51]. The B-Spline model results are consistent with those using the Beta distribution, peaking near $a \sim 0.2$, with 90% of BBH spins below $0.71^{+0.13}_{-0.14}$ at 90% credibility. The B-Spline model does not impose vanishing support at the extremal values like the Beta distribution, allowing it to probe the zero-spin question. We find broad support, with large variance, for non-zero probabilities at $a_i = 0$, but cannot confidently determine the presence of a significant non-spinning subpopulation, corroborating similar recent conclusions [146, 147, 148, 164]. We repeat the same analysis with independent B-Spline distributions for each spin magnitude component. In figure 5.4 we show the inferred primary (orange), and secondary (olive) spin magnitude distributions inferred when relaxing the IID assumption. We find no signs that the spin magnitude distributions are not

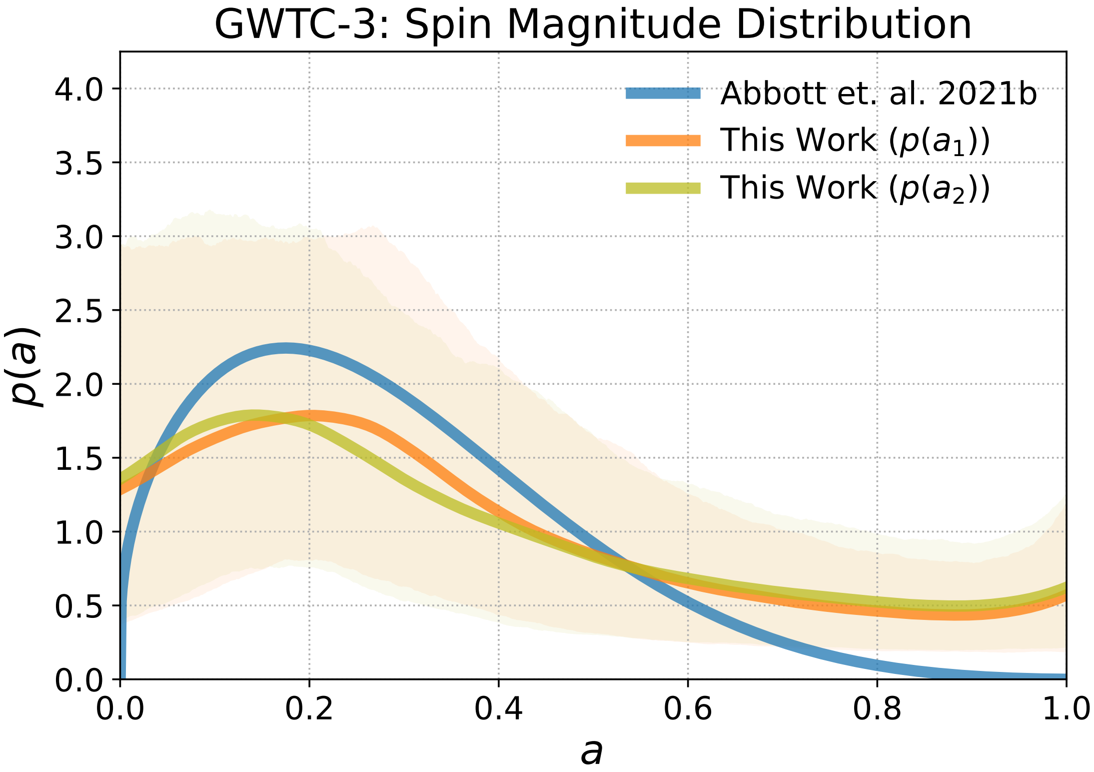

FIGURE 5.4. The primary (orange) and secondary (olive) spin magnitude distributions inferred with the B-Spline model with 16 knots spaced linearly from 0 to 1. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. For comparison, we show the inferred PPD from the Default (blue) model from The LIGO Scientific Collaboration et al. [51], the LVK's GWTC-3 population analyses.

IID but that the primary spin magnitude distribution peaks slightly higher, at $a \sim 0.25$, than the IID B-Spline model in figure 5.3, but with similar support at near vanishing spins. The secondary spin magnitude distribution is more uncertain due to the higher measurement uncertainty when inferring the secondary spins of BBH systems [165, 166]. The PPD of the secondary distribution peaks at smaller spin magnitudes $a \sim 0.15$, than the primary distribution or B-Spline IID spin magnitude distribution in figure 5.3, although the distributions are broadly consistent with each other considering the large uncertainties. A population of compact binary mergers with component spin magnitude distributions that are uniquely distributed can be produced through mass-ratio reversal in isolated binary evolution [164], and could be uncovered in the future with larger catalog sizes.

5.5.2.2. Spin Orientation

The Default spin model (used in The LIGO Scientific Collaboration et al. [51], Abbott et al. [53]) also assumes the spin orientation of both components are identical and independently distributed (IID), with a mixture model over an aligned and an isotropic component. The aligned component is modeled with a truncated Gaussian distribution with mean at $\cos\theta = 1$ and variance a free hyperparameter to be fit [51, 53, 161, 162]. This provides a simple 2-parameter model motivated by simple distributions expected from the two main formation scenario families, allowing for a straightforward interpretation of results. One possible limitation however, is that by construction this distribution is forced to peak at perfectly aligned spins, i.e. $\cos\theta = 1$. While this may be a reasonable assumption, Vitale et al. [149] recently extended the model space of parametric descriptions used to model the spin orientation distribution and found considerable evidence that the distribution peaks away from $\cos\theta = 1$. Again, this provides a clear use-case where data-driven models can help us understand the population.

Figure 5.5 shows the inferred spin orientation distribution with the IID spin B-Spline model, compared with the Default model from The LIGO Scientific Collaboration et al. [51]. The B-Spline inferences have large uncertainties but start to show the same features as found and discussed in Vitale et al. [149]. We find a distribution that instead of intrinsically peaking at $\cos\theta = 1$, is found to

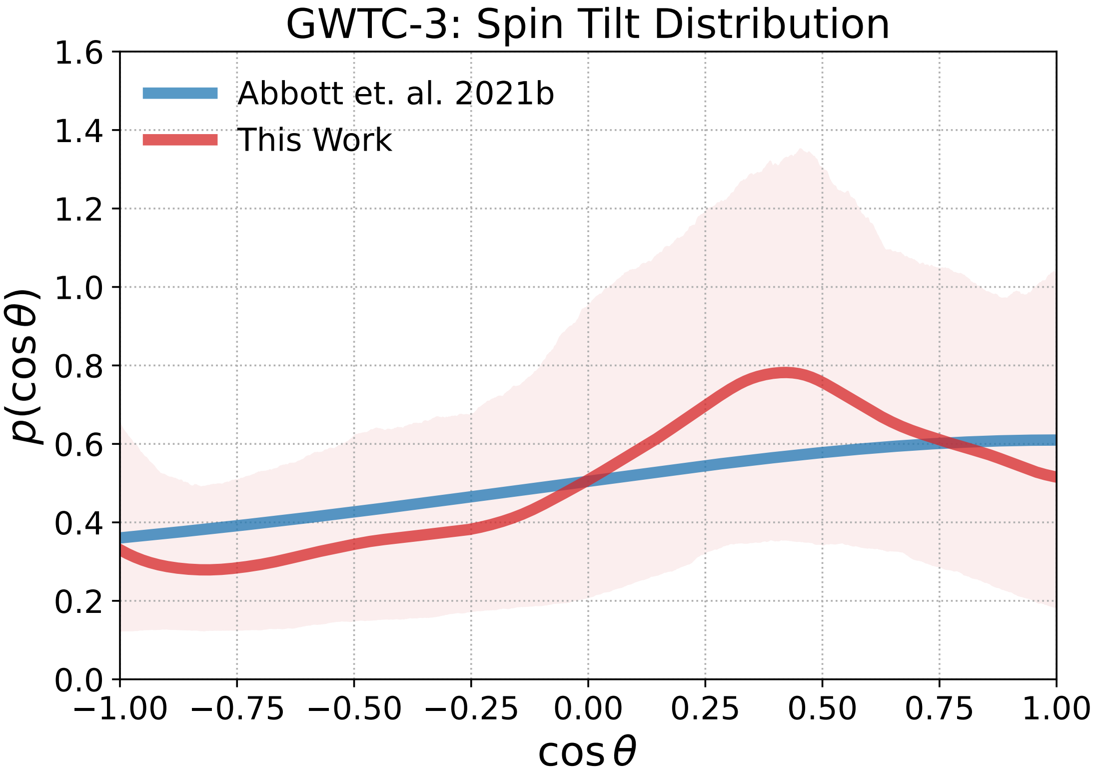

FIGURE 5.5. The spin orientation distribution inferred with the B-Spline model (red) with 16 knots spaced linearly from -1 to 1, and assuming the components are IID. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. For comparison, we show the inferred PPD from the Default (blue) model from The LIGO Scientific Collaboration et al. [51], the LVK's GWTC-3 population analyses.

peak at: $\cos\theta\ =0.44^{+0.56}_{-0.53}$, at 90% credibility. We find less, but still considerable support for misaligned spins (i.e. $\cos\theta\ <\ 0$), consistent with other recent studies [51, 53, 148]. Specifically we find that the fraction of misaligned systems is $f_{\cos\theta<0}\ =0.35^{+0.11}_{-0.11}$, compared to $f_{\cos\theta<0}\ =0.44^{+0.052}_{-0.12}$ with the Default model from The LIGO Scientific Collaboration et al. [51]. This implies the presence of an isotropic component as expected by dynamical formation channels, albeit less than with the Default model. To quantify the amount of isotropy in the tilt distribution we calculate $\log_{10} Y$, where Y is the ratio of nearly aligned tilts to nearly anti-aligned, introduced in Vitale et al. [149] and defined as:

$$Y \equiv \frac{\int_{0.9}^{1.0} d\cos\theta p(\cos\theta)}{\int_{-1.0}^{-0.9} d\cos\theta p(\cos\theta)}. \tag{5.1}$$

The log this quantity, $\log_{10} Y$, is 0 for tilt distribution that is purely isotropic, negative when anti-aligned values are favored, and positive when aligned tilts are favored. We find a $\log_{10} Y\ = 0.24^{+0.46}_{-0.46}$, exhibiting a slight preference for aligned tilts.

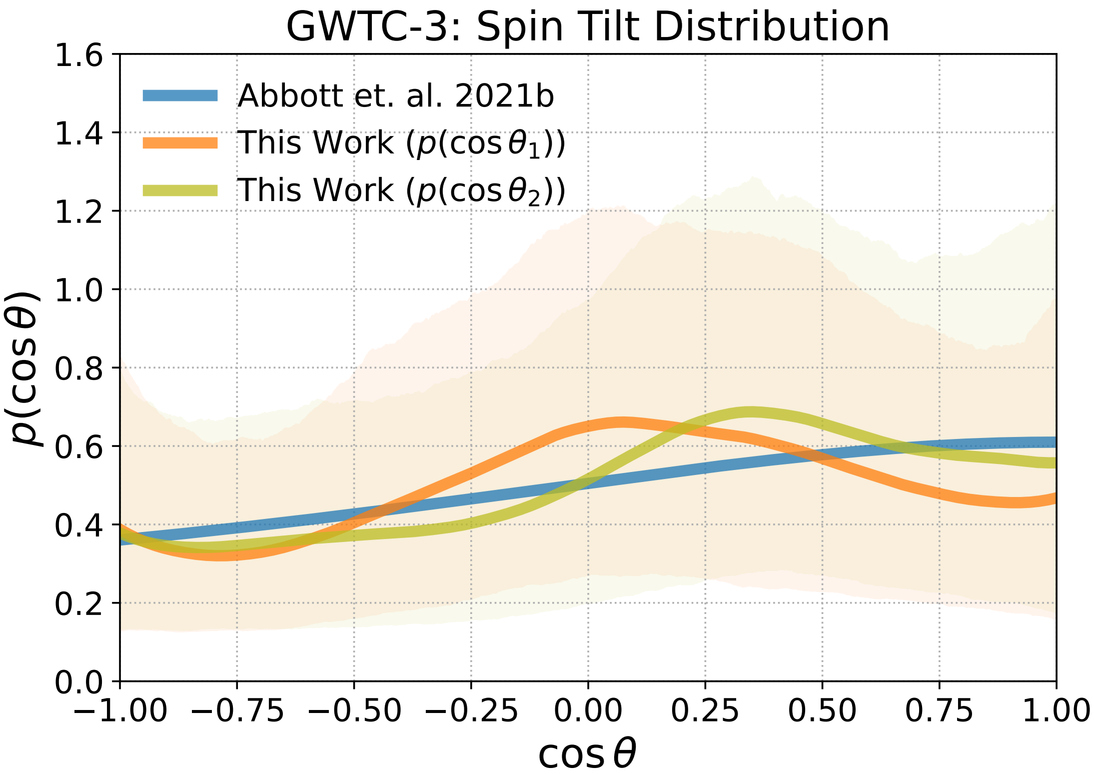

FIGURE 5.6. The primary (orange) and secondary (olive) spin orientation distributions inferred with the B-Spline model with 16 knots spaced linearly from -1 to 1. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. For comparison, we show the inferred PPD from the Default (blue) model from The LIGO Scientific Collaboration et al. [51], the LVK's GWTC-3 population analyses.

