May 2023

Cosmology, Lensing, and Modified Gravity with Gravitational Waves

Ignacio Magana Hernandez
University of Wisconsin-Milwaukee

COSMOLOGY, LENSING, AND MODIFIED GRAVITY WITH GRAVITATIONAL WAVES

by

Ignacio Magaña Hernandez

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin-Milwaukee

May 2023

ABSTRACT

COSMOLOGY, LENSING, AND MODIFIED GRAVITY WITH GRAVITATIONAL WAVES

by

Ignacio Magaña Hernandez

The University of Wisconsin-Milwaukee, 2023
Under the Supervision of Professor Patrick Brady, PhD

Since the first detection of gravitational waves (GWs) from the merger of two stellar-mass black holes in 2015, the LIGO-Virgo-KAGRA (LVK) Collaboration has accumulated over 90 observations of mergers involving neutron stars and black holes. With the upcoming observing runs for the LVK network of GW detectors, many more binary mergers are expected to be detected. The increasing size of gravitational wave catalogs has enabled the study of their population, its cosmic expansion history, signatures of gravitational wave lensing, and how well these observations agree with general relativity. In this dissertation, I will discuss my contributions to gravitational wave cosmology. I will focus on the development of the dark siren methodology and how it was used with the GWTC-1 catalog to provide a first joint measurement of the Hubble constant from GW standard sirens. I will also discuss a joint parameter estimation framework developed to identify and characterize pairs of strongly lensed GWs from binary black hole mergers and how it was was applied to GWTC-1 observations. Finally, I will explore two topics related to testing modifications to general relativity. First, I will describe how we used the strong lensing joint parameter estimation framework to measure the presence of alternative polarization modes in GW signals from BBH mergers, including a fully mixed tensor, vector, and scalar mode model. Second, I will discuss my work on modified GW propagation in the context of gravitational leakage models, which predict the existence of large extra spacetime dimensions. Using the latest GWTC-3 catalog, we provided the first constraints on such extra-dimension models using BBH mergers.

To my parents

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

EOS	equation of state
GRB	gamma-ray burst
GR	general relativity
GW	gravitational wave
LIGO	Laser Interferometer Gravitational-Wave Observatory
LISA	Laser Interferometer Space Antenna

ACKNOWLEDGMENTS

I would like to thank my Ph.D. advisor Patrick Brady for his guidance and extensive support over the past six years. Your mentorship made me an independent researcher as well as enabled me to focus on the bigger picture—I find that invaluable. I also want to express immense gratitude for providing more than plenty of financial support to conferences and research-related visits—these were critical in my growth as a researcher. I would also like to thank Jolien Creighton, (technically my co-supervisor at this point) for always being present and willing to listen and discuss some of my random ideas which eventually led to productive and stimulating science.

Chapter 1

Introduction

In 1915, Albert Einstein revolutionized our understanding of gravity with his development of the General Theory of Relativity (GR). This theory provided a comprehensive and modern explanation of gravity, including the first prediction of gravitational waves. According to GR, gravity is the result of the curvature of four-dimensional spacetime, rather than an attractive force between masses. The curvature of spacetime is a direct consequence of mass and energy warping the fabric of the universe, and the resulting effect is what we experience as gravity. This curvature can be described by the Einstein field equations, which explain how matter and energy interact with spacetime to create the force of gravity. As John Archibald Wheeler famously stated, "Space tells matter how to move, matter tells space how to curve."

Gravitational waves are ripples in the fabric of spacetime, caused by the acceleration of massive objects. These waves were predicted by Einstein, a year after his discovery of GR. Their discovery, however, took nearly a century to be directly observed and confirmed by the Laser Interferometer Gravitational-Wave Observatory (LIGO) in 2015. Since then, gravitational waves have provided new insights into the nature of the universe, including the detection of compact binary mergers from objects involving black holes and neutron stars.

In the sections that follow we provide a brief description of the generation, propagation, and detection of GWs from accelerated masses, and in particular, we pay special attention to compact binary coalescences. For an in-depth description of GWs, we refer the reader to Creighton & Anderson (2011).

1.1 GENERAL RELATIVITY

1.1.1 Basics

In General Relativity, we can write the gravitational action (S) in terms of the Einstein-Hilbert action (S_E) and the matter action (S_M) so that $S = S_E + S_M$.

In general, the Einstein-Hilbert action can be written in the following functional form:

$$S_E = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} R \tag{1.1}$$

where R is the usual Ricci scalar and g is the determinant of the spacetime metric. The matter action variation under a change of the metric $g_{\mu\nu} = g_{\mu\nu} \to \delta g_{\mu\nu}$ is,

$$\delta S_M = \frac{1}{2c} \int d^4x \sqrt{-g} T^{\mu\nu} \delta g_{\mu\nu} \tag{1.2}$$

where $T^{\mu\nu}$ is the stress-energy tensor describing the matter distribution of the spacetime. Now, taking the variation with respect to $g_{\mu\nu}$ of the total action S yields the Einstein field equations,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.3}$$

One of the most important properties of GR is that it is invariant under a large group of symmetries. We can use this gauge freedom and perform the following transformations on the metric,

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x) \tag{1.4}$$

1.1.2 Linearized Gravity

The first step towards an understanding of metric perturbations ($h_{\mu\nu}$) involves the expansion of the Einstein field equations around a flat spacetime background ($\eta_{\mu\nu}$). This is the so-called linearization of the Einstein field equations and so we consider the following,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| << 1. \tag{1.5}$$

Next, consider the gauge transformation described in Equation 1.4. In particular, we consider the following transformation: $x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)$. We can then write to first order:

$$h_{\mu\nu}(x) \to h'_{\mu\nu}(x') = h_{\mu\nu}(x) - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) \tag{1.6}$$

where we require that $|\partial_\mu \xi_\nu| << 1$ (as required by the condition $|h_{\mu\nu}| << 1$).

Our goal now is to linearize the LHS of the Einstein field equation. The Riemann tensor (and consequently the Ricci tensor and scalar) can also be expanded to first order on $h_{\mu\nu}$. Leaving the details to the reader, one obtains while using $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$ with $h = \eta^{\mu\nu}h_{\mu\nu}$ (to simplify the equations):

$$\Box \bar{h}_{\mu\nu} + \eta^{\mu\nu}\partial^\rho\partial^\sigma \bar{h}_{\rho\sigma} - \partial^\rho\partial_\nu \bar{h}_{\mu\rho} - \partial^\rho\partial_\mu \bar{h}_{\nu\rho} = -\frac{16\pi G}{c^4}T_{\mu\nu} \tag{1.7}$$

Finally, we can impose the gauge freedom allowed by Equation 1.6 that is, we impose the *Lorentz gauge* (or harmonic gauge):

$$\partial \bar{h}_{\mu\nu} = 0, \tag{1.8}$$

thus allowing us to write the linearized Einstein field equations as:

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu} \tag{1.9}$$

1.2 GRAVITATIONAL WAVES

1.2.1 Generation

The Einstein field equations, which describe the relationship between the curvature of spacetime and the distribution of matter and energy, can be simplified when small perturbations due to accelerating matter produce gravitational waves. Specifically, when considering a source of GWs in a vacuum and assuming the short-wavelength approximation, we can use the linearized field equations (defined in Equation 1.9) to understand GW generation. The following is a solution to the equations:

$$h_{\mu\nu}(t, \mathbf{x}) = \frac{2\pi G}{c^4} \int d^3\mathbf{y} \frac{T_{\mu\nu}(t_r, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \tag{1.10}$$

where we have defined the *retarded time* as $t_r = t - |\mathbf{x} - \mathbf{y}|$ This equation describes how the metric perturbation $h_{\mu\nu}$ at time t and position $\mathbf{x}$ depends on the stress-energy tensor at the retarded time t_r and position $\mathbf{y}$.

For an isolated, far-field non-relativistic source of GW radiation, we can use a multiple expansion to understand GW generation. The zeroth order term, which is proportional to the mass of the source, is time-independent due to the conservation of energy for an isolated source. The next-order term, the mass dipole term, is also time-independent and does not contribute to GW radiation due to the conservation of momentum. The leading contribution to the radiation is therefore the quadrupolar term, which depends on the mass quadrupole moment $M_{ij}(t)$:

$$h_{ij}(t, \mathbf{x}) = \frac{2G}{rc^4} \frac{d^2 M_{ij}(t_r)}{dt^2} \tag{1.11}$$

This result is the well know *quadrupole formula* describing the leading order contribution for GW generation.

1.2.2 Propagation

Gravitational waves once generated can be described by a propagation equation in a vacuum,

$$\Box \bar{h}_{\mu\nu} = 0 \tag{1.12}$$

since the stress-energy tensor $T_{\mu\nu} = 0$. Assuming the plane wave approximation, we can write a general solution for the propagation of gravitational waves,

$$\bar{h}_{\mu\nu}(t, \mathbf{x}) = \text{Re}\left[A_{\mu\nu} e^{i x_\alpha k^\alpha}\right] \tag{1.13}$$

Here, $A_{\mu\nu}$ is the amplitude of the wave and k^α are the wave vectors.

We can then insert this approximation back into the vacuum propagation equation, expand in powers of k, and obtain the leading-order result,

$$k_\mu k^\mu = g^{\mu\nu} k_\mu k_\nu = 0 \tag{1.14}$$

This result shows that the wave vectors for gravitational waves are null, which means they propagate along the null geodesics of the background spacetime described by the metric $g^{\mu\nu}$. In other words, gravitational waves propagate at the speed of light. Another thing to note is that,

$$k^\mu A_{\mu\nu} = 0, \tag{1.15}$$

meaning that the gravitational waves are transverse, that is, they propagate orthogonal to the GW propagation direction.

1.2.3 Detection

To understand how gravitational waves can be detected, we need to first consider how the propagating GWs affect the motion of test particles. We use the geodesic deviation equation, which describes how nearby geodesics deviate from each other due to the background curved spacetime. In general, the geodesic deviation for two particles with four velocities U^μ describes the evolution of their separation S^μ,

$$\frac{D^2 S^\mu}{d\tau^2} = R^\mu_{\alpha\beta\nu} U^\alpha U^\beta S^\nu \tag{1.16}$$

where τ is the proper time and $R^\mu_{\alpha\beta\nu}$ is the Riemann tensor describing the curvature of the background spacetime. We can write the geodesic deviation (to linear order) and in the *transverse-traceless* gauge as,

$$\frac{\partial^2 S^\mu}{\partial t^2} = \frac{1}{2} S^\nu \frac{\partial^2 h^\mu_\nu}{\partial t^2}, \tag{1.17}$$

whereas usual, h^μ_ν, is the linearized metric perturbation representing the propagating gravitational wave. Only the transverse components of the separation vector will feel any effects due to the passing gravitational wave as was discussed in the previous section.

Gravitational wave detectors such as LIGO and Virgo use interferometry to measure the differential motion of test masses along orthogonal L-shaped arms. By measuring this oscillation, the differential displacement between the test masses due to the gravitational wave can be determined. The plot in Figure 1.1 shows a cartoon depiction of a set of circularly placed test masses and how they will oscillate over time as a gravitational wave passes through them.

The ability of gravitational wave detectors to detect and measure astrophysical sources of GWs, such as those from CBCs, is determined by their sensitivity to the small variations in distance between the test masses along the detector arms caused by the passing GW. For second-generation GW detectors, the typical strain amplitude for CBCs is around $h \approx 10^{-21}$. So for an L-shaped Advanced LIGO-like interferometer with arms in the order of a few kilometers, we can measure

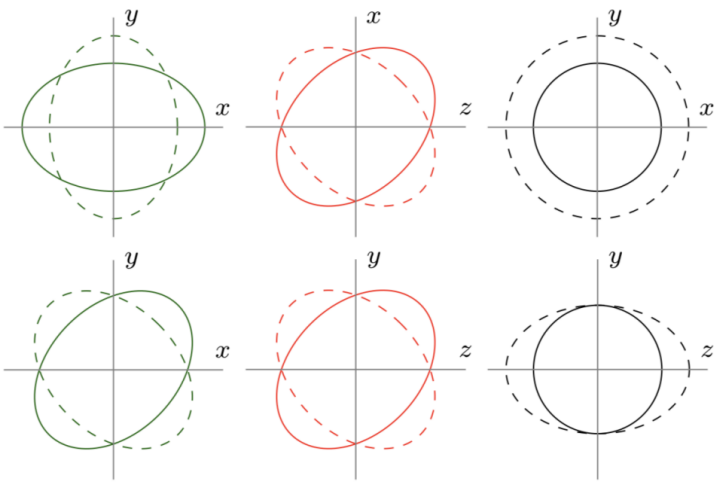


Figure 1.1: The impact of gravitational wave polarizations on a ring of test particles is demonstrated. The tensor modes are represented in green, the vector modes in red, and the scalar modes in gray. Image credit: Max Isi.

displacements on the order of

$$\delta L = \frac{1}{2} h L \approx 2 \times 10^{-18} \text{ m} \tag{1.18}$$

where we have used $L = 4000$ m. This is smaller than the size of typical atomic nuclei by at least a factor of 10^3. The way gravitational wave detectors reach this level of precision relies on Fabry-Perot cavities which effectively bounce the light coherently along the detector arms reaching gains in the order of a couple hundred effectively increasing the length of the interferometer arms. However, the design of the actual interferometers is much more complicated than what is described here and we point the reader to Creighton & Anderson (2011) for a thorough introduction.

1.3 GRAVITATIONAL WAVES IN THE ADVANCED DETECTOR ERA

1.3.1 O1 and the discovery of GW150914

The first direct detection of gravitational waves took place on September 14, 2015. The event, namely GW150914, was generated by the merger of two black holes, each with a mass of approximately 30 times that of the sun, located about 1.3 billion light-years away from Earth. As the two black holes spiraled towards each other, they emitted strong gravitational waves that were detectable by both the LIGO Hanford and Livingston interferometers. The detection of GW150914 marked a major milestone in astrophysics and a triumph for Einstein's theory of general relativity. It provided the first direct observation of black holes, confirming their existence. Throughout the remaining of the first observing run of LIGO, two more BBH mergers were discovered. These being GW151012 and GW151226 providing even more evidence that the GWs that LIGO was detecting were in fact coming from the astrophysical merger of black holes. The first discoveries of such cataclysmic events opened up the field of GW astronomy enabling the astronomical community to study the universe in ways that weren't possible before.

1.3.2 O2 and the discovery of GW170817

The second observing run of LIGO and Virgo was even more productive than the first. The third binary black hole merger event, GW170104, was detected early on in the run when only the LIGO detectors were operational during the early part of O2. Later on, the Advanced Virgo detector joined the observing run, enabling the detection of six additional BBH mergers, two of which were seen by all three detectors, namely GW170814 and GW170818. This allowed for much more precise sky localizations than what was possible with just the two LIGO detectors.

On August 17, 2017, the three-detector network made a groundbreaking discovery — the merger of two neutron stars. This event, famously known as GW170817, was localized to an area

of around 40 squared degrees and occurred at a distance of about 40 Mpc. Remarkably, roughly two seconds after the GW170817 merger, a short gamma-ray burst was detected by the FERMI Gamma Ray Burst Monitor (GBM) and the INTEGRAL detectors.

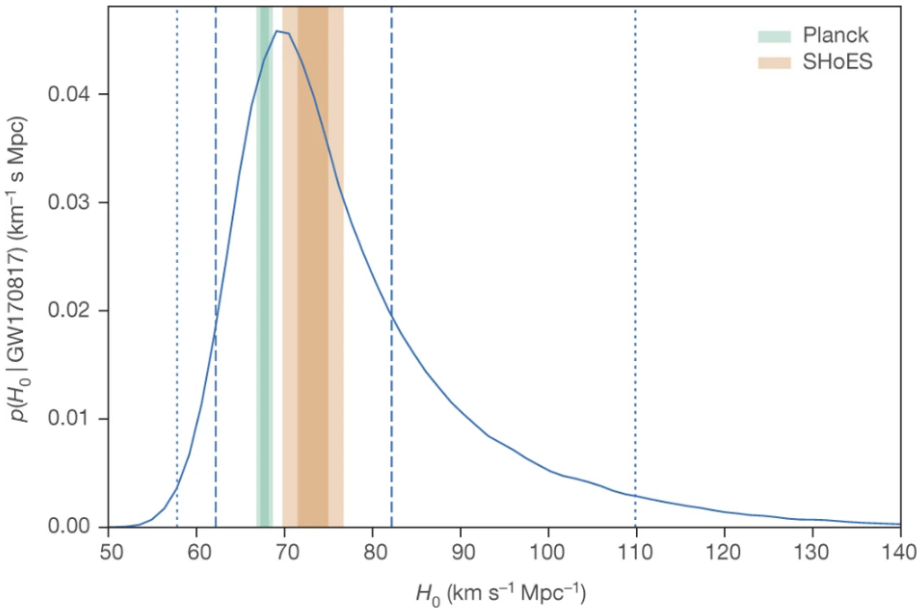


Figure 1.2: Posterior distribution on the Hubble constant from GW170817 and its uniquely determined host galaxy NGC4993. Credit: LVC Collaboration.

The precise localization of GW170817 enabled the astronomical community to follow up the transient with various electromagnetic observatories. This ultimately led to the discovery of the theoretically expected kilonova emission, allowing for the unique identification of the host galaxy of GW170817, namely NGC4993. Given the unique identification of the host galaxy for GW170817, the first standard siren measurement of the Hubble constant was performed. The corresponding posterior distribution on H_0 is shown in Figure 1.2 (Abbott et al., 2017a). The discovery of GW170817 was a major milestone in gravitational wave astronomy, as it not only provided the first direct detection of gravitational waves arising from the merger of two neutron stars but also, it allowed for multimessenger astronomy with gravitational waves, opening up a new era in astrophysics and cosmology.

1.3.3 O3, many BBHs and NSBH discovery

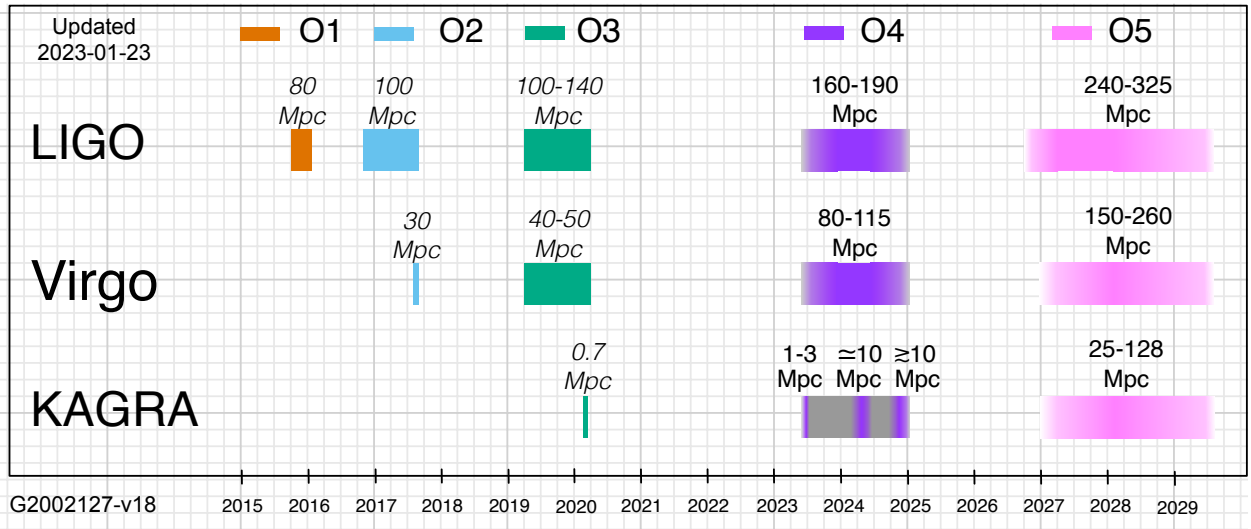


Figure 1.3: The current proposed long-term observing schedule for the LIGO, Virgo and KAGRA instruments is outlined above. The expected BNS range is indicated for each observing run. It is important to note that the O5 start dates, durations, and sensitivities presented here are current best estimates. Credit: LVK Collaboration.

The third observation run (O3) was a period of time between April 1, 2019 and March 27, 2020. During O3, LIGO detected an additional 50 potential gravitational wave events bringing the total number of events to a total of 69 confidently detected GW candidates (GWTC3). The O3 run included the second detection of gravitational waves from a binary neutron star merger (GW190425), which was observed by LIGO and Virgo on April 25, 2019. Most notably, O3 brought the first detection of two neutron star black hole candidates – GW200105 and GW200115 — completing the set of discoveries involving the merger of compact binaries involving neutron stars or black holes. The third observing run also brought two interesting candidates that challenge the astrophysical black hole population expectations: 1) a low mass, high mass ratio merger: GW190814 – challenging the lower end of the BH mass spectrum and the potential lower end mass gap. 2) The merger of two high-mass black holes — at 85 and 66 solar masses respectively, which combined

to form a BH of 142 solar masses, leading to the first indirect discovery of an intermediate-mass black hole, placing it likely above the upper PISN mass gap.

1.3.4 O4 and Beyond

The fourth observation run (O4) for the LIGO-Virgo-KAGRA Collaboration is set to start on May 24th, 2023 at the time of writing. O4 is expected to last 18 months in duration with a month break in between. We also expect a one-month engineering run prior to the start of O4 for detector commissioning purposes. O4 will see the KAGRA detector join observations along with the LIGO and Virgo detectors for the first time with an expected initial BNS range of 1 Mpc and ramping up to 10 Mpc by the end of the run. A detailed timeline for proposed O4 start dates and future observing runs is shown in Figure 1.3.

After the conclusion of O4, we expect the total number of CBC candidates to grow to $\mathcal{O}(100-500)$. Summarizing the predictions[1], the expected annual number of public GW alerts for candidates during O4 is 260^{+330}_{-150}, 36^{+49}_{-22}, and 6^{+11}_{-5} for BBH, BNS, and NSBH mergers respectively. The increasing size of the gravitational wave catalog is expected to allow for tighter constraints on the population of compact binary mergers and to use these to measure the underlying cosmological expansion and to test General Relativity.

A few of the most interesting prospects for discoveries in O4 include 1) The detection of an NSBH merger with an associated EM counterpart, 2) A second (or more) detections of BNS mergers with either sGRB, Kilonova, or other EM emissions, 3) The identification of a strongly lensed BBH merger, 4) outliers to the BBH population such as more detections of events like GW190814 and GW190521.

[1] Obtained from https://emfollow.docs.ligo.org/userguide/

1.4 INTRODUCTION TO THIS DISSERTATION

This dissertation is structured as follows. In Chapter 2, we discuss the Bayesian framework for bright and dark siren cosmology. We elaborate on the methodology by validating the `gwcosmo` library using simulated GW observations and simulated EM data in the form of mock galaxy catalogs. These mock data challenges are designed to rigorously test aspects of the methodology in a cumulative fashion, ensuring its robustness and reliability. In Chapter 3, we apply the Bayesian framework developed in Chapter 2 to real GW observations from the GWTC-1 catalog using `gwcosmo`. Notably, this marks the first combined dark siren measurement, a significant advancement in the field of GW cosmology. We thoroughly analyze the results and provide insight into potential systematics. In Chapter 4, we provide an introduction to the basics of GW lensing, with a specific focus on strong lensing. We develop a model selection framework to identify potentially lensed GW events and apply it to the GWTC-1 catalog. We identify a potential pair, namely GW170104, and GW170814, which moderately favors the lensing hypothesis. In Chapter 5, we extend the statistical framework for strong lensing GW analysis, building upon the methodology developed in the previous chapter. This led to the development of the `HANABI` library, which has been utilized in performing strong lensing analyses within the LVK. In Chapter 6, we leverage the joint parameter estimation framework developed in Chapters 4 and 5 to measure the polarization amplitudes of strongly lensed GW events. The focus of the work provides a scientific case for strong lensing, in particular, to enhance our ability to test General Relativity. In Chapter 7, we provide the first constraints on the number of allowed spacetime dimensions within modified gravity models, specifically, within gravitational leakage models, using the GWTC-3 catalog. Finally, in Chapters 8, we summarize the key findings and contributions presented in this dissertation. We provide an overview of the research conducted, highlight the main results, and discuss their implications. Furthermore, we outline future directions and set the stage for potential follow-up studies building upon the findings of this work.

1.5 DECLARATION

This thesis represents the culmination of six years of research, during which various themes and topics related to gravitational wave science have been investigated. Most of the findings presented in this work have already been published or made available as preprints elsewhere. As of May 2023, a comprehensive list of publications relevant to this dissertation is,

[1] *Cosmological inference using gravitational wave standard sirens: A mock data analysis*
R. Gray, **I. Magaña Hernandez** et al.
ApJ 908 97 (2021) (arXiv:1908.06050)

[2] *A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo*
LIGO Scientific and Virgo Collaborations (including **I. Magaña Hernandez**)
ApJ 909 218 (2021) (arXiv:1908.06060)

[3] *Identifying strong gravitational-wave lensing during the second observing run of Advanced LIGO and Advanced Virgo*
X. Liu, **I. Magaña Hernandez** and J.D.E. Creighton
ApJ 908 97 (2021) (arXiv:2009.06539)

[4] *A Bayesian statistical framework for identifying strongly-lensed gravitational-wave signals*
R.K.L. Lo and **I. Magaña Hernandez**
Accepted to PRD (arXiv:2104.09339)

[5] *Measuring the polarization content of gravitational waves with strongly lensed binary black hole mergers*
I. Magaña Hernandez
Submitted to PRD (arXiv:2211.01272)

[6] *Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers*
I. Magaña Hernandez
PRD 107, 084033 (2023) (arXiv:2112.07650)

Each of these works has contributed differently to the content of the thesis. A brief summary of the relevant publications with each chapter of this dissertation is as follows:

- Chapter 2 is mainly based on the simulated study of [1] for which I contributed extensively along with the first author. The methods are written in the style of the LVC publication [2].
- Chapter 4 is based on the results and discussion of the LVC publication [2] for which I was a member of the paper writing team and an analyst that produced and contributed to the majority of the shown results.
- Chapter 5 is adapted from my coauthor work on [3], the introduction is based on my work of [4].
- Chapter 6 is based for the most part on the methods and discussion section of [4] with slight modifications.
- Chapter 7 is based on my single-author work on GW polarization measurements with strongly lensed events [5].
- Chapter 8 is based on my single-author work on spacetime dimensions constraints with GWs [6].
- Appendices A, B, C, and D are based on [1], [2], [5], and [6] respectively.

My research has been supported by the NSF Graduate Research Fellowship Program under Grant No. DGE-17247915. This work was also supported by NSF Award No. PHY-1912649 and PHY-2207728. I am also grateful for computational resources provided by the Leonard E Parker Center for Gravitation, Cosmology and Astrophysics at the University of Wisconsin-Milwaukee. I thank LIGO and Virgo Collaboration for providing the data for this work. 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.

Part I

Cosmology

Chapter 2

Gravitational-wave cosmology

2.1 INTRODUCTION

More than thirty years ago, Bernard Schutz proposed that the detection of Gravitational wave (GW) signals could provide a means for inferring cosmological parameters, such as Hubble constant (H_0) (Schutz, 1986). The key to this method is the use of GW signals from compact binary coalescences (CBCs) as standard sirens. This means that they provide a self-calibrated luminosity distance to the source that can be directly obtained from the GW signal itself, independent of the cosmic distance ladder. By combining this distance information with the redshift of each source, we can obtain the necessary inputs for cosmological inference and thus a direct measurement of H_0.

This chapter specifically addresses the scenario where no electromagnetic (EM) counterpart is observed in association with a GW observation, which results in a lack of direct access to the host galaxy and therefore the necessary redshift information. The *statistical* or *dark siren* method proposed by Schutz in 1986 (Schutz, 1986; Del Pozzo, 2012; Soares-Santos et al., 2019b) addresses this issue by using galaxy catalogs as prior information on the potential host galaxies within the GW sky-localization volumes. The dark siren approach involves marginalizing over the possible host galaxies of each GW detection to account for the uncertainty in identifying the true host galaxy. By combining information from multiple GW events, contributions from the true host galaxies will statistically accumulate, resulting in a constraint on H_0 and potentially other cosmological parameters.

In this Chapter, we provide a brief introduction to GW cosmology and how the current Hubble tension motivates the study of GW sources as standard distance indicators. We introduce the

Bayesian framework for cosmological inference based on the `gwcosmo`[1] library. We formulate and validate on simulated GW and EM data using both the bright siren and dark siren approaches to measuring cosmological parameters.

2.1.1 The Hubble Constant Tension

Cosmology is the field that studies the properties of the universe at the largest of scales. Such properties involve the rate at which the universe is expanding as well as accelerated expansion. The expansion rate of the universe is a well-measured quantity and is given by the Hubble constant H_0. Current methods for measuring the Hubble constant are two-fold: 1) measuring the luminosity distance and redshift to type 1A supernova explosions using the so-called cosmic distance ladder for calibration; these are local (low redshift) measurements of cosmology (Riess et al., 2019) 2) measuring the light from primordial early universe (high redshift) fluctuations imprinted in the cosmic microwave background (CMB) radiation and fitting these to a cosmological model (Planck Collaboration, 2018). These two measurements provide an independent way of determining the Hubble constant at the percent level, however, there is currently a tension between these measurements, as they disagree with a statistical significance larger than 5-sigma.

In order to potentially resolve this tension, one ideally needs a third independent cosmological probe. Gravitational wave sources such as merging binary black holes and neutron stars provide an independent measurement for the Hubble constant. Gravitational wave sources are so-called standard sirens, that is, they provide an absolute measurement for the luminosity distance to the source without the need for the cosmic distance ladder as a calibrating step. To do cosmology with gravitational wave sources we thus require an independent measurement of the source redshift. For *bright sirens* such as binary neutron star mergers, we can attempt to detect the associated electromagnetic counterpart e.g. a kilonova, and measure the redshift directly. For *dark sirens* such as

[1]Publicly available at: https://git.ligo.org/lscsoft/gwcosmo

binary black hole mergers or neutron star mergers without a detectable (or missed) electromagnetic counterpart we can determine the redshift statistically using prior information coming from galaxy surveys (Schutz, 1986). The LIGO and Virgo detectors have now seen tens of binary black hole mergers and a few neutron star binaries, thus the joint cosmological inference using the population of these sources is not only important but also possible.

2.1.2 Brief review of cosmology

In a Friedmann-Lemaître-Robertson-Walker universe, the late-time cosmological expansion can be characterized by the Hubble parameter $H(z)$ as a function of redshift z. The expression for $H(z)$ is given by,

$$H(z) = H_0 \sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_{\rm k}(1+z)^2 + \Omega_\Lambda}\,, \tag{2.1}$$

where H_0 is the Hubble constant, $\Omega_{\rm m}$ and Ω_Λ are the matter and dark energy densities respectively, and $\Omega_{\rm k}$ is the curvature energy density. The constraint $\Omega_{\rm m} + \Omega_{\rm k} + \Omega_\Lambda = 1$ holds for all the components contributing to the energy density content of the Universe at the present epoch.

The redshift-distance relation maps the expansion history of the universe to the luminosity distance $D_L(z)$,

$$D_L(z) = \frac{c\,(1+z)}{H_0} \int_0^z \frac{H_0}{H(z')}\, dz'\,. \tag{2.2}$$

which depends explicitly on $H(z)$. Independent measurements of $D_L(z)$ and z from standard sirens allow us to determine $H(z)$ leading to a direct measurement of the underlying cosmology, namely measurements of H_0 and other cosmological parameters.

It is worth noting that, although the GW detections provide an alternative way to measure H_0, the measurements are not strictly independent of the other cosmological parameters that come from

EM observations. At low redshifts $z \ll 1$, the linear Hubble relation is given by,

$$D_L(z) = \frac{cz}{H_0} + \mathcal{O}(z^2) \tag{2.3}$$

and provides an approximate description of the redshift-distance relation, which is independent of the other cosmological parameters. Thus, any measurement of H_0 with GW sources based on this approximate relation would be independent of the values of the other cosmological parameters.

2.1.3 Standard Sirens

The amplitude of the observed GW strain is inversely proportional to the luminosity distance to the GW source. For compact binary sources, to the leading order (see, e.g. Sathyaprakash & Schutz (2009)),

$$A_{\rm GW} \propto \frac{\mathcal{M}_z^{5/3}}{D_L} \left[f(\mathcal{M}_z, t)\right]^{2/3} g(\iota)\,, \tag{2.4}$$

where $A_{\rm GW}$ is the observed GW amplitude, $\mathcal{M}_z \equiv \mathcal{M}(1+z)$ is the detector frame chirp mass, $f(\mathcal{M}_z, t)$ is the frequency evolution of the GW phase and $g(\iota)$ is a function of the inclination angle of the binary that depends on the polarization mode that is considered. The redshifted chirp mass, $\mathcal{M}_z$, can be estimated from the observed frequency evolution of the CBC.

The luminosity distance, D_L, can then be obtained directly from the measured amplitude of the signal. However, the inclination angle dependence introduces significant uncertainty in the measurement of D_L, this is the well-known *distance-inclination* degeneracy. With the aforementioned caveat, this makes compact binaries self-calibrated luminosity distance indicators or *standard sirens* unlike EM distance indicators which need to undergo calibration through the cosmic distance ladder. The redshift of the GW source, also required for cosmological inference, however, remains degenerate with the source frame masses, contained within $\mathcal{M}_z$, and needs to be measured through independent observations such as associated EM counterparts.

2.1.4 Electromagnetic Observations

There are multiple ways in which EM observations can provide the necessary redshift information for cosmological inference. A bright siren, such as a binary neutron star (BNS) event may be detected in coincidence with an EM counterpart, which can then potentially be uniquely associated with a host galaxy to provide a direct measurement of the redshift of the GW source. More generically, a GW event may not have a detected EM counterpart, or even expected EM emission, as is the case for binary black hole (BBH) mergers or so-called dark sirens. To be able to perform cosmological inference with dark sirens, one can use the dark siren method Schutz (1986) and use potential host galaxies within the event's sky localization region as the redshift information for the source.

While using galaxy catalogs to provide the prior redshift information, the possibility that the host galaxy lies beyond the reach of the catalog must be taken into account. EM observatories are typically flux-limited and can be reasonably modeled as having an apparent magnitude limit. This means that galaxy catalogs are inherently biased towards containing objects which are brighter and/or nearby. These are EM selection effects and must be compensated for in order to perform an unbiased measurement of cosmological parameters.

In either case, the uncertainty associated with each galaxy's redshift must also be taken into account, including the redshift error due to the galaxy's peculiar velocity, v_p. This is especially true for nearby galaxies, for which this contribution will dominate. The effect of v_p on the measurement of H_0 may be small if there are a large number of potential host galaxies in the GW events sky-localization, but for a small number of galaxies, and for the counterpart case in particular, this effect is particularly noticeable (see the treatment in Abbott et al. (2017a)).