We also model each component's orientation distribution with an independent B-Spline model as done above, and show the inferred primary (orange), and secondary (olive) distributions in figure 5.6. The orientation distributions are broadly consistent with each other and the DEFAULT model's PPD given the wide credible intervals. We find the two distributions to peak at: $\cos\theta_1 = 0.16^{+0.84}_{-0.84}$ and $\cos\theta_2 = 0.38^{+0.62}_{-1.0}$, showing that the primary distribution peak is inferred further away from the assumed $\cos\theta = 1$ with the DEFAULT model. There is also significant (albeit uncertain) evidence of spin misalignment in each distribution, finding the fraction of misaligned primary and secondary components as: $f_{\cos\theta_1<0} = 0.43^{+0.19}_{-0.16}$ and $f_{\cos\theta_2<0} = 0.38^{+0.18}_{-0.15}$. We again calculate $\log_{10} Y$ for each component distribution and find: $\log_{10} Y_1 = 0.12^{+0.5}_{-0.53}$ and $\log_{10} Y_2 = 0.18^{+0.53}_{-0.54}$.

Model	$\chi_{\mathrm{eff,peak}}$	$f_{\chi_{\mathrm{eff}}<0}$	$f_{\chi_{\mathrm{eff}}<-0.3}$	f_{dyn}	f_{HM}
B-Spline IID	$0.039^{+0.034}_{-0.038}$	$0.34^{+0.11}_{-0.11}$	$0.019^{+0.021}_{-0.012}$	$0.69^{+0.22}_{-0.22}$	$0.12^{+0.13}_{-0.074}$
B-Spline Ind	$0.023^{+0.034}_{-0.034}$	$0.41^{+0.083}_{-0.088}$	$0.035^{+0.027}_{-0.018}$	$0.82^{+0.17}_{-0.18}$	$0.22^{+0.17}_{-0.11}$
DEFAULT [51]	$0.017^{+0.034}_{-0.022}$	$0.43^{+0.059}_{-0.13}$	$0.013^{+0.017}_{-0.0095}$	$0.87^{+0.12}_{-0.26}$	$0.081^{+0.11}_{-0.059}$
GAUSSIAN [51]	$0.06^{+0.029}_{-0.037}$	$0.28^{+0.15}_{-0.13}$	$0.00024^{+0.0081}_{-0.00024}$	$0.55^{+0.3}_{-0.26}$	$0.0015^{+0.051}_{-0.0015}$

TABLE 5.2. Summary of the effective spin distributions inferred with the B-Spline model variations, along with the DEFAULT and GAUSSIAN models from The LIGO Scientific Collaboration et al. [51].

5.5.3. The Effective Spin Dimension

While the component spin magnitudes and tilts are more directly tied to formation physics, they are typically poorly measured. The best-measured spin quantity, which enters at the highest post-Newtownian order, is the effective spin: $\chi_{\mathrm{eff}} = \frac{a_1\cos\theta_1 + q a_2 \cos\theta_2}{1+q}$. There is additionally an effective precessing spin parameter, $\chi_{\mathrm{p}} = \max\left[a_1 \sin\theta_1, \frac{3+4q}{4+3q} q a_2 \sin\theta_2\right]$, that quantifies the amount of spin precession given the systems mass ratio and component spin magnitudes and orientation.

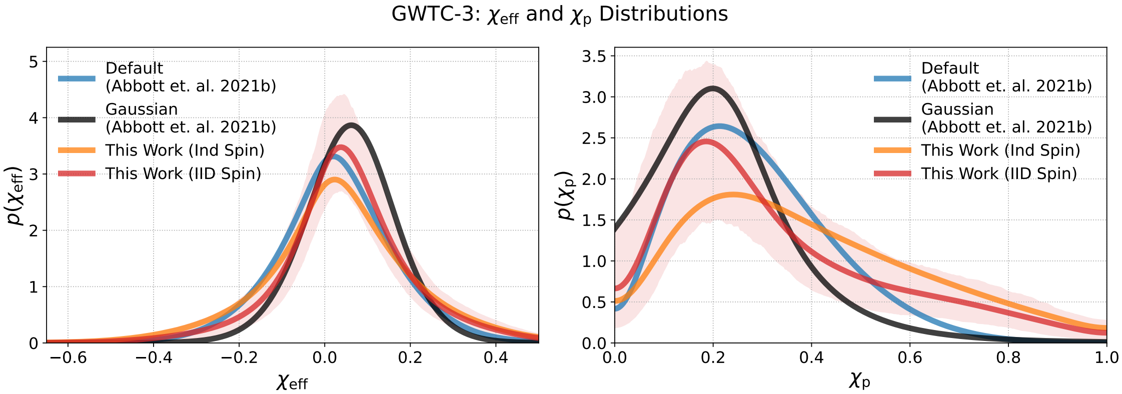

FIGURE 5.7. The effective (left) and precessing (right) spin distributions inferred with the B-Spline IID spin model (red). The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. We show the inferred PPDs from the independent component spin B-Spline model (purple), and both the Default (blue) model and the Gaussian (green) model from The LIGO Scientific Collaboration et al. [51], the LVK's GWTC-3 population analyses.

Figure 5.7 shows the inferred effective spin and precessing spin distributions with the two versions of our B-Spline models (red and purple), along with results on the Default [161] and Gaussian [167] models from The LIGO Scientific Collaboration et al. [51]. We find considerable agreement among the effective spin distributions, but the more flexible B-Spline models in component spins more closely resemble results from the Default model, also using the component spins. The B-Spline model finds very similar shapes to the other models, with a single peak centered at $\chi_{\rm eff} = 0.039^{+0.034}_{-0.038}$, compared to $\chi_{\rm eff} = 0.017^{+0.034}_{-0.022}$ with the Default model and $\chi_{\rm eff} = 0.06^{+0.029}_{-0.037}$ with the Gaussian $\chi_{\rm eff}$ models from The LIGO Scientific Collaboration et al. [51]. As for spin misalignment, we calculate the fraction of systems with effective spins that are misaligned (i.e. $\chi_{\rm eff} < 0$) and find similar agreement with previous work [51, 53, 148]. We find for the B-Spline model $f_{\chi_{\rm eff}<0} = 0.34^{+0.11}_{-0.11}$, compared to $f_{\chi_{\rm eff}<0} = 0.43^{+0.059}_{-0.13}$ and $f_{\chi_{\rm eff}<0} = 0.28^{+0.15}_{-0.13}$ with the Default and Gaussian models from The LIGO Scientific Collaboration et al.

[51]. The precessing spin distributions inferred with the B-Spline models exhibit a similar shape to the Default model, but with a much fatter tail towards highly precessing systems, driven by the extra support for highly spinning components seen in figures 5.3 and 5.3.

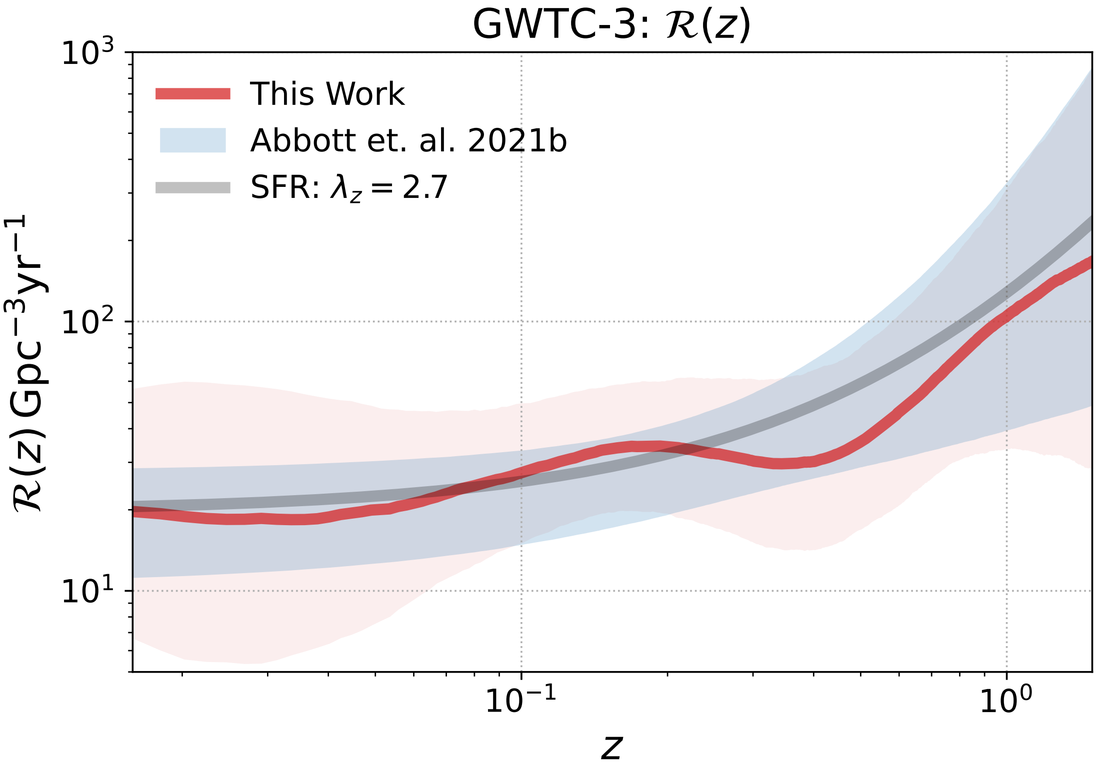

FIGURE 5.8. The BBH merger rate as a function of redshift. We show the B-Spline model (red) with 16 knots spaced linearly in $\log(z)$, from the minimum to the maximum observed redshifts. The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. We show the inferred 90% credible interval from the PowerlawRedshift model from the LVK's GWTC-3 population analyses in blue and a power law with exponent of 2.7 in gray, representing the expected star formation rate [51, 168].

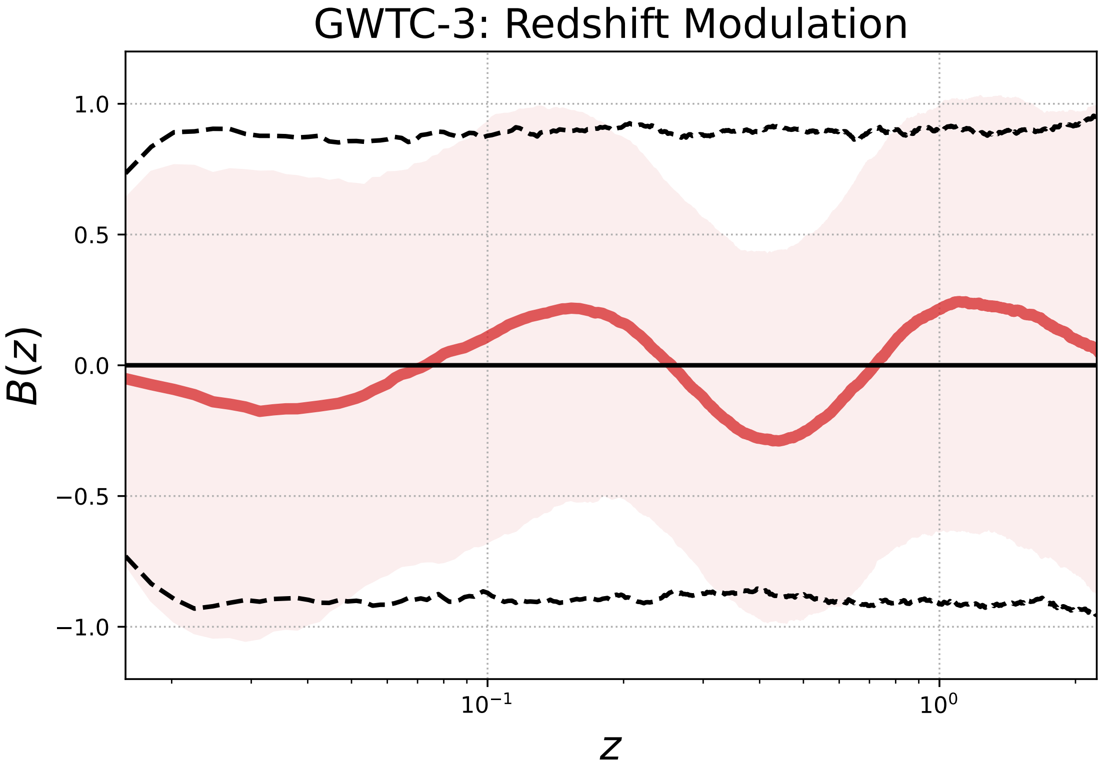

FIGURE 5.9. The B-spline modulation to the underlying power law in redshift, (red). The solid line shows the population predictive distribution (PPD), and the shaded region the 90% credible interval. We show the 90% credible interval of the prior predictive distribution in dashed black lines.

5.5.4. Merger Rate Evolution with Redshift

Recent analysis of the GWTC-3 BBH population has shown evidence for an increasing merger rate with redshift, nearly ruling out a merger rate that is constant with co-moving volume [51, 88]. When extending the power law form of the previously used model to have a modulation that we model with B-Splines, the merger rate as a function of redshift in figure 5.8 shows mild support for features departing from the underlying power law. In particular, we see a small increase in merger rate from $z \sim 0.09$ to $z \sim 0.2$ (where we best constrain the rate), followed by a plateau in the rate from $z \sim 0.2$ to $z \sim 0.4$. At larger redshifts, where we begin to have sparse observations, we see no sign of departure from the power-law as the rate continues to increase with redshift. The underlying power-law slope of our B-spline modulated model is consistent with the GWTC-3 results with the underlying model by itself: the PowerlawRedshift model found $\lambda_z = 2.7^{+1.8}_{-1.9}$ when inferred with the PowerlawPeak mass, and Default spin models. Our more flexible model infers a power law slope of $\lambda_z = 2.1^{+2.2}_{-2.5}$. We show the basis spline modulations or departure from the power law in 5.9 compared to the prior – showing where we cannot constrain any significant deviations from the simpler parametric power law model. The extra freedom of our model does inflate the uncertainty in its rate estimates, especially at $z \sim 0$ where there are not any observations in the catalog. We find a local ($z = 0$) merger rate of $\mathcal{R}_0 = 20^{+29}_{-12}\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ using the B-Spline modulation model which compares to $\mathcal{R}_0 = 17^{+10}_{-6.7}\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ for the GWTC-3 result.

5.6. Astrophysical Implications

The collective distribution of BBH source properties provides a useful probe of the complex and uncertain astrophysics that govern their formation and evolution until merger [109, 126, 127]. Our analyses with the newly constructed B-spline models uncover hints new features in the population (e.g., in mass ratio and redshift) and corroborates important conclusions of recent work, and provides a robust data-driven framework for future population studies.

The results presented in section 5.5.1 illustrate a wider mass distribution than inferred with power-law based models in The LIGO Scientific Collaboration et al. [51], and a suppressed merger rate at low primary masses (i.e. $\leq 8\mathrm{M}_\odot$), showing possible signs of binary selection effects or the purported low mass gap between neutron stars and black holes [98, 119, 169]. While isolated formation is able to predict the $10\mathrm{M}_{\odot}$ peak [170], cluster and dynamical formation scenarios struggle to predict a peak in the BH mass distribution less than $15 - 20\mathrm{M}_{\odot}$ [105, 171]. Globular cluster formation is expected to produce more top-heavy mass distributions than isolated and recent studies have shown suppressed BBH merger rates at lower ($m \leq 15\mathrm{M}_{\odot}$) masses when compared to predictions from the isolated channel [60, 105, 144, 172]. BBHs that form near active galactic nuclei (AGN) can preferentially produce higher mass black holes [173, 174, 175]. We do not find any evidence for a truncation or rapid decline in the merger rate as a function of mass, that stellar evolution theory predicts due to pair-instability supernovae (PISNe) [54, 55, 56, 59, 62]. The original motivation for the peak in the PowerlawPeak model [67] was to represent a possible "pileup" of masses just before such truncation, since massive stars just light enough to avoid PISN will shed large amounts of mass in a series of "pulses" before collapsing to BHs in a process called pulsational pair-instability supernova (PPISN) [57, 58, 61]. While the predictions of the mass scale where pair-instability kicks in are uncertain and depend on poorly understood physics like nuclear reaction rates of carbon and oxygen in the core of stars, models have a hard time producing this peak lower than $m \sim 40\mathrm{M}_{\odot}$ [60, 61, 63, 103, 176]. The lack of a truncation could point towards a higher prevalence of dynamical processes that can produce black holes in mass ranges stellar collapse cannot, such as hierarchical mergers of BHs [66, 69, 70, 75, 94, 177], very low metallicity population III stars [64, 72], new beyond-standard-model physics[77, 106], or black hole accretion of BHs in gaseous environments such as AGNs [73, 74, 78].

Our constraints on the mass ratio distribution are not yet precise enough to claim definitive departures from power law behavior, but do suggest possible plateaus in the rate at several mass ratios, including equal mass. These features should sharpen (or resolve) with future updates to the catalog.

Section 5.5.2 focused on inferences of the spin distributions of black holes, observing evidence of spin misalignment, spin anti-alignment, and suppressed support for exactly aligned systems. These point towards a significant contribution to the population from dynamical formation processes, agreeing with conclusions drawn about the mass distribution inference of section 5.5.1. While field formation is expected to produce systems with preferentially aligned spins due to tidal interactions, observational evidence suggests that tides may not be able to re-align spins in all systems as some isolated population models assume. Additionally, because of uncertain knowledge of supernovae kicks, isolated formation can produce systems with negative but small effective spins. Consistent with recent studies we report an effective spin distribution that is not symmetric about zero, disfavoring a scenario in which all BBHs are formed dynamically [51, 53, 148]. Following the rules in Fishbach et al. [177], we place conservative upper bounds on the fraction of hierarchical mergers f_{HM} and fraction of dynamically formed BBHs, f_{dyn} with the B-spline χ_{eff} model constraining f_{HM} <0.058 and f_{dyn} <0.52 at 90% credibility. This is consistent with the 90% credible interval found from the GWTC-2 analysis, $0.25 \leq f_{\mathrm{dyn}} \leq 0.93$ [53].

Finally, section 5.5.4 shows potentially interesting evolution of the BBH merger rate with redshift. Though uncertainties are still large, we may be seeing the first signs of departure from following the star formation rate, which could help in distinguishing different subpopulations should they exist [178]. Again, we expect these features to be resolved with future catalogs.

5.7. Conclusions

Non-parametric and data-driven statistical modeling methods have been put to use with great success across the ever-growing field of gravitational waves [8, 25, 31, 96, 97, 111, 112, 113, 114, 140, 151, 179]. We presented a case study exploring how basis splines make for an especially powerful and efficient data driven method of characterizing the binary black hole population observed with gravitational waves, along with the associated open source software GWInferno, that implements the models described in this paper and performs hierarchical Bayesian inference with NumPyro and Jax [180, 181, 182]. Our study paves the way as the first completely non-parametric compact object population study, employing data driven models for each of the hierarchically modeled population distributions. A complete understanding of the population properties of compact objects will help to advance poorly understood areas of stellar and nuclear astrophysics and provide a novel independent cosmological probe. With the coming influx of new data with the LVK's next observing run, development of model-agnostic methods, such as the one we proposed here, will become necessary to efficiently make sense of the vast amounts of data and to extract as much information as possible from the population.

5.8. Acknowledgements

We thank Tom Callister, Will Farr, Maya Fishbach, Salvatore Vitale, and Jaxen Godfrey for useful discussions during the preparation of this manuscript and/or helpful comments on early drafts. Z.D. also acknowledges support from the CIERA Board of Visitors Research Professorship. This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (`https://www.gw-openscience.org/`), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This work was supported in part by the National Science Foundation under Grant PHY-2146528 and benefited from access to the University of Oregon high performance computer, Talapas. This material is based upon work supported in part by the National Science Foundation under Grant PHY-1807046 and work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

5.9. Additional Analysis Details

5.9.1. Basis Splines

A common non-parametric method used in many statistical applications is basis splines. A spline function of order k, is a piece-wise polynomial of order k polynomials stitched together from defined "knot" locations across the domain. They provide a useful and cheap way to interpolate generically smooth functions from a finite sampling of "knot" heights. Basis splines of order k are a set of order k polynomials that form a complete basis for any spline function of order k. Therefore, given an array of knot locations, $\mathbf{t}$ or knot vector, there exists a single unique linear combination of basis splines for every possible spline function interpolated from $\mathbf{t}$. To construct a basis of n components and knots, t_0, t_1,...,t_{i+k}, we use the Cox-de Boor recursion formula [6, 183]. The recursion starts with the $k = 0$ (constant) case and recursively constructs the basis components of higher orders. The base case and recursion relation that generates this particular basis are defined as:

$$B_{i,0}(x|\mathbf{t}) = \begin{cases} 1, & \text{if } t_i \leq x < t_{i+1} \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

$$B_{i,k+1}(x|\mathbf{t}) = \omega_{i,k}(x|\mathbf{t}) B_{i,k}(x|\mathbf{t}) + \big[1 - \omega_{i+1,k}(x|\mathbf{t})\big] B_{i+1,k}(x|\mathbf{t}) \tag{5.3}$$

$$\omega_{i,k}(x|\mathbf{t}) = \begin{cases} \frac{x-t_i}{t_{i+k}-t_i}, & t_{i+k} \neq t_i \\ 0, & \text{otherwise} \end{cases} \tag{5.4}$$

This is known as the "B-Spline" basis after it's inventor de Boor [6]. The power of basis splines comes from the fact that one only has to do the somewhat-expensive interpolation once for each set of points at which the spline is evaluated. This provides a considerable computational speedup as each evaluation of the spline function becomes a simpler operation: a dot product of a matrix and a vector. This straightforward operation is also ideal for optimizations from the use of GPU accelerators, enabling our Markov chain Monte Carlo (MCMC) based analyses, often with hundreds of parameters, to converge in an hour or less. Basis splines can easily be generalized to their two-dimensional analog, producing tensor product basis splines that, with this computational advantage, allow for high fidelity modeling of two-dimensional spline functions.

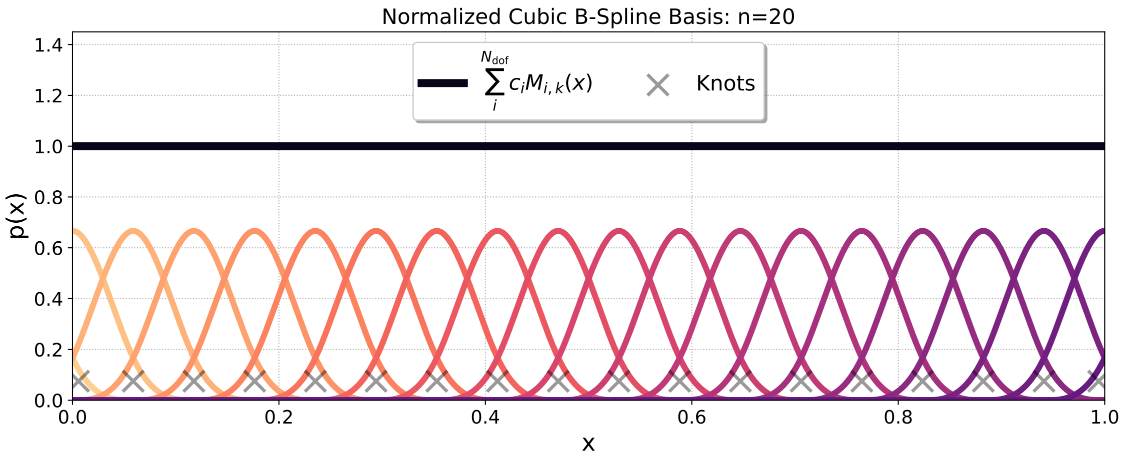

FIGURE 5.10. Plot showing a “proper” (see appendix 5.9.2) normalized B-Spline basis of order 3 (cubic) with 20 degrees of freedom and equal weights for each component. In black, we show the resulting spline function given equal weights and denote the location of the knots with gray x’s.

Another important feature of basis splines is that under appropriate prior conditions, one can alleviate sensitivities to arbitrarily chosen prior specifications that splines commonly struggle with. Previous studies using splines had to perform multiple analyses, varying the number of spline knots, then either marginalized over the models or used model comparisons to motivate the best choice [97]. We can avoid this step with the use of penalized splines (or P-Splines) [7, 152, 153], where one adds a smoothing prior comprised of Gaussian distributions on the differences between neighboring basis spline coefficients. This allows for knots to be densely populated across the domain without the worry of extra variance in the inferred spline functions. When also fitting the scale of the smoothing prior (i.e. the width of the Gaussian distributions on the differences), the data will inform the model of the preferred the scale of smoothing required. We discuss the details of our smoothing prior implementation in more detail in the next section,

Appendix 5.9.2, following with our specific prior and basis choices for each model in Appendix 5.9.4.

5.9.2. Penalized Splines and Smoothing Priors

Spline functions have been shown to be sensitive to the chosen number of knots, and their locations or spacing [6]. Adding more knots increases the a priori variance in the spline function, while the space between knots can limit the resolution of features in the data the spline is capable of resolving. To ensure your spline based model is flexible enough one would want to add as many knots as densely as possible, but this comes with unwanted side effect of larger variance imposed by your model. This can be fixed with the use of penalized splines (P-Spline) in which one applies a prior or regularization term to the likelihood based on the difference of adjacent knot coefficients [7]. The linear combination of spline basis components or the resulting spline function is flat when the basis coefficients are equal (see Figure 5.10). By penalizing the likelihood as the differences between adjacent knot coefficients get larger, one gets a smoothing effect on the spline function [7]. With hierarchical Bayesian inference as our statistical framework, we formulate the penalized likelihood of Eilers and Marx [7]'s P-Splines with their Bayesian analog [152]. The Bayesian P-Spline prior places Gaussian distributions over the r-th order differences of the coefficients [152, 153]. This is also sometimes referred to as a Gaussian random walk prior, and is similar in spirit to a Gaussian process prior used to regularize or smooth histogram bin heights as done in other non-parametric population studies [51, 113]. For a spline basis with n degree's of freedom, and a difference penalty of order of r (see Eilers and Marx [7]), the smoothing prior on our basis spline coefficients, $\boldsymbol{c}$ is defined as:

$$\boldsymbol{c} \sim \mathcal{N}(0, \sigma) \tag{5.5}$$

$$p(\boldsymbol{c}|\tau_\lambda) \propto \exp\big[-\frac{1}{2}\tau_\lambda \boldsymbol{c}^{\mathrm{T}} \boldsymbol{D}_r^{\mathrm{T}} \boldsymbol{D}_r \boldsymbol{c}\big] \tag{5.6}$$

Above $\boldsymbol{D}_r$ is the order-r difference matrix, of shape $(n - r \times n)$, and $\mathcal{N}(0, \sigma)$ a Gaussian distribution with zero mean and standard deviation, σ. This smoothing prior removes the strong dependence on number and location of knots that arises with using splines. The τ_λ controls the "strength" of the smoothing, or the inverse variance of the Gaussian priors on knot differences. We place uniform priors on τ_λ marginalize over this smoothing scale hyperparameter to let the data inform the optimal scale needed. When there are a very large number of knots, such that your domain is densely populated with basis coefficients, this allows the freedom for the model to find the smoothing scale that the data prefers.