2.2 THE DARK SIREN FRAMEWORK

In this section, we provide a complete summary of the Bayesian analysis described in Gray et al. (2019) to compute the posterior probability density on H_0 from a set of GW and EM observations (For detailed derivations, we refer the reader to the Appendix of Gray et al. (2019)). The posterior probability density on H_0, given a set $\{D_{\rm GW}\}$ of $N_{\rm det}$ detections and the associated GW data $\{x_{\rm GW}\}$ is given by,

$$\begin{aligned} p(H_0|\{x_{\rm GW}\},\{D_{\rm GW}\}) &\propto p(H_0)p(N_{\rm det}|H_0)\prod_i^{N_{\rm det}} p(x_{{\rm GW}i}|D_{{\rm GW}i},H_0) \\ &\approx p(H_0)\prod_i^{N_{\rm det}} p(x_{{\rm GW}i}|D_{{\rm GW}i},H_0) \end{aligned} \tag{2.5}$$

where in the second line we have marginalized over the astrophysical merger rate R with a prior $p(R) \propto R^{-1}$ (Fishbach et al., 2018) since the number of detected events is some fraction of the total number of expected sources $N_{\rm exp}$, and this fraction depends on H_0. Explicitly, then $p(N_{\rm det}|H_0) = \int p(N_{\rm det}|H_0,R)\,p(R)\,dR$ loses its dependence on H_0. We use $D_{{\rm GW}i}$ to indicate that the event i was confidently detected as a GW and denote by $p(H_0)$ as the prior on H_0 used in the Bayesian analysis.

The final term factorizes into the individual likelihoods for each GW detection. In the following, we write out the expressions for a single GW event i, omitting the subscript i for brevity of notation,

$$p(x_{\rm GW}|D_{\rm GW},H_0) = \frac{p(D_{\rm GW}|x_{\rm GW},H_0)p(x_{\rm GW}|H_0)}{p(D_{\rm GW}|H_0)}\,. \tag{2.6}$$

The denominator, $p(D_{\rm GW}|H_0)$, is evaluated as an integral over all possible $x_{\rm GW}$ (Abbott et al., 2017a; Chen et al., 2018; Mandel et al., 2019):

$$p(D_{\rm GW}|H_0) = \int p(D_{\rm GW}|x_{\rm GW},H_0)\,p(x_{\rm GW}|H_0)\,dx_{\rm GW}\,, \tag{2.7}$$

where $p(D_{\rm GW}|x_{\rm GW}, H_0) = 1$ in the case where the signal-to-noise ratio (SNR) of $x_{\rm GW}$ passes some detection threshold, and 0 in the case where it does not.

2.2.1 The electromagnetic counterpart case

In the presence of an EM counterpart, there is additional information in the EM data, $x_{\rm EM}$, which appears as an EM likelihood term. Together with this, there is an underlying assumption that there has been an EM detection; we denote this assumption as $D_{\rm EM}$. Thus, for a single event with an EM counterpart,

$$p(x_{\rm GW}, x_{\rm EM}|D_{\rm GW}, D_{\rm EM}, H_0) = \frac{p(x_{\rm GW}|H_0)p(x_{\rm EM}|H_0)}{p(D_{\rm EM}|D_{\rm GW}, H_0)p(D_{\rm GW}|H_0)}\,. \tag{2.8}$$

We assume that the detectability of an EM counterpart is dependent on luminosity distance (as opposed to redshift) because it is flux-limited. As GW detectability is also a function of luminosity distance, we expect $p(D_{\rm EM}|D_{\rm GW}, H_0)$ to be a constant that does not depend on H_0. This leads to

$$p(x_{\rm GW}, x_{\rm EM}|D_{\rm GW}, D_{\rm EM}, H_0) \approx \frac{p(x_{\rm GW}|H_0)p(x_{\rm EM}|H_0)}{p(D_{\rm GW}|H_0)}. \tag{2.9}$$

2.2.2 The galaxy-catalog case

In the absence of an EM counterpart, the analogous data comes from galaxy catalogs which provide a set of galaxies and their associated sky locations, redshifts, and apparent magnitudes. As we are in the regime where the detectability of GW sources extends beyond the distance to which current catalogs are complete, the possibility that the GW host galaxy is not contained in the catalog, because it is too faint, has to be taken into account. This is done by marginalizing over the cases

where the host is in the catalog (denoted G), and where it is not (denoted $\bar{G}$):

$$\begin{aligned} p(x_{\rm GW}|D_{\rm GW}, H_0) = \; & p(x_{\rm GW}|G, D_{\rm GW}, H_0) p(G|D_{\rm GW}, H_0) \\ & + p(x_{\rm GW}|\bar{G}, D_{\rm GW}, H_0) p(\bar{G}|D_{\rm GW}, H_0) \, . \end{aligned} \tag{2.10}$$

In the case where the host galaxy is in the catalog, the EM data enters in the form of information available in the galaxy catalog, this assumption is included in our conditioning on G. The EM information is used to modify our priors on galaxy redshift, sky location, and (apparent) magnitude, which are common among all GW events using the same catalog. This differs from the counterpart case where the EM data enters the expression as a likelihood term, $x_{\rm EM}$, a transient that is informative for only one GW event.

In the case where the host galaxy is not in the catalog, the complementary condition $\bar{G}$ implies that we include the information about the limitations of the EM survey. Following Gray et al. (2019), we model the galaxy catalog as having an apparent magnitude threshold, $m_{\rm th}$, since galaxy catalogs are flux-limited. This magnitude threshold, alongside the intrinsic (absolute) brightness of a galaxy and its luminosity distance, determines the probability that the galaxy is inside or outside the galaxy catalog. In the current formalism from Gray et al. (2019), the (in)completeness of the galaxy catalog follows naturally from the parameters of the underlying EM survey(s).

The quantities appearing on the right in Eq. 2.10 can be written out explicitly as follows.

$$p(x_{\rm GW}|G, D_{\rm GW}, H_0) = \frac{\sum_{j=1}^{N_{\rm gal}} \int p(x_{\rm GW}|z_j, \Omega_j, H_0) p(s|M(z_j, m_j, H_0)) p(z_j) dz_j}{\sum_{j=1}^{N_{\rm gal}} \int p(D_{\rm GW}|z_j, \Omega_j, H_0) p(s|M(z_j, m_j, H_0)) p(z_j) dz_j} \, . \tag{2.11}$$

Here, $N_{\rm gal}$ is the total number of galaxies in the galaxy catalog, Ω_j and m_j are respectively the sky coordinates and apparent magnitude for galaxy j, and $p(z_j)$ is a distribution representing the redshift of galaxy j. This quantity, $p(z_j)$, which enters as a prior for our analysis, is the posterior distribution on the galaxy redshift obtained following analysis of EM data which includes the

measurement uncertainty of individual galaxies in the catalog. We usually model this as a Gaussian distribution with mean corresponding to the redshift z_j and width corresponding to its measurement uncertainty.

The quantity $M(z_j, m_j, H_0)$ is the absolute magnitude (for a given H_0), and $p(s|M(z_j, m_j, H_0))$ is the probability of a galaxy with these parameters to host a GW source during the observation time, relative to other galaxies. Formally, s is the statement that a GW has been *emitted* (as opposed to being *detected*); the previous expressions are all implicitly conditioned on the assumption of s. In writing $p(s|M)$, we make the approximation that the probability of a galaxy hosting a source depends only on the intrinsic luminosity of the galaxy. In essence, this term allows for weighting galaxies by their luminosities $L(M_j(H_0))$ as

$$p(s|M(z_j, m_j, H_0)) \propto \begin{cases} \text{constant} & \text{if unweighted} \\ L(M_j(H_0)) & \text{if luminosity-weighted.} \end{cases} \tag{2.12}$$

The likelihood when the host galaxy is not in the catalog, $p(x_{\rm GW}|\bar{G}, D_{\rm GW}, H_0)$, is a ratio of marginalized integrals:

$$p(x_{\rm GW}|\bar{G}, D_{\rm GW}, H_0) = \frac{\iiint_{z(m_{\rm th},M,H_0)}^{\infty} p(x_{\rm GW}|z,\Omega,H_0)p(z)p(\Omega)p(M|H_0)p(s|M)dzd\Omega dM}{\iiint_{z(m_{\rm th},M,H_0)}^{\infty} p(D_{\rm GW}|z,\Omega,H_0)p(z)p(\Omega)p(M|H_0)p(s|M)dzd\Omega dM}\,. \tag{2.13}$$

Here the fact that the terms are conditioned on $\bar{G}$ is incorporated into the redshift limits as a function of the apparent magnitude threshold $m_{\rm th}$ of the galaxy catalog. Finally, the prior probabilities that a given GW detection has or does not have support in the galaxy catalog are respectively

$$p(G|D_{\rm GW}, H_0) = \frac{\iiint_{0}^{z(m_{\rm th},M,H_0)} p(D_{\rm GW}|z,\Omega,H_0)p(z)p(\Omega)p(M|H_0)p(s|M)dzd\Omega dM}{\iiint_{0}^{\infty} p(D_{\rm GW}|z,\Omega,H_0)p(z)p(\Omega)p(M|H_0)p(s|M)dzd\Omega dM}, \tag{2.14}$$

and

$$p(\bar{G}|D_{\rm GW}, H_0) = 1 - p(G|D_{\rm GW}, H_0). \tag{2.15}$$

In Eqs. (2.13) and (2.14), $p(z)$ is the prior on the redshift of host galaxies of GW events, taken to be of the form

$$p(z) \propto \frac{1}{1+z}\frac{dV_{\rm c}(z)}{dz}R(z)\,. \tag{2.16}$$

Here $V_{\rm c}(z)$ is the comoving volume as a function of redshift and the factor $(1+z)^{-1}$ converts the merger rate from the source frame to the detector frame. The merger rate density may in general be a function of redshift as it may track the star-formation rate e.g. $R(z) \propto (1+z)^{2.7}$ for $z \ll 1$, however, other models might also be viable. The prior on the GW sky location $p(\Omega)$ is taken to be uniform across the sky.

The term $p(M|H_0)$ is the prior on absolute magnitudes for all the galaxies in the universe (not just those inside the galaxy catalog), which we set to follow the Schechter luminosity function:

$$p(M|H_0) \propto 10^{-0.4(\alpha+1)(M-M^*(H_0))} \exp\left[-10^{-0.4(M-M^*(H_0))}\right]. \tag{2.17}$$

More complex models for $p(M|H_0)$ can be used; in fact, we expect the luminosity distribution of galaxies to also evolve with redshift (Caditz & Petrosian, 1989), as well as to depend on galaxy type and color (Madgwick et al., 2002).

This concludes the main ingredients for the galaxy catalog method for cosmological inference using GW data. In the following section, we apply this formalism to a set of simulated observations of GW and EM galaxy catalogs to test and validate the methodology.

2.3 MOCK DATA CHALLENGE

In this section, we describe a series of mock data challenges (MDCs) that we use to test the Bayesian formalism described in Section 2.2 and their ability to infer the posterior on H_0 under different conditions. For each case, the MDC consists of (i) simulated GW data, and (ii) a corresponding mock galaxy catalog.

For each of the MDCs we use an identical set of simulated BNS events from The First Two Years of Electromagnetic Follow-Up with Advanced LIGO and Virgo dataset (Singer et al., 2014; Berry et al., 2015). The set of BNS events comes from an end-to-end simulation of approximately 50,000 *injected* events in detector noise corresponding to a sensitivity similar to what was achieved during second observing run (O2). Only a subset (approximately 500 events) were *detected* by a network of two or three detectors with the `GstLAL` matched filter detection pipeline (Messick et al., 2017). From the above detections, 249 events were randomly selected (in a way that no selection bias was introduced), and these events underwent full Bayesian parameter estimation using the `LALInference` software library (Veitch et al., 2015) to obtain gravitational wave posterior samples and skymaps. It is these 249 events of the First Two Years dataset and the associated GW data that we use for our analysis.

The galaxy catalogs for each iteration of the MDC described below are designed to test a new part of the `gwcosmo` methodology in a cumulative fashion, starting with GW selection effects, adding in EM selection effects, and finally testing the ability to utilize the information available in the observed brightness of host galaxies, by weighting the galaxies with a function of their intrinsic luminosities.

The starting point for the galaxy catalogs is to take all 50,000 injected events from the First Two Years dataset and simulate a mock universe, which contains a galaxy corresponding to each injected event's sky location and luminosity distance, where the latter is converted to a redshift using a fiducial H_0 value of 70 km s^{-1} Mpc^{-1}. The First Two Years data was originally simulated

in a universe where GW events followed a D_L^2 distribution, and there was no distinction between the source frame and the (redshifted) detector frame masses. Though not ideal, this data reasonably mimics a low redshift universe ($z \ll 1$) in which the linearized Hubble relation of Eq. 2.3 holds, and galaxies follow a z^2 distribution. To be consistent, we use the same linearized relation for the generation of each of the MDCs.

The first few columns of Table 2.1 summarize the characteristics of each of the galaxy catalogs created and how they correspond to each MDC. We give a brief description of each of the cases below.

MDC	Host galaxy preference	Completeness[2]	$m_{\rm th}$	Analysis assumption	H_0 (km s^{-1} Mpc^{-1})	Percentage Error
0	Known host	-	-	direct counterpart	$69.08^{+0.79}_{-0.80}$	1.13%
1	equal weights	100%	-	unweighted catalog	$68.91^{+1.36}_{-1.22}$	1.84%
2a	equal weights	75%	19.5	unweighted catalog	$69.69^{+1.66}_{-1.44}$	2.21%
2b	equal weights	50%	18	unweighted catalog	$69.76^{+1.79}_{-1.65}$	2.46%
2c	equal weights	25%	16	unweighted catalog	$69.64^{+2.44}_{-2.10}$	3.24%
3a	luminosity weighted	50%	14	weighted catalog	$70.38^{+3.49}_{-2.64}$	4.38%
3b	luminosity weighted	50%	14	unweighted catalog	$68.95^{+4.41}_{-3.54}$	5.68%

Table 2.1: A summary of the main results from each of the MDCs. For H_0 we quote the mode and the 68.3% error region values with their 1σ error. The percentage error in H_0 is defined as the half-width of the 68.3% highest probability interval divided by 70 km s^{-1} Mpc^{-1}.

2.3.1 MDC0: Known Associated Host Galaxies

MDC0 is the simplest version of the direct counterpart case, in which we identify with certainty the host galaxy for each GW event. As the galaxies are generated with no redshift uncertainties or peculiar velocities, these results will be (very) optimistic for the counterpart case. Analysis of this MDC provides the *best possible* constraint on H_0 using the 249 events we consider, which then allows for comparison with the other MDCs.

2.3.2 MDC1: Complete Galaxy Catalog

The MDC1 universe consists of the full set of 50,000 galaxies out to $z \approx 0.1$ ($D_L \approx 428\,\mathrm{Mpc}$) in the original First Two Years dataset. This gives a galaxy number density of ~ 1 per $7000\,\mathrm{Mpc}^3$, which is ~ 35 times sparse compared to the actual density of galaxies in the local universe (Dálya et al., 2018a). Additional galaxies are generated beyond the edge of the dataset universe, uniformly across the sky and uniformly in comoving volume, thereby extending the universe out to a radius of 2000 Mpc ($z = 0.467$ for $H_0 = 70$ km s^{-1} Mpc^{-1}). This means that even allowing H_0 to be as large as 200 km s^{-1} Mpc^{-1}, the edge of the MDC1 is more than twice the highest redshift associated with the farthest detection (which is at ~ 270 Mpc). Each of the 249 detected BNS has a unique associated host galaxy contained within the MDC1 catalog. This catalog is thus *complete* in the sense that it contains every galaxy in the simulated universe. MDC1 is designed to further test the GW selection effects by ensuring that, given a set of sources and access to a complete catalog, our methodology and analysis produce a result consistent with the simulated value of H_0.

2.3.3 MDC2: Incomplete Galaxy Catalog

MDC2 is designed to test EM selection effects, by introducing an apparent magnitude threshold on their construction, such that a certain fraction of the host galaxies are not contained in it. This is a necessary consideration, given that we are in the regime where GW signals are being detected beyond the distance to which the current galaxy catalogs can be considered to be complete.

In order to create the catalog for MDC2, we start with the initial MDC1 universe and assign luminosities to each of the galaxies within it. We assume a luminosity distribution that follows a Schechter function of the form (see e.g. Binney & Tremaine (1987))

$$\phi(L)\,dL = n^* \left(\frac{L}{L^*}\right)^{\alpha} e^{-L/L^*}\,\frac{dL}{L^*}\,, \tag{2.18}$$

where L denotes a given galaxy luminosity and $\phi(L)\,dL$ is the number of galaxies within the luminosity interval $[L, L + dL]$[3]. These luminosities are then converted to apparent magnitudes using $m \equiv 25 - 2.5\log_{10}(L/L^*) + 5\log_{10}(D_L/\text{Mpc})$, and an apparent magnitude threshold m_{th} is applied as a very crude characterization of the selection function of an optical telescope observing only objects with $m < m_{\text{th}}$.

MDC2 is broken into three sub-MDCs, in order to test our ability to handle different levels of galaxy catalog completeness dictated by different telescope sensitivity thresholds. In each case, the catalog completeness is defined as the ratio of the number of galaxies inside the catalog relative to the number of galaxies inside the MDC, out to a reference fiducial distance D_L,

$$f_{\text{completeness}}(D_L) = \frac{\sum_j^{\text{MDC2}}(D_{L_j} < D_L)}{\sum_k^{\text{MDC1}}(D_{L_k} < D_L)}\,, \tag{2.19}$$

where the numerator is a sum over the galaxies contained within the MDC2 catalog out to some reference distance D_L, and the denominator is a sum over the galaxies in the MDC1 catalog.

Apparent magnitude thresholds of $m_{\text{th}} = 19.5$, 18, and 16 are chosen for the three sub-MDCs, which correspond to cumulative number completeness fractions of 75%, 50% and 25% respectively, evaluated at a distance of $D_L = 115\,\text{Mpc}$, chosen such that given the luminosity distance distribution of detected BNSs, the completeness fraction for the sub-MDC to this distance is roughly indicative of the percentage of host galaxies which remain inside the galaxy catalog. The left panel of Fig. 2.1 shows how the completeness of each of the MDC2 catalogs drop off as a function of distance.

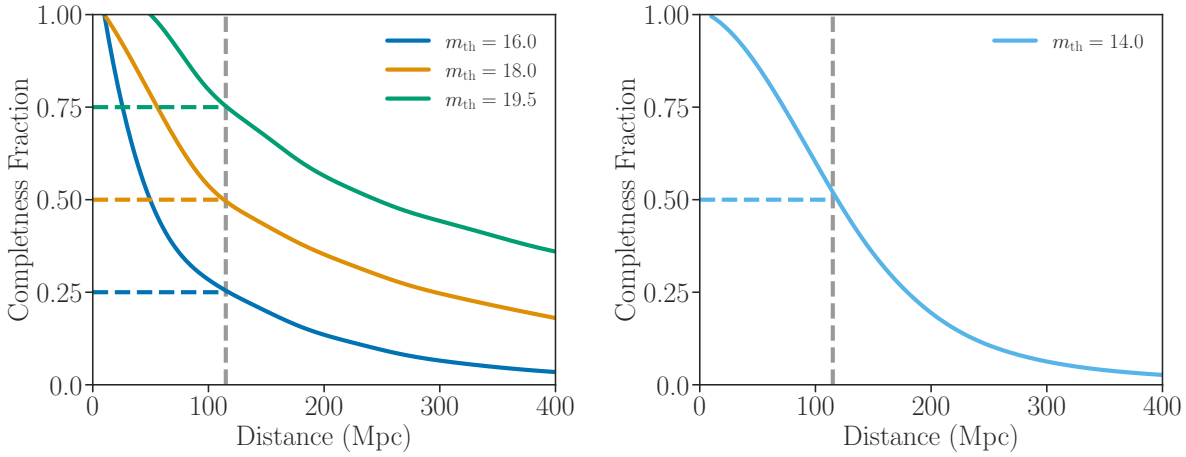


Figure 2.1: (left) Galaxy catalog completeness fractions for MDC2. Galaxy number completeness fraction defined in Eq. (2.19) as a function of luminosity distance for the three MDC2 sub-catalogs. The lines in blue, orange, and green correspond to the catalogs with $m_{\rm th} = 16$, 18, and 19.5 respectively; these correspond to completeness fractions of 25%, 50% and 75% out to a fiducial reference distance of 115 Mpc (shown as a vertical grey line). (right) Galaxy catalog completeness fractions for MDC3. The galaxy luminosity completeness fraction defined in Eq. (2.23) as a function of luminosity distance for the MDC3 catalog, with $m_{\rm th} = 14$. At the reference distance of 115 Mpc (vertical grey line), this corresponds to a completeness fraction of $\sim 50\%$.

2.3.4 MDC3: Luminosity Weighting

MDC3 is designed to test the effect of the weighting of galaxies with a function of their luminosity. It is likely that the more luminous galaxies are also more likely hosts for compact binary mergers; the luminosity in blue (B-band) is expected to be indicative of a galaxy's star formation rate, for example, while the luminosity in high infrared (K-band) a tracer of the stellar mass. The bulk of the host probability is expected to be contained within a smaller number of brighter galaxies, effectively reducing the number of galaxies that need to be considered. Additional information from

[3]The characteristic galaxy luminosity is given by $L^* = 1.2 \times 10^{10}\, h^{-2}\, L_\odot$ with solar luminosity $L_\odot = 3.828 \times 10^{26}$ W, and $h \equiv H_0/(100\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$, $\alpha = -1.07$ characterizes the exponential drop off of the luminosity function, and n^* denotes the number density of objects in the MDC universe (in practice, this only acts as a normalization constant). The integral of the Schechter function diverges at $L \to 0$, requiring a lower luminosity cut-off for the dimmest galaxies in the universe which we set to $L_{\rm lower} = 0.001 L^*$.

luminosity is thus expected to improve the constraint on H_0 by narrowing its posterior probability density.

For MDC3, the probability of a galaxy emitting a GW signal is taken to be proportional to the galaxy's luminosity. As with MDC2, the luminosity distribution of the galaxies in the universe is assumed to follow the Schechter luminosity function as in Eq. (2.18) (referred to from now on as $p(L)$). However, the joint probability of a single galaxy having luminosity L and containing an emitting source, s, is

$$\begin{aligned} p(L, s) &= p(s|L)\, p(L) \\ &\propto L\, p(L)\,, \end{aligned} \tag{2.20}$$

where we assume that the probability of a galaxy of luminosity L hosting a source is proportional to the luminosity itself. All host galaxies thus have luminosities sampled from $L\, p(L)$. In this context, we must consider all galaxies which emitted signals, not just those which were detected. With this in mind, the overall luminosity distribution has the following form:

$$p(L) = \beta \frac{L}{\langle L \rangle} p(L) + (1 - \beta)\, x(L) \tag{2.21}$$

where β is the fraction of emitting galaxies to total galaxies over the observed time period ($1 \geq \beta \geq 0$), $L/\langle L \rangle$ is the normalized luminosity, and $x(L)$ is the unknown luminosity distribution of galaxies which did not emit gravitational waves which we can sample for a given value of β.

Rearranging to obtain the only unknown, $x(L)$, gives

$$x(L) = \frac{p(L)}{1 - \beta} \left[1 - \beta \frac{L}{\langle L \rangle} \right], \tag{2.22}$$

and from this we see there is an additional constraint on β, because the term inside the brackets must be > 0. The maximum value that β can take is given by $\beta_{\text{max}} = \langle L \rangle / L_{\text{max}}$, where L_{max} is the maximum luminosity from the Schechter function, and $\langle L \rangle$ is the mean. From the Schechter

function parameters detailed in section 2.3.3, $\beta_{\max} \approx 0.015$. With these choices, for MDC3, the density of galaxies is increased by a factor of 100, meaning that MDC3 is not directly comparable with the previous MDC versions.

In order to include EM selection effects, an apparent magnitude cut $m_{\rm th}$ of 14 is applied, such that the completeness of the galaxy catalog is $\sim 50\%$ out to the same fiducial distance of $115\,$Mpc as in MDC2. In this case, completeness is however defined in terms of the fractional luminosity contained in the catalog, rather than in terms of the number of objects:

$$f_{\rm completeness}(D_L) = \frac{\sum_j^{\rm MDC3} L_j(d_{L_j} < D_L)}{\sum_k^{\rm complete} L_k(d_{L_k} < D_L)}\,, \tag{2.23}$$

where the numerator is summed over the galaxies inside the MDC3 apparent magnitude-limited catalog, and the denominator is summed over the galaxies in the whole MDC3 universe. This is shown in the right panel of Fig. 2.1. As the emitting galaxies are luminosity weighted, the cumulative luminosity completeness is representative of the percentage of BNS events inside the catalog.

2.4 RESULTS

In this section, we summarize the results for the mock data challenges described in Section 2.3. We show the combined posteriors on H_0 for each MDC, discuss the convergence to the simulated value of $H_0 = 70$ km s^{-1} Mpc^{-1}, and calculate the expected precision of the combined measurement under each set of conditions. In Table 2.1 we list the measured values of the Hubble constant for the combined 249 event posterior (maximum a posteriori and 68.3% highest density posterior intervals) all computed with a uniform prior in the range of $[20, 200]$ km s^{-1} Mpc^{-1}, as well as the corresponding fractional uncertainties for each of the MDCs.

We first consider the simple case where we identify the true host galaxy for every event and

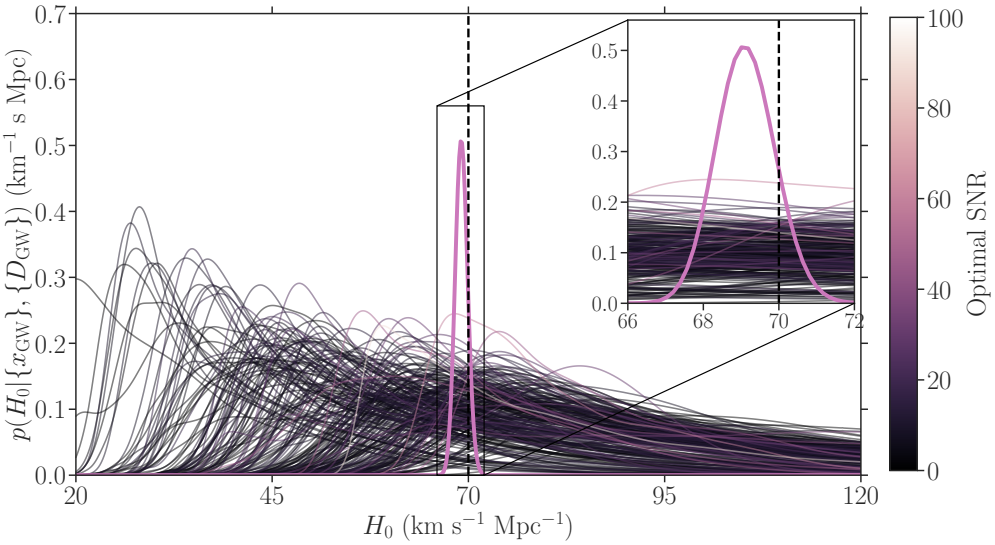


Figure 2.2: In purple we show the posterior probability on H_0 for MDC0. We also show the individual likelihoods of all events, color-coded according to their optimal SNR, and scaled up by an arbitrary value for easier viewing.

determine the resulting 249-event combined H_0 posterior. We show the corresponding single event and joint H_0 posteriors for MDC0 in Fig. 2.2 for reference. We measure a statistical estimate for the maximum a-posteriori value and 68.3% maximum-density credible interval for H_0 as $69.08^{+0.79}_{-0.80}\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$. When combining all the 249 events, we see that the final result is well-converged toward a symmetric Gaussian distribution. We note that most of the single event H_0 estimates have clearly defined modes given the uniquely associated host galaxy for each event. For subsequent MDCs this will not generally be the case and for reference, we refer the reader to Appendix A for single and combined results for other MDCs.

The result of MDC0 provides us with the best possible H_0 estimate given this set of GW detections, as this case corresponds to perfect knowledge of the host galaxy. This gives us a good benchmark against which other versions of the MDC can be compared. Since this is a best-case scenario, we have the least statistical uncertainty in the final result, making any systematic bias more apparent than for the later MDCs. For the final result with 249 events, the true value is contained within the posterior. Therefore we are confident that our method is free of systematics to

the required level for BNS counterpart measurements in observing runs in the Advanced-detector era.

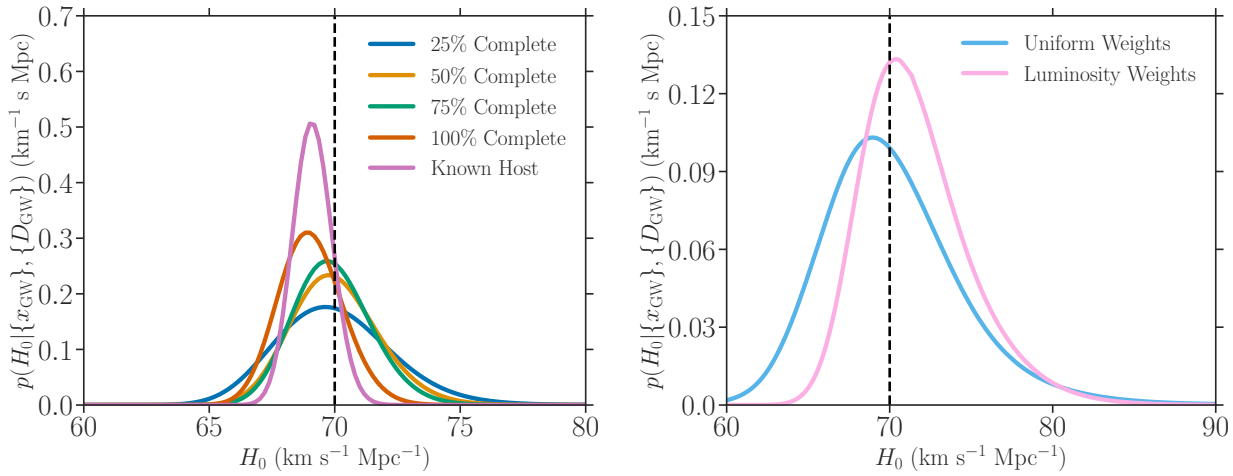


Figure 2.3: We show the joint posterior probability on H_0 using all 249 events for the MDCs considered in this chapter. (left) In purple, we show the posterior on H_0 for the case where we have identified each event's host galaxy (MDC0). We also show the posterior probability on H_0 for MDC1 and MDC2 where the number completeness of the galaxy catalog varies as a function of the chosen apparent magnitude limit. In red, green, yellow, and blue we show the 100%, 75%, 50%, and 25% completeness (out to 115 Mpc) cases respectively. (right) Posterior probability on H_0 for MDC3. By construction, the probability of any galaxy hosting a GW event is proportional to its luminosity. The pink curve gives the posterior for the case where we take the luminosity weighting into account. The blue curve gives the posterior for the case where we ignore the luminosity weighting effectively giving equal weights to each potential host.

The next most complex case is MDC1, where we assume no EM counterpart was observed, and use a galaxy catalog instead as EM information. However, MDC1 uses a *complete* galaxy catalog, so contains all potential hosts, and so EM selection effects are not present yet. Although this is an artificially optimistic scenario, we see that MDC1 produces a wider posterior on H_0 ($68.91^{+1.36}_{-1.22}$ km s^{-1} Mpc^{-1}) than MDC0, because of the uncertainty on the actual host galaxy. Our result is shown in the left panel of Fig. 2.3 as the red line. The introduction of this uncertainty means that the contributions from each event will be smoothed out, depending on the size of the

event's sky localization and the number of galaxies within it.

Moving to MDC2, the case with incomplete galaxy catalogs that are limited by an apparent magnitude threshold gives us the first case where accounting for EM selection effects is important. We consider three galaxy catalogs, with apparent magnitude thresholds of 16, 18, and 19.5, with respective completeness fractions of 25%, 50%, and 75% (see 2.3.3 for details). The combined 249-event posterior distributions on H_0 are shown also in the left panel of Fig. 2.3. As the catalogs become less complete (as defined by Eq. (2.19)), the combined H_0 posterior becomes wider. This is because the probability that the host galaxy is inside the catalog decreases. The contribution from the galaxies within the catalog is reduced, and the uninformative contribution from the out-of-catalog term in Eq. 2.10 increases. For the 25% complete case we estimate H_0 as $69.64^{+2.44}_{-2.10}$ km s^{-1} Mpc^{-1}, for 50% completeness we find $69.76^{+1.79}_{-1.65}$ km s^{-1} Mpc^{-1}, and for 75% completeness, we find $69.69^{+1.66}_{-1.44}$ all with a uniform prior on the range $[20, 200]$ km s^{-1} Mpc^{-1}.

Until now we have considered all galaxies in our catalog to be equally likely to host a gravitational-wave source. In MDC3 we analyze the case where this is no longer true by constructing a galaxy catalog such that the probability of any single galaxy hosting a GW source was directly proportional to its luminosity. MDC3 includes the same EM selection effects as MDC2, in that the catalog is magnitude limited. The completeness of this catalog, defined in terms of luminosity rather than numbers of galaxies, as defined in Eq. 2.23, is 50% out to 115 Mpc.

To investigate the importance of the luminosity weighting, MDC3 was analyzed twice under different assumptions. In the first, the analysis was matched to the known properties of the galaxy catalog, such that the probability of any galaxy hosting a GW event was proportional to its luminosity. In the second, we ran the analysis with the assumption that each galaxy was equally likely to be host to a GW event. The combined H_0 posteriors for both cases are shown in the right panel of Fig. 2.3. The estimated values of the Hubble constant are $70.38^{+3.49}_{-2.64}$ km s^{-1} Mpc^{-1}(assuming hosts are luminosity weighted), and $68.95^{+4.41}_{-3.54}$ km s^{-1} Mpc^{-1}(assuming equal weights).

By matching the analysis to the known qualities of the galaxy catalog, using the known luminosity weighting of host galaxies, the constraint on H_0 is improved in that the uncertainty narrows by a factor of 1.3, compared to the case in which equal weights are assumed. Both results are consistent with the simulated H_0 value of 70 km s^{-1} Mpc^{-1}. In the limit of a far greater number of events, one might expect to see a bias emerge in the case in which the assumptions in the analysis do not match those with which the catalog was simulated. However, for the 249 BNS events considered here, the final posteriors are too broad to be able to detect any kind of bias.

2.4.1 Comparison between the MDCs

The 249 GW events that we consider in this work allow us to study the convergence for the combined Hubble posterior as we combine events. We calculate the intermediate combined posteriors as a function of the number of events and show the resulting convergence in Fig. 2.4. We plot the fractional H_0 uncertainty (defined here as the width of the 68.3% credible interval divided by H_0, $\Delta^{68.3\%}_{H0}/H_0$), against the number of events we include in a randomly-selected group. The scatter between realizations of the group is indicated by the error bars, which encompass 68.3% of their range. There is a considerable variation between different realizations, for the incomplete catalogs. For example, of the 100 realizations we used, for 25% completeness and 40 events, there are groups that give $\sim 10\%$ precision, but others that give $\sim 70\%$ precision.