This prior is imparting a natural attraction of the coefficients closer to each other in order to smooth the spline function, so one must ensure that the spline function is in fact flat given all equal coefficients. There needs to be $n + k + 1$ knots to construct an order-k basis with n degrees of freedom. Some studies place knots on top of each other at hard parameter boundaries [6, 183], which may seem motivated, but this violates the above condition necessary for the P-Spline prior. We follow the distinction in Eilers and Marx [7] that such a smoothing prior is only valid with "proper" spline bases. A proper basis is where all $n + k + 1$ knots are evenly and equally spaced, see Figure 5.10, as opposed to stacking them at the bounds.

5.9.3. Hierarchical Bayesian Inference

We use hierarchical Bayesian inference to infer the population properties of compact binaries. We want to infer the number density of merging compact binaries in the universe and how this can change with their masses, spins, etc. Often times it is useful to formulate the question in terms of the merger rates which is the number of mergers per Gpc^3 co-moving volume per year. For a set of hyperparameters, Λ, λ, and overall merger rate, $\mathcal{R}$, we write the overall number density of BBH mergers in the universe as:

$$\frac{dN(\theta, z|\mathcal{R}, \Lambda, \lambda)}{d\theta dz} = \frac{dV_c}{dz}\left(\frac{T_{\text{obs}}}{1+z}\right)\frac{d\mathcal{R}(\theta, z|\mathcal{R}_0, \Lambda, \lambda)}{d\theta} = \mathcal{R}p(\theta|\Lambda)p(z|\lambda) \tag{5.7}$$

where up above, we denote the co-moving volume element as dV_c [117], and T_{obs} as the observing time period that produced the catalog with the related factor of $1+z$ converting this detector-frame time to source-frame. We assume a Lambda CDM cosmology using the cosmological parameters from Planck Collaboration et al. [120]. We model the merger rate evolving with redshift following a power law distribution: $p(z|\lambda) \propto \frac{dV_c}{dz}\frac{1}{1+z}(1+z)^{\lambda}$ [88]. When integrating equation 5.7 across all θ and out to some maximum redshift, z_{max}, we get the total number of compact binaries in the universe out to that redshift. We follow previous notations, letting $\{d_i\}$ represent the set of data from N_{obs} compact binaries observed with gravitational waves. The merger rate is then described as an inhomogeneous Poisson process and after imposing the usual log-uniform prior on the merger rate, we marginalize over the merger rate, $\mathcal{R}$, and arrive at the posterior distribution of our hyperparameters, Λ [83, 111].

$$p\left(\Lambda, \lambda|\{d_i\}\right) \frac{p(\Lambda)p(\lambda)}{\xi(\Lambda, \lambda)^{N_{\rm obs}}} \prod_{i=1}^{N_{\rm obs}} \left[\frac{1}{K_i} \sum_{j=1}^{K_i} \frac{p(\theta^{i,j}|\Lambda)p(z^{i,j}|\lambda)}{\pi(\theta, z^{i,j})}\right] \tag{5.8}$$

where above, we replaced the integrals over each event's likelihood with ensemble averages over K_i posterior samples [154]. Above, j indexes the K_i posterior samples from each event and $\pi(\theta, z)$ is the default prior used by parameter estimations that produced the posterior samples for each event. In the analyses of GWTC-3, either the default prior used was uniform in detector frame masses, component spins and Euclidean volume or the posterior samples were re-weighted to such a prior before using them in our analysis. The corresponding prior evaluated in the parameters we hierarchically model, i.e. source frame primary mass, mass ratio, component spins and redshift is:

$$\pi(m_1, q, a_1, a_2, \cos\theta_1, \cos\theta_2, z) \propto \frac{1}{4} m_1 (1+z)^2 D_L^2(z) \frac{dD_L}{dz} \tag{5.9}$$

Above, D_L is the luminosity distance. To carefully incorporate selection effects to our model we need to quantify the detection efficiency, $\xi(\Lambda, \lambda)$, of the search pipelines that were used to create GWTC-3, at a given population distribution described by Λ and λ.

$$\xi(\Lambda, \lambda) = \int d\theta dz P_{\rm det}(\theta, z) p(\theta|\Lambda) p(z|\lambda) \tag{5.10}$$

To estimate this integral we use a software injection campaign where gravitational waveforms from a large population of simulated sources. These simulated waveforms are put into real detector data, and then this data is evaluated with the same search pipelines that were used to produce the catalog we are analyzing. With these search results in hand, we use importance sampling and evaluate the integral with the Monte Carlo sum estimate μ, and its corresponding variance and effective number of samples:

$$\xi(\Lambda, \lambda) \approx \mu(\Lambda, \lambda) \frac{1}{N_{\mathrm{inj}}} \sum_{i=1}^{N_{\mathrm{found}}} \frac{p(\theta^i|\Lambda)p(z^i|\lambda)}{p_{\mathrm{inj}}(\theta, z^i)} \tag{5.11}$$

$$\sigma^2(\Lambda, \lambda) \equiv \frac{\mu^2(\Lambda, \lambda)}{N_{\mathrm{eff}}} \simeq \frac{1}{N_{\mathrm{inj}}^2} \sum_{i=1}^{N_{\mathrm{found}}} \left[\frac{p(\theta|\Lambda)p(z|\lambda)}{p_{\mathrm{inj}}(\theta, z)}\right]^2 - \frac{\mu^2(\Lambda, \lambda)}{N_{\mathrm{inj}}} \tag{5.12}$$

where the sum is only over the N_{found} injections that were successfully detected out of N_{inj} total injections, and $p_{\mathrm{inj}}(\theta, z)$ is the reference distribution from which the injections were drawn. We use the LVK released injection sets that describe the detector sensitivities over the first, second and third observing runs [184]. Additionally, we follow the procedure outlined in Farr [82] to marginalize the uncertainty in our estimate of $\xi(\Lambda, \lambda)$, in which we verify that N_{eff} is sufficiently high after re-weighting the injections to a given population (i.e. $N_{\mathrm{eff}} > 4N_{\mathrm{obs}}$). The total hyper-posterior marginalized over the merger rate and the uncertainty in the Monte Carlo integral calculating $\xi(\Lambda, \lambda)$ [82], as:

$$\log p\left(\Lambda, \lambda|\{d_i\}\right) \propto \sum_{i=1}^{N_{\mathrm{obs}}} \log\left[\frac{1}{K_i} \sum_{j=1}^{K_i} \frac{p(\theta^{i,j}|\Lambda)p(z^{i,j}|\lambda)}{\pi(\theta^{i,j}, z^{i,j})}\right] - N_{\mathrm{obs}} \log \mu(\Lambda, \lambda) + \frac{3N_{\mathrm{obs}} + N_{\mathrm{obs}}^2}{2N_{\mathrm{eff}}} + \mathcal{O}(N_{\mathrm{eff}}^{-2}). \tag{5.13}$$

We explicitly enumerate each of the models used in this work for $p(\theta|\Lambda)$, along with their respective hyperparameters and prior distributions in the next section. To calculate draw samples of the hyperparameters from the hierarchical posterior distribution shown in equation 5.13, we use the NUTS

Model	**Parameter**	**Description**	**Prior**
Primary Mass Model Parameters			
B-Spline Primary	$\boldsymbol{c}$	Basis coefficients	$\sim$ Smooth$(\tau_\lambda, \sigma, r, n)$
	τ_λ	Smoothing Prior Scale	$\sim$ U$(2, 1000)$
	r	order of the difference matrix for the smoothing prior	2
	σ	width of Gaussian priors on coefficients in smoothing prior	6
	n	number of knots in the basis spline	64
Mass Ratio Model Parameters			
B-Spline Ratio	$\boldsymbol{c}$	Basis coefficients	$\sim$ Smooth$(\tau_\lambda, \sigma, r, n)$
	τ_λ	Smoothing Prior Scale	$\sim$ U$(1, 100)$
	r	order of the difference matrix for the smoothing prior	2
	σ	width of Gaussian priors on coefficients in smoothing prior	4
	n	number of knots in the basis spline	18
Redshift Evolution Model Parameters			
PowerLaw+B-Spline	λ	slope of redshift evolution power law $(1+z)^\lambda$	$\sim \mathcal{N}(0, 3)$
	$\boldsymbol{c}$	Basis coefficients	$\sim$ Smooth$(\tau_\lambda, \sigma, r, n)$
	τ_λ	Smoothing Prior Scale	$\sim$ U$(1, 10)$
	r	order of the difference matrix for the smoothing prior	2
	σ	width of Gaussian priors on coefficients in smoothing prior	1
	n	number of knots in the basis spline	18
Spin Distribution Model Parameters			
B-Spline Magnitude	$\boldsymbol{c}$	Basis coefficients	$\sim$ Smooth$(\tau_\lambda, \sigma, r, n)$
	τ_λ	Smoothing Prior Scale	$\sim$ U$(1, 10)$
	r	order of the difference matrix for the smoothing prior	2
	σ	width of Gaussian priors on coefficients in smoothing prior	1
	n	number of knots in the basis spline	18
B-Spline Tilt	$\boldsymbol{c}$	Basis coefficients	$\sim$ Smooth$(\tau_\lambda, \sigma, r, n)$
	τ_λ	Smoothing Prior Scale	$\sim$ U$(1, 10)$
	r	order of the difference matrix for the smoothing prior	2
	σ	width of Gaussian priors on coefficients in smoothing prior	1
	n	number of knots in the basis spline	18

TABLE 5.3. All hyperparameter prior choices for each of the newly introduced basis spline models from this manuscript. See appendix 5.9.1 and 5.9.2 for more detailed description of basis spline or smoothing prior parameters.

Hamiltonian Monte Carlo sampler in NumPyro and Jax to calculate likelihoods [180, 181, 182].

5.9.4. Model and Prior Specification

For each of the distributions with basis spline distributions, we have 2 fixed hyperparameters to specify. The number of degrees of freedom, n, and the difference penalty order for the smoothing prior, r. Additionally, one must choose a prior distribution on the smoothing prior scale hyperparameter, τ_λ, which we take to be Uniform. For the primary mass distribution we model the log probability with a B-Spline interpolated in $\log(m_1)$ space. We follow a similar

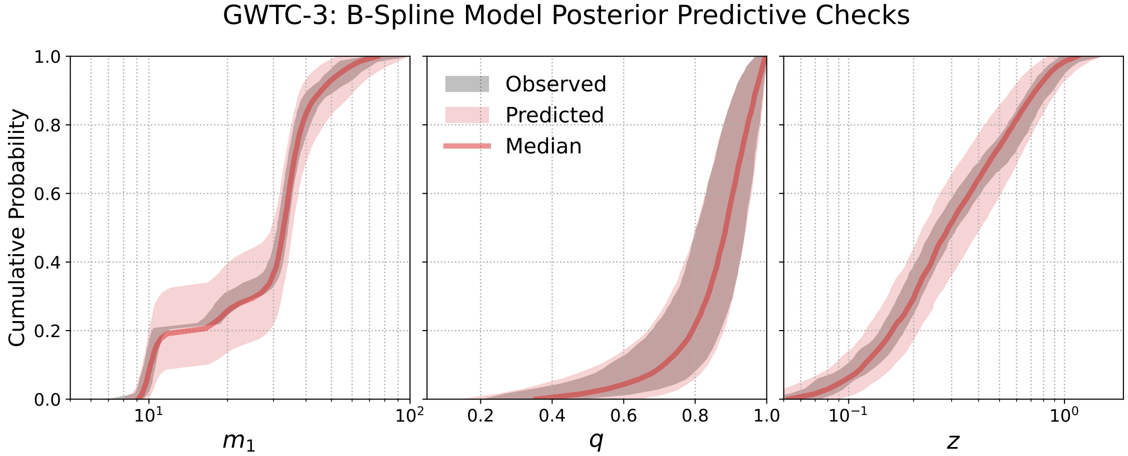

FIGURE 5.11. Posterior predictive checks showing the CDFs of the observed (black) and predicted (red) distributions of GWTC-3 sized catalogs for each posterior sample of the IID spin B-Spline model. The shaded regions show 90% credible intervals and the solid red line is the median of the predicted distribution.

scheme for the models in mass ratio and spin, except we model the log probability with B-Splines that are interpolated in q, a_i or $\cos\theta_i$ space. We adopt a minimum black hole mass of $5\mathrm{M}_{\odot}$, and maximum of $100\mathrm{M}_{\odot}$ with the equally spaced in this range. The knots for the mass ratio B-Spline are equally spaced from $\frac{m_{\min}}{m_{\max}} = 0.05$ to 1. There is motivation for the evolution of the merger rate with redshift to follow a power law form since it should be related to the star formation rate [168], motivating our adoption of a semi-parametric approach where we use B-Splines to model modulations to the simpler underlying PowerlawRedshift model [88, 97]. We model modulations to the underlying probability density with the multiplicative factor, $e^{B(\log z)}$, where $B(\log z)$ is the B-Spline interpolated from knots spaced linearly in $\log z$ space. We enumerate each of our specific model hyperparameter and prior choices in table 5.3.