With a sufficiently large number of events, we expect a $1/\sqrt{N}$ scaling of the number of events. To check whether this behavior is indeed true, we fit the results for each MDC to the expected scaling, obtaining the coefficient of $1/\sqrt{N}$ by maximizing its likelihood given the fractional uncertainties and their variances from the different realizations. The coefficient of the scaling is automatically dominated by the fractional uncertainties at large N where the variances are small. We show this scaling for each MDC as a set of dashed lines in Fig. 2.4. It can be seen that for each MDC the data converges to the expected $1/\sqrt{N}$ scaling. The number of events required before this

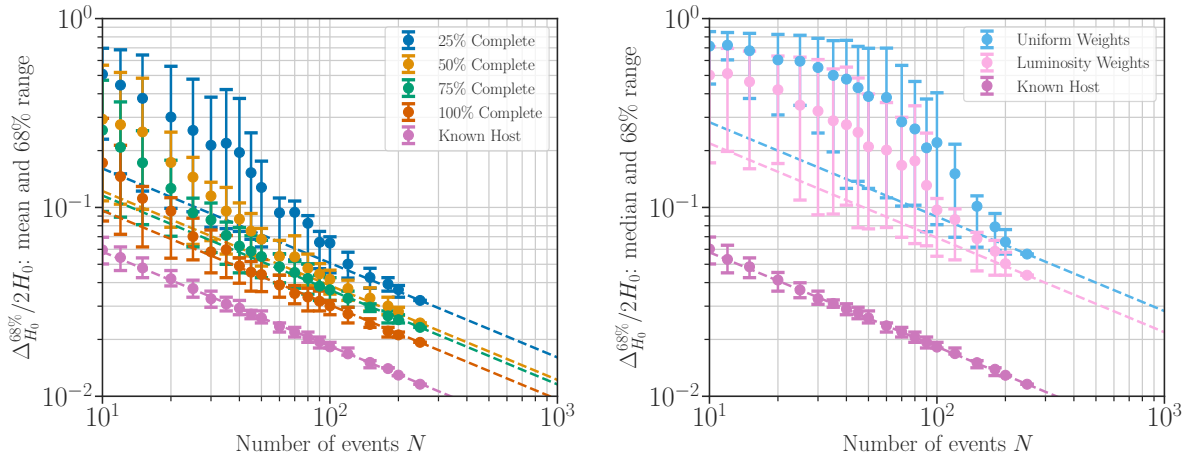


Figure 2.4: Fractional uncertainty in H_0 as a function of the number N of the events for the combined H_0 posteriors. The fractional uncertainty in H_0 is defined as the width of the 68.3% highest probability interval divided by 70 km s^{-1} Mpc^{-1}, and is shown as the plotted dots for all cases. (left) In purple, red, green, yellow, and blue we show the associated host galaxy case (MDC0), complete galaxy catalog (MDC1) case, and the 75%, 50%, and 25% completeness cases; we find a fractional H_0 uncertainty of 1.13%, 1.84%, 2.21%, 2.46% and 3.24% respectively for the combined H_0 posterior from 249 events. (right) convergence for MDC3 (event probability proportional to galaxy luminosity), analyzed with the luminosity-weighted likelihood (pink) or equally-weighted likelihood (light blue). We find fractional H_0 uncertainties of 4.38% and 5.68% respectively. MDC0 (purple) is included for reference. We plot the expected $1/\sqrt{N}$ scaling behavior for large values of N for all cases with the dashed lines. This scaling behavior is met by all MDCs as the number of events reaches 249, but for the less informative, lower completeness MDCs the trend is slower to emerge. This is even more evident in MDC3, where the density of galaxies is 100 times greater, producing more potential hosts for each event. This is mitigated somewhat by the effect of luminosity-weighting the potential hosts (pink).

behavior is reached is dependent on the amount of EM information available on average for each event. The direct counterpart case is always on the trend, after $\mathcal{O}(10)$ events. However, for the least complete catalog (25%) it appears to have reached this behavior by the time all 249 events are combined. As the density of galaxies in MDC3 was increased by 2 orders of magnitude over MDCs 1 and 2, the final posteriors cannot be directly compared between MDCs. However, by plotting the equivalent convergence figure for MDC3 (including the "known host" case as a reference,

see Fig. 2.4), one can see the impact of increasing the density of galaxies in the universe on the speed at which the posterior converges on the $1/\sqrt{N}$ behavior. As there are more host galaxies, the results are overall less precise and take longer to reach the $1/\sqrt{N}$ trend indicating a convergence on a Gaussian posterior. As expected, using luminosity-weighting of potential host galaxies as an assumption in the analysis concentrates the probability in a smaller number of galaxies, leading to a more precise result.

2.5 CONCLUSIONS AND OUTLOOK

We performed a series of MDCs to validate our method using 249 GW events. For each of the MDCs analyzed, the final posterior on H_0 is found to be consistent with the value of $H_0 = 70\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$ used to simulate the MDC galaxy catalogs, demonstrating that our method can produce sufficiently unbiased results.

GW selection effects are inherent in every version of the MDC and were fully taken into account in our formalism. EM selection effects are addressed in MDCs 2 and 3 by the out-of-catalog terms (incompleteness corrections) in our implementation. For each the MDCs, despite having an apparent magnitude-limited galaxy catalog, we are able to accurately infer H_0 without any bias. MDC2 further demonstrates our ability to account for missing host galaxies down to a level where only 25% of events have host galaxies within the catalog. Even in this case, we converge to the correct H_0 value, to the level of precision that could be reached by 249 events.

MDC3 demonstrates a clear tightening of the posterior distribution when we can assume that GW events trace the galaxy luminosities, compared to the case in which we treat all galaxies as equally likely hosts. The equal weights analysis of this MDC remains consistent with the simulated H_0 value, leaving us unable to conclude whether an incorrect assumption would lead to a biased result, as one might expect. Since for 249 BNS events, both analyses remain consistent, and it is only with a much greater number of events that any underlying bias would be detected. Future

work will be needed to expand these studies to include much greater numbers of GW events.

Chapter 3

A measurement of the Hubble constant following the Second Observation Run of LIGO and Virgo

3.1 INTRODUCTION

In this chapter, we report the first joint GW estimate of H_0 from detections during O1 and O2, the first and second observing runs of the Advanced LIGO and Virgo detector network. For our final result, along with the BNS GW170817, we choose the six O1 and O2 BBH detections which satisfy our selection criterion of a network signal-to-noise ratio SNR > 12 in at least one search pipeline for GW detections. The detections for which we have significant galaxy catalog support are GW150914, GW151226, GW170608, and GW170814. For these events, we expect the inferred H_0 posteriors to be driven by the additional EM information. The two remaining BBHs which satisfy the selection criterion, GW170104, and GW170809, need to be retained in the analysis for consistency with the assumed population model. These events for which a significant fraction of potential host galaxies are missing in the associated galaxy catalog, can also potentially contribute to the H_0 measurement, since there is information about cosmology present in the observed source distribution (Taylor et al., 2012; Taylor & Gair, 2012; Farr et al., 2019). The distribution of the observed source parameters, including the observed luminosity distance distribution, is driven by H_0 in addition to the intrinsic astrophysical source distributions.

The main result of our analysis—a posterior distribution on H_0—is dominated by the contribution from GW170817 with its optical counterpart, with a modest improvement from the inclusion of the O1 and O2 BBHs. These results, possibly refined and marginalized over the aforementioned assumptions, can be used as a prior for future GW estimates of H_0. The analysis performed in this

work thus serves as a precursor of future analyses for the third and subsequent observing runs of the Advanced detector network.

Throughout this chapter, we assume a ΛCDM cosmology and use the best-fit Planck 2015 values of $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$, respectively for the fractional matter and dark energy densities in the present epoch (Ade et al., 2016a). Although these parameters enter the redshift-distance relationship central to the method for Bayesian inference of H_0, we have verified that our results are robust with regard to a variation of their values within the current measurement uncertainties.

3.2 DATA

3.2.1 Gravitational-wave data

The GW searches performed during the first and the second observation runs of Advanced LIGO and Virgo have led to the identification of ten BBH and one BNS mergers (Abbott et al., 2019b). The BNS event GW170817, well-localized and at a nearby distance of 40^{+10}_{-10} Mpc, helped discover the electromagnetic transient from the merger, and was subsequently associated with host galaxy NGC4993. The BBHs span a large range of distances from 320^{+120}_{-110} Mpc to 2840^{+1400}_{-1360} Mpc and are distributed over the sky with 90% credible regions as low as 39 deg^2 to as high as 1666 deg^2. For the main results presented in Section 3.3 of this work, we choose the events which meet the selection criterion of network SNR > 12 in at least one of the two pipelines for modeled searches, namely PyCBC and GstLAL. A summary of the relevant parameters of all the GW detections are given in Table 3.1.

3.2.2 Galaxy Catalogs

The analysis with BBHs is performed in conjunction with appropriate galaxy catalogs. In the following, we describe in more detail the galaxy catalogs that we use, quantify the probability that the host galaxy for each event is in the galaxy catalog that is used for its analysis and discuss the

Event	SNR	$\Delta\Omega/\mathrm{deg}^2$	$d_\mathrm{L}/\mathrm{Mpc}$	z_event	V/Mpc^3	Galaxy catalog	Number of galaxies	m_th	$p(G\|z_\mathrm{event}, D_\mathrm{GW})$
GW150914	24.4	182	440^{+150}_{-170}	$0.09^{+0.03}_{-0.03}$	3.5×10^6	GLADE	3910	17.92	0.42
GW151012	10.0	1523	1080^{+550}_{-490}	$0.21^{+0.09}_{-0.09}$	5.8×10^8	GLADE	78195	17.97	0.01
GW151226	13.1	1033	450^{+180}_{-190}	$0.09^{+0.04}_{-0.04}$	2.4×10^7	GLADE	27677	17.93	0.41
GW170104	13.0	921	990^{+440}_{-430}	$0.20^{+0.08}_{-0.08}$	2.4×10^8	GLADE	42221	17.76	0.01
GW170608	15.4	392	320^{+120}_{-110}	$0.07^{+0.02}_{-0.02}$	3.4×10^6	GLADE	6267	17.84	0.60
GW170729	10.8	1041	2840^{+1400}_{-1360}	$0.49^{+0.19}_{-0.21}$	8.7×10^9	GLADE	77727	17.82	< 0.01
GW170809	12.4	308	1030^{+320}_{-390}	$0.20^{+0.05}_{-0.07}$	9.1×10^7	GLADE	18749	17.62	< 0.01
GW170814	16.3	87	600^{+150}_{-220}	$0.12^{+0.03}_{-0.04}$	4.0×10^6	DES-Y1	31554	23.84	> 0.99
GW170817	33.0	16	40^{+7}_{-15}	$0.01^{+0.00}_{-0.00}$	227	–	–	–	–
GW170818	11.3	39	1060^{+420}_{-380}	$0.21^{+0.07}_{-0.07}$	1.5×10^7	GLADE	1059	17.51	< 0.01
GW170823	11.5	1666	1940^{+970}_{-900}	$0.35^{+0.15}_{-0.15}$	3.5×10^9	GLADE	117680	17.98	< 0.01

Table 3.1: Relevant parameters of the O1 and O2 detections: network signal-to-noise ratio (SNR) for the search pipeline (PyCBC/GstLAL) in which the signal is the loudest, 90% sky localization region $\Delta\Omega$ (deg^2), luminosity distance d_L (Mpc, median with 90% credible intervals), and estimated redshift z_event (median with 90% range assuming Planck 2015 cosmology) from Abbott et al. (2019b). In the remaining columns, we report the corresponding 90% 3D localization comoving volumes, the number of galaxies within each volume for public catalogs which we find to be the most complete, and the apparent magnitude threshold, m_th, of the galaxy catalog associated with the corresponding sky region (as described in Section 3.2.3). The final column gives the probability that the host galaxy is inside the galaxy catalog for each event, $p(G|z_\mathrm{event}, D_\mathrm{GW})$, also evaluated at the median redshift for each event.

assessment of the completeness over the relevant localization volume for the best-localized events. Finally, we quantify the uncertainties associated with the photometric measurement of redshifts in some of these catalogs. For extensive details in the selection criteria and construction of the galaxy samples we refer the reader to (LVC, 2019).

In Table 3.1 we summarize the galaxy catalogs that we use for our analysis for each of the detections, along with the number of galaxies in the 90% error volume calculated from 3D skymaps constructed from posterior samples associated with the data release of Abbott et al. (2019b),and estimates of the probability that the host galaxy is in the catalog, evaluated at the median redshift for each detection assuming a Planck 2015 cosmology.

3.2.2.1 *GLADE*

We use the GLADE version 2.4 galaxy catalog (Dálya et al., 2018b),[1] to construct the observed redshift distributions for the majority of the detected BBHs. The GLADE catalog has an all-sky coverage (Fig. 1 of Dálya et al., 2018b) since it is constructed from the GWGC (White et al., 2011), 2MPZ (Bilicki et al., 2014), 2MASS XSC (Skrutskie et al., 2006), HyperLEDA (Makarov et al., 2014) and SDSS-DR12Q (Pâris et al., 2017) catalogs. The GLADE catalog is complete (in B-band luminosity) out to 37 Mpc and has an estimated completeness of 50% out to 91 Mpc (Fig. 2 of Dálya et al., 2018b). At low redshifts ($\lesssim 0.05$), we expect to be dominated by the peculiar velocity field. GLADE reports peculiar-velocity-corrected redshifts following the reconstruction of Carrick et al. (2015). GLADE provides apparent magnitudes in the B-band, which we can use directly (without any photometric transformations) for the luminosity weighting of the galaxies.

3.2.2.2 *DES Year 1*

The Dark Energy Survey (DES) is a five-year survey mapping ≈ 300 million galaxies in five filters (*g*rizY) over 5000 deg^2. It is worth noting that the GW170814 sky localization is fully enclosed within the footprint of the DES (Drlica-Wagner et al., 2018; Abbott et al., 2018b) Year 3 (Y3) gold catalog. An estimate of H_0 from the GW170814 distance and the Y3 catalog of the DES has been carried out (Soares-Santos et al., 2019a). In this work, we use the publicly available DES-Y1 catalog (Abbott et al., 2018b),[2] to compute the H_0 posterior for GW170814. Approximately 87% of the probability region for the GW170814 sky localization is enclosed within the DES-Y1 catalog. Analysis with a different catalog provides a parallel measurement of H_0 with GW170814, and (given the catalog differences) can potentially be indicative of systematic effects in the catalogs, such as the treatment of redshift uncertainties (provided that a similar set of galaxies are present in both catalogs). See Appendix B for more details.

[1] GLADE is publicly available at: http://glade.elte.hu
[2] DES-Y1 is available at: https://des.ncsa.illinois.edu/releases/y1a1

3.2.2.3 *Redshift uncertainties*

An important source of measurement uncertainty with galaxy catalogs is in the galaxy redshifts, which are often photometric estimates due to a lack of spectroscopic measurements out to large redshifts. In our formalism, the uncertainty in redshift is reflected in the term $p(z_j)$ in Eq. (2.11). We model the individual galaxy redshift probability distributions, $p(z_j)$, as a Gaussian distribution with mean set to the quoted median photometric redshift z_{photo}, and with a standard deviation of $\sigma_{z_{\text{photo}}}$ (both obtained from the galaxy catalog data). Photometric redshift estimates are often not approximated by Gaussian distributions, and we make this assumption only due to limited information present in some of the public galaxy catalogs which we use. A rigorous quantification of the effect of this assumption is beyond the scope of this work. In practice, we implement the redshift uncertainty by the process of a Monte Carlo integration: the integral over z_j in Eq. (2.11) becomes an additional sum over N_{photo} random samples. This number varies between the different galaxies and is always high enough to ensure that the final event likelihood is independent of the set of random draws from each galaxy.

3.2.3 Probability that the host galaxy is in the catalog

In this work, we assume that we can characterize the completeness of a galaxy catalog using an apparent magnitude threshold (limiting magnitude) m_{th}. We estimate m_{th} by calculating the median value from the apparent magnitude distribution of all the galaxies within the sky localization of each event. For GLADE, this choice allows us to account for some of the larger changes in completeness across the sky, which come from it being a composite catalog, comprised of many surveys of differing depths. Galaxy catalogs are directional, and a more sophisticated analysis would involve calculating the limiting magnitude for a given line of sight. Obtaining the H_0 posterior distribution would thus require a joint estimate of m_{th} along the lines of sights within an event's sky localization. We leave this for future work.

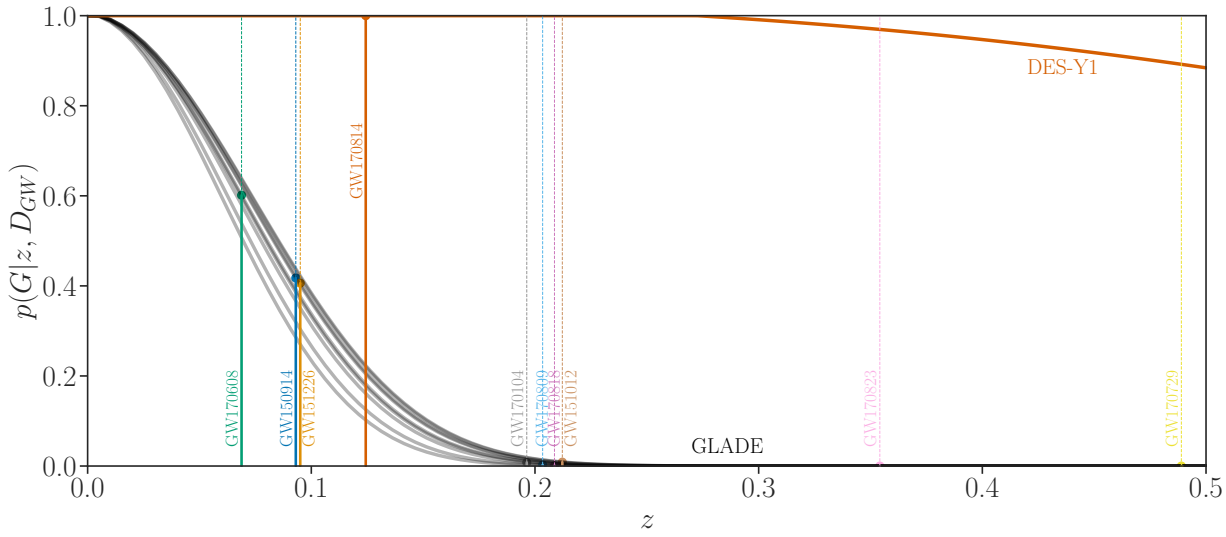


Figure 3.1: The probability that the host galaxy is inside the galaxy catalog, shown for GLADE (gray curves) and DES-Y1 (orange curve), as a function of redshift. For GLADE this quantity is calculated for each individual event, using the completeness estimated within each event's sky localization. For DES-Y1 the curve is only valid in the patch of sky covering GW170814. Each curve is independent of the value of H_0. The vertical lines show the median redshift (assuming a Planck 2015 cosmology) for each event as in Table 3.1. These lines are thick and solid up to the intercept with the galaxy catalog they are used with, and thin and dashed above. Reproduced with permission from LVC (2019)

For now, we use the $m_{\rm th}$ estimated as described above, and show in Fig.3.1 the probability of a host galaxy being inside the catalog $p(G|z, D_{\rm GW})$ as a function of redshift z, for each of the galaxy catalogs under consideration. For GLADE this quantity is calculated for each event using the $m_{\rm th}$ calculated for each event's sky localization. For DES-Y1, the curve is for the patch of sky covering GW170814. These probabilities are calculated using the expressions in Eq. (2.14), but as a function of z, and are therefore independent of the choice of H_0. We additionally show as the vertical lines in Fig. 3.1 the median redshift for each event $z_{\rm event}$ (calculated assuming a Planck 2015 cosmology). In the final two columns of Table 3.1, we report the $m_{\rm th}$ of the relevant catalog within the sky localization of each event, and the probability of the host galaxy being in the catalog at the median redshift of each event.

3.3 RESULTS

We apply the method described above to obtain a measurement of the Hubble constant using GW standard sirens only. We carry out our analysis with a prior on H_0 uniform in the interval of $[20,\ 140]$ km s^{-1} Mpc^{-1}; we report our final results also using a flat-in-log prior $p(H_0) \propto H_0^{-1}$ in the same interval for ease of comparison with previous studies. We use the marginalized distance likelihood and skymaps constructed from the posterior samples of Abbott et al. (2019b).For the BBHs, we choose all galaxies in the 99.9% sky region of the corresponding catalog and we further weight the galaxies in proportion to their $B(g)$-band luminosities. In order to calculate the term $p(D_{\rm GW}|H_0)$ in the denominator, we use a Monte Carlo integration, sampling parameters which affect an event's detectability (masses, sky location, inclination angle, and polarisation) from chosen priors. We choose a power-law mass distribution for BBHs with $p(m_1) \propto m_1^{-\alpha}$ and m_2 uniform in its range with $5M_\odot < m_2 < m_1 < 100M_\odot$ in the source frame, and a distribution of merger rates that does not evolve with redshift; for the power-law index α, we choose $\alpha = 1.6$ (which is supported by Model B of Abbott et al., 2019a). For BNSs, we use a Gaussian mass distribution with a mean of $1.35M_\odot$ and a standard deviation of $0.15M_\odot$ (Kiziltan et al., 2013).The remaining GW parameters are marginalized over their natural distributions: uniform in the sky, uniform on the sphere for orientation, uniform in polarization. We use the time-averaged power-spectral-density of detector noise for the corresponding observation run from Aasi et al. (2015), and for the selection criterion, we use a network SNR threshold of 12 in at least one search pipeline. The O1 and O2 BBHs which pass this criterion are GW150914, GW151226, GW170104, GW170608, GW170809, and GW170814 (see Table 3.1).

Our result for these O1 and O2 BBHs is shown in Fig. 3.2. The detections for which there is considerable support from the galaxies present in the catalog show features of the galaxy catalog in their H_0 posterior distribution. The GW170814 estimate is qualitatively similar to the result in Soares-Santos et al. (2019a) with analogous peaks in the posterior distribution. The differences in

peak locations can be attributed to a difference in the redshift distribution for the DES-Y3 catalog used in Soares-Santos et al. (2019a) versus that for the public DES-Y1 catalog used in this work. For the detections for which the galaxy catalogs are relatively empty, we see the features of the assumptions on mass distribution and redshift evolution of binary merger rate that have entered our analysis. The more distant events such as GW170809 lead to H_0 estimates pushed to the lower end of the prior. This is due to the following reason. In a universe where host galaxies are distributed uniformly in comoving volume, for lower values of H_0, the expected distribution of detected GW events favors relatively higher luminosity distances. Thus events at high luminosity distances have more support for smaller values of H_0, while relatively nearby events, namely GW150914, GW151226, GW170608 and GW170814, correspondingly have lower support at smaller values of H_0. The information present in the observed luminosity distance distribution would thereby potentially contribute to the H_0 measurement, independent of, or even in absence of information in galaxy catalogs, if the underlying distributions of source parameters were known. In the absence of knowledge of the astrophysical distribution of BBH source parameters, a thorough treatment would involve a marginalization over all possible mass distributions and rate models. The following section discusses the systematic differences that the assumptions on the assumed population model could lead to.

For our final result, we combine the contribution of the BBHs with the result from GW170817 obtained using the low spin prior samples from Abbott et al. (2019b) and an estimated Hubble velocity of $v_H \equiv cz = 3017 \pm 166\,\mathrm{kms}^{-1}$ (where c is the speed of light) for NGC4993 from Abbott et al. (2017a). Our final combined result is shown in Fig. 3.3, with the posterior distribution plotted assuming a uniform H_0 prior: we obtain $H_0 = 69.6^{+20.4}_{-8.6}$ km s^{-1} Mpc^{-1} (68.3% highest density posterior interval). To compare with values in the literature, we also use a flat-in-log prior, $p(H_0) \propto H_0^{-1}$, and calculate $H_0 = 68.7^{+17.0}_{-7.8}$ km s^{-1} Mpc^{-1}, which corresponds to an improvement by a factor of 1.04 (about 4%) over the GW170817-only value of $68.7^{+17.5}_{-8.3}$ km s^{-1} Mpc^{-1}.

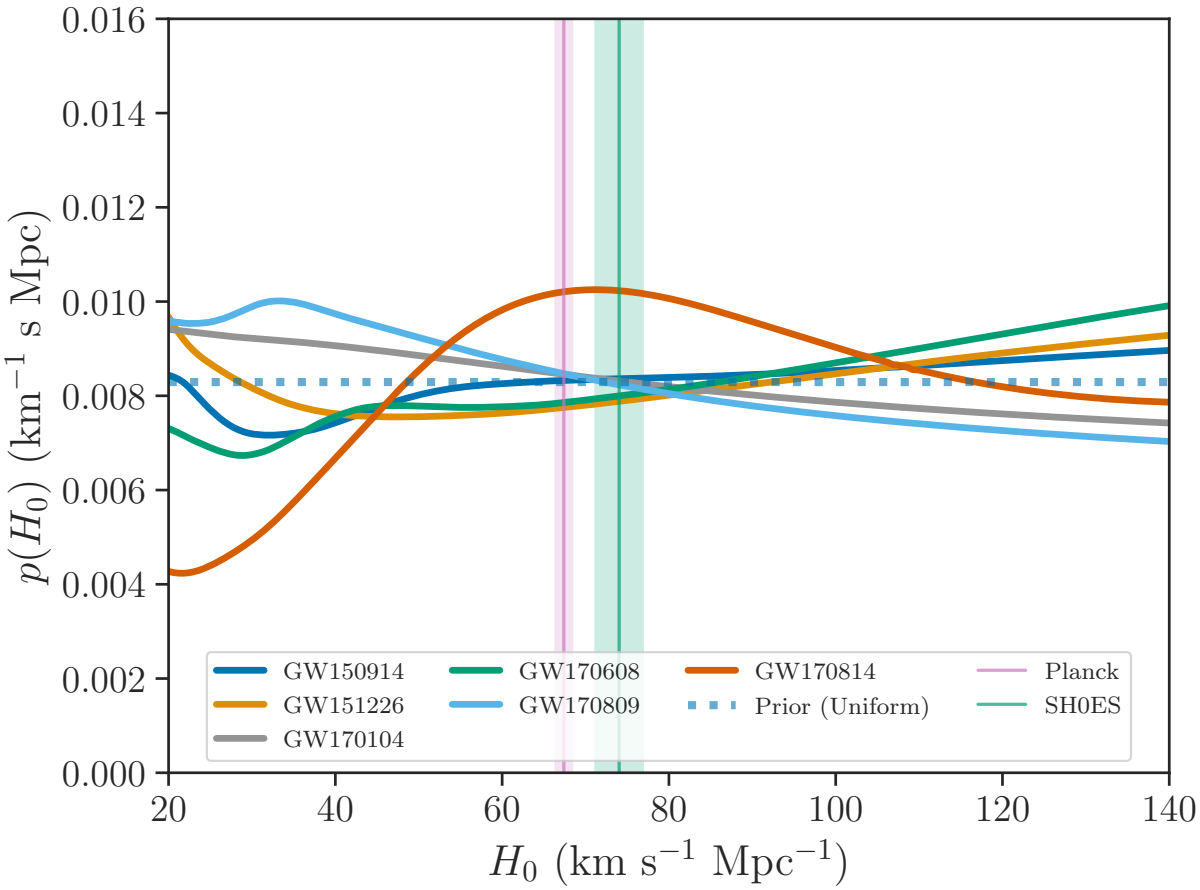


Figure 3.2: Individual estimates of H_0 from the six binary black hole detections which satisfy the selection criterion of network SNR > 12 in at least one search pipeline. These results assume a $m^{-1.6}$ power-law distribution on masses and a non-evolving rate model. All results assume a prior on H_0 uniform in the interval $[20, 140]$ km s^{-1} Mpc^{-1}(dotted blue). We also show the estimates of H_0 from CMB (Planck: Planck Collaboration, 2018) and supernova observations (SH0ES: Riess et al., 2019). Reproduced with permission from LVC (2019).

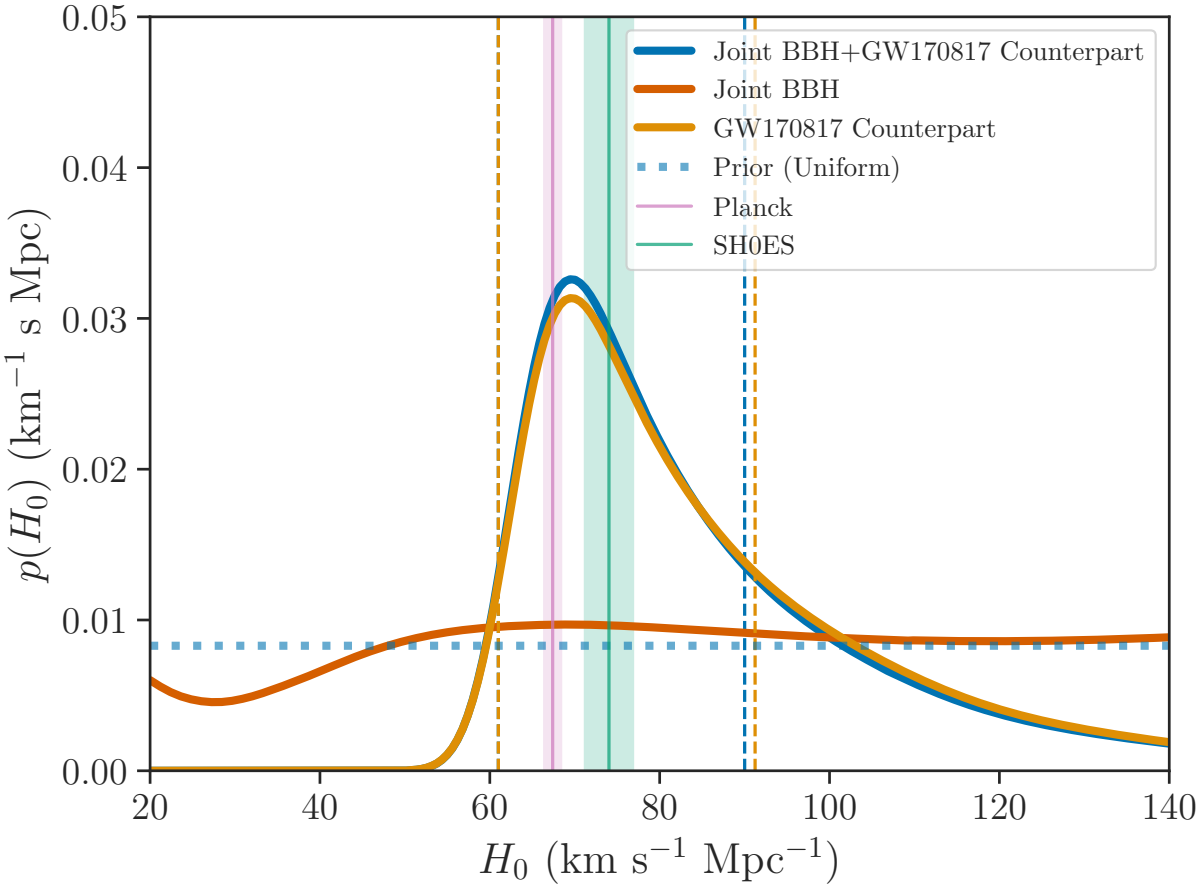


Figure 3.3: The gravitational-wave measurement of H_0 (dark blue) from the detections in the first two observing runs of Advanced LIGO and Virgo. The GW170817 estimate (orange) comes from the identification of its host galaxy NGC4993 (Abbott et al., 2017a). The additional contribution comes from binary black holes in association with appropriate galaxy catalogs; for GW170814 we use the DES-Y1 galaxy catalog, while for the remaining five BBHs, GW150914, GW151226, GW170104, GW170608, and GW170809, we use the GLADE catalog. The 68% maximum a-posteriori intervals are indicated with the vertical dashed lines. All results assume a prior on H_0 uniform in the interval $[20, 140]$ km s^{-1} Mpc^{-1}(dotted blue). We also show the estimates of H_0 from CMB (Planck: Planck Collaboration, 2018) and supernova observations (SH0ES: Riess et al., 2019). Reproduced with permission from LVC (2019).

3.4 SYSTEMATIC EFFECTS

3.4.1 Mass Distribution

We test the sensitivity to our assumptions regarding the assumed BBH mass distribution. In addition to the power-law mass distribution with $\alpha = 1.6$ (median inferred value using Model B of Abbott et al., 2019a), we choose a shallower flat-in-log mass distribution with $\alpha = 1$, and a steeper distribution with $\alpha = 2.3$ (both within the support of the inferred range in Abbott et al., 2019a). Our results are shown in the left panel of Fig. 3.4, and they demonstrate that the systematic differences due to the choice of power-law slope α are insignificant.

As a test case, we vary the upper cut-off for the mass distribution, $M_{\rm max}$. For our default analysis, $M_{\rm max}$ was chosen to be $100\,M_\odot$, consistent with all the considered BBHs for all values of H_0 within the prior range. Reducing this cut-off to a slightly restrictive $M_{\rm max} = 50\,M_\odot$ (e.g., motivated by the pair instability supernova process, Fowler & Hoyle, 1964), we see a significant difference (right panel of Fig. 3.4). Lowering $M_{\rm max}$ corresponds to a closer GW detection horizon. This systematically leads each event to prefer slightly lower values of H_0 than in the main result, for the reasons outlined in Section 3.3 – namely the relationship between the predicted event distribution (from our GW selection effects) and the real detected event distribution.

3.4.2 Merger Rate Evolution

Next, we relax our assumption on the evolution of the rate of binary mergers with redshift. A constant merger rate density, $R(z) =$ constant, implicit in the previous treatment, assumed that the merger rate traces the comoving volume. In addition, we repeat our analysis using a merger rate $R(z) \propto (1+z)^3$, which traces the star formation rate at low redshifts ($z < 2.5$) (Saunders et al., 1990), as well as a merger rate $R(z) \propto (1+z)^{-3}$ which could arise if typical delay times are very long, as may be expected from the chemically homogeneous evolution formation channel (Mandel

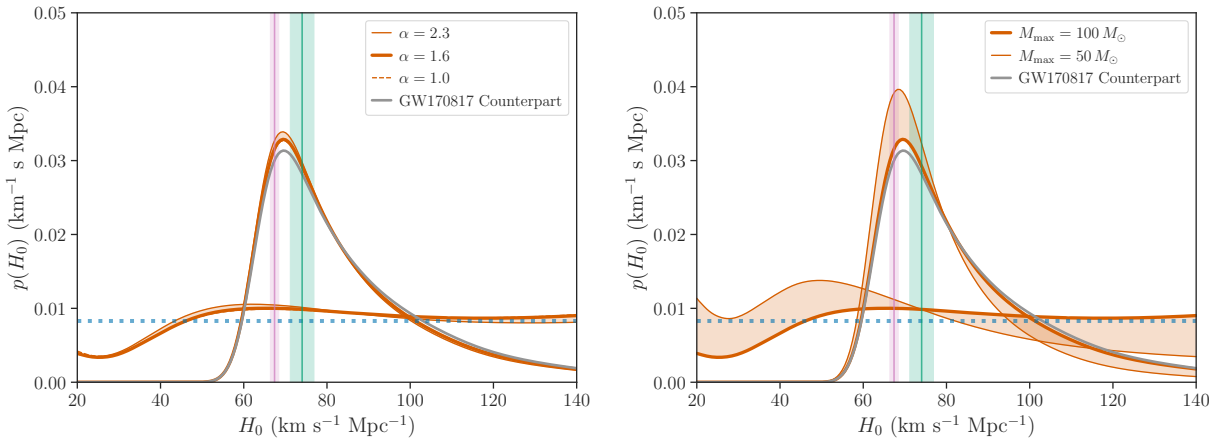


Figure 3.4: Sensitivity of the results to the assumed mass distribution model. Left panel: Variation of the results with three different choices of the power-law index for the mass distribution, $\alpha = 1.6$ (thick solid), $\alpha = 2.3$ (thin solid) and $\alpha = 1$ (thin dashed) assuming a constant intrinsic astrophysical merger rate, $R(z) =$ constant and $M_{\rm max} = 100\,M_\odot$. Right panel: Variation of the results with two different choices for the allowed black hole maximum mass, $M_{\rm max} = 100\,M_\odot$ (thick solid), and $M_{\rm max} = 50\,M_\odot$ (thin solid), both assuming $R(z) =$ constant and $\alpha = 1.6$.

& de Mink, 2016). These relaxed assumptions thus cover a large fraction of physically viable and inferred population models (Abbott et al., 2019a). We show our results for the different assumed redshift evolution models in Fig. 3.5. The model in which the merger rate traces star formation shows a significant difference, with a tendency to prefer lower values of H_0, compared to the other two models which are quite similar. This is the only systematic effect that leads to a significant difference in the results at this time.

3.4.3 Luminosity weighting

The results in the previous section assumed a weighting of galaxies by their luminosities in the B-band, which are indicative of galaxies' star formation rates. In order to quantify the difference likely to be caused by alternate ways of weighting the galaxies, we repeat our analysis with no luminosity weighting. These results are shown in Fig. 3.6.

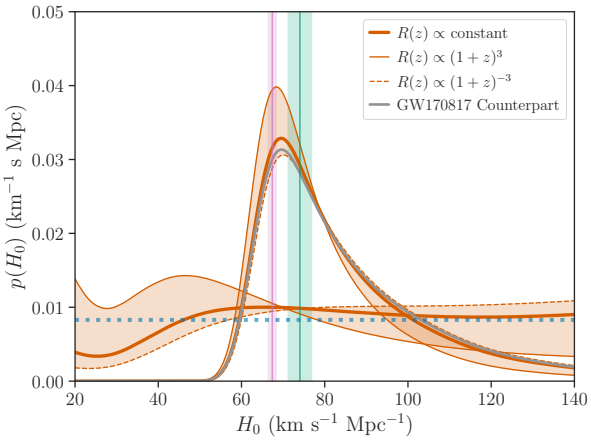


Figure 3.5: Variation of the results with two different choices for the rate evolution, $R(z) =$ constant (thick solid), and $R(z) \propto (1+z)^3$ (thin solid) and $R(z) \propto (1+z)^{-3}$ (thin dashed) for $\alpha = 1.6$ and $M_{max} = 100 M_\odot$.

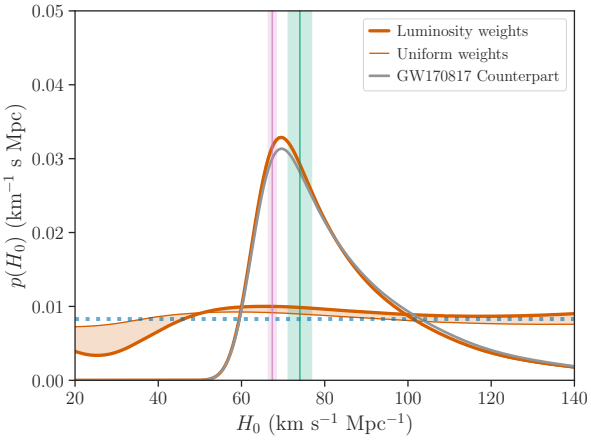


Figure 3.6: Sensitivity of the results to luminosity weighting. We show how the results vary when we weight the galaxies in the catalog by their B(g)-band luminosity (thick solid) as well as with constant (uniform) weights (thin solid), both assuming a power-law index for the mass distribution, $\alpha = 1.6$ and constant intrinsic astrophysical merger rate, $R(z) =$ constant.

With uniform luminosity weights, we obtain a result on a joint binary black hole estimate which is close to flat (thin orange line in Fig. 3.6). This can be understood as follows: 1) The out-of-catalog terms in Eq. (2.10) take into account the lack of galaxies beyond the apparent magnitude threshold $m_{\rm th}$ of the catalog in a uniform way. When galaxies are unweighted, the probability of the host galaxy being inside the catalog is reduced compared to the luminosity-weighted case. More weight is given to the uniform out-of-catalog contribution, and the events become less informative. 2) The catalogs used in this analysis contain high numbers of low-luminosity galaxies. The contribution from these more evenly distributed dim galaxies, in the unweighted case, again reduces the informativeness of each event and flattens the final result. This is also in agreement with our expectations from Fishbach et al. (2019a) and Gray et al. (2019), where weighting by luminosities enhances the features in the posterior distribution coming from the galaxy catalog.

3.4.4 Selection criterion

We look into the sensitivity of our results to the GW selection criterion. We reduce the threshold on the network SNR from 12 chosen in the previous section to 11 in at least one search pipeline. The computation of the GW selection effects is adjusted accordingly. Fig. 3.7 shows the results with the two sets of assumptions. Reducing the SNR threshold to 11 introduces two additional events in our analysis, namely GW170818 and GW170823, neither of which have a significant in-catalog contribution. Differences are expected due to the fact that additional low-SNR events are introduced, and also because the individual likelihoods change slightly with a different SNR threshold used in the GW selection term. In the regime of a large number of events, these two effects are expected to cancel, provided that the additional low-SNR events do not have significant in-catalog support. For our result, this difference is not significant; however, the small variation is a reminder that we are still in the regime of a low number of events.

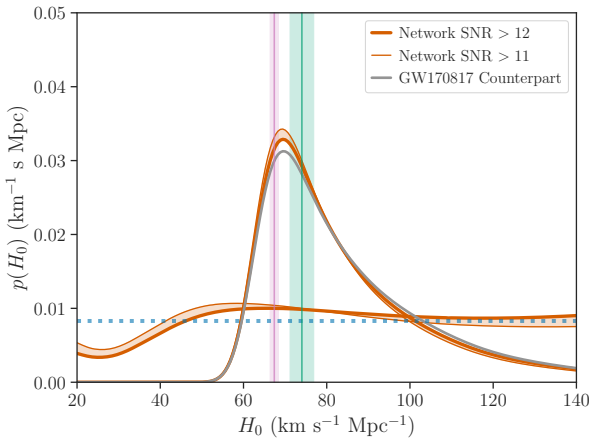


Figure 3.7: Sensitivity of the results to the selection criterion on the signal-to-noise ratio (SNR). The final joint result as well as the contribution from all the BBHs which satisfy the selection criterion are shown for a threshold network SNR of 12 (thick orange) and 11 (thin orange); the variation is shown as a shaded band. The GW170817 counterpart result (gray) is added to guide the eye. Six BBHs (GW150914, GW151226, GW170104, GW170608, GW170809, and GW170814) pass the selection criterion with SNR > 12. Two additional BBHs (GW170818 and GW170823) are included with SNR > 11. These results assume the default $m^{-1.6}$ power-law distribution on masses and a non-evolving rate model.

3.5 CONCLUSION AND OUTLOOK

In this work we have presented the first measurement of the Hubble constant using multiple GW observations. Our result reanalyzes and combines the posterior probability distribution obtained from the BNS event GW170817 using the redshift of the host galaxy inferred from the observed EM counterpart (Abbott et al., 2017a), along with constraints using galaxy catalogs for the BBH events observed by Advanced LIGO and Virgo in their first and second observing runs. We measure $H_0 = 68.7^{+17.0}_{-7.8}$ km s^{-1} Mpc^{-1} (68.3% highest density posterior interval with a flat-in-log prior). This result is mainly dominated by the information from GW170817 with its counterpart, but does show a modest improvement with the inclusion of the BBHs. The BBHs contribute both from associated galaxy catalogs as well as via their observed luminosity distance distribution. Since the latter contribution is sensitive to the assumptions on the mass distribution and rate evolution, a

more thorough treatment requires a marginalization over these unknown population parameters.

The contribution from events without counterparts is dominated by GW170814, for which the associated galaxy catalog is highly complete. This highlights the importance of deeper surveys and of dedicated EM follow-up of sky regions following GW triggers for a better H_0 measurement. With numerous anticipated detections in the upcoming observing runs with improved detector sensitivities (Aasi et al., 2015; Abbott et al., 2016b,a, 2017b, 2019b,a), these results pave the road towards an era of precision multimessenger cosmology to be performed with a multitude of sources, including both neutron star and black hole mergers, with or without transient EM counterparts.

Part II

Lensing

Chapter 4

Strong gravitational-wave lensing

Similar to electromagnetic waves, gravitational waves can be gravitationally lensed by intervening matter, resulting in multiple instances of the same astrophysical signal arriving at different times and with different apparent luminosity distances due to the magnifications induced by lensing. If left unaccounted for, these lensed gravitational wave signals can mimic mergers that appear louder than expected, with inferred systems appearing more massive and originating from closer distances than anticipated. This chapter presents a Bayesian statistical framework for identifying strongly-lensed gravitational wave signals, which incorporates astrophysical prior information and accounts for gravitational wave selection effects. The confirmation of lensed gravitational waves would provide a tool for probing the Universe at higher redshifts, enabling the study of cosmology and testing of gravity in previously inaccessible regimes.

4.1 INTRODUCTION

As gravitational waves propagate through the Universe, they can be deflected, or *lensed*, by intervening matter such as galaxies or galaxy clusters acting as gravitational lenses, just like electromagnetic waves. For visible light coming from a distant background source, in the case of *strong lensing* where the deflection is sufficiently large, multiple images of the source will be formed, typically separated by only several arcseconds and distorted compared to the unlensed image, see Figure 4.1. For transients, the time variation of the images are correlated and *delayed*, where the time delays can range from days to months depending on the mass of the lens (Schneider et al., 1992). For transient gravitational-wave (GW) signals, such as those emitted from the coalescence of compact binary systems, multiple images refer to multiple GW triggers registered at different

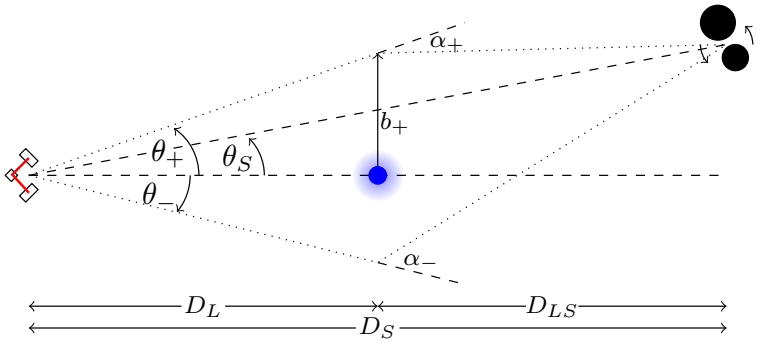


Figure 4.1: Diagram depicting the strong lensing scenario, with two strongly lensed GW images. The observer (detector configuration, left) receives emission from a GW source (compact binary, right), through a lens (blue circle, middle). Two strongly lensed images are formed, with angular positions θ_+ and θ_-, which are received and detected by the observer (from Ezquiaga et al. (2020)).

times by GW detectors. For an in-depth review of the strong lensing of explosive transients across the spectrum, see for example, Oguri (2019).

To investigate the effects of strong lensing to a GW signal, we first briefly describe how the observed GW strain $h(t)$ depends on some parameters θ when the signal is not affected by strong lensing. The GW strain observed by a ground-based GW detector is given by

$$h(t - t_{\rm c}; \theta) = \frac{1}{D_{\rm L}(z)} \sum_{\rm pol=+,\times} F_{\rm pol}(\alpha, \delta, \psi; t_{\rm c}) h_{\rm pol}(t - t_{\rm c}; \theta), \tag{4.1}$$

where the detected GW strain is a projection along the arms of the detector where the response of the detector to the two polarization states of the GW is defined by the detector's beam pattern functions $F_{+,\times}(\alpha, \delta, \psi)$, where α is the right ascension, δ is the declination, and ψ is the polarization angle of the source respectively. The luminosity distance to the source $D_{\rm L}(z)$ is a function of the redshift z and depends explicitly on cosmology. The time and phase at coalescence of the signal are denoted by $t_{\rm c}$ and $\phi_{\rm c}$ respectively.

For the case of GWs from a binary black hole (BBH) merger, the set of intrinsic parameters $\theta = \{\mathcal{M}_{\rm c}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2\}$ where $\mathcal{M}_{\rm c} \equiv (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ is the chirp mass of the binary,

$q \equiv m_2/m_1 \leq 1$ is its mass ratio, both in terms of the binary component masses $m_1 \geq m_2$. The vectors $\boldsymbol{\chi}_1$, $\boldsymbol{\chi}_2$ are the dimensionless spin vectors for the binary components. Note that this is just one particular parametrization, and other ways of specifying the intrinsic parameters are also possible, for example using the total mass of the binary $M_{\rm tot} \equiv m_1 + m_2$ instead of the chirp mass $\mathcal{M}_{\rm c}$. To account for the effect of expanding universe, one can do so by simply replacing the (source frame) masses m_i with the (detector frame) redshifted masses $m_i^{\rm det} = (1+z)m_i$. Similarly, we define the redshifted chirp mass as $\mathcal{M}_{\rm c}^{\rm det} = (1+z)\mathcal{M}_{\rm c}$, while the mass ratio remains unchanged.

Working in the geometric optics limit where the wavelength is much shorter than the lens length scale, for a majority of the time, strongly-lensed GW signals from a binary system will have the same morphology with different amplitudes (corresponding to different magnifications) and arrive at different times. A given image has an *absolute magnification* μ which can be defined in terms of the true luminosity distance to the source $D_{\rm L}$ and the observed apparent luminosity distances for each lensed images $D_{\rm L}^{(i)}$ as

$$D_{\rm L}^{(i)} = \frac{D_{\rm L}^{\rm src}}{\sqrt{\mu^{(i)}}}, \tag{4.2}$$

where the bracketed superscript indexes the images. The lensed images arrive at the (center of the) Earth at different times because of the geometrical time delay, as they follow different null geodesics with an additional time delay due to the gravitational potential. We define the relative time delay Δt between two images as $\Delta t \equiv t_{\rm c}^{(2)} - t_{\rm c}^{(1)}$, where $t_{\rm c}$ is the GW arrival (trigger) time for each image. We refer to the image that arrives first as the first image, and vice versa, such that $\Delta t > 0$. We can also define the relative magnification $\mu_{\rm rel}$, which is simply the ratio of the two absolute magnifications as

$$\mu_{\rm rel} \equiv \frac{\mu^{(2)}}{\mu^{(1)}}. \tag{4.3}$$

The strong lensing of GWs can also induce non-trivial effects on the GW waveform other than a change in the amplitude of the signal and a simple shift in time, such as a Hilbert transform

compared to the unlensed waveform. In general, the lensed waveform $\tilde{h}_{\rm pol}^{\rm lensed}(f)$ is related to the unlensed waveform $\tilde{h}_{\rm pol}^{\rm unlensed}(f)$ by a frequency-dependent amplification factor $F(f)$ as (Schneider et al., 1992; Takahashi & Nakamura, 2003)

$$\tilde{h}_{\rm pol}^{\rm lensed}(f) = F(f)\ \tilde{h}_{\rm pol}^{\rm unlensed}(f), \tag{4.4}$$

for $f > 0$. The negative frequency components can be obtained via the reality condition $\tilde{h}_{\rm pol}(-f) = \tilde{h}_{\rm pol}^{*}(f)$. In the geometric optics limit, the amplification factor is given by (Schneider et al., 1992; Takahashi & Nakamura, 2003)

$$F(f) = \sum_j \sqrt{\mu^{(j)}} \exp(2\pi i f \Delta t^{(j)} - i n^{(j)} \pi/2), \tag{4.5}$$

for $f > 0$ and $n^{(j)} \in \{0, 1, 2\}$ is known as the Morse index of the j-th image (Schneider et al., 1992). The factor $\sqrt{\mu}$ causes the apparent luminosity distance to differ from the luminosity distance of the source, and the phase factor $\exp(2\pi i f \Delta t)$ causes the aforementioned time delay. The phase shift from $\exp(-in\pi/2)$ is degenerate with a shift in the phase at coalescence when we consider GW signals from non-precessing binaries with contributions only from the quadrupole $\ell = |m| = 2$ modes (Dai & Venumadhav, 2017; Ezquiaga et al., 2020). The discrete shifts in the coalescence phase ϕ_c by an integer multiple of $\pi/4$ model the lensing Morse phase, which is related to the lensing phase shift $\Delta\phi = 2\Delta\phi_c$ (Dai & Venumadhav, 2017). To explore this effect, we examine all possible relative shifts $\Delta\phi = \{0, \pi/4, \pi/2, 3/pi/4\}$ between the two GW lensed candidates. The amplification factor expression becomes much more complicated than Eq. (4.5) and encodes additional gravitational lens information when the geometric optics approximation breaks down and the full wave optics treatment is required (Schneider et al., 1992; Takahashi & Nakamura, 2003).

Previous studies have shown that the detection rate for lensed GWs could be, optimistically,

5^{+5}_{-3} yr^{-1} (Ng et al., 2018) for Advanced LIGO (Aasi et al., 2015) and Advanced Virgo (Acernese et al., 2015a) operating at their design sensitivities, while others predicted more pessimistic rates, ranging from 0.58 yr^{-1} (Oguri, 2018) to 1.20 yr^{-1} (Li et al., 2018) depending on the source population model assumed. The detection rate for lensed GWs can also be constrained from the detection, and more surprisingly the non-detection, of stochastic gravitational-wave background from individually unresolvable binaries (Buscicchio et al., 2020; Mukherjee et al., 2021a).

Searches on the first observing run (O1) and the second observing run (O2) data for strongly-lensed GW signals were performed (Hannuksela et al., 2019; Dai et al., 2020; Liu et al., 2021), and it was concluded that there is no significant evidence that any of the eleven detected GW events during O1 and O2 are lensed, while Broadhurst et al. (2019) suggested that GW170809 and GW170814 could be lensed images of each other due to the similarity of the waveforms for these two events. We will discuss in more detail the results of Liu et al. (2021) in the remainder of this chapter.

4.2 JOINT PARAMETER ESTIMATION

Our goal is to determine whether two observed detector strains $d_1(t)$ and $d_2(t)$ with confirmed GW detections described by parameters θ_1 and θ_2 are lensed images of each other due to the same astrophysical system or independent GW detections. To accomplish this, we need to compare two hypotheses: the lensed hypothesis $\mathcal{H}_{\mathrm{L}}$, suggesting that the two signals are lensed images of the same astrophysical source, and the unlensed hypothesis $\mathcal{H}_{\mathrm{NL}}$, which assumes that the two signals come from independent astrophysical sources.

Assuming the lensed hypothesis, we first introduce a set of common parameters for the potentially lensed events, that is, we define $\theta = \{m_1, m_2, a_1, a_2, \iota, \alpha, \delta, \psi\}$. Here, m_1 and m_2 are the component source frame masses, a_1 and a_2 are the component spins, ι is the inclination angle of the binary, α is the right ascension, δ is the declination, and ψ is the polarization angle. We also

introduce lensing-dependent parameters $\gamma = \{D_L, \phi_c, t_c\}$, where D_L is the luminosity distance to the source, ϕ_c is the coalescence phase, and t_c is the time at coalescence. Therefore, under the lensed hypothesis, we expect the common parameters θ to be identical for the two events, with only the lensing-dependent parameters γ differing. Hence, we can express the likelihood under the lensed hypothesis as:

$$\mathcal{L}(d_1, d_2|\theta_1, \theta_2, \mathcal{H}_{\rm L}) = \mathcal{L}(d_1|\theta, \gamma_1, \mathcal{H}_{\rm L})\mathcal{L}(d_2|\theta, \gamma_2, \mathcal{H}_{\rm L}) \tag{4.6}$$

where, d_1, γ_1, and d_2, γ_2 are the data and independent parameters for the first and second images respectively. We will refer to the above likelihood as the strong lensing joint likelihood.

To determine whether a pair of GW events is lensed, we first consider the unlensed hypothesis $\mathcal{H}_{\rm NL}$, where the parameters of the two events are independent of each other. The likelihood under $\mathcal{H}_{\rm NL}$ is the product of the likelihoods of the two events, given by:

$$\mathcal{L}(d_1, d_2|\theta_1, \theta_2, \mathcal{H}_{\rm NL}) = \mathcal{L}(d_1|\theta_1, \mathcal{H}_{\rm NL})\mathcal{L}(d_2|\theta_2, \mathcal{H}_{\rm NL}) \tag{4.7}$$

Here, θ_1 and θ_2 denote the parameters for the first and second GW events, respectively.

To compare the two hypotheses, we compute the lensed and unlensed evidences, denoted by Z_L and Z_U, respectively:

$$Z_L = \int \mathcal{L}(d_1|\theta, \gamma_1, \mathcal{H}_{\rm L})\mathcal{L}(d_2|\theta, \gamma_2, \mathcal{H}_{\rm L})p(\theta, \gamma_1, \gamma_2|\mathcal{H}_{\rm L})d\theta d\gamma_1 d\gamma_2 \tag{4.8}$$

and,

$$Z_U = \int \mathcal{L}(d_1|\theta_1, \mathcal{H}_{\rm NL})\mathcal{L}(d_2|\theta_2, \mathcal{H}_{\rm NL})p(\theta_1, \theta_2|\mathcal{H}_{\rm NL})d\theta_1 d\theta_2 \tag{4.9}$$

Here, $p(\theta, \gamma_1, \gamma_2|\mathcal{H}_{\rm L})$ and $p(\theta_1, \theta_2|\mathcal{H}_{\rm NL})$ are the prior distributions under the lensed and unlensed hypothesis respectively. The ratio of the evidences, also known as the strong lensing Bayes factor,

is then computed as:

$$B_U^L = \frac{P(d_1, d_2|\mathcal{H}_{\rm L})}{P(d_1, d_2|\mathcal{H}_{\rm NL})} = \frac{Z_L}{Z_U}. \tag{4.10}$$

This ratio measures the relative support for the strong lensing hypothesis versus the unlensed hypothesis.

However, to subject the strong lensing hypothesis $\mathcal{H}_{\rm L}$ to a more stringent test, we impose a condition on the strong lensing Bayes factor. Specifically, we consider the fact that for a pair of lensed GW images, there should be a relationship between the differences in the GW phases at coalescence, as described in Eq. (4.5). The Morse index, denoted by n_j, determines the type of lensing image (type-I/II/III) and the corresponding expected phase shift. We can introduce a phase difference for the pair, denoted by $\Delta\phi$, by conditioning the prior under $\mathcal{H}_{\rm L}$ as $p(\theta, \gamma_1, \gamma_2|\mathcal{H}_{\rm L}, \Delta\phi)$. Accordingly, we compute $B_U^L|_{\Delta\phi}$ instead.

The time delay between any two events can also be used to compute a corresponding timing Bayes factor (Haris et al., 2018; Hannuksela et al., 2019),

$$B_t = \frac{P(\Delta t|\mathcal{H}_{\rm L})}{P(\Delta t|\mathcal{H}_{\rm NL})} \tag{4.11}$$

We estimate the probability distribution $P(\Delta t|\mathcal{H}_{\rm L})$ through simulation following the methodology of (Haris et al., 2018). We compute $P(\Delta t|\mathcal{H}_{\rm NL})$, by assuming that independent (unlensed) events are Poisson distributed. To obtain the odds ratio for the lensed and unlensed hypothesis we compute,

$$O_U^L|n_j = \frac{P(\mathcal{H}_{\rm L})}{P(\mathcal{H}_{\rm NL})} B_U^L|_{\Delta\phi} B_t \tag{4.12}$$

where the ratio $P(\mathcal{H}_{\rm L})/P(\mathcal{H}_{\rm NL})$ is the prior odds for lensing compared to the unlensed event model.

We estimate the prior odds based on our belief in the probability of lensing for any two events, which can be inferred from simulations and electromagnetic observations. This is computed as the

ratio of the expected lensed event rate to the independent event rate. For galaxy lenses, the relative lensed event rate at design sensitivity has been estimated to be around $p(\mathcal{H}_{\rm L})/p(\mathcal{H}_{\rm NL}) \approx 10^{-3}$ (Li et al., 2018; Oguri, 2018), while for galaxy cluster lensing, Smith et al. (2018) finds the relative rate to be $p(\mathcal{H}_{\rm L}) \approx 10^{-5}$ yr^{-1} at O1 sensitivity, which we expect to be somewhat higher at O2 sensitivity.

4.3 RESULTS

We examine potential pairs of lensed events from the second observing run of Advanced LIGO and Virgo. To reduce the computational cost of the parameter estimation, we select pairs of events with similar sky localizations. We then use `LALInference` (LIGO Scientific Collaboration, 2018) to estimate the evidences for both the lensed and unlensed models.

In our analysis, we use the `IMRPhenomD` waveform approximant (Husa et al., 2016; Khan et al., 2016), a non-precessing and spin-aligned (22-mode only) frequency-domain BBH waveform. This waveform allows us to test different coalescence phase shifts caused by various image types. Since there is no evidence of precession in any of the events detected in O1 and O2 (Abbott et al., 2019), we expect that the IMRPhenomD model is suitable for this analysis.

In a Bayesian analysis, a more complex model is penalized by the Bayesian evidence. Therefore, when comparing two models, the one with fewer parameters or a smaller prior volume is preferred. This Occam factor is automatically accounted for in Bayesian inference and is well-established in the literature (Thrane & Talbot, 2019). Our lensing model, which shares parameters, has fewer parameters than the alternative model in our analysis. The magnitude of the Bayes factor is determined by the prior volume, particularly when the signal is consistent with the lensing hypothesis. A larger prior volume can increase the Bayes factor by orders of magnitude, while a smaller prior volume reduces it. To reduce the difference in prior volumes between the two models, we use a uniform prior in $\log_{10}(m_1)$ and $\log_{10}(m_2)$, instead of the typical uniform priors on m_1

and m_2 (in the detector frame). We apply the prior within the mass range of $1 - 100M_{\odot}$. Using a uniform prior in log space can reduce the difference in prior volumes by a factor of $10^2 - 10^3$. We acknowledge that an astrophysically motivated mass prior, such as the power-law model used by the LVC (Abbott et al., 2019a), could be used. However, we opted for the uniform in log prior due to the reasons explained above.

To account for selection effects in our joint parameter estimation, we consider that gravitational-wave detectors are not capable of detecting all binaries in the prior parameter space. Therefore, we incorporate a selection function directly into the parameter estimation process. This selection function keeps a sample if it is above the detection SNR threshold, and rejects it otherwise. We emphasize that considering selection effects is particularly crucial when one or both events are below the detection threshold (Li et al., 2019; McIsaac et al., 2020). This consideration is particularly important when performing joint parameter estimation using the potential third image, GWC170620 (Dai et al., 2020). We set the network SNR threshold to 10 for single events and 14 for joint events, except for the sub-threshold event GWC170620.

Figure 4.2 illustrates the Bayes factors $\log_{10} B^L_U|_{\Delta\phi}$ and $\log_{10} B_t$ for pairs of events detected during O2. We evaluate the lensed model for each pair of events using four different possible coalescence phase shifts, while for the unlensed model, we independently sample the phases for each event. The pair GW170104-GW170814, with a coalescence phase shift of $\pi/2$, has the largest Bayes factor $\log_{10} B^L_U|_{\Delta\phi} \approx 4.3$, favoring the lensed hypothesis in the absence of prior probability. Nonetheless, a high Bayes factor is not necessarily indicative of lensing, as the two events could be from independent sources with similar parameters. It is intriguing, however, that the event still favors the lensing hypothesis even when taking all of the binary parameters into account. We note that the GW170809-GW170814 pair, previously suggested as a lensed event by Broadhurst et al. (2019), is clearly disfavored by the model selection.

Figure 4.4 displays the sky localization posterior for the GW170104-GW170814 pair. The joint

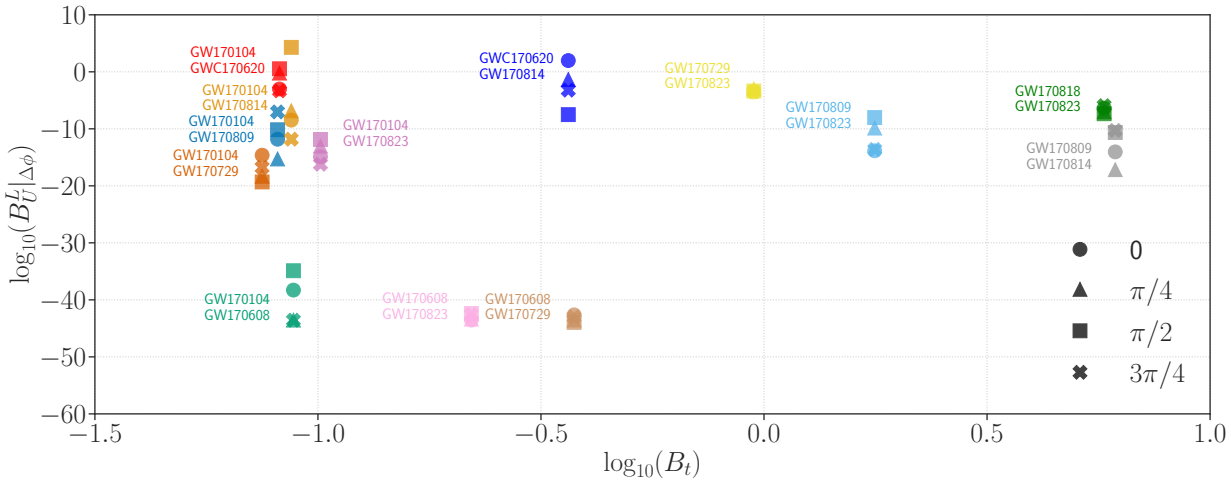


Figure 4.2: The logarithm of the strong lensing Bayes factor $\log_{10} B^L_U|\Delta\phi$ and the time delay Bayes factor $\log 10 B_t$ are shown for pairs of events detected and analyzed in O2, with coalescence phase shifts of 0, $\pi/4$, $\pi/2$, and $3\pi/4$. The pair GW170104-GW170814, with a coalescence phase shift of $\pi/2$, has the largest Bayes factor $\log_{10} B^L_U|_{\Delta\phi} \approx 4.3$, favoring the lensed hypothesis in the absence of prior probability. The time delay between GW170104 and GW170814 is approximately 7 months, the resulting timing Bayes factor is, therefore, $\log_{10} B_t \approx -1.1$.

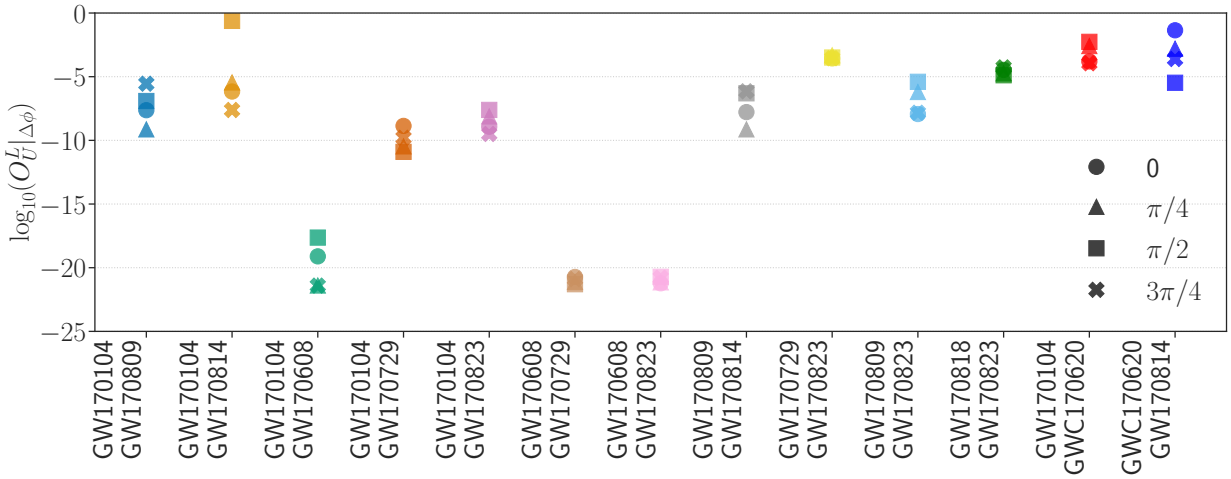


Figure 4.3: The logarithm of the odds ratio $\log_{10} O^L_U|_{\Delta\phi}$ with 0, $\pi/4$, $\pi/2$, and $3\pi/4$ coalescence phase shifts for pairs of events detected in O2. The odds ratios are computed using Eq. 4.12. The only pair of events that moderately prefers the lensing hypothesis, even with prior information included, is GW170104 and GW170814 with a coalescence phase shift of $\pi/2$, with an odds ratio of $\log_{10} O^L_U|_{\Delta\phi} \sim 1.3$.

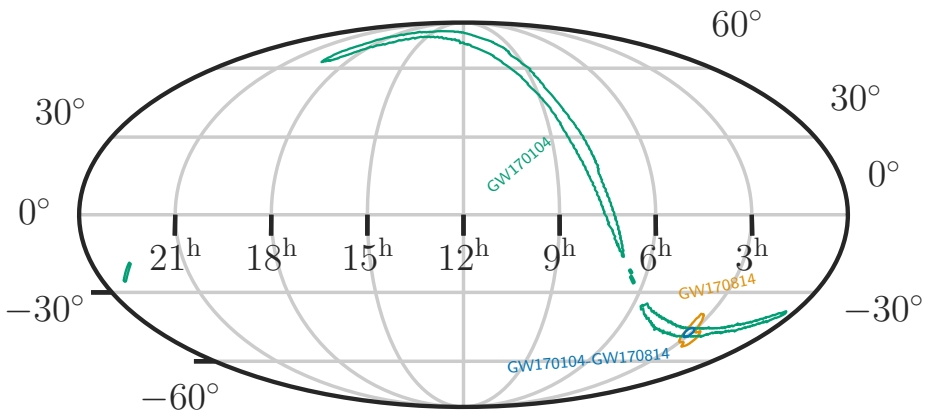


Figure 4.4: Sky localizations posteriors of GW170104, GW170814 (treated as independent events) as well as the joint GW170104-GW170814 sky localization posterior inferred under the lensing hypothesis with a coalescence phase shift of $\Delta\phi = \pi/2$. The contours show 90% confidence regions for all the sky localization posteriors shown.

parameter estimation under the lensing hypothesis provides a better-constrained 90 percent confidence region, which is located within the overlap region of the independent parameter estimation runs of GW170104 and GW170814. This improvement is due to the higher SNR in the joint run and the additional detector baselines for localization resulting from the rotation of the Earth and the different times of arrival of each image. The posterior distributions for parameters that are expected to remain unchanged under lensing are shown in Fig. 4.5, and similarly, the joint parameter estimation under the lensing hypothesis yields better constrained parameter estimates compared to the independent runs.