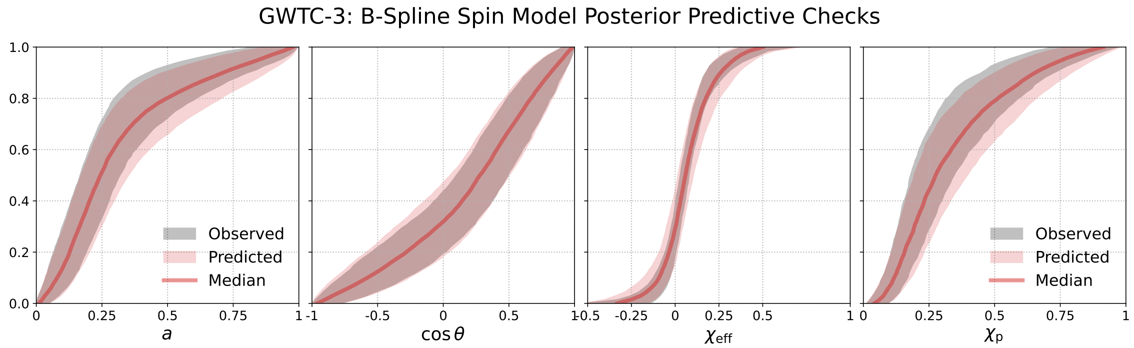

FIGURE 5.12. Posterior predictive checks showing the CDFs of the observed (black) and predicted (red) distributions of GWTC-3 sized catalogs for each posterior sample of the IID spin B-Spline model. The shaded regions show 90% credible intervals and the solid red line is the median of the predicted distribution.

5.9.5. Posterior Predictive Checks

We follow the posterior predictive checking procedure done in recent population studies to validate our models inferences [53, 97]. For each posterior sample describing our model's inferred population we reweigh the observed event samples and the found injections to that population and draw a set 69 (size of GWTC-3 BBH catalog) samples to construct the observed and predicted distributions we show in figure 5.11 and figure 5.12. When the observed region stays encompassed within the predicted region the model is performing well, which we see across each of the fit parameters.

5.9.6. Reproducibility

In the spirit of open source and reproducible science, this study was done using the reproducibility software ShowYourWork [185], which leverages continuous integration to programmatically download the data from zenodo.org, create the figures, and compile the manuscript. Each figure caption contains two links that point towards the dataset (stored on zenodo) used in the corresponding figure, and to the script used to make the figure (at the commit corresponding to the current build of the manuscript). The git repository associated to this study is publicly available at `https://github.com/bruce-edelman/CoveringYourBasis`, which allows anyone to re-build the entire manuscript. The datasets and all analysis or figure generating scripts are all stored on zenodo.org at `https://zenodo.org/record/7566301` [186].

REFERENCES CITED

[1] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, and et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. *arXiv e-prints*, art. arXiv:2111.03606, November 2021. doi: 10.48550/arXiv.2111.03606.

[2] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. *prd*, 100(2):024059, July 2019. doi: 10.1103/PhysRevD.100.024059.

[3] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa. Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes. *prd*, 103(10):104056, May 2021. doi: 10.1103/PhysRevD.103.104056.

[4] Serguei Ossokine, Alessandra Buonanno, Sylvain Marsat, Roberto Cotesta, Stanislav Babak, Tim Dietrich, Roland Haas, Ian Hinder, Harald P. Pfeiffer, Michael Pürrer, Charles J. Woodford, Michael Boyle, Lawrence E. Kidder, Mark A. Scheel, and Béla Szilágyi. Multipolar effective-one-body waveforms for precessing binary black holes: Construction and validation. *prd*, 102(4): 044055, August 2020. doi: 10.1103/PhysRevD.102.044055.

[5] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Davide Gerosa, Leo C. Stein, Lawrence E. Kidder, and Harald P. Pfeiffer. Surrogate models for precessing binary black hole simulations with unequal masses. *Physical Review Research*, 1(3):033015, October 2019. doi: 10.1103/PhysRevResearch.1.033015.

[6] Carl de Boor. *A Practical Guide to Splines*. Applied Mathematical Sciences. Springer, 1978. ISBN 978-1-4612-6333-3.

[7] P.H.C. Eilers and B.D. Marx. *Practical Smoothing: The Joys of P-splines*. Cambridge University Press, 2021. ISBN 9781108482950. URL `https://books.google.com/books?id=ez0QEAAAQBAJ`.

[8] Bruce Edelman, F. J. Rivera-Paleo, J. D. Merritt, Ben Farr, Zoheyr Doctor, Jeandrew Brink, Will M. Farr, Jonathan Gair, Joey Shapiro Key, Jess McIver, and Alex B. Nielsen. Constraining unmodeled physics with compact binary mergers from GWTC-1. *prd*, 103(4):042004, February 2021. doi: 10.1103/PhysRevD.103.042004.

[9] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Physical Review X*, 9(3):031040, July 2019. doi: 10.1103/PhysRevX.9.031040.

[10] Clifford M. Will. The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, 17(1):4, December 2014. doi: 10.12942/lrr-2014-4.

[11] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *prl*, 116(6):061102, February 2016. doi: 10.1103/PhysRevLett.116.061102.

[12] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Properties of the Binary Black Hole Merger GW150914. *prl*, 116(24):241102, June 2016. doi: 10.1103/PhysRevLett.116.241102.

[13] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Tests of General Relativity with GW150914. *prl*, 116(22):221101, June 2016. doi: 10.1103/PhysRevLett.116.221101.

[14] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. Tests of General Relativity with GW170817. *prl*, 123(1):011102, July 2019. doi: 10.1103/PhysRevLett.123.011102.

[15] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. Tests of general relativity with the binary black hole signals from the LIGO-Virgo catalog GWTC-1. *prd*, 100(10):104036, November 2019. doi: 10.1103/PhysRevD.100.104036.

[16] Abhirup Ghosh, Archisman Ghosh, Nathan K. Johnson-McDaniel, Chandra Kant Mishra, Parameswaran Ajith, Walter Del Pozzo, David A. Nichols, Yanbei Chen, Alex B. Nielsen, Christopher P. L. Berry, and Lionel London. Testing general relativity using golden black-hole binaries. *prd*, 94 (2):021101, July 2016. doi: 10.1103/PhysRevD.94.021101.

[17] Christopher J. Moore and Jonathan R. Gair. Novel Method for Incorporating Model Uncertainties into Gravitational Wave Parameter Estimates. *prl*, 113 (25):251101, December 2014. doi: 10.1103/PhysRevLett.113.251101.

[18] R. J. E. Smith, K. Cannon, C. Hanna, D. Keppel, and I. Mandel. Towards rapid parameter estimation on gravitational waves from compact binaries using interpolated waveforms. *prd*, 87(12):122002, June 2013. doi: 10.1103/PhysRevD.87.122002.

[19] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, W. Del Pozzo, F. Feroz, J. Gair, C. J. Haster, V. Kalogera, T. Littenberg, I. Mandel, R. O'Shaughnessy, M. Pitkin, C. Rodriguez, C. Röver, T. Sidery, R. Smith, M. Van Der Sluys, A. Vecchio, W. Vousden, and L. Wade. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *prd*, 91(4):042003, February 2015. doi: 10.1103/PhysRevD.91.042003.

[20] Katerina Chatziioannou, Roberto Cotesta, Sudarshan Ghonge, Jacob Lange, Ken K. Y. Ng, Juan Calderón Bustillo, James Clark, Carl-Johan Haster, Sebastian Khan, Michael Pürrer, Vivien Raymond, Salvatore Vitale, Nousha Afshari, Stanislav Babak, Kevin Barkett, Jonathan Blackman, Alejandro Bohé, Michael Boyle, Alessandra Buonanno, Manuela Campanelli, Gregorio Carullo, Tony Chu, Eric Flynn, Heather Fong, Alyssa Garcia, Matthew Giesler, Maria Haney, Mark Hannam, Ian Harry, James Healy, Daniel Hemberger, Ian Hinder, Karan Jani, Bhavesh Khamersa, Lawrence E. Kidder, Prayush Kumar, Pablo Laguna, Carlos O. Lousto, Geoffrey Lovelace, Tyson B. Littenberg, Lionel London, Margaret Millhouse, Laura K. Nuttall, Frank Ohme, Richard O'Shaughnessy, Serguei Ossokine, Francesco Pannarale, Patricia Schmidt, Harald P. Pfeiffer, Mark A. Scheel, Lijing Shao, Deirdre Shoemaker, Bela Szilagyi, Andrea Taracchini, Saul A. Teukolsky, and Yosef Zlochower. On the properties of the massive binary black hole merger GW170729. *prd*, 100(10):104015, November 2019. doi: 10.1103/PhysRevD.100.104015.

[21] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. Constraining the p -Mode-g -Mode Tidal Instability with GW170817. *prl*, 122(6):061104, February 2019. doi: 10.1103/PhysRevLett.122.061104.

[22] Abhirup Ghosh, Nathan K. Johnson-McDaniel, Archisman Ghosh, Chandra Kant Mishra, Parameswaran Ajith, Walter Del Pozzo, Christopher P. L. Berry, Alex B. Nielsen, and Lionel London. Testing general relativity using gravitational wave signals from the inspiral, merger and ringdown of binary black holes. *Classical and Quantum Gravity*, 35(1):014002, January 2018. doi: 10.1088/1361-6382/aa972e.

[23] Matteo Breschi, Richard O'Shaughnessy, Jacob Lange, and Ofek Birnholtz. IMR consistency tests with higher modes on gravitational signals from the second observing run of LIGO and Virgo. *Classical and Quantum Gravity*, 36 (24):245019, December 2019. doi: 10.1088/1361-6382/ab5629.

[24] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio. Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence. *prd*, 85(8):082003, April 2012. doi: 10.1103/PhysRevD.85.082003.

[25] Will M Farr, Ben Farr, and Tyson Littenberg. Modelling calibration errors in cbc waveforms. Technical Report LIGO-T1400682, LIGO Project, 2015. URL `https://dcc.ligo.org/LIGO-P1500262/public`.

[26] LIGO Scientific Collaboration, J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, April 2015. doi: 10.1088/0264-9381/32/7/074001.

[27] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, and et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, January 2015. doi: 10.1088/0264-9381/32/2/024001.

[28] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *prl*, 119(16):161101, October 2017. doi: 10.1103/PhysRevLett.119.161101.

[29] Salvatore Vitale, Walter Del Pozzo, Tjonnie G. F. Li, Chris Van Den Broeck, Ilya Mandel, Ben Aylott, and John Veitch. Effect of calibration errors on Bayesian parameter estimation for gravitational wave signals from inspiral binary systems in the advanced detectors era. *prd*, 85(6):064034, March 2012. doi: 10.1103/PhysRevD.85.064034.

[30] George Kimeldorf and Grace Wahba. A correspondence between bayesian estimation on stochastic processes and smoothing by splines. *Annals of Mathematical Statistics*, 41:495–502, 1970.

[31] Tyson B. Littenberg and Neil J. Cornish. Bayesian inference for spectral estimation of gravitational wave detector noise. *prd*, 91(8):084034, April 2015. doi: 10.1103/PhysRevD.91.084034.

[32] Laura Sampson, Nicolás Yunes, Neil Cornish, Marcelo Ponce, Enrico Barausse, Antoine Klein, Carlos Palenzuela, and Luis Lehner. Projected constraints on scalarization with gravitational waves from neutron star binaries. *prd*, 90(12): 124091, December 2014. doi: 10.1103/PhysRevD.90.124091.

[33] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika Linde. The Deviance Information Criterion: 12 Years on. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(3):485–493, 04 2014. ISSN 1369-7412. doi: 10.1111/rssb.12062. URL `https://doi.org/10.1111/rssb.12062`.

[34] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002. doi: 10.1111/1467-9868.00353. URL `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00353`.

[35] Francesco Pannarale, Emanuele Berti, Koutarou Kyutoku, Benjamin D. Lackey, and Masaru Shibata. Aligned spin neutron star-black hole mergers: A gravitational waveform amplitude model. *prd*, 92(8):084050, October 2015. doi: 10.1103/PhysRevD.92.084050.

[36] Benjamin D. Lackey, Koutarou Kyutoku, Masaru Shibata, Patrick R. Brady, and John L. Friedman. Extracting equation of state parameters from black hole-neutron star mergers: Nonspinning black holes. *prd*, 85(4):044061, February 2012. doi: 10.1103/PhysRevD.85.044061.

[37] Tim Dietrich, Sebastian Khan, Reetika Dudi, Shasvath J. Kapadia, Prayush Kumar, Alessandro Nagar, Frank Ohme, Francesco Pannarale, Anuradha Samajdar, Sebastiano Bernuzzi, Gregorio Carullo, Walter Del Pozzo, Maria Haney, Charalampos Markakis, Michael Pürrer, Gunnar Riemenschneider, Yoshinta Eka Setyawati, Ka Wa Tsang, and Chris Van Den Broeck. Matter imprints in waveform models for neutron star binaries: Tidal and self-spin effects. *prd*, 99(2):024029, January 2019. doi: 10.1103/PhysRevD.99.024029.

[38] Lionel London, Sebastian Khan, Edward Fauchon-Jones, Cecilio García, Mark Hannam, Sascha Husa, Xisco Jiménez-Forteza, Chinmay Kalaghatgi, Frank Ohme, and Francesco Pannarale. First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries. *prl*, 120(16):161102, April 2018. doi: 10.1103/PhysRevLett.120.161102.

[39] Sebastian Khan, Frank Ohme, Katerina Chatziioannou, and Mark Hannam. Including higher order multipoles in gravitational-wave models for precessing binary black holes. *prd*, 101(2):024056, January 2020. doi: 10.1103/PhysRevD.101.024056.