Based on the simulations in Haris et al. (2018) and assuming a 9-month observation time for O2, we estimate $P(\Delta t|L)$. For the unlensed case, we adopt a Poisson distribution for the detected event rate, with $P(\Delta t|U) = 2(T - \Delta t)/T^2$, where T is the observation time. In Fig. 4.2, we present the Bayes factor $\log_{10} B_t$ for the lensing model with four different phase shifts, compared to the unlensed model. As the time delay between GW170104 and GW170814 is approximately

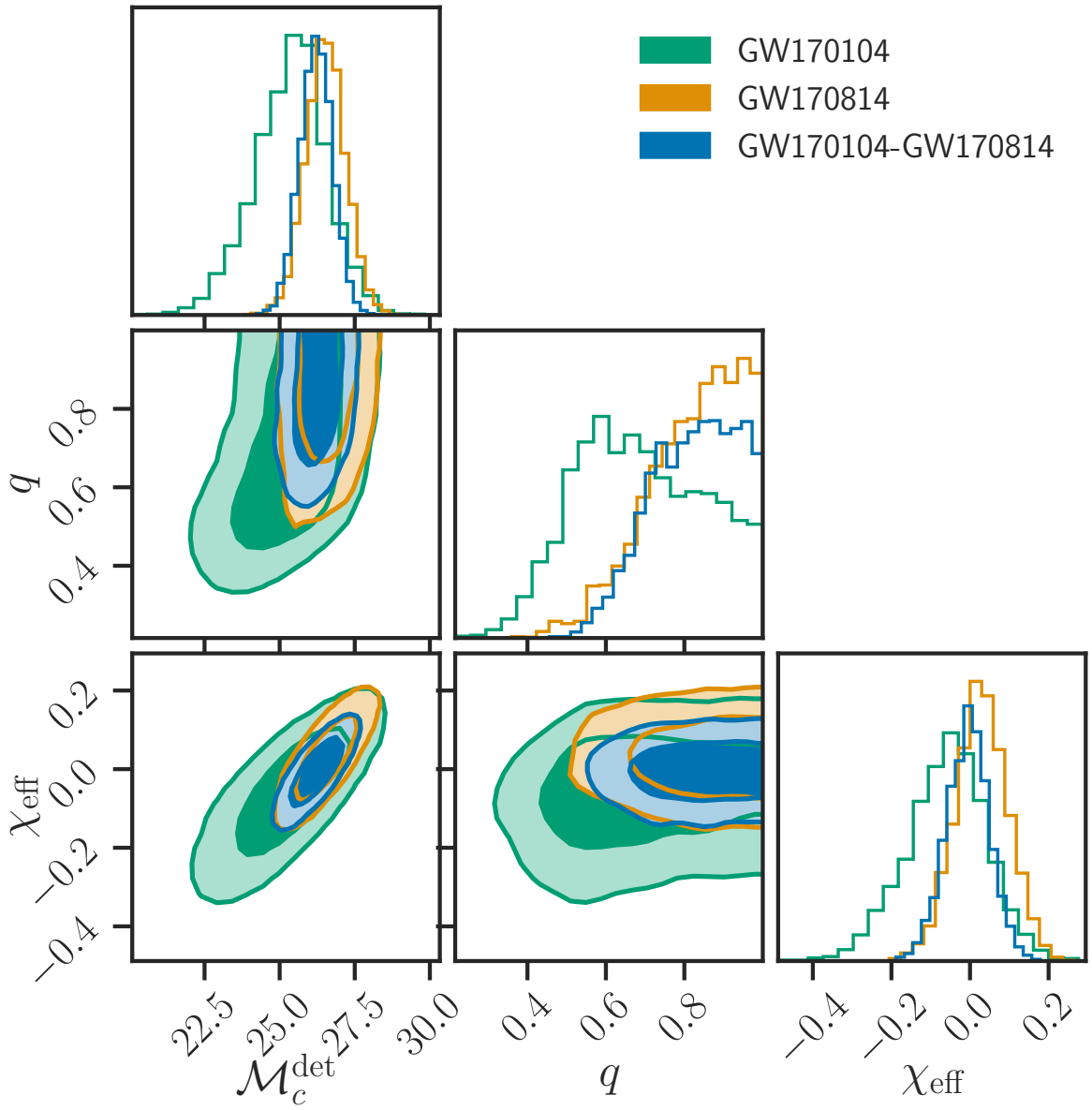


Figure 4.5: Corner plot showing the posterior distributions of the detector-frame chirp mass $\mathcal{M}_c^{\rm det}$, mass ratio $q = m_2/m_1$, and the effective spin parameter χ_{eff}. We show the 68% and 98% credible regions for the independent GW170104, and GW170814 posteriors as well the jointly inferred posteriors for the GW170104-GW170814 pair under the lensing hypothesis (with a coalescence phase shift of $\pi/2$).

7 months, the resulting Bayes factor is $\log_{10} B_t \approx -1.1$. However, for pairs with a time delay of only a few days, such as GW170809-GW170814 and GW170818-GW170823, the Bayes factor can be as large as $\log_{10} B_t \approx 0.78$.

The prior odds for lensing, which can be combined with the Bayes factors $\log_{10} B^L_U|_{\Delta\phi}$ and $\log_{10} B_t$ to compute the odds ratio (Eq. 4.12), are estimated to be around 0.009, based on the eight events detected during O2 in nine months of observing time (Ng et al., 2018; Hannuksela et al., 2019). The lensed event rate during O2 is estimated to be about 10^{-1} yr^{-1}. In Fig. 4.3, the odds ratios for events detected in O2 are shown. The only pair of events that moderately prefers the lensing hypothesis, even with prior information included, is GW170104 and GW170814 with a coalescence phase shift of $\pi/2$, with an odds ratio of $\log_{10} O^L_U|_{\Delta\phi} \approx 1.3$. For other pairs detected during O2, we do not see any evidence of lensing, as the odds ratios are much less than 1.

Our estimations for $\log_{10} B_t$ and prior odds are based on the assumption of galaxy lensing. However, it is possible that the long time delay observed in the GW170104-GW170814 pair is due to lensing by a galaxy cluster, which is expected to be rare at the sensitivity level of O2. The prior probability of galaxy cluster lensing is estimated to be 10^{-5} yr^{-1} (Smith et al., 2018). Furthermore, the fact that the phase shift corresponds to π, or a type-III image as pointed out in Dai et al. (2020), further disfavors the lensing hypothesis. The probability of observing type-III images is expected to be very low, and indeed, they are rarely observed in the electromagnetic band.

4.4 DISCUSSION AND CONCLUSIONS

To utilize Bayesian model selection for identifying lensed events in the second observing run of Advanced LIGO and Virgo, we conducted tests on the lensing model with coalescence phase shifts of 0, $\pi/4$, $\pi/2$, and $3\pi/4$, while accounting for selection effects. The most significant event pair, GW170104-GW170814, displayed a very high Bayes factor. However, due to their similar parameters and possible independent sources, it remains unclear whether lensing is involved. Additionally,

based on our understanding of lens configurations, BBH, and lens populations, this event is highly disfavored as a lensed candidate. Indeed, observing lensing configurations with these time delays and image configurations is improbable for both galaxies and galaxy clusters. While the estimates of lensed rates and the relative contributions of galaxies and galaxy clusters may vary, current predictions suggest that galaxy cluster lensing should not become prominent at O2 sensitivity.

Assuming that the event was lensed, it would require extraordinary evidence to support such a claim. We must consider several factors, including the relative contribution of galaxy cluster lensing, the merger rate density of binary black holes, and the prominence of type-III images for gravitational-wave sources. Other studies (Smith et al., 2018, 2017, 2019; Robertson et al., 2020) have studied galaxy cluster lensing and argued that highly magnified events have historically been observed more prominently lensed by galaxy cluster scale lenses. However, current estimates suggest that galaxy cluster lensing should not become observable at O2 sensitivity (Oguri, 2018).

It is also crucial to note that the Belczynski distribution, which is often used to model the merger-rate density, predicts merger rates that are not high enough to observe galaxy cluster lensing at O2 sensitivity (Belczynski et al., 2016). Additionally, type-III images are more prominent for gravitational-wave sources than they are for electromagnetic sources.

In the absence of clear-cut evidence to the contrary, we must conclude that there is no sufficient evidence to claim that the event is lensed, in agreement with (Dai et al., 2020) and (Hannuksela et al., 2019). Extraordinary claims require extraordinary evidence, and based on the prior probability of lensing by these types of systems, we advise the reader to be cautious in interpreting the results.

However, there are two pieces of evidence that could potentially determine if the events were lensed. First, injection campaigns must be performed to determine the probability of a non-lensed event. A similar study was conducted in (Dai et al., 2020) with a false alarm probability between 10^{-4} and 10^{-2} for O2 events. Secondly, cross-verification through searches in the electromagnetic

channels could be an intriguing possibility. If the events were lensed, then their host galaxy must also be lensed (Hannuksela et al., 2020). Ref. (Hannuksela et al., 2020) demonstrated that such a search is possible for galaxies, and the search is expected to be even more powerful for galaxy clusters due to their rarity. In the case of doubly lensed events such as the GW170104-GW170814 pair, the single time-delay estimate may be quite degenerate with the lens parameters and the source alignment.

Chapter 5

Strong Lensing Joint Parameter Estimation with `HANABI`

5.1 INTRODUCTION

An overarching approach in searches for lensed GW signals is the use of Bayesian statistics (Haris et al., 2018; **?**; McIsaac et al., 2020; Liu et al., 2021; Dai et al., 2020), where a statistic, either called a "Bayes factor" in the usual context of Bayesian hypothesis testing or a ranking score, is calculated. In this paper, we present a Bayesian statistical framework for identifying strongly-lensed GW signals that utilizes hierarchical Bayesian modeling. By modeling the data generation processes when the observed GW signals are lensed and not lensed respectively, we develop a framework that allows us to compute a Bayes factor, and hence a posterior odds, that incorporates astrophysical information directly and accounts for selection effects. We argue that in order to interpret the Bayes factor properly as a ratio of normalized probability densities of the observed data, selection effects cannot be ignored and must be accounted for in order to normalize the probability densities. The ability to directly incorporate astrophysical information, both on the GW sources as well as the gravitational lenses, serves to better bridge the astrophysical modeling community and the GW data analysis community. In addition, we argue that whether a GW signal is lensed or not depends also on the astrophysical models assumed, making the prior astrophysical information an indispensable ingredient of the analysis.

This chapter is structured as follows: Sec. 5.2 presents the hierarchical Bayesian framework for identifying strongly-lensed GW signals in a general setting, and the technique to marginalize over the source redshift separately and infer the unbiased source parameters. In Sec. 5.3 we apply and showcase the statistical framework to analyze strongly-lensed GW signals from BBH mergers

when we analyze two GW signals jointly. Throughout the paper, we assume a flat ΛCDM cosmology with $H_0 = 67.7\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_\mathrm{m} = 0.307$ from the Planck 2015 results (Ade et al., 2016a).

5.2 STATISTICAL FRAMEWORK

In order to differentiate strongly-lensed GW signals from GW signals that are not lensed, we adopt a Bayesian statistical framework where we introduce two models/hypotheses[1] that we want to compare, namely the lensed hypothesis $\mathcal{H}_\mathrm{L}$ and the not-lensed hypothesis $\mathcal{H}_\mathrm{NL}$.[2] In the following, we will elaborate on what we mean by the two hypotheses, and model the data generation processes under the two different hypotheses using hierarchical Bayesian modeling.

5.2.1 The not-lensed hypothesis $\mathcal{H}_\mathrm{NL}$

Suppose we have N distinct GW events under consideration, i.e. we have a set of N time series data $D = \{D^{(i)}\}_{i=1}^{i=N}$, where the bracketed superscript indexes the events. The not-lensed hypothesis means that the observed N events are N *independent realizations* of a population distribution of GW source p_src, parametrized by some parameters λ that *control only the shape of the distribution* and the total number of sources N_src in that population.[3] Note that we have assumed all N of them are of astrophysical origins. Simply put, the N events are just N different systems, with the event-level parameters $\theta^{(i)}$ (such as component masses and spins) describing the i-th event being randomly drawn from a source population distribution $p_\mathrm{src}(\theta|\lambda)$, where λ might be for example the maximum mass of a black hole in that population. These N signals will have different source redshifts z drawn from the distribution $p_z(z^{(i)}|\mathcal{R})$, where $\mathcal{R}$ is the merger rate density, and with different extrinsic parameters such as the sky location drawn from the distribution p_ext. A concise

[1] We will use the word model and hypothesis interchangeably throughout the paper.

[2] The not-lensed hypothesis $\mathcal{H}_\mathrm{NL}$ is often referred as the unlensed hypothesis, denoted by $\mathcal{H}_\mathrm{U}$, in literature. Here we reserve the meaning of *unlensed* for effects due to lensing being reverted.

[3] Note that the source population distribution $p_\mathrm{src}(\theta|\lambda)$ is normalized such that $\mathrm{d}N_\mathrm{src}/\mathrm{d}\theta = 1/N_\mathrm{src}\ p_\mathrm{src}(\theta|\lambda)$.

way of expressing this is that $\theta^{(i)} \sim p_{\rm pop}(\theta^{(i)})$ where $p_{\rm pop} = p_{\rm src}(\theta^{(i)}|\lambda)p_z(z^{(i)}|\mathcal{R}, \mathcal{H}_{\rm NL})p_{\rm ext}$ is the population-informed prior distribution under the not-lensed hypothesis. The event-level parameters $\theta^{(i)}$ then in turn "generate" the data $D^{(i)}$ that we observed for the i-th event. Figure 5.1 shows a graphical representation of this data generation process. Although we are *not making any inference* on the population-level parameters of GW sources and instead we *fix them* in our analysis (i.e. choosing λ and $\mathcal{R}$ a priori), we see that the problem of identifying strongly-lensed signals can be naturally framed as a population analysis. Moreover, we can re-use a lot of the results from the usual rates and population analysis.

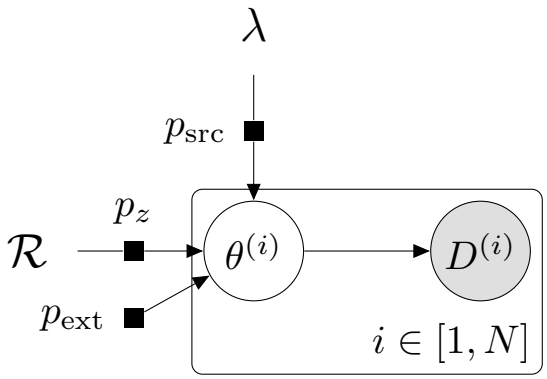


Figure 5.1: Data generation process for the N observed data under the not-lensed hypothesis $\mathcal{H}_{\rm NL}$. Each data $D^{(i)}$ can be described by the event-level parameters $\theta^{(i)}$ which were drawn from the population-informed prior distribution $p_{\rm pop} = p_{\rm src}(\theta^{(i)}|\lambda)p_z(z^{(i)}|\mathcal{R}, \mathcal{H}_{\rm NL})p_{\rm ext}$ with λ controlling the shape of the source population distribution, $\mathcal{R}$ being the merger rate density, and $p_{\rm ext}$ describing the distribution of the extrinsic parameters except for the redshift.

5.2.2 The lensed hypothesis $\mathcal{H}_{\rm L}$

For the lensed hypothesis, suppose we also have the same N events under consideration. However, the lensed hypothesis means that these N events are actually N strongly-lensed images of the *same* source. Instead of drawing N independent realizations from the population distribution $p_{\rm src}$, now we only have *one* realization of this source population distribution as the images correspond to the

same GW source. In addition to the source population distribution, we will need to introduce the *lens population distribution* $p_{\rm lens}$, parametrized by some parameters γ, that describes for example the joint probability distribution of the absolute magnification of lensed images. Furthermore, we partition the event-level parameters $\theta^{(i)}$ into two disjoint sets: common parameters $\theta^{(i)}_{\rm com}$ and independent parameters $\theta^{(i)}_{\rm ind}$. For the common parameters $\theta^{(i)}_{\rm com}$ we expect them to be the same across the N signals, for example, the masses and spins of the source binary system, as the N events correspond to the same source. In addition to the source parameters, we also expect the redshift $z^{(i)}$ of each image to be the same as the source redshift as *strong lensing is achromatic*, leaving the redshift unchanged. For extrinsic parameters, we can also assume them to be the same except for the (apparent) luminosity distance and the time of arrival. While it is true that strong lensing will deflect a GW signal from its original null trajectory, the typical deflection angle for gravitational lensing due to a galaxy or a galaxy cluster is only of the order of arcseconds and arcminutes respectively (Schneider et al., 1992; Singer et al., 2019), which is much smaller than the typical uncertainty in the source localization of a GW signal. Therefore, it is valid to assume that the N images share the same sky location. We also expect the difference in the polarization angle ψ to be negligible (Hou et al., 2019). In summary, the common parameters $\theta^{(i)}_{\rm com}$ are one random draw of the distribution $p_{\rm pop,com} = p_{\rm src}(\theta^{(i)}_{\rm com}|\lambda)p_z(z^{(i)}|\mathcal{R},\mathcal{H}_{\rm L})p_{\rm ext}$, where $p_{\rm pop,com}$ is the population-informed prior for the common parameters $\theta_{\rm com}$ under the lensed hypothesis.

As for the independent parameters $\theta^{(i)}_{\rm ind}$, we expect them to be different for each event. For example, the absolute magnification μ and the arrival time $t_{\rm c}$ of each image would be different. Note that the dimension of the event-level parameters θ under the lensed hypothesis *can be different* than that under the not-lensed hypothesis. For example, different lensed images can be classified into three types where each type of an image will have a different phasing effect to the lensed waveform (for example see Eq. (4.5)). The number of lensed images produced by a gravitational lens can also inform us on the type of lens that produces the images. Here we do not use this information since

it is possible (and often the case) that we are only analyzing a subset of lensed images coming from a particular source and lens, either deliberately or simply because we did not observe all of the lensed images. In short, each image will take different values for the independent parameters $\theta_{\rm ind}^{(i)}$ where each of them is a random realization of the distribution $p_{\rm pop,ind} = p_{\rm lens}(\theta_{\rm ind}^{(i)}|\gamma)$.

Figure 5.2 shows a graphical representation of this data generation process. Again one should note that we are not making any inference on the population-level parameters of the GW sources and lenses. Instead we consider them as given in our analysis. Next, we will use our knowledge of the data generation processes under the two hypotheses to construct a statistic that would allow us to evaluate whether some GW signals are lensed or not.

5.2.3 Model comparison

The standard approach to perform a Bayesian model comparison is to compute the posterior odds $\mathcal{O}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}}$, which is defined as[4]

$$\mathcal{O}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}} = \frac{p(\mathcal{H}_{\rm L}|D,\lambda,\mathcal{R},\gamma)}{p(\mathcal{H}_{\rm NL}|D,\lambda,\mathcal{R})}. \tag{5.1}$$

Note that for both models we fix the population-level parameters and the merger rate density, therefore we will not write them out explicitly when there is no ambiguity. Using Bayes' theorem, we can easily re-write the posterior odds into a product of two terms, namely the Bayes factor and the prior odds as

$$\mathcal{O}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}} = \underbrace{\frac{p(D|\mathcal{H}_{\rm L})}{p(D|\mathcal{H}_{\rm NL})}}_{\text{Bayes factor } \mathcal{B}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}}} \underbrace{\frac{p(\mathcal{H}_{\rm L})}{p(\mathcal{H}_{\rm NL})}}_{\text{Prior odds } \mathcal{P}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}}}. \tag{5.2}$$

We first focus on getting an expression for evaluating the Bayes factor $\mathcal{B}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}}$ from the set of N observed data. And later we will discuss the evaluation of the prior odds $\mathcal{P}_{\mathcal{H}_{\rm NL}}^{\mathcal{H}_{\rm L}}$.

[4]We will abuse the notation and use p to denote both probability and probability density when the context is clear.

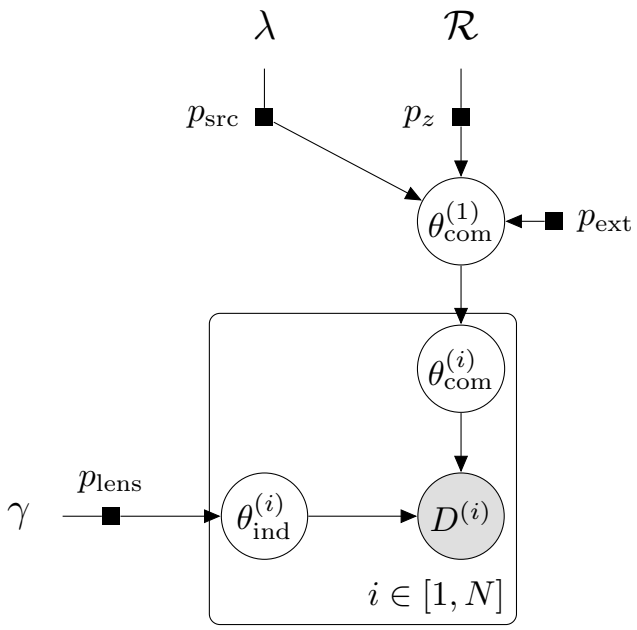


Figure 5.2: Data generation process for the N observed data under the lensed hypothesis $\mathcal{H}_{\rm L}$. Each data $D^{(i)}$ can be described by the event-level parameters $\theta^{(i)}$, which are partitioned into two disjoint sets: $\theta^{(i)}_{\rm com}$ which are assumed to be the same across the N signals and $\theta^{(i)}_{\rm ind}$ which can be different for each signal. Without loss of generality, we assume that $\theta^{(i)}_{\rm com} = \theta^{(1)}_{\rm com}$ in the graph. The common parameters $\theta^{(1)}_{\rm com}$ are one realization of the distribution $p_{\rm pop,com} = p_{\rm src}(\theta^{(i)}_{\rm com}|\lambda)p_z(z^{(i)}|\mathcal{R}, \mathcal{H}_{\rm L})p_{\rm ext}$, while the independent parameters $\theta^{(i)}_{\rm ind}$ for the N signals are N realizations of the distribution $p_{\rm pop,ind} = p_{\rm lens}(\theta^{(i)}_{\rm ind}|\gamma)$.

5.2.4 The Bayes factor $\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}}$

The Bayes factor $\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}}$, defined as

$$\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} = \frac{p(D|\mathcal{H}_{\rm L})}{p(D|\mathcal{H}_{\rm NL})}, \tag{5.3}$$

is a ratio of the *normalized* probability densities of observing the data set D assuming the two hypotheses under consideration. In the Appendix of Lo & Magaña Hernandez (2021) we give the full derivation for the expressions evaluating the normalized probability densities of observing the

data set D under each of the hypotheses. Here we will outline the derivation. The core idea is to use the graphs that describe the data generation processes for the two hypotheses in Figure 5.1 and 5.2 to write down the desired expressions for the probability densities, and that the likelihood functions (which are the probability densities viewed as functions of the event-level parameters) can be factorized under both the hypotheses.

For the not-lensed hypothesis, since the N signals are independent, we have

$$p(D|\mathcal{H}_{\rm NL}) = \prod_{i=1}^{N} p(D^{(i)}|\mathcal{H}_{\rm NL}). \tag{5.4}$$

Combining this with the data generation process described in Figure 5.1 we have

$$p(D|\mathcal{H}_{\rm NL}) \propto \prod_{i=1}^{N} \int {\rm d}\theta^{(i)}\ p(D^{(i)}|\theta^{(i)}, \mathcal{H}_{\rm NL}) p_{\rm pop}(\theta^{(i)}), \tag{5.5}$$

where the expression on the right-hand side is also known as the (unnormalized) marginal likelihood under the not-lensed hypothesis. Note that we need to make sure that the probability density $p(D^{(i)}|\mathcal{H}_{\rm NL})$ is normalized over all observable data, accounting for selection effects (Mandel et al., 2019). This can be done by evaluating the proper normalization constant α, where

$$\alpha = \int_{\text{all obs. data}} {\rm d}D^{(i)}\ p(D^{(i)}|\mathcal{H}_{\rm NL}). \tag{5.6}$$

Therefore the expression for the normalized $p(D|\mathcal{H}_{\rm NL})$ is given by

$$p(D|\mathcal{H}_{\rm NL}, \lambda, \mathcal{R}) = \frac{1}{\alpha^N} \prod_{i=1}^{N} \int {\rm d}\theta^{(i)}\ p(D^{(i)}|\theta^{(i)}, \mathcal{H}_{\rm NL}) p_{\rm pop}(\theta^{(i)}, \lambda, \mathcal{R}). \tag{5.7}$$

As for the lensed hypothesis, unfortunately the probability density $p(D|\mathcal{H}_{\rm L})$ cannot be factorized like Eq. (5.4). However, the likelihood functions can still be factorized if we assume that the noise realizations for the N events are independent and that a signal is deterministic given a

set of parameters $\boldsymbol{\theta}$ that describe the waveform. Marginalizing the joint likelihood function with parameters according to Figure 5.2, we have

$$p(D|\mathcal{H}_{\rm L},\lambda,\mathcal{R},\gamma) \propto \int \mathrm{d}\theta^{(1)}_{\rm com}\ \mathrm{d}\theta^{(1)}_{\rm ind}\ \cdots \mathrm{d}\theta^{(N)}_{\rm ind} \left[\prod_{j=1}^{N} p(D^{(j)}|\theta^{(j)}_{\rm ind},\theta^{(1)}_{\rm com})\right] \times p_{\rm pop,ind}(\theta^{(1)}_{\rm ind},...,\theta^{(N)}_{\rm ind}|\gamma) p_{\rm pop,com}(\theta^{(1)}_{\rm com}|\lambda,\mathcal{R}), \tag{5.8}$$

where the expression on the right-hand side is known as the (unnormalized) marginal likelihood under the lensed hypothesis. Again, we will need to compute the normalized probability density $p(D|\mathcal{H}_{\rm L})$ in order to compute a meaningful Bayes factor, and take selection effects into account. The proper normalization constant β in this case, is given by

$$\beta \propto \int_{\text{all obs. data set}} \mathrm{d}D^{(1)}\cdots\mathrm{d}D^{(N)} \int \mathrm{d}\theta^{(1)}_{\rm com}\ \mathrm{d}\theta^{(1)}_{\rm ind}\ \cdots \mathrm{d}\theta^{(N)}_{\rm ind} \left[\prod_{j=1}^{N} p(D^{(j)}|\theta^{(j)}_{\rm ind},\theta^{(1)}_{\rm com})\right] \times p_{\rm pop,ind}(\theta^{(1)}_{\rm ind},...,\theta^{(N)}_{\rm ind}) p_{\rm pop,com}(\theta^{(1)}_{\rm com}). \tag{5.9}$$

Therefore, the expression for the normalized $p(D|\mathcal{H}_{\rm L})$ is given by

$$p(D|\mathcal{H}_{\rm L}) = \frac{1}{\beta}\int \mathrm{d}\theta^{(1)}_{\rm com}\ \mathrm{d}\theta^{(1)}_{\rm ind}\ \cdots \mathrm{d}\theta^{(N)}_{\rm ind} \left[\prod_{j=1}^{N} p(D^{(j)}|\theta^{(j)}_{\rm ind},\theta^{(1)}_{\rm com})\right] \times p_{\rm pop,ind}(\theta^{(1)}_{\rm ind},...,\theta^{(N)}_{\rm ind}) p_{\rm pop,com}(\theta^{(1)}_{\rm com}). \tag{5.10}$$

Finally, we have the expression that we can use to evaluate the Bayes factor for the lensed hypothesis versus the not-lensed hypothesis, namely

$$\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} = \frac{\alpha(\lambda,\mathcal{R})^N}{\beta(\lambda,\mathcal{R},\gamma)} \underbrace{\frac{p(D|\mathcal{H}_{\rm L},\lambda,\mathcal{R},\gamma)}{p(D|\mathcal{H}_{\rm NL},\lambda,\mathcal{R})}}_{\text{coherence ratio } \mathcal{C}} . \tag{5.11}$$

One can interpret the second factor in Eq. (5.11), which is the ratio of unnormalized marginal likelihoods under the two hypotheses, as a measurement of how well the data set D of N signals can be jointly fit by a set of common parameters versus N sets of independent parameters, which we call it the *coherence ratio* $\mathcal{C}$ to differentiate it with the Bayes factor. While a negative log coherence ratio means that the lensed hypothesis, that is setting some of the parameters to be the same across events, fails to fit the N signals jointly, *a positive log coherence ratio however does not mean that the N signals are lensed.* This is the Occam's razor at play. Assuming that the lensed hypothesis and the not-lensed hypothesis fit the data set D equally well, the lensed hypothesis will be favored by the Bayesian model selection framework because it has fewer free parameters, and hence a smaller prior volume. For GW signals from high-mass BBH mergers, this issue will be more apparent as they produce shorter signals detectable in the interferometers, and we usually make less precise measurements of the masses for these high-mass systems (Ghosh et al., 2016). This is partially alleviated by incorporating the population information that they are rarer compared to lighter systems. It also brings out an important point that the *Bayes factor, or generally any probabilistic statement, that some GW signals are strongly lensed depends on the source population* one is considering.

We can think of the factor $\beta(\lambda,\mathcal{R},\gamma)/\alpha(\lambda,\mathcal{R})^N$ in Eq. (5.11) as a *population-averaged* scale of the coherence ratio accounting for selection effects, which affect the two hypotheses differently. If the coherence ratio is greater than the population typical value for β/α^N, then the Bayes factor will indicate that the lensed hypothesis is favored by the observed data. In fact, the normalization constant under the not-lensed hypothesis α can be interpreted as the detectable fraction of sources (Mandel et al., 2019). Similarly, we can interpret the normalization constant under the lensed

hypothesis β as the fraction of sources that would produce N detectable lensed signals. We expect that the order of magnitude for β would be similar to that for α. Therefore, essentially selection effects penalize the lensed hypothesis by a factor of roughly α^{N-1}, counteracting the Occam's razor.

5.2.5 The prior odds $\mathcal{P}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}}$

The Bayes factor we derived above in Eq. (5.11) *only compares the coherence of the data set* with each hypothesis, but not the probability in which each hypothesis would occur. We know *a priori* that strong lensing causing at least N images occurs less frequently than observing N independent GW events with each coming from a different source. We can easily incorporate our knowledge about the rate in the form of prior odds $\mathcal{P}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}}$, which is defined as

$$\mathcal{P}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} = \frac{p(\mathcal{H}_{\rm L})}{p(\mathcal{H}_{\rm NL})}. \tag{5.12}$$

We can then compute the posterior odds $\mathcal{O}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}}$ using Eq. (5.2) from the Bayes factor in Eq. (5.11) and the prior odds in Eq. (5.12).

One can assign the prior odds simply as the ratio of the rate of observing N lensed images from a single source over the rate of observing N GW signals coming from N independent sources. Obtaining this will require detailed modeling of GW sources and lenses. In particular these numbers should be computed under the the chosen source and lens population models for an analysis. However, one can argue that irrespective of the population models chosen, the prior odds is very small with the current sensitivities of GW detectors, i.e. $\mathcal{P}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} \approx 10^{-2} - 10^{-4}$ (Ng et al., 2018; Oguri, 2018; Li et al., 2018; Buscicchio et al., 2020; Mukherjee et al., 2021a).

5.2.6 Marginalization over redshift

With the expression for $p(D|\mathcal{H}_{\rm L})$ under the lensed hypothesis in Eq. (5.10), one can estimate the integral using a stochastic sampling algorithm such as nested sampling (Skilling, 2004) by sampling over $\left\{\theta^{(1)}_{\rm com}, \theta^{(1)}_{\rm ind}, \cdots, \theta^{(N)}_{\rm ind}\right\}$ with a prior $p_{\rm pop,ind}(\theta^{(1)}_{\rm ind}, ..., \theta^{(N)}_{\rm ind}) p_{\rm pop,com}(\theta^{(1)}_{\rm com})$ and a *joint* likelihood $\prod_{j=1}^{N} p(D^{(j)}|\theta^{(j)}_{\rm ind}, \theta^{(1)}_{\rm com})$. However, a direct sampling will be inefficient because of the degeneracy between the absolute magnification and the luminosity distance, and hence the redshift of the source. Under the not-lensed hypothesis, we can infer the source redshift since we can infer the luminosity distance of the source $d_{\rm L}^{\rm src}$, and by assuming a particular cosmology we can compute the redshift $z^{\rm src} = z(d_{\rm L}^{\rm src})$ from the luminosity distance. Under the lensed hypothesis, each image will be, in general, magnified by a different factor. In fact, we can only measure the apparent luminosity distance for each image as in Eq. (4.2). Therefore, we will not be able to infer the absolute magnification for each image and the source redshift at the same time. For example, a signal with a said redshift of $z \approx 0.363$ and an absolute magnification of $\mu = 4$ would have the same apparent luminosity distance of 1 Gpc as a signal with a redshift of $z \approx 0.780$ and an absolute magnification of $\mu = 25$.

In order to explore the degenerate parameter space more efficiently, we can marginalize over the source redshift separately. In fact, the source redshift z stands out from the rest of the parameters. This is because with a given redshift, one can figure out the prior distribution of the apparent luminosity distance $d_{\rm L}^{(i)}$ given the prior distribution of the absolute magnification $p(\mu^{(i)})$ by

$$\begin{aligned} p(d_{\rm L}^{(i)}) &= p(\mu^{(i)} = \left(\frac{d_{\rm L}^{\rm src}(z)}{d_{\rm L}^{(i)}}\right)^2 |z) \left|\frac{\partial \mu^{(i)}}{\partial d_{\rm L}^{(i)}}\right| \\ &= \frac{2\mu^{(i)}}{d_{\rm L}^{(i)}} p(\mu^{(i)} = \left(\frac{d_{\rm L}^{\rm src}(z)}{d_{\rm L}^{(i)}}\right)^2 |z), \end{aligned} \tag{5.13}$$

and similarly for the prior distribution of the redshifted/detector-frame masses given the distribu-

tion of source-frame masses and the redshift as

$$\begin{aligned} p(m_{1,2}^{\rm det}) &= p(m_{1,2}^{\rm src} = \frac{m_{1,2}^{\rm det}}{1+z}|z)\left|\frac{\partial m_{1,2}^{\rm src}}{\partial m_{1,2}^{\rm det}}\right| \\ &= \left(\frac{1}{1+z}\right)^2 p(m_{1,2}^{\rm src} = \frac{m_{1,2}^{\rm det}}{1+z}|z). \end{aligned} \tag{5.14}$$

Therefore, we can rewrite Eq. (5.10) as a 1D integral over the redshift as

$$p(D|\mathcal{H}_{\rm L}) \propto \int {\rm d}z\ \mathcal{L}_{\rm marg}(z)\ p_z(z|\mathcal{H}_{\rm L}), \tag{5.15}$$

where $\mathcal{L}_{\rm marg}(z)$ is given by

$$\begin{aligned} \mathcal{L}_{\rm marg}(z) = \int {\rm d}\theta_{\rm com}^{(1)} \setminus \{z\}\ {\rm d}\theta_{\rm ind}^{(1)}\ \cdots {\rm d}\theta_{\rm ind}^{(N)} \times \left[\prod_{j=1}^{N} p(D^{(j)}|\theta_{\rm ind}^{(j)}, \theta_{\rm com}^{(1)} \setminus \{z\})\right] \\ \times\ p_{\rm pop,ind}(\theta_{\rm ind}^{(1)}, ..., \theta_{\rm ind}^{(N)}) p_{\rm pop,com}(\theta_{\rm com}^{(1)} \setminus \{z\}). \end{aligned} \tag{5.16}$$

The marginalized likelihood $\mathcal{L}_{\rm marg}(z)$, which is a function of z only, can be obtained via the conventional Monte Carlo methods (such as Markov Chain Monte Carlo method and nested sampling) by *sampling over redshifted/detector-frame parameters without the redshift*. This will alleviate the degeneracy problem, as well as open up the possibility of computing Eq. (5.10) by reusing computations done with the not-lensed hypothesis assumed, without re-exploring the joint parameter space. It also lends itself to the interpretation of treating the redshift as a hyper-parameter of a sub-population of *signals* sharing the same intrinsic parameters (and some of the extrinsic parameters).

Given a merger rate density $\mathcal{R}(z) \equiv {\rm d}N_{\rm src}/\left({\rm d}V_{\rm c}{\rm d}t\right)$, which is the number density of mergers per co-moving volume $V_{\rm c}$ per unit time t in the source frame, one can compute the probability density of the source redshift z as

$$p(z) \propto \frac{{\rm d}V_{\rm c}}{{\rm d}z}\frac{1}{1+z}\mathcal{R}(z). \tag{5.17}$$

Using the product rule, we can write down the prior distribution for the redshift z under the lensed hypothesis as

$$\begin{aligned} p_z(z|\mathcal{H}_{\rm L}) &= \frac{1}{C}\frac{p(\mathcal{H}_{\rm L}|z)p(z)}{p(\mathcal{H}_{\rm L})} \\ &= \frac{1}{C}\frac{\tau(z)p(z)}{p(\mathcal{H}_{\rm L})}, \end{aligned} \tag{5.18}$$

where C is the normalization constant, and $\tau(z) \equiv p(\mathcal{H}_{\rm L}|z)$ is the *optical depth* of strong lensing at redshift z. Similarly, under the not-lensed hypothesis, the prior distribution for the redshift $p_z(z|\mathcal{H}_{\rm NL})$ is given by

$$\begin{aligned} p_z(z|\mathcal{H}_{\rm NL}) &= \frac{1}{C'}\frac{p(\mathcal{H}_{\rm NL}|z)p(z)}{p(\mathcal{H}_{\rm L})} \\ &= \frac{1}{C'}\frac{[1-\tau(z)]\,p(z)}{p(\mathcal{H}_{\rm NL})}, \end{aligned} \tag{5.19}$$

where the normalization constant $C^{'}$ is defined accordingly. Figure 5.3 shows the prior distribution of redshift z under the lensed (solid blue line) and not-lensed hypothesis (solid green line), using the optical depth model in **?** and a merger rate density tracking the star formation rate in Belczynski et al. (2017); Oguri (2018). The peak of the prior distribution under the lensed hypothesis shifts to a higher value of $z \sim 3$ compared to that under the not-lensed hypothesis, which peaks at roughly $z \sim 2$ because of the optical depth (grey dotted line) being higher at higher redshifts.

As a by-product of evaluating Eq. (5.15), we also get a set of posterior samples of z, which are distributed according to

$$p(z|D,\mathcal{H}_{\rm L}) = \frac{\mathcal{L}_{\rm marg}(z)p_z(z|\mathcal{H}_{\rm L})}{\int {\rm d}z\ \mathcal{L}_{\rm marg}(z)\ p_z(z|\mathcal{H}_{\rm L})}. \tag{5.20}$$

In the next sub-section, we describe how to reconstruct the unbiased (but degenerate) source parameters using Gibbs sampling.

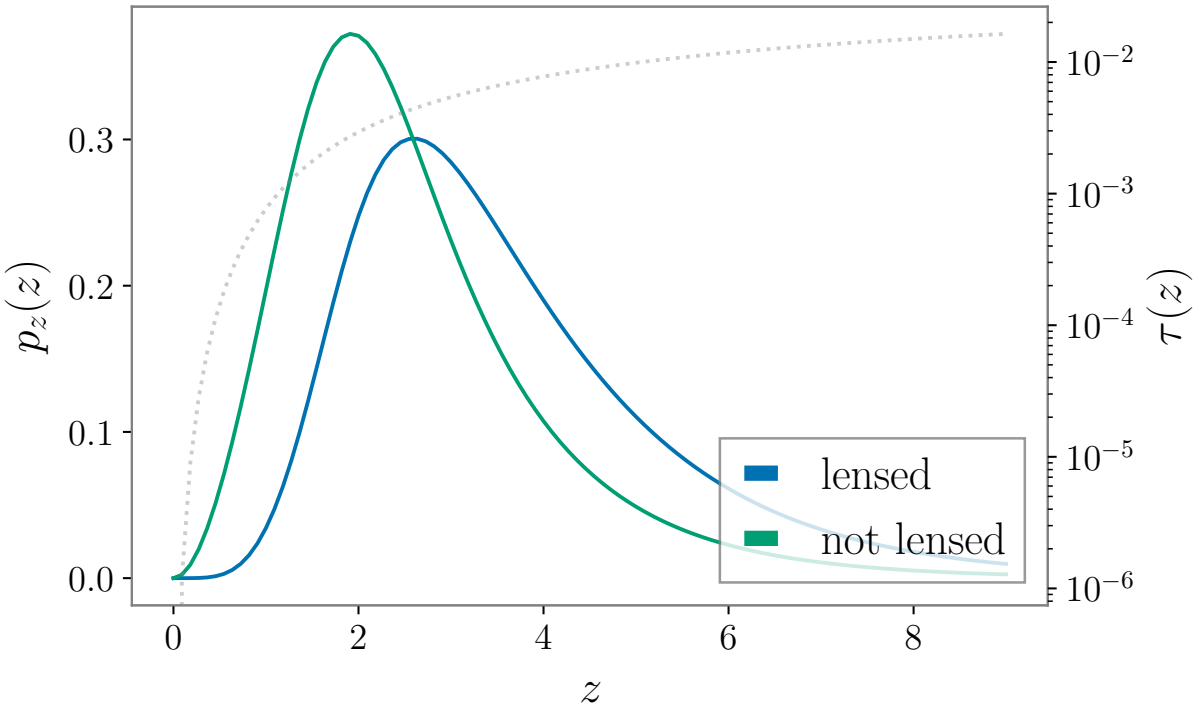


Figure 5.3: The probability densities $p_z(z)$ of the source redshift z under the lensed and not-lensed hypothesis. The grey dotted line shows the optical depth $\tau(z)$. As the optical depth increases with the redshift, the peak of the density p_z under the lensed hypothesis shifts to a higher value of $z \sim 3$ compared to the density under the not-lensed hypothesis.

5.2.7 Inferring unbiased source parameters using Gibbs sampling

Ultimately we want a set of joint posterior samples $\{z, \theta\}$ describing the source of the observed lensed signals. As a by-product of the marginalization over the redshift calculation using nested sampling, we obtain a set of posterior samples of the redshift $z \sim p(z|D, \mathcal{H}_\mathrm{L})$ marginalized over the parameters θ. Using Gibbs sampling, we can obtain the desired joint posterior samples from samples drawn from the conditional probability distributions $p(z|D, \mathcal{H}_\mathrm{L})$ from the marginalization step and $p(\theta|z, D, \mathcal{H}_\mathrm{L})$ from the inference step. This is because

$$p(z, \theta|D, \mathcal{H}_\mathrm{L}) \propto p(\theta|z, D, \mathcal{H}_\mathrm{L}) p(z|D, \mathcal{H}_\mathrm{L}). \tag{5.21}$$

5.3 STRONG LENSING OF GRAVITATIONAL WAVES FROM A BINARY BLACK HOLE MERGER: OBSERVING A PAIR OF LENSED SIGNALS

Now that we have developed the statistical framework in a general setting, here we want to apply the framework to analyze two particular cases and discuss the technical subtleties involved, namely for the case of strong lensing of a GW signal from a BBH merger with a pair of lensed images (i.e. $N = 2$) observed, and with only *one* image (i.e. $N = 1$) observed. In this section, we focus on the former case first.

5.3.1 Under the not-lensed hypothesis

Suppose we write the event-level parameters for each of the BBH mergers under the not-lensed hypothesis as

$$\theta^{(i)} = \{\underbrace{M_{\rm tot}^{\rm det}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2}_{\text{intrinsic parameters}}, \underbrace{d_{\rm L}, \alpha, \delta, \psi, \iota, \phi_{\rm c}, t_{\rm c}}_{\text{extrinsic parameters}}\}, \tag{5.22}$$

and these are the parameters that are being sampled over during the inference step. As derived in Eq. (5.7) with $N = 2$, under the not-lensed hypothesis we have

$$p(\{D^{(1)}, D^{(2)}\}|\mathcal{H}_{\rm NL}) = \frac{1}{\alpha^2} p(D^{(1)}|\mathcal{H}_{\rm NL}) p(D^{(2)}|\mathcal{H}_{\rm NL}). \tag{5.23}$$

Figure 5.4 shows a graphical representation of the data generation process under the not-lensed hypothesis for signals from BBH mergers using the parametrization in Eq. (5.22). Here we use $\mathbf{\Phi}$ to denote the set of extrinsic parameters $\{\alpha, \delta, \psi, \iota, \phi_c\}$ that are distributed according to the distribution $p_{\rm ext}$. As for the time of arrival $t_{\rm c}$, we treat it separately and hence it is not shown in Figure 5.4. From matched-filtering pipelines that scan through all the data looking for GW triggers, we know roughly the time of arrival for each trigger. Let us write $t_{\rm c}^{(1)} = t_1 + \delta t_{\rm c}^{(1)}$ and

$t_{\rm c}^{(2)} = t_2 + \delta t_{\rm c}^{(2)}$, where t_1 and t_2 are the point estimates of the arrival times given by a pipeline for the two triggers respectively. Instead of sampling over $t_{\rm c}^{(1)}$ and $t_{\rm c}^{(2)}$, we sample over $\delta t_{\rm c}^{(1)}$ and $\delta t_{\rm c}^{(2)}$ with a small prior range (typically ~ 0.2 s) and t_1, t_2 taken to be known. Mathematically, this means

$$p(t_{\rm c}^{(1)}, t_{\rm c}^{(2)}|\mathcal{H}_{\rm NL}){\rm d}t_{\rm c}^{(1)}{\rm d}t_{\rm c}^{(2)} = p(\delta t_{\rm c}^{(1)}, \delta t_{\rm c}^{(2)}|t_1, t_2, \mathcal{H}_{\rm NL})p(t_1, t_2|\mathcal{H}_{\rm NL}){\rm d}\delta t_{\rm c}^{(1)}{\rm d}\delta t_{\rm c}^{(2)}. \tag{5.24}$$

Suppose we order the two events by their times of arrival, i.e. $t_2 > t_1$, and define the time delay $\Delta t \equiv (t_2 - t_1) > 0$. After this transformation, there is an extra factor in the prior that accounts for the probability of having two random events separated by a time delay of Δt under the not-lensed hypothesis. If we model the arrival of events by a Poisson process, the prior probability density that *any random pair* of events having a time delay of Δt, *given that* there are $N_{\rm obs}$ events during the time interval of $(0, T_{\rm obs}]$, is given by

$$p(\Delta t|\mathcal{H}_{\rm NL}) = \frac{2}{T_{\rm obs}}\left(1 - \frac{\Delta t}{T_{\rm obs}}\right), \tag{5.25}$$

for a detailed derivation see Lo & Magaña Hernandez (2021). This can be considered as part of the time-delay Bayes factor in Haris et al. (2018) from the not-lensed hypothesis.

Therefore, the full expression for $p(\{D^{(1)}, D^{(2)}\}|\mathcal{H}_{\rm NL})$ now reads

$$\begin{aligned} p(\{D^{(1)}, D^{(2)}\}|\mathcal{H}_{\rm NL}) = \frac{1}{\alpha^2}p(\Delta t|\mathcal{H}_{\rm NL})\prod_{i=1}^{2}\int {\rm d}\{\underbrace{M_{\rm tot}^{\rm det}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2, d_{\rm L}, \alpha, \delta, \psi, \iota, \phi_{\rm c}, \delta t_{\rm c}}_{\theta^{(i)}}\} \\ \times\, p(D^{(i)}|\theta^{(i)})p_{\rm pop}(\theta^{(i)}|\mathcal{H}_{\rm NL}), \end{aligned} \tag{5.26}$$

where under the not-lensed hypothesis there is a one-to-one mapping between $d_{\rm L}$ and z, and hence one will only need to convert Eq. (5.19) by multiplying the proper Jacobian without the need of a separate marginalization of the source redshift.

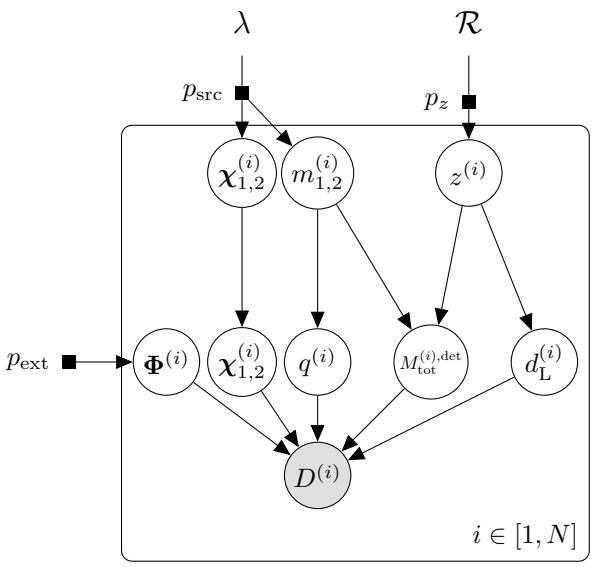


Figure 5.4: Data generation process for the N observed data under the not-lensed hypothesis $\mathcal{H}_{\rm NL}$. This is similar to Fig. 5.1 but with the event-level parameters $\theta^{(i)}$ written out explicitly. Here we use $\boldsymbol{\Phi}$ to denote the set of extrinsic parameters $\{\alpha, \delta, \psi, \iota, \phi_c\}$ that are distributed according to the distribution $p_{\rm ext}$.

5.3.2 Under the lensed hypothesis

Under the lensed hypothesis, we write the event-level parameters differently, namely we let the common parameters $\theta_{\rm com}^{(i)} = \{M_{\rm tot}^{\rm det}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \alpha, \delta, \psi, \iota, \phi_{\rm c}\}$. As for the independent parameters, we write $\theta_{\rm ind}^{(i)} = \{\delta t_{\rm c}^{(i)}, d_{\rm L}^{(i)}, \Xi^{(i)}\}$, where we perform the same transformation to the time of arrival as in the case under the not-lensed hypothesis, and Ξ denotes the *type of an image* which can be either {I, II, III}.

Each strongly lensed image can be classified into three types (I, II or III), where each image type corresponds to a Morse index of $\{0, 1, 2\}$ respectively, inducing a different phase shift as shown in Eq. (4.5) to the image because of the interaction of the lensed image with the caustic. One would expect the image that arrives at the Earth first to be of type I since type I images correspond to local minima of the Fermat time-of-arrival potential. However, the signal that we called the first image in an analysis might not actually be the first image that had arrived the Earth

since, for example, the GW detectors might be offline. Various arguments on the type of images one would see can be made if we know *a priori* the geometry of the gravitational lens but this is not known prior to the analysis. Therefore, we assume that the type of a lensed image to follow a discrete uniform distribution, namely

$$p_{\rm lens}(\Xi^{(1)}, \Xi^{(2)}) = p_{\rm lens}(\Xi^{(1)}) p_{\rm lens}(\Xi^{(2)}), \tag{5.27}$$

where

$$p_{\rm lens}(\Xi^{(i)}) = \begin{cases} 1/3 \ \text{ when } \Xi^{(i)} = \text{I} \\ 1/3 \ \text{ when } \Xi^{(i)} = \text{II} \\ 1/3 \ \text{ when } \Xi^{(i)} = \text{III} \end{cases} . \tag{5.28}$$

Figure 5.5 shows a graphical representation of the data generation process under the lensed hypothesis for BBH signals. Similar to Figure 5.4, we use $\mathbf{\Phi}$ to denote the set of extrinsic parameters $\{\alpha, \delta, \psi, \iota, \phi_{\rm c}\}$ that are distributed according to the distribution $p_{\rm ext}$, and that we treat the time of arrival $t_{\rm c}$ separately. Unlike the not-lensed case, here we assume that $\theta^{(i)}_{\rm com} = \{M^{\rm det}_{\rm tot}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \alpha, \delta, \psi, \iota, \phi_{\rm c}\}$ are the same across the signals (hence we dropped the superscript in the graph). Also, even though we sample the apparent luminosity distance for each image, there is no one-to-one mapping between it and the true source redshift since the apparent luminosity distance is also related to the absolute magnification of a lensed image. As discussed in Section 5.2.6, we perform the marginalization over the source redshift separately.

For the time of arrival $t^{(i)}_{\rm c}$, we can perform the same transformation as in the case for the not-lensed hypothesis (similar to Eq. (5.24)), and sample $\delta t^{(i)}_{\rm c}$ that has a much smaller range instead. However, instead of having an analytical expression for the time delay Δt, there is no analytically tractable expression for the time delay under the lensed hypothesis. That being said, we can obtain it readily from numerical simulations (for example, see Haris et al. (2018)). As a result, there is an

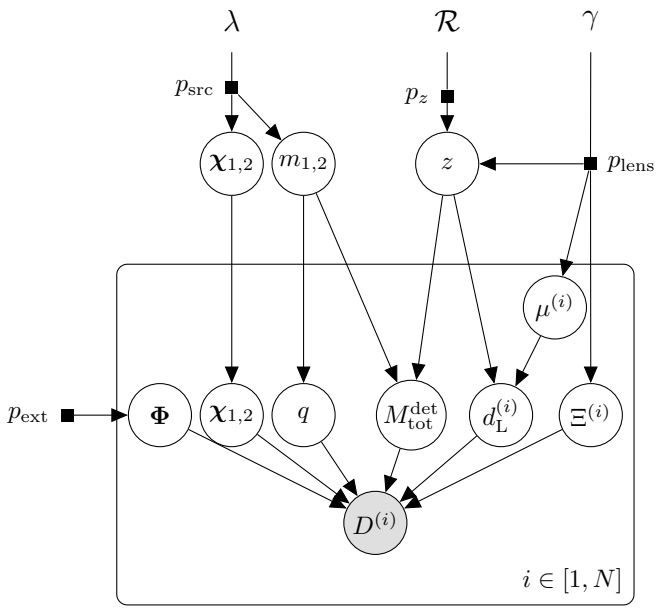


Figure 5.5: Data generation process for the N observed data under the lensed hypothesis $\mathcal{H}_{\rm L}$. This is similar to Fig. 5.2 but with the common parameters $\theta_{\rm com}^{(i)}$ and the independent parameters $\theta_{\rm ind}^{(i)}$ written out explicitly. Again, we use $\boldsymbol{\Phi}$ to denote the set of extrinsic parameters $\{\alpha, \delta, \psi, \iota, \phi_{\rm c}\}$ that are distributed according to the distribution $p_{\rm ext}$.

extra factor of $p(\Delta t|\mathcal{H}_{\rm L})$ in the prior that accounts for the probability of having two lensed images separated by a time delay of Δt.

Therefore, the full expression for $p(\{D^{(1)}, D^{(2)}\}|\mathcal{H}_{\rm L})$ now reads

$$
\begin{aligned}
p(\{D^{(1)}, D^{(2)}\}|\mathcal{H}_{\rm L}) = \frac{1}{\beta} p(\Delta t|\mathcal{H}_{\rm L}) \int \mathrm{d}z\; p_z(z|\mathcal{H}_{\rm L}) \Bigg[\int \mathrm{d}\{\underbrace{M_{\rm tot}^{\rm det}, q, \boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \alpha, \delta, \psi, \iota, \phi_{\rm c}}_{\theta_{\rm com}}\} \\
\times \int \mathrm{d}\{\underbrace{\delta t_{\rm c}^{(1)}, d_{\rm L}^{(1)}, \Xi^{(1)}}_{\theta_{\rm ind}^{(1)}}, \underbrace{\delta t_{\rm c}^{(2)}, d_{\rm L}^{(2)}, \Xi^{(2)}}_{\theta_{\rm ind}^{(2)}}\} p(D^{(1)}|\theta_{\rm com}, \theta_{\rm ind}^{(i)}) \\
\times p(D^{(2)}|\theta_{\rm com}, \theta_{\rm ind}^{(2)}) p_{\rm pop,ind}(\theta_{\rm ind}^{(1)}, \theta_{\rm ind}^{(2)}|\mathcal{H}_{\rm L}) p_{\rm pop,com}(\theta_{\rm com}|\mathcal{H}_{\rm L}) \Bigg],
\end{aligned}
\tag{5.29}
$$

where the bracketed expression would be identified as $\mathcal{L}_{\rm marg}(z)$ as discussed in Sec 5.2.6.

5.3.3 Demonstration

Here we demonstrate the framework with two examples. In the first example, we injected two GW signals with a redshifted total mass $M_{\rm tot}^{\rm det} = 280M_{\odot}$ into simulated data streams. With this example, we show explicitly how the source population model would change the Bayes factor. In the second example, we injected instead two GW signals with a redshifted total mass $M_{\rm tot}^{\rm det} = 60M_{\odot}$, which corresponds to typical stellar-mass BBH systems for the LIGO-Virgo detectors. In both examples, we use the waveform approximant `IMRPhenomXPHM` (Pratten et al., 2020), which models both the leading quadrupole ($\ell = 2$) radiation, as well as some of the non-quadrupole ($\ell > 2$) multipoles. By incorporating the higher order modes, we show that the image type of each lensed signal can also be inferred. All the results presented here were computed using the software package `hanabi`[5], which is built upon the package `bilby` (Ashton et al., 2019) and `parallel_bilby` (Smith et al., 2020). Also, we used the nested sampling algorithm implemented in the package `dynesty` (Speagle, 2020).

5.3.3.1 Example 1: two lensed signals from apparent intermediate-mass binary black hole mergers

In this example, we have two lensed GW signals injected into two simulated data streams with Gaussian noise recolored to match the Advanced LIGO (aLIGO) design noise curve (Abbott et al., 2018a). Table 5.1 summarizes some of the waveform parameters for the two signals. The two injected signals, when analyzed on their own, seem to originate from two separate mergers of an intermediate-mass binary black hole system.

To demonstrate how using different source population models would change one's interpretation of the two signals, as well as the numerical value of the Bayes factor using our framework, we

[5] https://github.com/ricokaloklo/hanabi

Parameter	Value
Redshifted total mass $M_{\rm tot}^{\rm det}$	$280M_\odot$
Mass ratio q	0.75
Redshifted primary mass $m_1^{\rm det}$	$160M_\odot$
Redshifted secondary mass $m_2^{\rm det}$	$120M_\odot$
Apparent luminosity distance for the first signal $d_{\rm L}^{(1)}$	3.11 Gpc
Apparent luminosity distance for the second signal $d_{\rm L}^{(2)}$	3.15 Gpc

Table 5.1: Summary of some of the injection parameters for Example 1 in Sec. 5.3.3.1. The two injected signals, when analyzed on their own, seem to originate from two separate mergers of an intermediate-mass binary black hole system.

first use a log-uniform distribution as the population model for the component masses, namely

$$p_{\rm src}(m_{1,2}^{\rm src}) \propto \begin{cases} 1/m_{1,2}^{\rm src} & \text{for } 5M_\odot \leq m_{1,2}^{\rm src} \leq 300M_\odot \\ 0 & \text{otherwise} \end{cases}. \tag{5.30}$$

For the component spins, we use a distribution that is uniform in the component spin magnitude, and isotropic in the spin orientation.

As for the merger rate density, here we use, for the sake of demonstration, an analytical fit from Oguri (2018) that tracks the population synthesis results from Belczynski et al. (2017) for population-I and population-II stars, namely

$$\mathcal{R}(z) = \frac{6.6 \times 10^3 \ \exp(1.6z)}{30 + \exp(2.1z)}. \tag{5.31}$$

For the absolute magnification, again for the purpose of demonstration, we use a simple power law

distribution that is independent of the time delay, namely

$$p_{\rm lens}(\mu^{(1)},\mu^{(2)}|\Delta t) = p_{\rm lens}(\mu^{(1,2)}) \propto \begin{cases} \mu^{-3} & \text{for } \mu \geq 2 \\ 0 & \text{otherwise} \end{cases}, \tag{5.32}$$

where it captures the general μ^{-3} scaling in the high-magnification regime, as well as the requirement that the absolute magnification has to exceed some threshold in order for multiple lensed images to be formed. However, it does not capture the correlation between the magnifications of the lensed images, and the correlation between the magnification and the time delay. For example, the relative magnification tends to unity if the lensed images are highly magnified (Schneider et al., 1992). In fact, one can derive a *poor-man's prior distribution* for the relative magnification, if we assume that the absolute magnification for each of the two images follows Eq. (5.32), with the form

$$p(\mu_{\rm rel}) = \begin{cases} \mu_{\rm rel} & \text{for } \mu_{\rm rel} \leq 1 \\ \mu_{\rm rel}^{-3} & \text{for } \mu_{\rm rel} > 1 \end{cases}. \tag{5.33}$$

In addition, we use a simple analytical model for the optical depth (Turner et al., 1984; **?**), which is the probability of strong lensing at a given redshift, with the form

$$\tau(z) = F\left(\frac{d_{\rm C}(z)}{d_{\rm H}}\right)^3, \tag{5.34}$$

where $d_{\rm C}(z)$ is the co-moving distance at z, and $d_{\rm H}$ is the Hubble distance. The empirical constant F is taken to be 0.0017 (**?**). A more realistic and detailed model for the merger rate density, the magnification distribution, as well as the optical depth, that impart more astrophysical information to an analysis would certainly help differentiating lensed signals.

With this set of population models, we obtained a log coherence ratio of $\log_{10}\mathcal{C} = 2.7$, and a log Bayes factor of $\log_{10}\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} = 1.1$ *without accounting for the time delay*. We see that with this

set of population models and the detector sensitivity, the selection effects down-weight the pair by a factor of ≈ 40. Figure 5.6 shows both the 1D and 2D marginalized posterior distributions for $\left\{M_{\rm tol}^{\rm src}, q, \mu^{(1)}, \mu^{(2)}, \mu_{\rm rel}, z\right\}$ obtained using the algorithm described in Sec. 5.2.7. The orange solid lines show the correct values for each of the parameters if the redshift z is set to 1. The plot shows that our two-step hierarchical procedure described in Sec. 5.2.6 is able to find the correct values describing the signals. From the plot we can also see the various degeneracies between parameters. For example, the degeneracy between the total mass $M_{\rm tot}^{\rm src}$ and the redshift z, where the blob in the lower left corner of Figure 5.6 corresponds to the redshifted total mass that we do measure. Note that we are able to infer the mass ratio q and the relative magnification $\mu_{\rm ref}$ as they are not degenerate with the redshift.

If we instead use a population model that asserts there are no black holes with mass greater than $60M_{\odot}$, referred as Model A in Abbott et al. (2019a), namely

$$p_{\rm src}(m_1^{\rm src}, m_2^{\rm src}|\alpha, \beta, m_{\rm min}, m_{\rm max}) = \frac{1-\alpha}{m_{\rm max}^{1-\alpha} - m_{\rm min}^{1-\alpha}}(m_1^{\rm src})^{-\alpha}\frac{1+\beta}{(m_1^{\rm src})^{1+\beta} - m_{\rm min}^{1+\beta}}(m_2^{\rm src})^{\beta} \quad (5.35)$$

valid over the range $m_{\rm min} \leq m_2^{\rm src} \leq m_1^{\rm src} \leq m_{\rm max}$ and is zero otherwise. We use $\alpha = 1.8$, $\beta = 0$, $m_{\rm min} = 5\ M_{\odot}$, and $m_{\rm max} = 60\ M_{\odot}$ as fiducial parameters for the assumed population. Now, *both the log coherence ratio and the log Bayes factor are infinite, while the log evidence under the lensed hypothesis is finite*. This is a *smoking-gun evidence* that the two signals are lensed. This is not surprising because the two signals are impossible under the not-lensed hypothesis with this set of population models. Under the not-lensed hypothesis, we interpret the apparent luminosity distance as the true luminosity distance without any magnification bias, allowing us to infer the redshift directly from the measured luminosity distance. In this case, the redshift that corresponds to the apparent luminosity distance of the first signal is roughly $z \approx 0.53$, meaning that both the primary and secondary mass would be above the $60M_{\odot}$ maximum. This example, though extreme,

Parameter	Value
Redshifted total mass $M_{\rm tot}^{\rm det}$	$60M_{\odot}$
Mass ratio q	0.875
Redshifted primary mass $m_1^{\rm det}$	$32M_{\odot}$
Redshifted secondary mass $m_2^{\rm det}$	$28M_{\odot}$
Apparent luminosity distance for the first signal $d_{\rm L}^{(1)}$	811 Mpc
Apparent luminosity distance for the second signal $d_{\rm L}^{(2)}$	823 Mpc

Table 5.2: Summary of some of the injection parameters for Example 2 in Sec. 5.3.3.2. This example serves to represent typical scenarios for second-generation terrestrial GW detectors observing stellar-mass BBH systems.

clearly shows that the Bayes factor, and hence one's interpretation on the origin, of the signals would be sensitive to the population models that one assumes.

5.3.3.2 Example 2: two lensed signals from a stellar-mass binary black hole merger

In the second example, we also have two lensed GW signals injected into two simulated data streams with Gaussian noise. However, this time the two signals have a lower redshifted total mass ($M_{\rm tot}^{\rm det} = 60M_{\odot}$). Table 5.2 summarizes some of the waveform parameters. This example serves to represent typical scenarios for second-generation terrestrial GW detectors such as the two Advanced LIGO detectors (Aasi et al., 2015) and the Advanced Virgo detector (Acernese et al., 2015a) observing stellar-mass BBH systems, and demonstrate how would the Bayes factor change with different detector sensitivities. For the population models, we use the same set of models in the last subsection with the Model A mass model described in Eq. (5.35).

In Figure 5.7, we show the marginalized 1D and 2D posterior distributions for the parameters, $\left\{M_{\rm tol}^{\rm src}, q, \mu^{(1)}, \mu^{(2)}, \mu_{\rm rel}, z\right\}$, we recover when the two lensed signals were injected into data streams with simulated Gaussian noise recolored to match the aLIGO design sensitivity (Abbott et al., 2018a). From the plot we see similar degenerate structures between parameters as in Figure 5.6.

	O1+O2	O3a	aLIGO design
$\log_{10}\alpha$	-3.5	-3.1	-2.4
$\log_{10}\beta$	-4.1	-3.7	-2.5
$\log_{10}(\beta/\alpha^2)$	2.9	2.5	2.3

Table 5.3: The values of $\log_{10}\alpha(\lambda,\mathcal{R})$, $\log_{10}\beta(\lambda,\mathcal{R},\gamma)$, and $\log_{10}[\beta(\lambda,\mathcal{R},\gamma)/\alpha(\lambda,\mathcal{R})^2]$ with different detector sensitivities computed using the `pdetclassifier` library (Gerosa et al., 2020) for the population models described in Sec. 5.3.3.2.

To demonstrate the degeneracies more explicitly, we show the correct source parameters for this two signals if we assume the true source redshift is $z = 0.4$ (solid orange lines), as well as that if the true redshift is instead $z = 1$ (dotted grey lines). Note that both the mass ratio q and the relative magnification $\mu_{\rm rel}$ take the same value when different source redshifts are assumed. While we are not able to constrain the source parameters individually because of the aforementioned degeneracies, *we are capable of providing joint constraints for the source parameters* by properly incorporating information from both the detected signals and the astrophysical population models assumed. From Figure 5.7, we see that it is less likely for the signals to come from a binary system with a total mass of $M_{\rm tol}^{\rm src} = 30M_\odot$ at a redshift $z = 0.4$ under the lensed hypothesis because of the large absolute magnifications required are less probable under the lens model we assumed in the analysis.

For this example, we obtained a log coherence ratio of $\log_{10}\mathcal{C} = 5.2$ and a log Bayes factor of $\log_{10}\mathcal{B}^{\mathcal{H}_{\rm L}}_{\mathcal{H}_{\rm NL}} = 3.0$ when injecting the signals into simulated Gaussian noise recolored to match the aLIGO design sensitivity (Abbott et al., 2018a). Table 5.3 tabulates the values of $\log_{10}\alpha$, $\log_{10}\beta$, and $\log_{10}(\beta/\alpha^2)$ for this particular set of population models under different detector sensitivities, computed using `pdetclassifier` (Gerosa et al., 2020).

As expected, the values of α and β increase as the detector network becomes more sensitive and capable of detecting weaker signals. The difference between the values of α and β narrows

as the network increases in sensitivity, and that the selection effects penalize the lensed hypothesis to a lesser extend, roughly by a factor of $\sim \alpha$. While we did not perform the same injection test with simulated noise recolored to match the sensitivity during O1+O2 and O3a, we can reasonably expect the log coherence ratio increases with a more sensitive detector network as we can better measure the waveform parameters to a higher precision. Therefore, the log coherence ratio, as well as the log Bayes factor would increase with the detector sensitivity given the same set of lensed signals.

5.3.4 Identifying the image types

When we consider only the dominant $\ell = |m| = 2$ modes and a non-precessing binary system, the phasing effect due to strong lensing reduces to a shift in the observed phase at coalescence (or any reference orbital phase) (Dai & Venumadhav, 2017; Ezquiaga et al., 2020). For a GW signal from the merger of a precessing binary system with a significant contribution from higher order modes, for example, when the system is asymmetric in component masses and/or is inclined with respect to our line of sight, we can break the degeneracy between the phasing effect from strong lensing and the orbital phase. This allows us to identify the image type for each of the lensed signals. We demonstrate this by injecting signals with an asymmetric mass ratio $q \approx 0.3$ viewing at an angle of roughly $107\,\mathrm{deg}$ between the line of sight and the total angular momentum vector into simulated Gaussian noise at aLIGO design sensitivity using two different waveform models, `IMRPhenomXP` and `IMRPhenomXPHM` (Pratten et al., 2020). The former approximant, `IMRPhenomXP`, includes only the quadrupole ($\ell = 2$) radiation from a precessing binary system, while the latter approximant, `IMRPhenomXPHM`, includes both the quadrupole radiation and some of the higher multipoles ($\ell > 2$) from the precessing system. In both cases, the first injected lensed GW signal is of type I, while the second injected signal is of type II. Figure 5.8 shows the joint probability mass function of the image type inferred for the first signal $\Xi^{(1)}$ and that for the second

signal $\Xi^{(2)}$. We see that when there are measurable contributions from higher modes, we are able to pin-point the type of each lensed image from the phasing effect (left panel of Figure 5.8), breaking the degeneracy between the phasing effect from strong lensing and the shift in the orbital phase. This is in line with the findings reported in Wang et al. (2021), where one can tell type-II images apart individually for third-generation detectors.

5.3.5 Improvement in localizing the source in the sky

Since we expect the lensed GW signals coming from the same source to have approximately identical sky locations, the signals should be better localized when analyzed jointly compared to the case when they are analyzed individually. This is because we gain information about the shared sky location from two data streams instead of just one. We demonstrate this using the inference results from Example 1 in Sec. 5.3.3.1. Figure 5.9 shows the 90% credible regions of the localization of signals, when analyzed separately (blue and green) and when analyzed jointly (orange). In all cases, the credible regions enclose the true source location (grey crosshair). However, the area of the 90% credible region, a metric for the localization uncertainty, from the joint inference is only $17\,\mathrm{deg}^2$, which is roughly two times smaller than that when localizing the first image only ($31\,\mathrm{deg}^2$) and the second, fainter, image only ($80\,\mathrm{deg}^2$).

Combining the improved sky localization of the source with the joint constraints of the source parameters (such as the redshift), one will be more informed when trying to locate the gravitational lens and the source electromagnetically (see, for example, Seto (2004); Hannuksela et al. (2020)). Indeed, if we were able to identify the massive object responsible for the gravitational lensing and observe lensing of electromagnetic waves as well, that can serve as a cross-validation that the GW signals that were being analyzed are indeed strongly lensed.