[40] Chase Kimball, Christopher Berry, and Vicky Kalogera. What GW170729's Exceptional Mass and Spin Tells Us about Its Family Tree. *Research Notes of the American Astronomical Society*, 4(1):2, January 2020. doi: 10.3847/2515-5172/ab66be.

[41] Alex Nitz, Ian Harry, Duncan Brown, Christopher M. Biwer, Josh Willis, Tito Dal Canton, Collin Capano, Larne Pekowsky, Thomas Dent, Andrew R. Williamson, Gareth S Davies, Soumi De, Miriam Cabero, Bernd Machenschalk, Prayush Kumar, Steven Reyes, Duncan Macleod, Francesco Pannarale, Dfinstad, Thomas Massinger, Márton Tápai, Leo Singer, Sebastian Khan, Stephen Fairhurst, Sumit Kumar, Alex Nielsen, Shasvath, Iain Dorrington, Amber Lenon, and Hunter Gabbard. gwastro/pycbc: PyCBC Release 1.16.4. Zenodo, June 2020.

[42] Duncan Macleod, Alex L. Urban, Scott Coughlin, Thomas Massinger, Matt Pitkin, Paulaltin, Joseph Areeda, Eric Quintero, The Gitter Badger, Leo Singer, and Katrin Leinweber. gwpy/gwpy: 1.0.1. Zenodo, January 2020.

[43] John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science and Engineering*, 9(3):90–95, May 2007. doi: 10.1109/MCSE.2007.55.

[44] Daniel Foreman-Mackey. corner.py: Scatterplot matrices in Python. *The Journal of Open Source Software*, 1:24, June 2016. doi: 10.21105/joss.00024.

[45] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo during the First Half of the Third Observing Run. *Physical Review X*, 11(2):021053, April 2021. doi: 10.1103/PhysRevX.11.021053.

[46] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and et al. GW190521: A Binary Black Hole Merger with a Total Mass of 150 $M_{\odot}$. *prl*, 125(10): 101102, September 2020. doi: 10.1103/PhysRevLett.125.101102.

[47] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and et al. Properties and Astrophysical Implications of the 150 $M_\odot$ Binary Black Hole Merger GW190521. *apjl*, 900(1):L13, September 2020. doi: 10.3847/2041-8213/aba493.

[48] Maya Fishbach and Daniel E. Holz. Minding the Gap: GW190521 as a Straddling Binary. *apjl*, 904(2):L26, December 2020. doi: 10.3847/2041-8213/abc827.

[49] Alexander H. Nitz and Collin D. Capano. GW190521 May Be an Intermediate-mass Ratio Inspiral. *apjl*, 907(1):L9, January 2021. doi: 10.3847/2041-8213/abccc5.

[50] Bruce Edelman, Zoheyr Doctor, and Ben Farr. Poking Holes: Looking for Gaps in LIGO/Virgo's Black Hole Population. *apjl*, 913(2):L23, June 2021. doi: 10.3847/2041-8213/abfdb3.

[51] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, and et al. The population of merging compact binaries inferred using gravitational waves through GWTC-3. *arXiv e-prints*, art. arXiv:2111.03634, November 2021. doi: 10.48550/arXiv.2111.03634.

[52] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *apjl*, 882(2):L24, September 2019. doi: 10.3847/2041-8213/ab3800.

[53] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. *apjl*, 913(1):L7, May 2021. doi: 10.3847/2041-8213/abe949.

[54] S. E. Woosley, A. Heger, and T. A. Weaver. The evolution and explosion of massive stars. *Reviews of Modern Physics*, 74(4):1015–1071, November 2002. doi: 10.1103/RevModPhys.74.1015.

[55] A. Heger and S. E. Woosley. The Nucleosynthetic Signature of Population III. *apj*, 567(1):532–543, March 2002. doi: 10.1086/338487.

[56] A. Heger, C. L. Fryer, S. E. Woosley, N. Langer, and D. H. Hartmann. How Massive Single Stars End Their Life. *apj*, 591(1):288–300, July 2003. doi: 10.1086/375341.

[57] S. E. Woosley. Pulsational Pair-instability Supernovae. *apj*, 836(2):244, February 2017. doi: 10.3847/1538-4357/836/2/244.

[58] S. E. Woosley. The Evolution of Massive Helium Stars, Including Mass Loss. *apj*, 878(1):49, June 2019. doi: 10.3847/1538-4357/ab1b41.

[59] Mario Spera and Michela Mapelli. Very massive stars, pair-instability supernovae and intermediate-mass black holes with the sevn code. *mnras*, 470(4):4739–4749, October 2017. doi: 10.1093/mnras/stx1576.

[60] K. Belczynski, A. Heger, W. Gladysz, A. J. Ruiter, S. Woosley, G. Wiktorowicz, H. Y. Chen, T. Bulik, R. O'Shaughnessy, D. E. Holz, C. L. Fryer, and E. Berti. The effect of pair-instability mass loss on black-hole mergers. *aap*, 594:A97, October 2016. doi: 10.1051/0004-6361/201628980.

[61] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham. Mind the Gap: The Location of the Lower Edge of the Pair-instability Supernova Black Hole Mass Gap. *apj*, 887(1):53, December 2019. doi: 10.3847/1538-4357/ab518b.

[62] Simon Stevenson, Matthew Sampson, Jade Powell, Alejandro Vigna-Gómez, Coenraad J. Neijssel, Dorottya Szécsi, and Ilya Mandel. The Impact of Pair-instability Mass Loss on the Binary Black Hole Mass Distribution. *apj*, 882(2):121, September 2019. doi: 10.3847/1538-4357/ab3981.

[63] R. Farmer, M. Renzo, S. E. de Mink, M. Fishbach, and S. Justham. Constraints from Gravitational-wave Detections of Binary Black Hole Mergers on the $^{12}C(\alpha, \gamma)^{16}O$ Rate. *apjl*, 902(2):L36, October 2020. doi: 10.3847/2041-8213/abbadd.

[64] Krzysztof Belczynski. The Most Ordinary Formation of the Most Unusual Double Black Hole Merger. *apjl*, 905(2):L15, December 2020. doi: 10.3847/2041-8213/abcbf1.

[65] Jorick S. Vink, Erin R. Higgins, Andreas A. C. Sander, and Gautham N. Sabhahit. Maximum black hole mass across cosmic time. *mnras*, 504(1): 146–154, June 2021. doi: 10.1093/mnras/stab842.

[66] Maya Fishbach and Daniel E. Holz. Where Are LIGO's Big Black Holes? *apjl*, 851(2):L25, December 2017. doi: 10.3847/2041-8213/aa9bf6.

[67] Colm Talbot and Eric Thrane. Measuring the Binary Black Hole Mass Spectrum with an Astrophysically Motivated Parameterization. *apj*, 856(2): 173, April 2018. doi: 10.3847/1538-4357/aab34c.

[68] Isobel Romero-Shaw, Paul D. Lasky, Eric Thrane, and Juan Calderón Bustillo. GW190521: Orbital Eccentricity and Signatures of Dynamical Formation in a Binary Black Hole Merger Signal. *apjl*, 903(1):L5, November 2020. doi: 10.3847/2041-8213/abbe26.

[69] Chase Kimball, Colm Talbot, Christopher P. L. Berry, Matthew Carney, Michael Zevin, Eric Thrane, and Vicky Kalogera. Black Hole Genealogy: Identifying Hierarchical Mergers with Gravitational Waves. *apj*, 900(2):177, September 2020. doi: 10.3847/1538-4357/aba518.

[70] Z. Doctor, D. Wysocki, R. O'Shaughnessy, D. E. Holz, and B. Farr. Black Hole Coagulation: Modeling Hierarchical Mergers in Black Hole Populations. *apj*, 893(1):35, April 2020. doi: 10.3847/1538-4357/ab7fac.

[71] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepanczyk, I. Bartos, M. Campanelli, S. Klimenko, C. Lousto, and R. O'Shaughnessy. Eccentricity Estimate for Black Hole Mergers with Numerical Relativity Simulations. *arXiv e-prints*, art. arXiv:2009.05461, September 2020. doi: 10.48550/arXiv.2009.05461.

[72] Eoin Farrell, Jose H. Groh, Raphael Hirschi, Laura Murphy, Etienne Kaiser, Sylvia Ekström, Cyril Georgy, and Georges Meynet. Is GW190521 the merger of black holes from the first stellar generations? *mnras*, 502(1): L40–L44, March 2021. doi: 10.1093/mnrasl/slaa196.

[73] Amy Secunda, Jillian Bellovary, Mordecai-Mark Mac Low, K. E. Saavik Ford, Barry McKernan, Nathan W. C. Leigh, Wladimir Lyra, Zsolt Sándor, and Jose I. Adorno. Orbital Migration of Interacting Stellar Mass Black Holes in Disks around Supermassive Black Holes. II. Spins and Incoming Objects. *apj*, 903(2):133, November 2020. doi: 10.3847/1538-4357/abbc1d.

[74] B. McKernan, K. E. S. Ford, R. O'Shaugnessy, and D. Wysocki. Monte Carlo simulations of black hole mergers in AGN discs: Low χ_{eff} mergers and predictions for LIGO. *mnras*, 494(1):1203–1216, May 2020. doi: 10.1093/mnras/staa740.

[75] Chase Kimball, Colm Talbot, Christopher P. L. Berry, Michael Zevin, Eric Thrane, Vicky Kalogera, Riccardo Buscicchio, Matthew Carney, Thomas Dent, Hannah Middleton, Ethan Payne, John Veitch, and Daniel Williams. Evidence for Hierarchical Black Hole Mergers in the Second LIGO-Virgo Gravitational Wave Catalog. *apjl*, 915(2):L35, July 2021. doi: 10.3847/2041-8213/ac0aef.

[76] Juan Calderón Bustillo, Nicolas Sanchis-Gual, Alejandro Torres-Forné, and José A. Font. Confusing Head-On Collisions with Precessing Intermediate-Mass Binary Black Hole Mergers. *prl*, 126(20):201101, May 2021. doi: 10.1103/PhysRevLett.126.201101.

[77] Jeremy Sakstein, Djuna Croon, Samuel D. McDermott, Maria C. Straight, and Eric J. Baxter. Beyond the Standard Model Explanations of GW190521. *prl*, 125(26):261105, December 2020. doi: 10.1103/PhysRevLett.125.261105.

[78] Alejandro Cruz-Osorio, Fabio D. Lora-Clavijo, and Carlos Herdeiro. GW190521 formation scenarios via relativistic accretion. *jcap*, 2021(7):032, July 2021. doi: 10.1088/1475-7516/2021/07/032.

[79] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa. Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes. *prd*, 103(10):104056, May 2021. doi: 10.1103/PhysRevD.103.104056.

[80] Jose María Ezquiaga and Daniel E. Holz. Jumping the Gap: Searching for LIGO's Biggest Black Holes. *apjl*, 909(2):L23, March 2021. doi: 10.3847/2041-8213/abe638.

[81] Eric J. Baxter, Djuna Croon, Samuel D. McDermott, and Jeremy Sakstein. Find the Gap: Black Hole Population Analysis with an Astrophysically Motivated Mass Function. *apjl*, 916(2):L16, August 2021. doi: 10.3847/2041-8213/ac11fc.

[82] Will M. Farr. Accuracy Requirements for Empirically Measured Selection Functions. *Research Notes of the American Astronomical Society*, 3(5):66, May 2019. doi: 10.3847/2515-5172/ab1d5f.

[83] Ilya Mandel, Will M. Farr, and Jonathan R. Gair. Extracting distribution parameters from multiple uncertain observations with selection biases. *mnras*, 486(1):1086–1093, June 2019. doi: 10.1093/mnras/stz896.

[84] Gregory Ashton, Moritz Hübner, Paul D. Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, Jan Harms, Marcus E. Lower, Grant D. Meadors, Denyz Melchor, Ethan Payne, Matthew D. Pitkin, Jade Powell, Nikhil Sarin, Rory J. E. Smith, and Eric Thrane. BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. *apjs*, 241(2): 27, April 2019. doi: 10.3847/1538-4365/ab06fc.

[85] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, C. P. L. Berry, S. Coughlin, S. Galaudage, C. Hoy, M. Hübner, K. S. Phukon, M. Pitkin, M. Rizzo, N. Sarin, R. Smith, S. Stevenson, A. Vajpeyi, M. Arène, K. Athar, S. Banagiri, N. Bose, M. Carney, K. Chatziioannou, J. A. Clark, M. Colleoni, R. Cotesta, B. Edelman, H. Estellés, C. García-Quirós, Abhirup Ghosh, R. Green, C. J. Haster, S. Husa, D. Keitel, A. X. Kim, F. Hernandez-Vivanco, I. Magaña Hernandez, C. Karathanasis, P. D. Lasky, N. De Lillo, M. E. Lower, D. Macleod, M. Mateu-Lucena, A. Miller, M. Millhouse, S. Morisaki, S. H. Oh, S. Ossokine, E. Payne, J. Powell, G. Pratten, M. Pürrer, A. Ramos-Buades, V. Raymond, E. Thrane, J. Veitch, D. Williams, M. J. Williams, and L. Xiao. Bayesian inference for compact binary coalescences with BILBY: validation and application to the first LIGO-Virgo gravitational-wave transient catalogue. *mnras*, 499(3): 3295–3319, December 2020. doi: 10.1093/mnras/staa2850.

[86] Colm Talbot, Rory Smith, Eric Thrane, and Gregory B. Poole. Parallelized inference for gravitational-wave astronomy. *prd*, 100(4):043030, August 2019. doi: 10.1103/PhysRevD.100.043030.

[87] Joshua S. Speagle. DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *mnras*, 493(3):3132–3158, April 2020. doi: 10.1093/mnras/staa278.

[88] Maya Fishbach, Daniel E. Holz, and Will M. Farr. Does the Black Hole Merger Rate Evolve with Redshift? *apjl*, 863(2):L41, August 2018. doi: 10.3847/2041-8213/aad800.

[89] M. Renzo, R. Farmer, S. Justham, Y. Götberg, S. E. de Mink, E. Zapartas, P. Marchant, and N. Smith. Predictions for the hydrogen-free ejecta of pulsational pair-instability supernovae. *aap*, 640:A56, August 2020. doi: 10.1051/0004-6361/202037710.

[90] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass ~ 3.4 $M_\odot$. *apjl*, 892(1):L3, March 2020. doi: 10.3847/2041-8213/ab75f5.

[91] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and et al. GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. *apjl*, 896(2):L44, June 2020. doi: 10.3847/2041-8213/ab960f.

[92] Davide Gerosa and Emanuele Berti. Are merging black holes born from stellar collapse or previous mergers? *prd*, 95(12):124046, June 2017. doi: 10.1103/PhysRevD.95.124046.

[93] Maya Fishbach, Daniel E. Holz, and Ben Farr. Are LIGO's Black Holes Made from Smaller Black Holes? *apjl*, 840(2):L24, May 2017. doi: 10.3847/2041-8213/aa7045.

[94] Zoheyr Doctor, Ben Farr, and Daniel E. Holz. Black Hole Leftovers: The Remnant Population from Binary Black Hole Mergers. *apjl*, 914(1):L18, June 2021. doi: 10.3847/2041-8213/ac0334.

[95] Will M. Farr, Maya Fishbach, Jiani Ye, and Daniel E. Holz. A Future Percent-level Measurement of the Hubble Expansion at Redshift 0.8 with Advanced LIGO. *apjl*, 883(2):L42, October 2019. doi: 10.3847/2041-8213/ab4284.

[96] Vaibhav Tiwari and Stephen Fairhurst. The Emergence of Structure in the Binary Black Hole Mass Distribution. *apjl*, 913(2):L19, June 2021. doi: 10.3847/2041-8213/abfbe7.

[97] Bruce Edelman, Zoheyr Doctor, Jaxen Godfrey, and Ben Farr. Ain't No Mountain High Enough: Semiparametric Modeling of LIGO-Virgo's Binary Black Hole Mass Distribution. *apj*, 924(2):101, January 2022. doi: 10.3847/1538-4357/ac3667.

[98] L. A. C. van Son, S. E. de Mink, M. Renzo, S. Justham, E. Zapartas, K. Breivik, T. Callister, W. M. Farr, and C. Conroy. No Peaks without Valleys: The Stable Mass Transfer Channel for Gravitational-wave Sources in Light of the Neutron Star-Black Hole Mass Gap. *apj*, 940(2):184, December 2022. doi: 10.3847/1538-4357/ac9b0a.

[99] Amanda M. Farah, Bruce Edelman, Michael Zevin, Maya Fishbach, Jose María Ezquiaga, Ben Farr, and Daniel E. Holz. Things that might go bump in the night: Assessing structure in the binary black hole mass spectrum. *arXiv e-prints*, art. arXiv:2301.00834, January 2023. doi: 10.48550/arXiv.2301.00834.

[100] Yin-Jie Li, Yuan-Zhu Wang, Shao-Peng Tang, and Yi-Zhong Fan. Resolving the stellar-collapse and hierarchical-merger origins of the coalescing black holes. *arXiv e-prints*, art. arXiv:2303.02973, March 2023. doi: 10.48550/arXiv.2303.02973.

[101] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. GW150914: The Advanced LIGO Detectors in the Era of First Discoveries. *prl*, 116(13):131103, April 2016. doi: 10.1103/PhysRevLett.116.131103.

[102] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and et al. GW190412: Observation of a binary-black-hole coalescence with asymmetric masses. *prd*, 102(4):043015, August 2020. doi: 10.1103/PhysRevD.102.043015.

[103] Pablo Marchant, Mathieu Renzo, Robert Farmer, Kaliroe M. W. Pappas, Ronald E. Taam, Selma E. de Mink, and Vassiliki Kalogera. Pulsational Pair-instability Supernovae in Very Close Binaries. *apj*, 882(1):36, September 2019. doi: 10.3847/1538-4357/ab3426.

[104] Héctor Estellés, Sascha Husa, Marta Colleoni, Maite Mateu-Lucena, Maria de Lluc Planas, Cecilio García-Quirós, David Keitel, Antoni Ramos-Buades, Ajit Kumar Mehta, Alessandra Buonanno, and Serguei Ossokine. A Detailed Analysis of GW190521 with Phenomenological Waveform Models. *apj*, 924 (2):79, January 2022. doi: 10.3847/1538-4357/ac33a0.

[105] Carl L. Rodriguez, Michael Zevin, Pau Amaro-Seoane, Sourav Chatterjee, Kyle Kremer, Frederic A. Rasio, and Claire S. Ye. Black holes: The next generation—repeated mergers in dense star clusters and their gravitational-wave properties. *prd*, 100(4):043027, August 2019. doi: 10.1103/PhysRevD.100.043027.

[106] Djuna Croon, Samuel D. McDermott, and Jeremy Sakstein. New physics and the black hole mass gap. *prd*, 102(11):115024, December 2020. doi: 10.1103/PhysRevD.102.115024.

[107] Michela Mapelli, Filippo Santoliquido, Yann Bouffanais, Manuel Arca Arca Sedda, Maria Celeste Artale, and Alessandro Ballone. Mass and Rate of Hierarchical Black Hole Mergers in Young, Globular and Nuclear Star Clusters. *Symmetry*, 13(9):1678, September 2021. doi: 10.3390/sym13091678.

[108] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, C. Affeldt, M. Agathos, and et al. Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Reviews in Relativity*, 23(1):3, September 2020. doi: 10.1007/s41114-020-00026-9.

[109] Michael Zevin, Chris Pankow, Carl L. Rodriguez, Laura Sampson, Eve Chase, Vassiliki Kalogera, and Frederic A. Rasio. Constraining Formation Models of Binary Black Holes with Gravitational-wave Observations. *apj*, 846(1):82, September 2017. doi: 10.3847/1538-4357/aa8408.

[110] Ken K. Y. Ng, Salvatore Vitale, Will M. Farr, and Carl L. Rodriguez. Probing Multiple Populations of Compact Binaries with Third-generation Gravitational-wave Detectors. *apjl*, 913(1):L5, May 2021. doi: 10.3847/2041-8213/abf8be.

[111] Salvatore Vitale, Carl-Johan Haster, Ling Sun, Ben Farr, Evan Goetz, Jeff Kissel, and Craig Cahillane. Physical approach to the marginalization of LIGO calibration uncertainties. *prd*, 103(6):063016, March 2021. doi: 10.1103/PhysRevD.103.063016.

[112] Zoheyr Doctor, Ben Farr, Daniel E. Holz, and Michael Pürrer. Statistical gravitational waveform models: What to simulate next? *prd*, 96(12):123011, December 2017. doi: 10.1103/PhysRevD.96.123011.

[113] Ilya Mandel, Will M. Farr, Andrea Colonna, Simon Stevenson, Peter Tiňo, and John Veitch. Model-independent inference on compact-binary observations. *mnras*, 465(3):3254–3260, March 2017. doi: 10.1093/mnras/stw2883.

[114] Vaibhav Tiwari. VAMANA: modeling binary black hole population with minimal assumptions. *Classical and Quantum Gravity*, 38(15):155007, August 2021. doi: 10.1088/1361-6382/ac0b54.

[115] Pavel Kroupa. On the variation of the initial mass function. *mnras*, 322(2): 231–246, April 2001. doi: 10.1046/j.1365-8711.2001.04022.x.

[116] Salvatore Vitale, Davide Gerosa, Will M. Farr, and Stephen R. Taylor. Inferring the Properties of a Population of Compact Binaries in Presence of Selection Effects. In *Handbook of Gravitational Wave Astronomy*, page 45. 2022. doi: 10.1007/978-981-15-4702-7_45-1.

[117] David W. Hogg. Distance measures in cosmology. *arXiv e-prints*, art. astro-ph/9905116, May 1999. doi: 10.48550/arXiv.astro-ph/9905116.

[118] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, and et al. Observation of Gravitational Waves from Two Neutron Star-Black Hole Coalescences. *apjl*, 915(1):L5, July 2021. doi: 10.3847/2041-8213/ac082e.

[119] Maya Fishbach, Reed Essick, and Daniel E. Holz. Does Matter Matter? Using the Mass Distribution to Distinguish Neutron Stars and Black Holes. *apjl*, 899(1):L8, August 2020. doi: 10.3847/2041-8213/aba7b6.

[120] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *aap*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

[121] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002. doi: https://doi.org/10.1111/1467-9868.00353. URL `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00353`.

[122] Bruce Edelman, Ben Farr, and Zoheyr Doctor. Cover Your Basis: Comprehensive Data-driven Characterization of the Binary Black Hole Population. *apj*, 946(1):16, March 2023. doi: 10.3847/1538-4357/acb5ed.

[123] Jaxen Godfrey, Bruce Edelman, and Ben Farr. Cosmic Cousins: Identification of a Subpopulation of Binary Black Holes Consistent with Isolated Binary Evolution. *arXiv e-prints*, art. arXiv:2304.01288, April 2023. doi: 10.48550/arXiv.2304.01288.

[124] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Bae, and et al. Overview of KAGRA: Calibration, detector characterization, physical environmental monitors, and the geophysics interferometer. *Progress of Theoretical and Experimental Physics*, 2021(5):05A102, May 2021. doi: 10.1093/ptep/ptab018.

[125] Salvatore Vitale, Ryan Lynch, Riccardo Sturani, and Philip Graff. Use of gravitational waves to probe the formation channels of compact binaries. *Classical and Quantum Gravity*, 34(3):03LT01, February 2017. doi: 10.1088/1361-6382/aa552e.

[126] Carl L. Rodriguez, Michael Zevin, Chris Pankow, Vasilliki Kalogera, and Frederic A. Rasio. Illuminating Black Hole Binary Formation Channels with Spins in Advanced LIGO. *apjl*, 832(1):L2, November 2016. doi: 10.3847/2041-8205/832/1/L2.

[127] Will M. Farr, Simon Stevenson, M. Coleman Miller, Ilya Mandel, Ben Farr, and Alberto Vecchio. Distinguishing spin-aligned and isotropic black hole populations with gravitational waves. *nat*, 548(7667):426–429, August 2017. doi: 10.1038/nature23453.

[128] Ben Farr, Daniel E. Holz, and Will M. Farr. Using Spin to Understand the Formation of LIGO and Virgo's Black Holes. *apjl*, 854(1):L9, February 2018. doi: 10.3847/2041-8213/aaaa64.

[129] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, R. Abbott, H. Abe, F. Acernese, K. Ackley, N. Adhikari, R. X. Adhikari, V. K. Adkins, and et al. Constraints on the cosmic expansion history from GWTC-3. *arXiv e-prints*, art. arXiv:2111.03604, November 2021. doi: 10.48550/arXiv.2111.03604.

[130] Jose María Ezquiaga and Daniel E. Holz. Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries. *prl*, 129(6):061102, August 2022. doi: 10.1103/PhysRevLett.129.061102.

[131] Maria Okounkova, Will M. Farr, Maximiliano Isi, and Leo C. Stein. Constraining gravitational wave amplitude birefringence and Chern-Simons gravity with GWTC-2. *prd*, 106(4):044067, August 2022. doi: 10.1103/PhysRevD.106.044067.

[132] Andreas Finke, Stefano Foffa, Francesco Iacovelli, Michele Maggiore, and Michele Mancarella. Modified gravitational wave propagation and the binary neutron star mass function. *Physics of the Dark Universe*, 36:100994, June 2022. doi: 10.1016/j.dark.2022.100994.

[133] Michele Mancarella, Edwin Genoud-Prachex, and Michele Maggiore. Cosmology and modified gravitational wave propagation from binary black hole population models. *prd*, 105(6):064030, March 2022. doi: 10.1103/PhysRevD.105.064030.

[134] Macarena Lagos, Maya Fishbach, Philippe Landry, and Daniel E. Holz. Standard sirens with a running Planck mass. *prd*, 99(8):083504, April 2019. doi: 10.1103/PhysRevD.99.083504.

[135] Ken K. Y. Ng, Otto A. Hannuksela, Salvatore Vitale, and Tjonnie G. F. Li. Searching for ultralight bosons within spin measurements of a population of binary black hole mergers. *prd*, 103(6):063010, March 2021. doi: 10.1103/PhysRevD.103.063010.

[136] Ken K. Y. Ng, Salvatore Vitale, Otto A. Hannuksela, and Tjonnie G. F. Li. Constraints on Ultralight Scalar Bosons within Black Hole Spin Measurements from the LIGO-Virgo GWTC-2. *prl*, 126(15):151102, April 2021. doi: 10.1103/PhysRevLett.126.151102.

[137] Ken K. Y. Ng, Gabriele Franciolini, Emanuele Berti, Paolo Pani, Antonio Riotto, and Salvatore Vitale. Constraining High-redshift Stellar-mass Primordial Black Holes with Next-generation Ground-based Gravitational-wave Detectors. *apjl*, 933(2):L41, July 2022. doi: 10.3847/2041-8213/ac7aae.