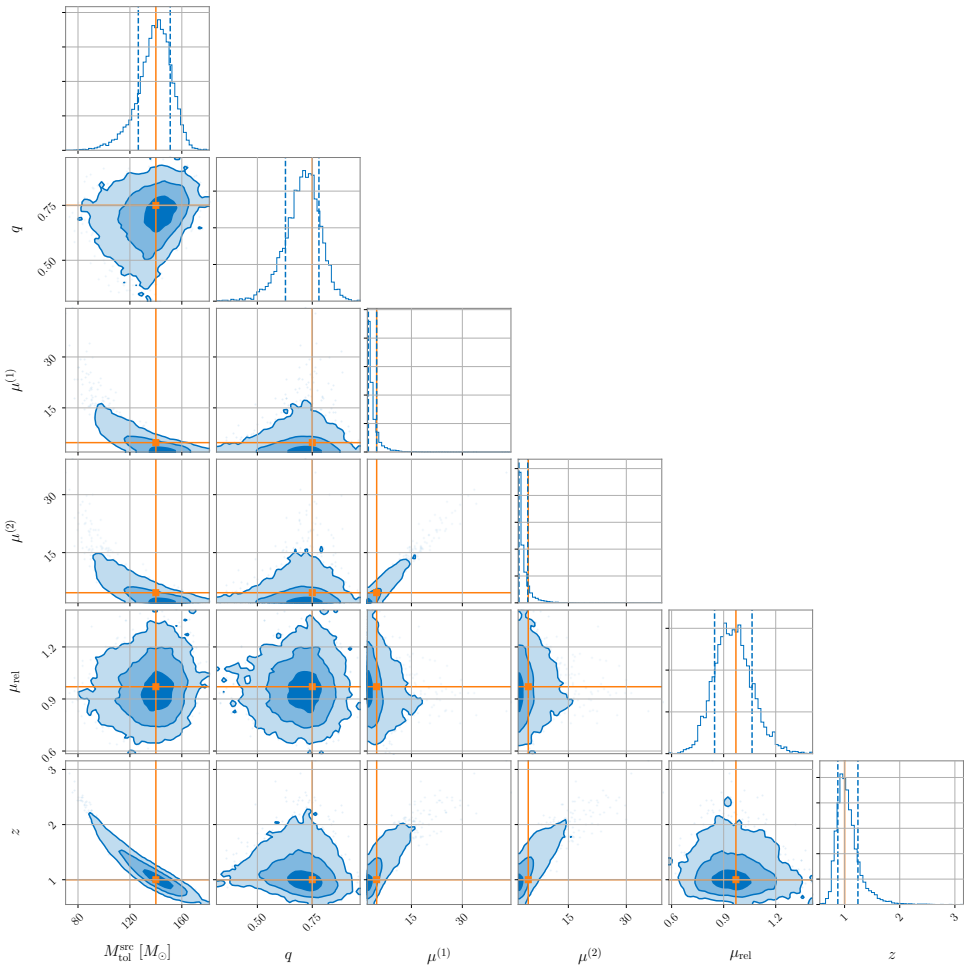


Figure 5.6: The 1D and 2D marginalized posterior distributions of $\left\{M_{\rm tol}^{\rm src}, q, \mu^{(1)}, \mu^{(2)}, \mu_{\rm rel}, z\right\}$ for Example 1 (c.f. Sec. 5.3.3.1) obtained using the algorithm described in Sec. 5.2.7. The orange solid lines show the correct values for each of the parameters if the redshift z is set to 1. The plot shows that our two-step hierarchical procedure described in Sec. 5.2.6 is able to find the correct values describing the signals. From the plot we can also see the various degeneracies between parameters. For example, the degeneracy between the total mass $M_{\rm tot}^{\rm src}$ and the redshift z, where the blob in the lower left corner corresponds to the redshifted total mass that we do measure. Note that we are able to infer the mass ratio q and the relative magnification $\mu_{\rm ref}$ as they are not degenerate with the redshift.

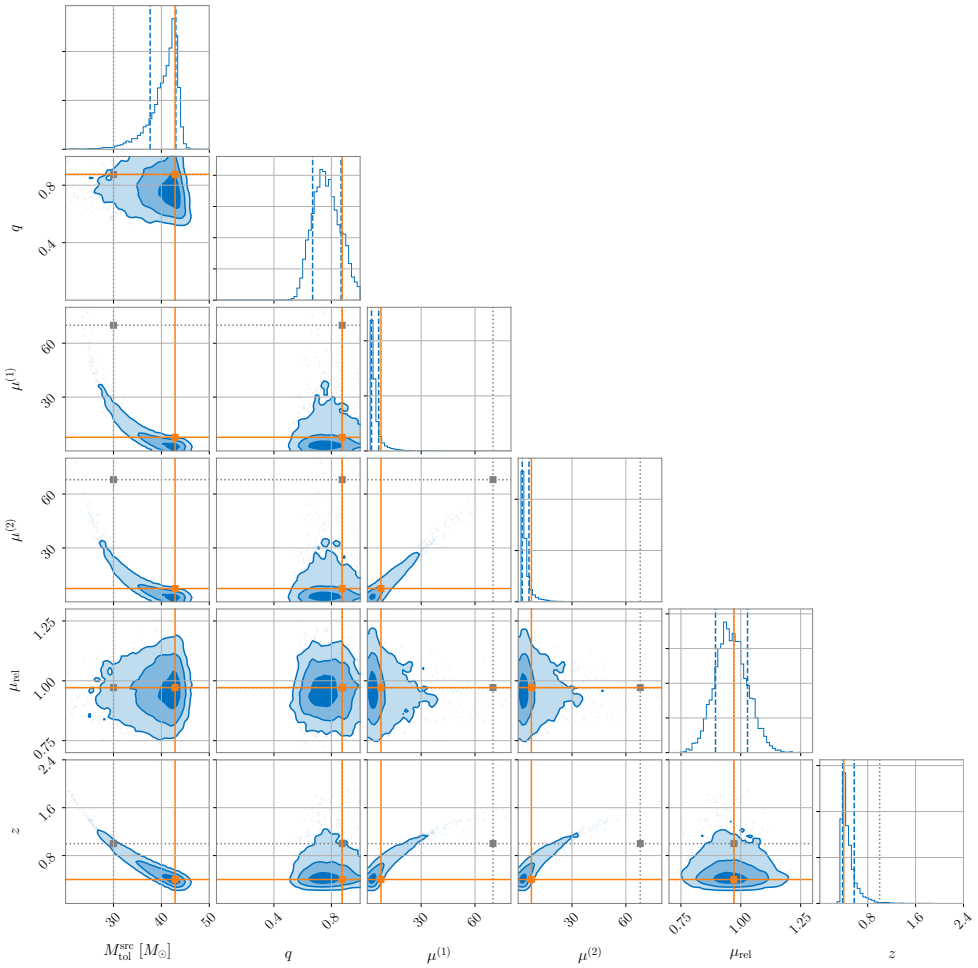
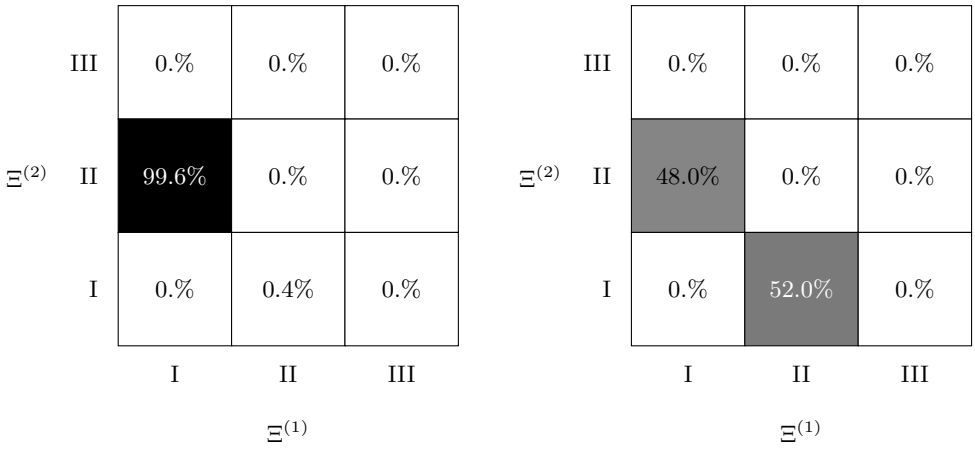


Figure 5.7: The 1D and 2D marginalized posterior distributions of $\left\{M_{\rm tol}^{\rm src}, q, \mu^{(1)}, \mu^{(2)}, \mu_{\rm rel}, z\right\}$ for Example 2 (c.f. Sec. 5.3.3.2) obtained using the algorithm described in Sec. 5.2.7. The orange solid lines show the correct values for each of the parameters if the redshift is set to $z = 0.4$, while the grey dotted lines show the correct values for the parameters if the redshift is instead set to $z = 1$. Note that both the mass ratio q and the relative magnification $\mu_{\rm rel}$ take the same value when different source redshifts are assumed. While we are not able to constrain the source parameters individually because of the degeneracies, we are capable of providing joint constraints for the source parameters by properly incorporating information from both the detected signals and the astrophysical population models assumed. We see that it is less likely for the signals to come from a binary system with a total mass of $M_{\rm tol}^{\rm src} = 30\ M_\odot$ at a redshift $z = 0.4$ under the lensed hypothesis because of the large absolute magnifications required are less probable under the lens model we assumed in the analysis.



(a) Using `IMRPhenomXPHM` Pratten et al. (2020) waveform model. It includes some higher modes ($\ell > 2$) other than the quadrupole ($\ell = 2$) radiation.

(b) Using `IMRPhenomXP` Pratten et al. (2020) waveform model. It only models the quadrupole ($\ell = 2$) radiation.

Figure 5.8: The joint posterior probability mass function of the (discrete) image type for the first signal $\Xi^{(1)}$ and that for the second signal $\Xi^{(2)}$ in an injection test. In the test, we injected a type-I signal into the first data stream, and a type-II signal into the second data stream. We see that when there are measurable contributions from higher modes, we are able to pin-point the type of each lensed image, breaking the degeneracy between the phasing effect from strong lensing and a shift in the orbital phase. This is in line with the findings in Wang et al. (2021).

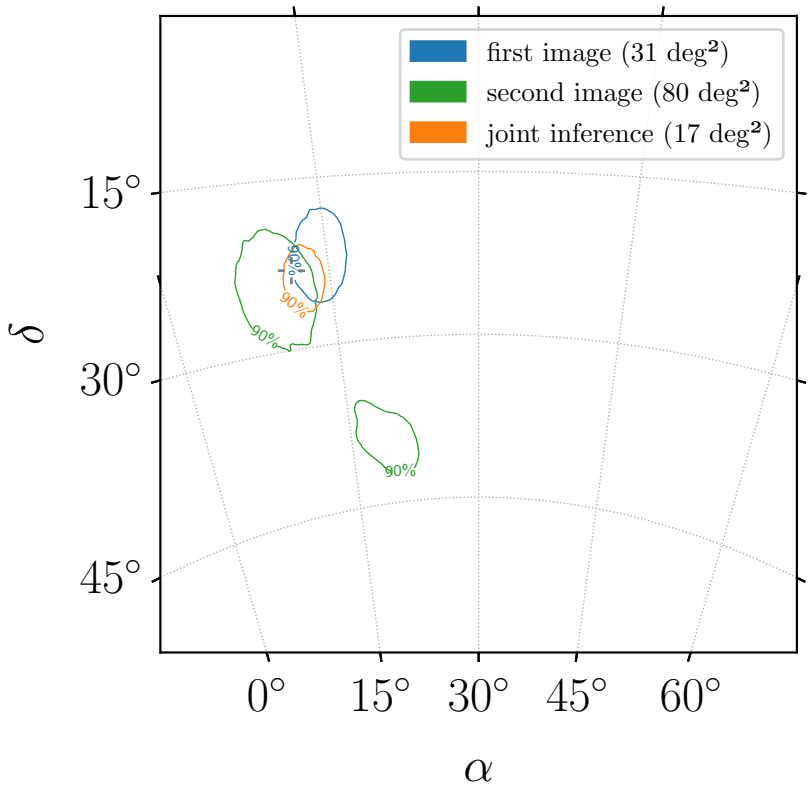


Figure 5.9: The sky localizations when two simulated lensed GW signals are analyzed jointly and when they are analyzed individually. The grey crosshair shows the injected values for the right ascension α and the declination δ. The signals are better localized when analyzed jointly (area of the 90% credible region: $17\,\mathrm{deg}^2$) compared to the case when they are analyzed individually (area of the 90% credible region: $31\,\mathrm{deg}^2$ for the brighter image, $80\,\mathrm{deg}^2$ for the fainter image) as expected (Seto, 2004) since we gain information about the shared sky location from two data streams instead of just one.

Part III

Modified Gravity

Chapter 6

Measuring the polarization content of gravitational waves with strongly lensed binary black hole mergers

Alternative modified theories of gravity predict up to six distinct polarization modes for gravitational-wave (GW) sources. In order to measure the relative amplitudes for each mode, we require at least six linearly-independent GW detectors, as they measure the projection of the GW signal onto their geometry. This projection is encoded in the antenna pattern functions of the instruments, which modulate the detectability of each mode as a function of time due to Earth's rotation. Strong gravitational lensing of gravitational waves allows us to probe the polarization content of these signals by effectively increasing the number of possible observations from the same astrophysical source. Given that the lensed images will arrive at different times, each measures a different projection of the GW waveform originating from the same astrophysical system, effectively doubling the number of detectors that observe the same event (for a pair of lensed events) and allowing us to measure the relative amplitudes of additional polarization modes. To measure these amplitudes, we jointly fit the lensed image observations to a single GW signal model, taking into account the image magnifications, time delays, and polarization mode amplitudes. We show that for specific GW signals from binary black hole mergers, we can make a measurement of the relative mode amplitudes for strongly lensed events with at least two detectable images.

6.1 INTRODUCTION

The latest set of gravitational–wave (GW) observations released by the LIGO Scientific (**?**), Virgo (Acernese et al., 2015b) and KAGRA (Aso et al., 2013) Collaboration (LVK) as part of The third

Gravitational-wave Transient Catalog (GWTC-3) catalog (Abbott et al., 2019b, 2021c,d,e) contains 69 confident binary black hole (BBH) detections as well as both confident detections for binary neutron star and neutron star black hole mergers. As a consequence, the increasing size of gravitational wave catalogs has allowed for in-depth studies of the binary black hole population properties (Abbott et al., 2019a, 2021f,j), cosmic expansion history (Abbott et al., 2021a,b) as well as tests of general relativity in the strong field regime (Abbott et al., 2019c, 2021h,i) including a search for gravitational wave lensing signatures (Abbott et al., 2021g).

When gravitational waves propagate and interact with intervening matter such as galaxies or dense galaxy clusters, there is a change for strong gravitational lensing and for multiply lensed GW images to be produced with time delays ranging from minutes to months (Takahashi & Nakamura, 2003; **?**; Dai et al., 2017). Over the upcoming years, ground-based GW detectors such as Advanced LIGO, Advanced Virgo and KAGRA are expected to find 0.1 to 1 pairs of strongly lensed GW signals per year originating from binary black hole mergers at their corresponding design sensitivities (Ng et al., 2018; Abbott et al., 2021g; Xu et al., 2022; Çalışkan et al., 2022; Mukherjee et al., 2021a). In fact, the first search for signatures of lensing (including strongly lensed pairs) was performed in **?** using the 10 BBH events of the GWTC-1 catalog (Abbott et al., 2019b). No conclusive evidence for a strongly lensed pair was found, however, the pair with the highest evidence favoring the lensing hypothesis was GW170104/GW170814 as pointed in (**?**McIsaac et al., 2020). Subsequent studies followed up the pair with a fully Bayesian joint parameter estimation study over the lensed images and arrived at similar conclusions disfavoring the lensing hypothesis (Liu et al., 2021; Dai et al., 2020). The most comprehensive study to date using the first half of LIGO-Virgo's third observation run observations has also yielded no substantial evidence for lensing (Abbott et al., 2021e,g)

Alternative metric theories of gravity predict up to six distinct polarization modes for GW emission, besides the two tensorial modes allowed by general relativity (Isi et al., 2017; Chatzi-

ioannou et al., 2012). In order to probe the presence (or lack off) for these alternative polarizations, a network of six linearly independent detectors is needed. Future ground based detector networks will allow for some statements about the relative amplitudes for each mode, however, discerning the full polarization content would be difficult for most systems (Chatziioannou et al., 2021). The most recent observational results using the full GWTC-3 catalog have placed stringent constrains on alternative polarizations being present (Abbott et al., 2019c, 2021h,i). The strongest of such constraints disfavor the presence of vector or scalar modes being present individually when compared to the expected GR tensor modes. However, the presence of tensor modes as well as either vector or scalar modes (or both) as a fully mixed model has yet to be constrained strongly.

In this work, we explore constraints on alternative GW polarizations with simulated pairs of strongly lensed GW signals. We parameterize the GW model as a fully mixed tensor, vector, and scalar mode model with up to 5 degrees of freedom allowing us to make statements about the relative amplitudes for each mode. The difference in arrival times for each strongly lensed image probes the same GW signal arriving at different times. The rotation of the Earth imposes the time dependence of the antenna beam pattern functions, allowing us to see a different projection for the GW signal at each detector (essentially doubling the number of detectors for a pair of lensed events) (Goyal et al., 2021). We measure the relative amplitudes for each mode by jointly fitting the detected lensed image pairs using the framework described in (Liu et al., 2021; Lo & Magaña Hernandez, 2021) and show that for some systems the polarization mode amplitude degeneracies can be broken with a single pair of lensed events.

This paper is organized as follows. In Section 2 we describe the alternative (non-tensorial) polarization modes for gravitational-wave signals. In Section 3, we summarize the effect of strong lensing in detected GW signals, focusing on pairs of lensed events. In Section 4, we present the main results of this paper and in Section 5, we provide a summary of this work. We use the a flat Planck 2015 cosmological model (Ade et al., 2016b) throughout this paper, that is, $H_0 =$

$67.8\ \mathrm{km/s/Mpc}$ and $\Omega_m = 0.308$.

6.2 NONTENSOR POLARIZATIONS

Alternative metric theories of gravity (beyond general relativity) may allow up to six distinct polarization modes on the GW waveform, including the two tensor $+$ and $\times$ modes expected in general relativity (GR). These additional polarization modes are the two vector modes x and y, as well as two scalar modes b and l (breathing and longitudinal respectively). The GW perturbation can thus be written as,

$$h_{ij} = \sum_A h_A \, e^A_{ij} \,, \tag{6.1}$$

where e^A_{ij} is the polarization tensor for mode A and h_A are the corresponding polarization mode amplitudes. The GW perturbation is thus a linearly independent weighted sum over modes, the most generic case corresponding to $A \in \{+, \times, x, y, b, l\}$.

In general, GW interferometers measure the projection of the perturbation given by Eq. (6.1) onto the detector arms. Thus the measured GW strain at detector I can be written as,

$$h_I(t) = \sum_A F^A_I(\alpha, \delta, \psi, t) \, h_A(t) \,, \tag{6.2}$$

with antenna beam pattern functions $F^A_I \equiv D^{ij}_I e^A_{ij}$ defined with respect to the detector tensor D^{ij}_I which encodes the geometry of the GW detector. The antenna pattern functions are in general functions of time and depend on the sky location of the GW source defined by its right ascension α and declination δ as well as the polarization angle ψ. It is worth noting that the breathing and longitudinal mode antenna pattern functions are identical (up to a constant) so that $F_b = -F_l$. This degeneracy makes each scalar mode contribution difficult to disentangle unless a specific modified theory of gravity is chosen a-priori, leading to model-dependent constraints. Following convention we pick the breathing mode as the scalar mode of interest, thus the sum over linearly

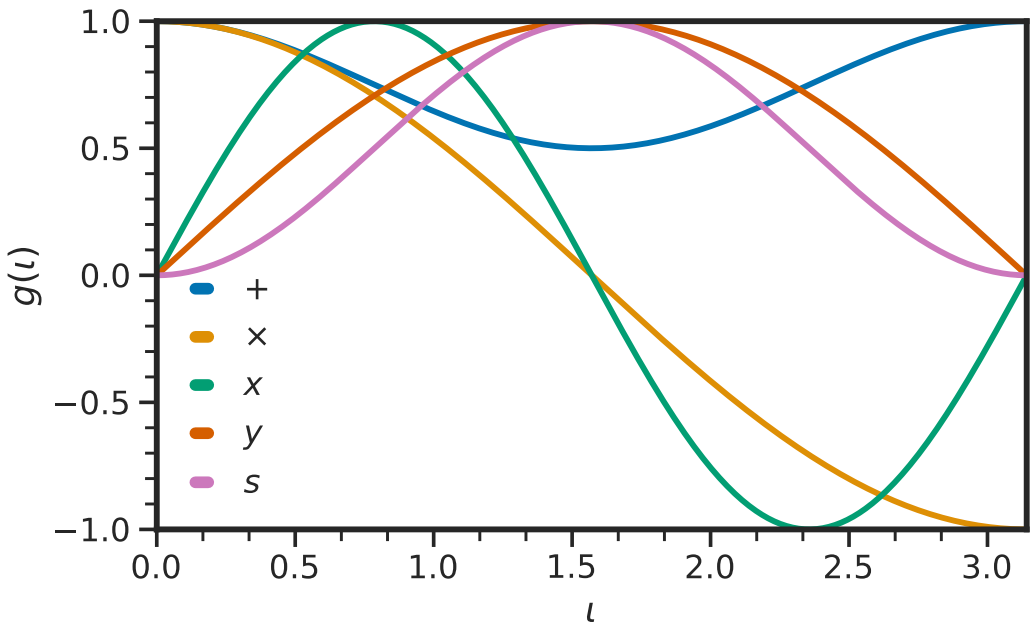


Figure 6.1: We show the values for $g(\iota)$ as a function of the inclination angle ι for each of the six polarization modes. We note that for face-on systems ($\iota = 0$), the vector and scalar modes will not be present in the GW strain data even if emitted. For edge-on systems ($\iota = \pi/2$), $\times$-mode and the vector-x will not be detectable. The optimal inclination angle for which we maximize over the presence of all modes in the GW data corresponds to $\iota_{\rm opt} \approx 0.87$.

independent modes in Eq. (6.1) reduces to a sum over five polarization modes, $A \in \{+, \times, x, y, s\}$ where we denoted the breathing mode (b) by (s) for convenience. For a detailed discussion on GW polarizations and the various polarization angle conventions, we refer the reader to (Isi et al., 2017; Isi, 2022).

For gravitational waves produced by a compact binary merger such as a pair of merging binary black holes, there is an additional inclination angle dependence for each polarization mode (Chatziioannou et al., 2012; Takeda et al., 2021). We define this dependence via the function $g_A(\iota)$, so that $g_+(\iota) = (1 + \cos^2\iota)/2$, $g_\times(\iota) = \cos\iota$, $g_x(\iota) = \sin 2\iota$, $g_y(\iota) = \sin\iota$ and $g_{b,l}(\iota) = \sin^2\iota$ where ι is the inclination angle of the binary. We can thus write the gravitational wave strain at

detector I as,

$$h_I(t) = \sum_A F_I^A(\alpha, \delta, \psi, t)\, g_A(\iota)\, h_A(t)\,. \tag{6.3}$$

From the above expression, we can see that the inclination angle dependence on the polarization modes is important since for a face-on system ($\iota = 0$) only the tensor modes will be present in the data while for an edge-on system ($\iota = \pi/2$), only the cross-polarization mode vanishes but all other modes are present. The inclination angle dependence is critical for 2nd generation ground-based detectors since we expect most mergers to be near the face-on limit. In Fig. 6.1, we plot the dependence on inclination for the mode amplitudes. Clearly, if all other parameters are fixed, then the optimal inclination would be $\iota_{\rm opt} \approx 0.87$.

6.3 STRONG GRAVITATIONAL WAVE LENSING

As gravitational waves propagate, there is a chance for strong gravitational lensing to occur due to intervening galaxies or larger cosmic structures such as galaxy clusters. The strong lensing of gravitational waves can give rise to multiple images of the same GW transient each with its own absolute magnification factor μ_k. When the GW images are detected, each will arrive at a different time $t_c^{(k)}$ and each might have a frequency-independent phase shift (Morse phase) $\Delta\phi_k = -\pi n_k/2$ with index $n_k = 0, 1, 2$ defining Type-I, Type-II and Type-III images respectively. The gravitational wave waveform for each lensed image is then given by,

$$h_L(f, \theta, \mu_k, t_c^{(k)}, \Delta\phi_k) = \sqrt{\mu_k}\exp\left(if\Delta\phi_k\right) h_U(f, \theta, t_c^{(k)}) \tag{6.4}$$

where h_U is the waveform without any strong lensing effects (unlensed)

and $\theta = \{m_1, m_2, a_1, a_2, \iota, \alpha, \delta, \psi\}$ where m_1 and m_2 are the primary and secondary masses of the binary in the source frame, a_1 and a_2 are the (aligned) component spin magnitudes. The set of parameters θ is common across all lensed images, including the sky location of the GW source due

to the expected order of arcsecond deflection angles for each image being much smaller than the typical localization regions for 2G detectors (Takahashi & Nakamura, 2003).

Now, for a source at luminosity distance D_L, the lensed images are magnified (de-magnified) by their corresponding magnification factors $\sqrt{\mu_k}$ as in Eq. (6.4) so that the observed distances correspond to,

$$D_{\rm obs}^{(k)} = D_L/\sqrt{\mu_k}\,, \tag{6.5}$$

clearly showing the degeneracy between the luminosity distance to the source and the absolute magnification factors for each lensed image. For a pair of lensed images, it is convenient to define the relative magnification factor μ as,

$$\mu = \left(\frac{D_{\rm obs}^{(1)}}{D_{\rm obs}^{(2)}}\right)^2 = \frac{\mu_2}{\mu_1}\,, \tag{6.6}$$

where we label the signal that is detected first by (1) and consequently the later arriving signal by (2). Finally, we can also define the lensing time delay for the pair of lensed images as $\Delta t = t_c^{(2)} - t_c^{(1)}$ which is always greater than zero.

The most important effect due to strong lensing and the production of multiple images is the different times of arrival for each lensed image. Due to Earth's rotation, the location of the network of detectors will change as a function of time relative to the sky location of the lensed signals. This means that the antenna pattern functions for each polarization mode will probe the polarization content of the arriving GW signal differently depending on the arrival time of each lensed image, allowing us to constrain the relative amplitudes for each polarization mode (Goyal et al., 2021). In principle, this leads to effectively doubling the number of detectors in the network for a pair of strongly lensed images. In order to illustrate this point we show the antenna pattern functions $F_A^2(t)$ in Fig. 6.2 at a fixed polarization angle $\psi = 0$ for two different sky locations over a period of two days. As mentioned in Section 6.1, the expected time delay between a pair of lensed events by

an intervening galaxy could range from hours to months. With respect to probing the polarization amplitudes for each mode, the expected time delay is not important but the relative time delay corresponding to the rotation of the Earth over a day.

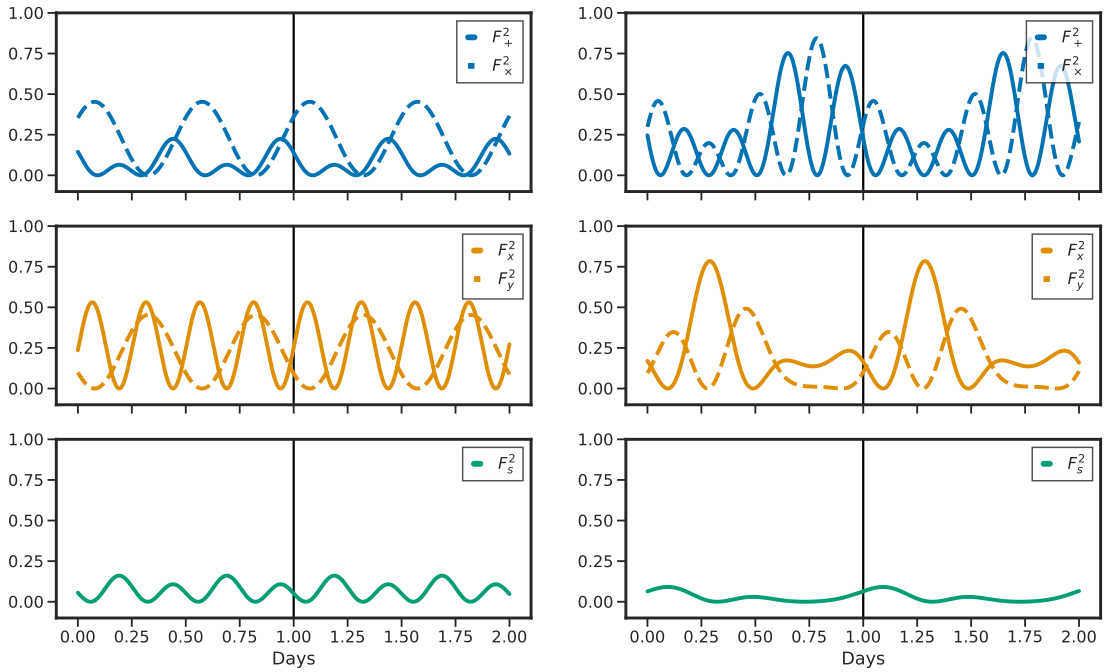


Figure 6.2: We show the values for $F_A^2(t)$ for the six polarization modes over the span of two days where we have fixed the polarization angle to $\psi = 0$ for convenience. In the left panel, we show an example where the sky location of the source is fixed at $(\alpha, \delta) = (0, 0)$. Similarly, we show another example but with $(\alpha, \delta) = (1.375, -1.211)$ to illustrate the complex behavior of the antenna beam pattern functions in terms of sky location and time.

6.4 JOINT PARAMETER ESTIMATION

Since strongly lensed systems leave the frequency evolution of the gravitational-wave binary unchanged and thus only induce an overall amplitude and phase difference amongst the detected images. We are thus able to jointly fit the lensed events by taking into account the predicted strong lensing effects on the GW waveform. We provide a summary of the joint parameter estimation

below. For the full derivations and detailed discussion of the framework, see (Liu et al., 2021; Lo & Magaña Hernandez, 2021).

Under the assumption that we have a confidently detected pair of strongly lensed GW events. We can perform joint parameter estimation by considering the strong lensing waveform model for each detected image in Eq. (6.4) and use the model with alternative polarizations as defined in Eq. (6.3) as the definition for $h_U(t, \theta)$. Additionally, we parameterize each polarization mode amplitude by a set of relative amplitude parameters $\{\epsilon_A\}$ which must satisfy the following constraint $\sum_A \epsilon_A = 1$. For GR, we must have $\epsilon_+ = \epsilon_\times = 0.5$ while the vector and scalar mode contributions are all zero.

Under the lensing hypothesis, for a pair of lensed events with measured strains d_1 and d_2, we jointly infer the binary parameters θ, the lensing observables $\{\mu, \Delta t\}$ (in this work we set $\Delta\phi_k = 0$ for all images) as well as the relative amplitudes for each polarization mode $\{\epsilon_A\}$ in terms of the observed distance and time of arrival of the first image,

$$\mathcal{L}(d_1, d_2|\theta, D_{\rm obs}^{(1)}, t_c^{(1)}, \mu, \delta t, \{\epsilon_A\}) = \mathcal{L}(d_1|\theta, D_{\rm obs}^{(1)}, t_c^{(1)}, \{\epsilon_A\})\mathcal{L}(d_2|\theta, \mu, \delta t, \{\epsilon_A\}), \tag{6.7}$$

where $\mathcal{L}(d_1, d_2|\ldots)$ is referred to as the strong lensing joint likelihood. We note that this can be generalized to an arbitrary number of lensed images and we refer the reader to (Lo & Magaña Hernandez, 2021) for more details. To obtain the posterior distribution over the parameters describing the joint likelihood function we use Bayes theorem and defer the details of our choice for the prior distribution to Section 6.5.

6.5 RESULTS

We perform joint parameter estimation to estimate the posterior distribution on the parameters defined through the strong lensing joint likelihood function as defined in Eq. (6.7). As an example,

we simulate a pair of lensed GW images from a non-spinning binary black hole merger with the following intrinsic parameters: $m_1^{\rm det} = 36M_\odot$, $m_2^{\rm det} = 29M_\odot$ and $a_1 = a_2 = 0$. The extrinsic parameters for the simulated system are $\psi = 2.659$, $\phi_c = 2.9$, $\alpha = 1.375$, $\delta = -1.2108$ and $\iota = \pi/4$. We have chosen a sky location for the merger consistent with the beam pattern functions as shown in the right panel of Fig. 6.2 and an inclination angle close to the value of $\iota_{\rm opt}$ in order to not suppress the extra polarization modes through the inclination dependence introduce via the $g_A(\iota)$ factors.

As discussed in 6.4, we sample over the observed distance to the first event $D_{\rm obs}^{(1)}$ and the relative amplification factor for the pair μ. The first lensed pair has $D_{\rm obs}^{(1)} = 1000$ Mpc and $\mu = 2$ (corresponding to $D_{\rm obs}^{(2)} = 500$ Mpc) with a time delay $\Delta t = 6$ hours. We set the Morse index for both images to zero (both Type-I) for simplicity. For the polarization mode amplitudes we choose, $\epsilon_+ = \epsilon_\times = 0.35$, $\epsilon_x = \epsilon_y = 0.15$ and $\epsilon_s = 0.05$.

We consider two examples, a 2-detector network composed of LIGO Hanford and LIGO Livingston (HL) and a 4-detector network, with Advanced Virgo and KAGRA as additional detectors (HLVK), all at their corresponding design sensitivities. We generate the GW waveform using the `TaylorF2` (Damour et al., 2001) waveform model for simplicity and inject the two lensed GW signals into simulated data streams with Gaussian noise and sample over the strong lensing joint likelihood using `Bilby` (Ashton et al., 2019; Romero-Shaw et al., 2020). For the 4-detector network, we show in Fig. 6.3 the marginalized posterior distribution on the relative polarization mode amplitudes, inclination angle, relative magnification factor and the observed distance to the first image (See Appendix C for our prior choices as well as full parameter estimation results in Fig. C.1 for the HLVK case and in Fig. C.2 for the HL case). It is evident from the posterior distribution shown in Fig. 6.3 that the relative polarization mode amplitudes can be measured with a single pair of lensed events using a 4-detector network at design sensitivity. For the example with a 2-detector network observing the same system, the polarization mode amplitudes cannot be fully

constrained due to the lack of linearly independent detectors (in principle four but both the Hanford and Livingston detectors are nearly co-aligned).

6.6 DISCUSSION

In this work, we have performed Bayesian joint parameter estimation on pairs of strongly lensed GW events in order to constrain the relative amplitudes for alternative polarization modes using simulated data. We have used a simplified signal model as a proxy for the signal morphology for the additional polarization modes, and have also made sure to include the expected inclination angle dependence for each mode for GWs emitted by a merging binary. We have shown that the relative amplitudes, as well as the amplitude-relevant parameters such as the observed distance, inclination angle, and relative amplification factor for the lensed pair, can be measured since the additional data from the same astrophysical system provides enough independent detectors to measure the aforementioned parameters.

Strongly lensed pairs of GW signals for binary black hole mergers are expected to be detected as early as O4 but more likely in O5. Once a confident detection has been established, the joint parameter estimation framework described in this work can be applied to a real lensed pair of GW signals. However, we do mention that a proper treatment of real GW data will involve the strong lensing joint likelihood with a model-independent framework to describe the GW signal morphology as explored in Chatziioannou et al. (2021) which used `bayeswave` to model the GW signal morphology using sine Gaussians. Given that, the results of this paper can be seen as being slightly more pessimistic than what they would be if any alternative polarization modes are present in the data with significantly different signal morphology. The varying morphology should allow for the relative mode amplitude degeneracy to be broken, however, using a specific modified gravity model that predicts additional polarization modes for the Bayesian inference would make the results model dependent.