[138] Jacob Golomb and Colm Talbot. Hierarchical Inference of Binary Neutron Star Mass Distribution and Equation of State with Gravitational Waves. *apj*, 926(1):79, February 2022. doi: 10.3847/1538-4357/ac43bc.

[139] Philippe Landry and Jocelyn S. Read. The Mass Distribution of Neutron Stars in Gravitational-wave Binaries. *apjl*, 921(2):L25, November 2021. doi: 10.3847/2041-8213/ac2f3e.

[140] Vaibhav Tiwari. Exploring Features in the Binary Black Hole Population. *apj*, 928(2):155, April 2022. doi: 10.3847/1538-4357/ac589a.

[141] Davide Gerosa, Emanuele Berti, Richard O'Shaughnessy, Krzysztof Belczynski, Michael Kesden, Daniel Wysocki, and Wojciech Gladysz. Spin orientations of merging black holes formed from the evolution of stellar binaries. *prd*, 98(8):084036, October 2018. doi: 10.1103/PhysRevD.98.084036.

[142] Michael Zevin and Simone S. Bavera. Suspicious Siblings: The Distribution of Mass and Spin across Component Black Holes in Isolated Binary Evolution. *apj*, 933(1):86, July 2022. doi: 10.3847/1538-4357/ac6f5d.

[143] Simone S. Bavera, Tassos Fragos, Ying Qin, Emmanouil Zapartas, Coenraad J. Neijssel, Ilya Mandel, Aldo Batta, Sebastian M. Gaebel, Chase Kimball, and Simon Stevenson. The origin of spin in binary black holes. Predicting the distributions of the main observables of Advanced LIGO. *aap*, 635:A97, March 2020. doi: 10.1051/0004-6361/201936204.

[144] Simone S. Bavera, Tassos Fragos, Michael Zevin, Christopher P. L. Berry, Pablo Marchant, Jeff J. Andrews, Scott Coughlin, Aaron Dotter, Konstantinos Kovlakas, Devina Misra, Juan G. Serra-Perez, Ying Qin, Kyle A. Rocha, Jaime Román-Garza, Nam H. Tran, and Emmanouil Zapartas. The impact of mass-transfer physics on the observable properties of field binary black hole populations. *aap*, 647:A153, March 2021. doi: 10.1051/0004-6361/202039804.

[145] Javier Roulet, Horng Sheng Chia, Seth Olsen, Liang Dai, Tejaswi Venumadhav, Barak Zackay, and Matias Zaldarriaga. Distribution of effective spins and masses of binary black holes from the LIGO and Virgo O1-O3a observing runs. *prd*, 104(8):083010, October 2021. doi: 10.1103/PhysRevD.104.083010.

[146] Shanika Galaudage, Colm Talbot, Tushar Nagar, Deepnika Jain, Eric Thrane, and Ilya Mandel. Building Better Spin Models for Merging Binary Black Holes: Evidence for Nonspinning and Rapidly Spinning Nearly Aligned Subpopulations. *apjl*, 921(1):L15, November 2021. doi: 10.3847/2041-8213/ac2f3c.

[147] Hui Tong, Shanika Galaudage, and Eric Thrane. The population properties of spinning black holes using Gravitational-wave Transient Catalog 3. *arXiv e-prints*, art. arXiv:2209.02206, September 2022. doi: 10.48550/arXiv.2209.02206.

[148] Thomas A. Callister, Simona J. Miller, Katerina Chatziioannou, and Will M. Farr. No Evidence that the Majority of Black Holes in Binaries Have Zero Spin. *apjl*, 937(1):L13, September 2022. doi: 10.3847/2041-8213/ac847e.

[149] Salvatore Vitale, Sylvia Biscoveanu, and Colm Talbot. Spin it as you like: The (lack of a) measurement of the spin tilt distribution with LIGO-Virgo-KAGRA binary black holes. *aap*, 668:L2, December 2022. doi: 10.1051/0004-6361/202245084.

[150] Jacob Golomb and Colm Talbot. Searching for structure in the binary black hole spin distribution. *arXiv e-prints*, art. arXiv:2210.12287, October 2022. doi: 10.48550/arXiv.2210.12287.

[151] Matthew C. Edwards, Renate Meyer, and Nelson Christensen. Bayesian nonparametric spectral density estimation using b-spline priors. *Statistics and Computing*, 29(1):67–78, jan 2018. doi: 10.1007/s11222-017-9796-9. URL `https://doi.org/10.1007%2Fs11222-017-9796-9`.

[152] S. Lang and Andreas Brezger. Bayesian p-splines, 2001. URL `http://nbn-resolving.de/urn/resolver.pl?urn=nbn:de:bvb:19-epub-1617-2`.

[153] Astrid Jullion and Philippe Lambert. Robust specification of the roughness penalty prior distribution in spatially adaptive bayesian p-splines models. *Comput. Stat. Data Anal.*, 51:2542–2558, 2007.

[154] LVK Collaboration. Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run — parameter estimation data release, November 2021. URL `https://doi.org/10.5281/zenodo.5546663`.

[155] Reed Essick, Amanda Farah, Shanika Galaudage, Colm Talbot, Maya Fishbach, Eric Thrane, and Daniel E. Holz. Probing Extremal Gravitational-wave Events with Coarse-grained Likelihoods. *apj*, 926(1):34, February 2022. doi: 10.3847/1538-4357/ac3978.

[156] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. *prl*, 113 (15):151101, October 2014. doi: 10.1103/PhysRevLett.113.151101.

[157] Yi Pan, Alessandra Buonanno, Andrea Taracchini, Lawrence E. Kidder, Abdul H. Mroué, Harald P. Pfeiffer, Mark A. Scheel, and Béla Szilágyi. Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism. *prd*, 89(8):084006, April 2014. doi: 10.1103/PhysRevD.89.084006.

[158] Andrea Taracchini, Alessandra Buonanno, Yi Pan, Tanja Hinderer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Geoffrey Lovelace, Abdul H. Mroué, Harald P. Pfeiffer, Mark A. Scheel, Béla Szilágyi, Nicholas W. Taylor, and Anil Zenginoglu. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *prd*, 89(6):061502, March 2014. doi: 10.1103/PhysRevD.89.061502.

[159] The LIGO Scientific Collaboration, the Virgo Collaboration, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, and et al. GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *arXiv e-prints*, art. arXiv:2108.01045, August 2021. doi: 10.48550/arXiv.2108.01045.

[160] LVK Collaboration. The population of merging compact binaries inferred using gravitational waves through gwtc-3 - data release, November 2021. URL `https://doi.org/10.5281/zenodo.5655785`.

[161] Colm Talbot and Eric Thrane. Determining the population properties of spinning black holes. *prd*, 96(2):023012, July 2017. doi: 10.1103/PhysRevD.96.023012.

[162] Daniel Wysocki, Jacob Lange, and Richard O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. *prd*, 100(4):043012, August 2019. doi: 10.1103/PhysRevD.100.043012.

[163] Jim Fuller and Linhao Ma. Most Black Holes Are Born Very Slowly Rotating. *apjl*, 881(1):L1, August 2019. doi: 10.3847/2041-8213/ab339b.

[164] Matthew Mould, Davide Gerosa, Floor S. Broekgaarden, and Nathan Steinle. Which black hole formed first? Mass-ratio reversal in massive binary stars from gravitational-wave data. *mnras*, 517(2):2738–2745, December 2022. doi: 10.1093/mnras/stac2859.

[165] Salvatore Vitale, Ryan Lynch, John Veitch, Vivien Raymond, and Riccardo Sturani. Measuring the Spin of Black Holes in Binary Systems Using Gravitational Waves. *prl*, 112(25):251101, June 2014. doi: 10.1103/PhysRevLett.112.251101.

[166] Salvatore Vitale, Ryan Lynch, Vivien Raymond, Riccardo Sturani, John Veitch, and Philip Graff. Parameter estimation for heavy binary-black holes with networks of second-generation gravitational-wave detectors. *prd*, 95(6): 064053, March 2017. doi: 10.1103/PhysRevD.95.064053.

[167] Simona Miller, Thomas A. Callister, and Will M. Farr. The Low Effective Spin of Binary Black Holes and Implications for Individual Gravitational-wave Events. *apj*, 895(2):128, June 2020. doi: 10.3847/1538-4357/ab80c0.

[168] Piero Madau and Mark Dickinson. Cosmic Star-Formation History. *araa*, 52: 415–486, August 2014. doi: 10.1146/annurev-astro-081811-125615.

[169] Amanda Farah, Maya Fishbach, Reed Essick, Daniel E. Holz, and Shanika Galaudage. Bridging the Gap: Categorizing Gravitational-wave Events at the Transition between Neutron Stars and Black Holes. *apj*, 931(2):108, June 2022. doi: 10.3847/1538-4357/ac5f03.

[170] Fabio Antonini and Mark Gieles. Merger rate of black hole binaries from globular clusters: Theoretical error bars and comparison to gravitational wave data from GWTC-2. *prd*, 102(12):123016, December 2020. doi: 10.1103/PhysRevD.102.123016.

[171] Jongsuk Hong, Enrico Vesperini, Abbas Askar, Mirek Giersz, Magdalena Szkudlarek, and Tomasz Bulik. Binary black hole mergers from globular clusters: the impact of globular cluster properties. *mnras*, 480(4):5645–5656, November 2018. doi: 10.1093/mnras/sty2211.

[172] Carl L. Rodriguez, Meagan Morscher, Bharath Pattabiraman, Sourav Chatterjee, Carl-Johan Haster, and Frederic A. Rasio. Binary Black Hole Mergers from Globular Clusters: Implications for Advanced LIGO. *prl*, 115 (5):051101, July 2015. doi: 10.1103/PhysRevLett.115.051101.

[173] K. E. Saavik Ford and Barry McKernan. Binary black hole merger rates in AGN discs versus nuclear star clusters: loud beats quiet. *mnras*, 517(4): 5827–5834, December 2022. doi: 10.1093/mnras/stac2861.

[174] Hiromichi Tagawa, Bence Kocsis, Zoltán Haiman, Imre Bartos, Kazuyuki Omukai, and Johan Samsing. Mass-gap Mergers in Active Galactic Nuclei. *apj*, 908(2):194, February 2021. doi: 10.3847/1538-4357/abd555.

[175] Y. Yang, I. Bartos, Z. Haiman, B. Kocsis, Z. Márka, N. C. Stone, and S. Márka. AGN Disks Harden the Mass Distribution of Stellar-mass Binary Black Hole Mergers. *apj*, 876(2):122, May 2019. doi: 10.3847/1538-4357/ab16e3.

[176] M. Renzo, R. J. Farmer, S. Justham, S. E. de Mink, Y. Götberg, and P. Marchant. Sensitivity of the lower edge of the pair-instability black hole mass gap to the treatment of time-dependent convection. *mnras*, 493(3): 4333–4341, April 2020. doi: 10.1093/mnras/staa549.

[177] Maya Fishbach, Chase Kimball, and Vicky Kalogera. Limits on Hierarchical Black Hole Mergers from the Most Negative χ_{eff} Systems. *apjl*, 935(2):L26, August 2022. doi: 10.3847/2041-8213/ac86c4.

[178] L. A. C. van Son, S. E. de Mink, T. Callister, S. Justham, M. Renzo, T. Wagg, F. S. Broekgaarden, F. Kummer, R. Pakmor, and I. Mandel. The Redshift Evolution of the Binary Black Hole Merger Rate: A Weighty Matter. *apj*, 931(1):17, May 2022. doi: 10.3847/1538-4357/ac64a3.

[179] Ethan Payne and Eric Thrane. Model exploration in gravitational-wave astronomy with the maximum population likelihood. *Physical Review Research*, 5(2):023013, April 2023. doi: 10.1103/PhysRevResearch.5.023013.

[180] Eli Bingham, Jonathan P. Chen, Martin Jankowiak, Fritz Obermeyer, Neeraj Pradhan, Theofanis Karaletsos, Rohit Singh, Paul A. Szerlip, Paul Horsfall, and Noah D. Goodman. Pyro: Deep universal probabilistic programming. *J. Mach. Learn. Res.*, 20:28:1–28:6, 2019. URL `http://jmlr.org/papers/v20/18-403.html`.

[181] Du Phan, Neeraj Pradhan, and Martin Jankowiak. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro. *arXiv e-prints*, art. arXiv:1912.11554, December 2019. doi: 10.48550/arXiv.1912.11554.

[182] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, and et. al., 2018. URL `http://github.com/google/jax`.

[183] J. O. Ramsay. Monotone Regression Splines in Action. *Statistical Science*, 3 (4):425 – 441, 1988. doi: 10.1214/ss/1177012761. URL `https://doi.org/10.1214/ss/1177012761`.

[184] LVK Collaboration. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run — O1+O2+O3 Search Sensitivity Estimates, November 2021. URL `https://doi.org/10.5281/zenodo.5636816`. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. The construction and operation of KAGRA are funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Society for the Promotion of Science (JSPS), National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS) and the Ministry of Science and Technology (MoST) in Taiwan.

[185] Rodrigo Luger, Megan Bedell, Daniel Foreman-Mackey, Ian J. M. Crossfield, Lily L. Zhao, and David W. Hogg. Mapping stellar surfaces III: An Efficient, Scalable, and Open-Source Doppler Imaging Model. *arXiv e-prints*, art. arXiv:2110.06271, October 2021. doi: 10.48550/arXiv.2110.06271.

[186] Bruce Edelman, Ben Farr, and Zoheyr Doctor. Cover Your Basis: Comprehensive Data-Driven Characterization of the Binary Black Hole Population, December 2022. URL `https://doi.org/10.5281/zenodo.7566301`.