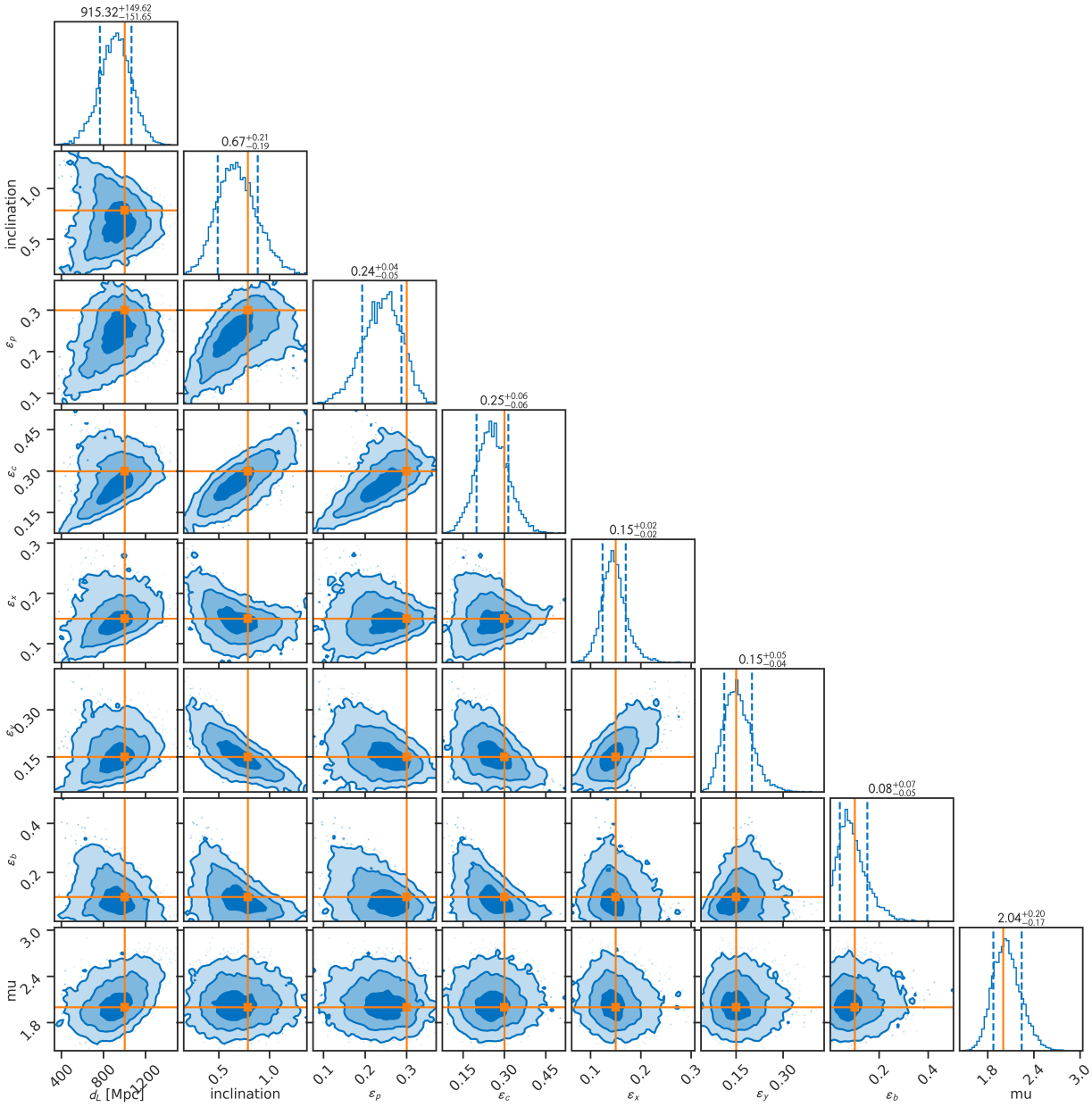


Figure 6.3: Marginalized posterior distribution for the pair of lensed images as described in 6.5 observed by four detectors (HLVK) on the relative amplitudes for each polarization mode, inclination angle, relative magnification factor and the observed distance of the first image. The simulated system has $\epsilon_+ = \epsilon_\times = 0.35$, $\epsilon_x = \epsilon_y = 0.15$ and $\epsilon_s = 0.05$ for the polarization mode amplitude, $D^{(1)}_{\rm obs} = 1000$ Mpc, $\mu = 2$ and $\iota = \pi/4$ (shown in orange) with a relative time delay of six hours.

Chapter 7

Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers

In modified gravity models that allow for additional noncompact spacetime dimensions, energy from gravitational waves can leak into these extra spacetime dimensions, leading to a reduction in the amplitude of the observed gravitational waves, and thus are a source of potential systematics in the inferred luminosity distances to gravitational wave sources. Since binary black hole (BBH) mergers are standard sirens, we use the pair-instability supernova mass gap and its predicted features to determine a mass scale in order to break the mass-redshift degeneracy and thus infer the redshift of the source. We simultaneously fit the BBH population and the extra spacetime dimensions parameters from gravitational leakage models using BBH observations from the recently released GWTC-3 catalog. We set constraints on the number of spacetime dimensions and find that $D = 3.93^{+0.08}_{-0.05}$ at 68% C.L. for models that are independent of a screening scale, finding that the GWTC-3 constraint is as competitive as that set from GW170817 and its electromagnetic counterpart. For models where gravity leaks below a certain screening scale R_c, we find $D = 4.49^{+1.63}_{-0.87}$ and $\log_{10} R_c/\mathrm{Mpc} = 4.75^{+0.86}_{-0.99}$ with a transition steepness $\log_{10} n = 0.85^{+0.76}_{-0.86}$ for the leakage, which for the first time are constrained jointly with the BBH population at cosmological distances. These constraints are consistent with General Relativity (GR), where gravitational waves propagate in $D = 3 + 1$ spacetime dimensions. Using the BBH population to probe modifications to standard cosmological models provides an independent test of GR that does not rely on any electromagnetic information but purely on gravitational wave observations.

7.1 INTRODUCTION

With the recent release of the GWTC-3 catalog (Abbott et al., 2019b, 2021c,d,e) from the LIGO Scientific, Virgo and KAGRA Collaborations(LVK), the number of gravitational wave detections from binary black hole mergers has increased to 90 confidently detected events (Aasi et al., 2015; Acernese et al., 2015b). The increasing size of GW catalogs has enabled the study of the binary black hole (BBH) population (Abbott et al., 2021f,j), its cosmic expansion history (Abbott et al., 2021a,b), signatures of gravitational wave lensing (Abbott et al., 2021g) and how well the population agrees with General Relativity (Abbott et al., 2021h,i).

Gravitational wave sources provide a direct measurement of their luminosity distance—they are standard sirens (Schutz, 1986; Holz & Hughes, 2005). Uniquely associated electromagnetic counterparts can constrain the redshift of the source and hence allow for an independent determination of the corresponding electromagnetic luminosity distance, as was the case for the first bright standard siren GW170817 (Abbott et al., 2017c). Combined gravitational wave (GW) and electromagnetic (EM) counterpart measurements have allowed constraints (to mention a few) of cosmological parameters such as the Hubble constant (Abbott et al., 2017a), the number of extra spacetime dimensions allowed under gravitational leakage models (Pardo et al., 2018) and modified gravitational wave propagation due to a running Planck mass (Lagos et al., 2019).

For gravitational wave events without expected EM counterparts (dark sirens), including the numerous binary black hole mergers, in absence of a uniquely identified host galaxy, a galaxy survey can be used as prior information on the potential host galaxies of the event in combination with its gravitational wave localization volume (Del Pozzo, 2012; Nair et al., 2018; Chen et al., 2018; Fishbach et al., 2019b; Gray et al., 2019; Soares-Santos et al., 2019c; Abbott et al., 2021a; Palmese et al., 2020; Mukherjee et al., 2021b; Diaz & Mukherjee, 2022). The dark siren methodology has been used to constrain cosmology (Soares-Santos et al., 2019c; Palmese et al., 2020; Abbott et al., 2021a,b; Palmese et al., 2023) as well as to measure modified gravitational wave

propagation (Finke et al., 2021; Mukherjee et al., 2021c).

However, even without electromagnetic information, one can still make a statistical measurement of redshift using the features of the population distribution of compact binary mergers. Since we measure redshifted "detector-frame" masses, $m_1^{\rm det} = m_1(1+z)$, one can model the expected source frame mass distribution (given our current understanding of BBH formation channels) on m_1 to estimate the redshift z. This idea was first explored in Taylor et al. (2012) in the context of binary neutron star mergers and 3G detectors to constrain cosmology, where the "known" galactic neutron star mass distribution (a sharply peaked Gaussian with mean around 1.4 $M_\odot$) was used as a feature to break the mass-redshift degeneracy.

More recently, Farr (2019), applied this methodology to BBH mergers by using the predicted cut-off and excess of black holes with masses $M \approx 45 M_\odot$ in the BH mass spectrum due to the pair-instability supernovae (PISNe) mass gap (Woosley et al., 2002; Heger & Woosley, 2002; Heger et al., 2003; Woosley, 2017, 2019). Recent work in Ezquiaga (2021), used the same methodology to constrain the value of c_M (in general relativity, $c_M = 0$), a parameter that allows for modifications of ΛCDM due to a time-varying Planck mass (Lagos et al., 2019) as well as other modified gravity parametrizations (Mancarella et al., 2022). The LVK collaboration, subsequently applied the methodology of Farr (2019) to the GWTC-3 catalog to constrain the cosmic expansion history of BBH mergers using astrophysically motivated source frame mass models (Abbott et al., 2021b).

In this work, we set constraints on gravitational leakage models, specifically, on the number of extra non-compact spacetime dimensions where gravity may leak, using the observed population of BBH mergers in GWTC-3. In order to determine the redshift for these events, we fit the BBH population with an astrophysically motivated PISNe mass model together with modifications to the gravitational wave luminosity distance induced by the gravitational leakage models we consider.

This Chapter is organized as follows. In Section 7.2 we describe the gravitational leakage models and how these relate to the damping of the gravitational wave amplitude. In Section 7.3.1,

we summarize the hierarchical Bayesian framework used in our analysis and provide the details for the BBH population model that we use. In Section 7.4, we present the main results of this paper, and in Section 7.5, we provide a summary of this work. We use the Planck 2015 Ade et al. (2016b) cosmological model throughout this paper, that is, a spatially flat ΛCDM model with $H_0 = 67.8\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $\Omega_m = 0.308$.

7.2 GRAVITATIONAL LEAKAGE MODELS

In this section, we describe the different gravitational leakage models and how these relate to higher spacetime dimensional theories. We consider how these models modify the gravitational waveform amplitude, and hence correspondingly can bias the observed luminosity distance to the source. This section relies heavily on the work of Deffayet & Menou (2007); Pardo et al. (2018); Corman et al. (2022).

In General Relativity the gravitational wave strain is proportional to the luminosity distance d_L^{GW} to the source as,

$$h_{\mathrm{GR}} \propto \frac{1}{d_L^{\mathrm{GW}}}, \tag{7.1}$$

For a higher-dimensional spacetime theory where there is some leakage of gravity, one would expect damping of the gravitational waveform in the form of a power-law due to flux conservation, so the simplest phenomenological model to consider is (Pardo et al., 2018; Corman et al., 2022):

$$d_L^{\mathrm{GW}} = d_L \left(\frac{d_L}{1\ \mathrm{Mpc}} \right)^{(D-4)/2}, \tag{7.2}$$

where D is the number of spacetime dimensions, d_L is the D-dimensional luminosity distance, and d_L^{GW} is the measured GW luminosity distance (from parameter estimation analyses).

However, one usually parameterizes this model so that below a certain length scale the space-time becomes four-dimensional (Corman et al., 2022). So following the standard phenomenology,

we define a screening scale R_c for the gravitational leakage as well as the overall transition steepness n, which determines the strength of the leakage. So more generally, following (Pardo et al., 2018; Corman et al., 2022):

$$d_L^{\rm GW} = d_L \left[1 + \left(\frac{d_L}{R_c}\right)^n\right]^{(D-4)/(2n)}, \tag{7.3}$$

this relation reduces to model in Equation 7.2 for $d_L \gg R_c$ and to $d_L^{\rm GW} = d_L$ for $d_L \ll R_c$.

To understand how the leakage model of Equation 7.3 affects the detectability of a population of BBH mergers, we compute $p(z|\text{det})$, the probability of detecting a BBH at a given redshift z, since the corresponding luminosity distance $d_L^{\rm GW}$ will change as a function of D, R_c and n. By fixing the BBH population to a fiducial set of parameters that correspond to the "POWER LAW + PEAK" model (see Sec. 7.3.2) and determining the values of d_L (given $d_L^{\rm GW}$) as a function of D while fixing R_c and n constant, we show how the detectable fraction of mergers depends on the number of spacetime dimensions in the top panel of Figure 7.1. At fixed $R_c = 100$ Mpc and $R_c = 1000$ Mpc with $n = 2$, and we note that as D increases from its GR value ($D = 4$), the maximum detectable redshift, as well as its peak value, decreases for the detectable population. This makes sense, as energy from GW sources leaks for higher dimensional theories, leading to a reduction in their measured amplitude and subsequently a lower detectable horizon.

Similarly, we show the dependence of detectability as a function of R_c in the bottom panel of Figure 7.1, while we fix the spacetime dimensions to $D = 5$ and consider $n = 2$ and $n = 50$ as examples for the transition steepness. As R_c increases, the maximum and peak redshift for the detectable distribution increases, where for $R_c \approx 10^4$ Mpc the model resembles GR-like behavior for $D = 5$ and is independent of the value for n. If there is gravitational leakage, we expect the size of R_c to be of cosmological scales, since small screening scales ($R_c < 20$ Mpc) have been ruled out by the analysis of GW170817 in Pardo et al. (2018) irrespective of the value for n.

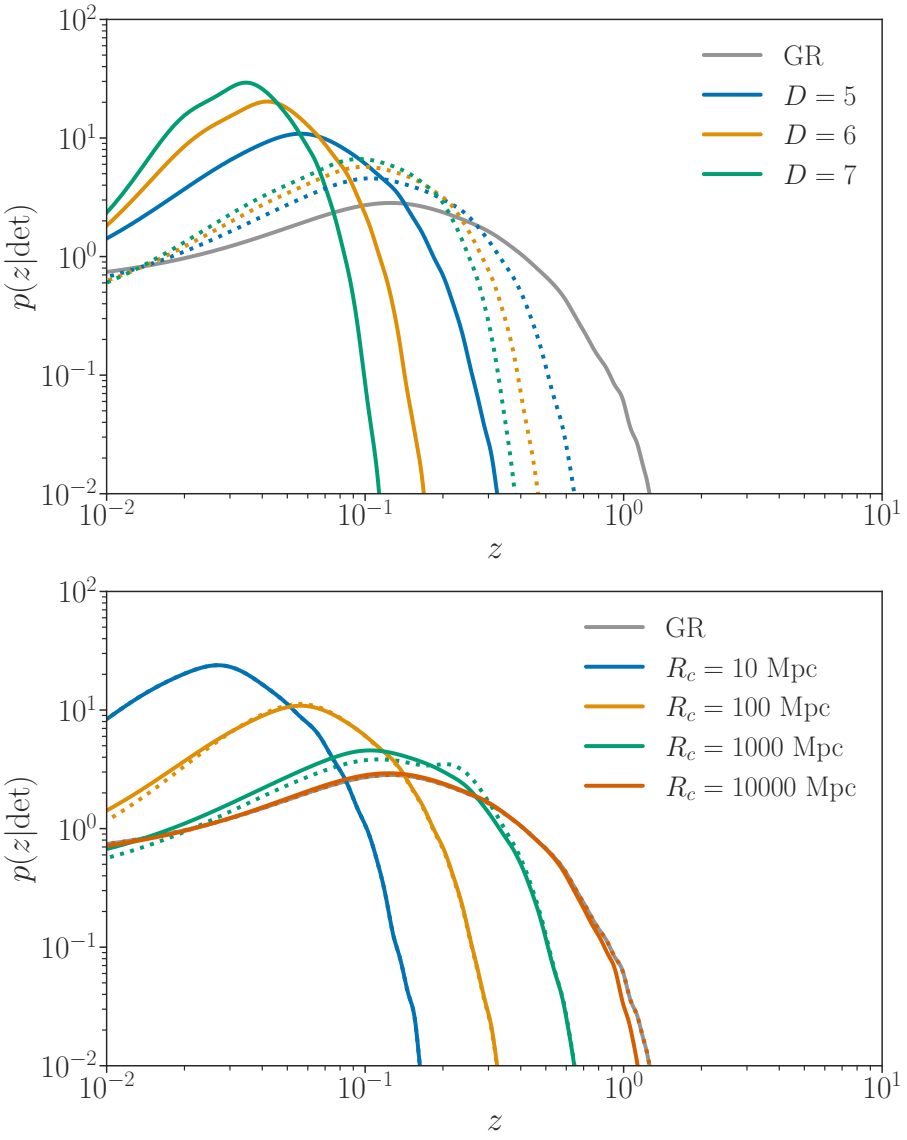


Figure 7.1: Top panel: Detectability of BBH mergers as a function of redshift z when varying the spacetime dimension D at fixed R_c and with $n = 1$. Solid lines have a screening scale $R_c = 10^2$ Mpc, while dashed lines have $R_c = 10^3$ Mpc for comparison. Bottom panel: Detectability as a function of redshift given the screening scale R_c at fixed spacetime dimension $D = 5$ and transition steepness $n = 2$ (solid-lines) and $n = 500$ (dashed-lines). We fix the population of BBH mergers to the "POWER LAW + PEAK" model with parameters $m_{\rm min} = 5M_\odot$, $m_{\rm max} = 70M_\odot$, $\alpha = 3$, $\beta = 1.5$, $m_{\rm pp} = 35M_\odot$, $\sigma_{\rm pp} = 2M_\odot$ and $f_{\rm pp} = 0.03$ and to the SFR-like redshift evolution model with parameters $\gamma = 2.7$, $\kappa = 3$ and $z_{\rm p} = 2.4$.

Finally, it is also worth investigating theories in which the graviton has a finite lifetime, meaning that it decays away as it travels cosmological distances. In this case, the modified luminosity distance of the GW would scale as (Pardo et al., 2018):

$$d_L^{\rm GW} = d_L e^{d_L/R_g} \tag{7.4}$$

where R_g is the "decay-length". Consequently, one can measure the graviton's "decay-time"' as long as we assume that they propagate at the speed of light, that is, $t_g = R_g/c$. We leave constraints from decaying graviton models as future work.

7.3 METHODS

7.3.1 Hierarchical Inference

We use hierarchical Bayesian inference to simultaneously infer the parameters for the population distribution of binary black hole mergers as well as the additional parameters that describe the gravitational leakage models described in Section 7.2. The binary black hole observations provide us with an estimate of their primary mass $m_1^{\rm det}$ and mass ratio q in the detector frame, as well as their luminosity distances d_L.

The number density of BBH events as a function of detector frame quantities $(m_1^{\rm det}, m_2^{\rm det}, d_L)$ is related to the source frame parameters (m_1, q, z) by,

$$\frac{dN(m_1^{\rm det}, q, d_L|\Lambda)}{dm_1^{\rm det} m_2^{\rm det} dd_L} = \frac{1}{m_1(1+z)^2}\frac{dz}{dd_L}\frac{dN(m_1, q, z|\Lambda)}{dm_1 dq dz} \tag{7.5}$$

with Λ being the population hyperparameters that we want to measure and the proportionality factor relating detector frame to source frame quantities is the Jacobian transformation relating these parametrizations.

Now we relate the BBH number density in terms of the BBH merger rate density:

$$\frac{dN(m_1, q, z|\Lambda)}{dm_1 dq dz} = \frac{dV_c}{dz}\left(\frac{T_{\rm obs}}{1+z}\right)\frac{d\mathcal{R}(m_1, q, z|\Lambda)}{dm_1 dq} \tag{7.6}$$

where the BBH merger rate $d\mathcal{R}$ over a range of primary mass, mass ratio, and redshift (assuming the BBH mass distribution is redshift independent) gives:

$$\frac{d\mathcal{R}(m_1, q, z|\Lambda)}{dm_1 dq} = \mathcal{R}_0 p(m_1, q|\Lambda)\psi(z|\Lambda), \tag{7.7}$$

where $\mathcal{R}_0$ is the local merger rate at $z = 0$. The BBH population is modeled through the normalized mass distribution $p(m_1, q|\Lambda) = p(m_1|\Lambda)p(q|m_1, \Lambda)$ and its redshift evolution $\psi(z|\Lambda)$, which is chosen such that $\psi(z = 0|\Lambda) = 1$. Here dV_c/dz is the differential uniform-in-comoving volume element, $T_{\rm obs}$ the total observation time, and the factor of $1/(1+z)$ converts the source-frame time to detector-frame time. By integrating the BBH number density across all primary masses and mass ratios, and out to a maximum redshift $z_{\rm max}$ we get the expected number of BBH within $z_{\rm max}$. In this work, we take $z_{\rm max} = 4$ throughout.

Given a set of $N_{\rm obs}$ gravitational wave observations $\{d_i\}$, we can calculate the posterior on Λ following e.g. Farr (2019) and Mandel et al. (2019):

$$p\left(\Lambda|\{d_i\}\right) \propto p(\Lambda)e^{-\mathcal{R}_0\xi(\Lambda)}\prod_{i=1}^{N_{\rm obs}}\left[\int \mathcal{L}\left(d_i|m_1^i, q^i, z^i\right)\frac{dN(m_1, q, z|\Lambda)}{dm_1 dq dz}dm_1 dq dz\right], \tag{7.8}$$

where $\mathcal{L}(d_i|m_1, q, z)$ is the single-event likelihood function for each event, and $\xi(\Lambda)$ is the detectable fraction of sources corresponding to a population determined by the population hyperparameters Λ.

Following Farr (2019) to estimate $\xi(\Lambda)$, we assume that sampling of ξ will follow a normal distribution (i.e. $\xi(\Lambda) \sim \mathcal{N}(\Lambda|\mu, \sigma)$), with μ the importance sample estimate of ξ and σ its as-

sociated uncertainty. Practically, we estimate the detectable fraction $\xi(\Lambda)$, by using the LVK's injection campaign of BBH events simulated from a broad BBH population and injected into real detector data, then searched for using the same analysis pipelines that found GWTC-3 and prior GW Transient catalogs (Abbott et al., 2021j).

Finally, we marginalize over the local BBH merger rate $\mathcal{R}_0$ using a log-Uniform prior (Abbott et al., 2021f), and neglect terms of $\mathcal{O}(N_{\rm eff}^{-2})$ (Farr, 2019). We also approximate the integral over the individual event likelihoods in Equation 7.8 with importance sampling over N_i single-event posterior samples generated from single event inference analysis with default event prior $\pi(m_1, q, z) \propto d_L^2 m_1 (1+z)^2 \frac{dd_L}{dz}$. [1]

7.3.2 Binary Black Hole Population Models

We model the black hole mass distribution in the source frame using the "POWER LAW + PEAK" model as described by the LVK second gravitational-wave transient catalog (GWTC-2) population analysis (Abbott et al., 2021f,j) and in Talbot & Thrane (2018). It is a mixture distribution consisting of a power law distribution truncated at a maximum mass to model the PISNe mass gap and of a Gaussian distribution that models the build-up of black holes due to the pulsational PISNe mass loss (Talbot & Thrane, 2018). For simplicity, we neglect the low-mass smoothing feature used in (Abbott et al., 2021f).

Under this model, the probability distribution for the primary mass is,

$$p(m_1|\alpha, m_{\rm min}, m_{\rm max}, f_{\rm pp}, m_{\rm pp}, \sigma_{\rm pp}) = f_{\rm pp} p_{\rm pp}(m_1|m_{\rm pp}, \sigma_{\rm pp}) + (1 - f_{\rm pp}) p(m_1|\alpha, m_{\rm min}, m_{\rm max}) \,, \tag{7.9}$$

where $f_{\rm pp}$ is a mixing fraction parameter that gives the weight of the Gaussian component (Abbott

[1] Note that we have absorbed the Jacobian transformation in Equation 7.5 in the definition for $\pi(m_1, q, z)$ as it is typically done.

et al., 2021f). The power law distribution is defined as,

$$p(m_1|\alpha, m_{\rm min}, m_{\rm max}) \propto (m_1)^{-\alpha}\,, \tag{7.10}$$

where α is the powerlaw index and $m_{\rm min}$ ($m_{\rm max}$) is the minimum (maximum) black-hole mass with the constraint that $m_{\rm min} < m_1 < m_{\rm max}$. The Gaussian component has a mean $m_{\rm pp}$ and a standard deviation $\sigma_{\rm pp}$ and is given by,

$$p_{\rm pp}(m_1|m_{\rm pp}, \sigma_{\rm pp}) \propto \exp\left[-\frac{(m_1 - m_{\rm pp})^2}{2\sigma_{\rm pp}^2}\right]\,. \tag{7.11}$$

Following Abbott et al. (2021f), we assume a powerlaw distribution for the mass ratio ($q = m_2/m_1 \leq 1$) (with powerlaw index β). So the conditional probability distribution on q given m_1 can be written as,

$$p(q|m_1, \beta, m_{\rm min}) \propto q^{\beta}. \tag{7.12}$$

and is defined in the range $m_{\rm min}/m_1 < q < 1$.

The redshift distribution for the binary black hole population is assumed to be a distribution that closely resembles the star formation rate but it is flexible enough to accommodate different shapes. Hence we follow the model used in Callister et al. (2020); Ezquiaga (2021),

$$\psi(z|\gamma, \kappa, z_p) \propto \frac{(1+z)^{\gamma}}{1 + \left(\frac{1+z}{1+z_p}\right)^{\gamma+\kappa}} \tag{7.13}$$

which peaks at z_p and where γ controls the low redshift rise and κ the high redshift tail of the distribution.

7.4 RESULTS

We present constraints on gravitational leakage models using GWTC-3 BBH mergers. We use only the events that pass the 1 per-year IFAR threshold as was done in Abbott et al. (2021j). In total, we analyze 68 BBH mergers and exclude outlier BBH events such as GW190814 or any NSBH and BNS candidates.

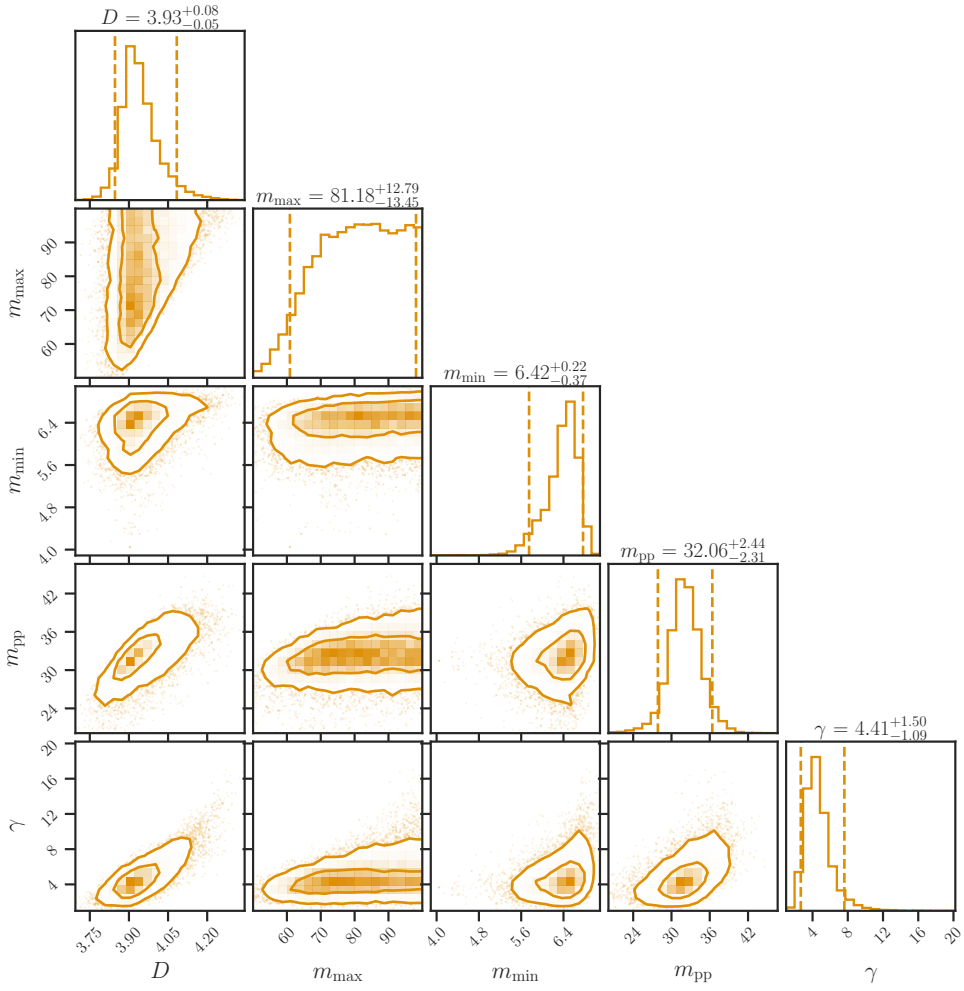


Figure 7.2: Posterior distributions for the number of spacetime dimensions and the BBH population parameters most strongly correlated with D for the gravitational leakage model defined in Equation 7.2. We show the BBH maximum mass $m_{\rm max}$, minimum mass $m_{\rm max}$, peak of the Gaussian (PISN) component $m_{\rm pp}$ and the slope of the merger rate evolution γ.

First, we show constraints for the gravitational leakage model that only depends on the number of spacetime dimensions D (see Equation 7.2) and so allows for gravitational leakage at all scales. The posterior distribution on the number of spacetime dimensions D as well as the hyperparameters that correlate strongly with D are shown in Figure 7.2 (See Appendix D for full posterior distributions). We find that we can place tight constraints on the number of spacetime dimensions with BBH observations alone to $D = 3.93^{+0.08}_{-0.05}$ at 68% C.L. The constraint presented here is as competitive to the GW170817 constraint with its associated electromagnetic counterpart of $D = 3.98^{+0.07}_{-0.09}$ at 68% C.L. (Pardo et al., 2018). The BBH population hyperparameters are broadly consistent with constraints placed in the GWTC-3 population analysis by the LVK Collaboration. However, we do find a broadening and shift to higher mass for the allowed BBH maximum mass $m_{\rm max}$. This is consistent with the analysis of (Ezquiaga & Holz, 2021), as the additional parameters in the modified luminosity distance models due to gravitational leakage (or varying c_M) introduce uncertainty in the value of $m_{\rm max}$.

We also show results for the model of Equation 7.3 that allows for a varying screening scale and transition steepness. Posterior distributions on D and relevant population parameters are shown in Figure 7.3. Under this model, we find we find $D = 4.49^{+1.63}_{-0.87}$ at 68% C.L., which gives broader constraints compared to the previous model (due to the additional screening scale parameters) but are still consistent with GR. We find that D has strong correlations with the screening scale and transition steepness as shown in Figure D.2, which can also be seen through the detectability of the simulated population $p(z|\text{det})$ shown in Figure 7.1. We note that the BBH population hyperparameters under this model are similarly constrained with respect to the model of Equation 7.2.

Since we have used binary black hole mergers across a wide range of distances, we are also able to constrain the screening scale R_c and the transition steepness n jointly with the number of spacetime dimensions D for the first time, without having to fix any of the three gravitational leakage model parameters as has been done in other studies (see Appendix D for full poste-

rior distributions). We show in Figure 7.4, the posterior distribution on R_c and n and find that $\log_{10} R_c/\mathrm{Mpc} = 4.75^{+0.86}_{-0.99}$ with a transition steepness $\log_{10} n = 0.85^{+0.76}_{-0.86}$ both at 68% C.L., consistent with GW170817 (and GW190521) constraints (Pardo et al., 2018; Corman et al., 2022).

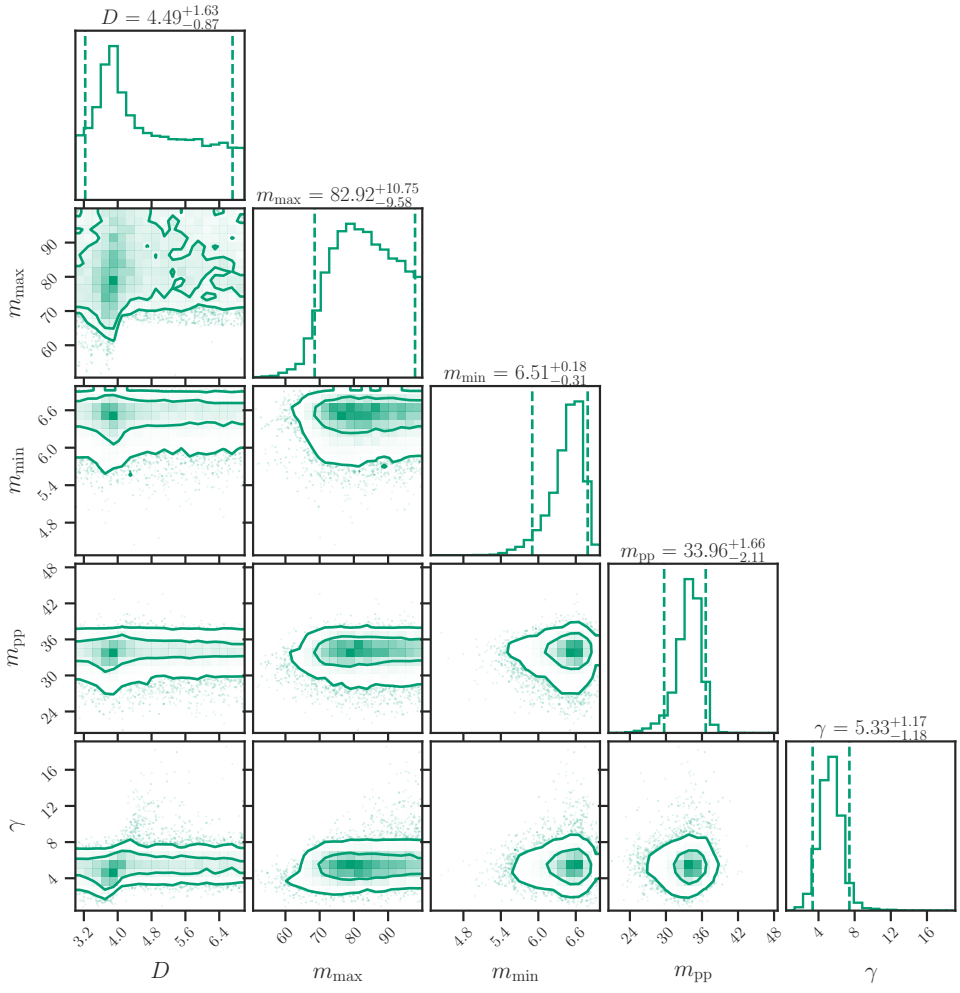


Figure 7.3: Posterior distribution for the leakage model with the number of space-time dimensions D, screening scale R_c and transition steepness n as defined in Equation 7.3 and the BBH population parameters most strongly correlated with these: the BBH maximum mass $m_{\rm max}$, minimum mass $m_{\rm max}$, peak of the Gaussian (PISN) component $m_{\rm pp}$ and the slope of the merger rate evolution γ.

These constraints are also consistent with General Relativity since in the limit of large R_c we expect GR-like behavior, which in the limit of large R_c becomes independent of the value of the transition steepness n.

We also report Bayes factors $\ln \mathcal{B}^{D}_{\rm GR}$, to do hypothesis testing while fixing the number of spacetime dimensions $D = 4$ (GR) compared to $D \neq 4$, under the leakage model of Equation 7.3. We report the following Bayes factors, $\ln \mathcal{B}^{D=5}_{\rm GR} = -1.21$, $\ln \mathcal{B}^{D=6}_{\rm GR} = -1.36$, $\ln \mathcal{B}^{D=7}_{\rm GR} = -0.85$ and $\ln \mathcal{B}^{D=8}_{\rm GR} = -1.45$. As expected, a higher number of spacetime dimensions with respect to $D = 4$ are disfavored by the data.

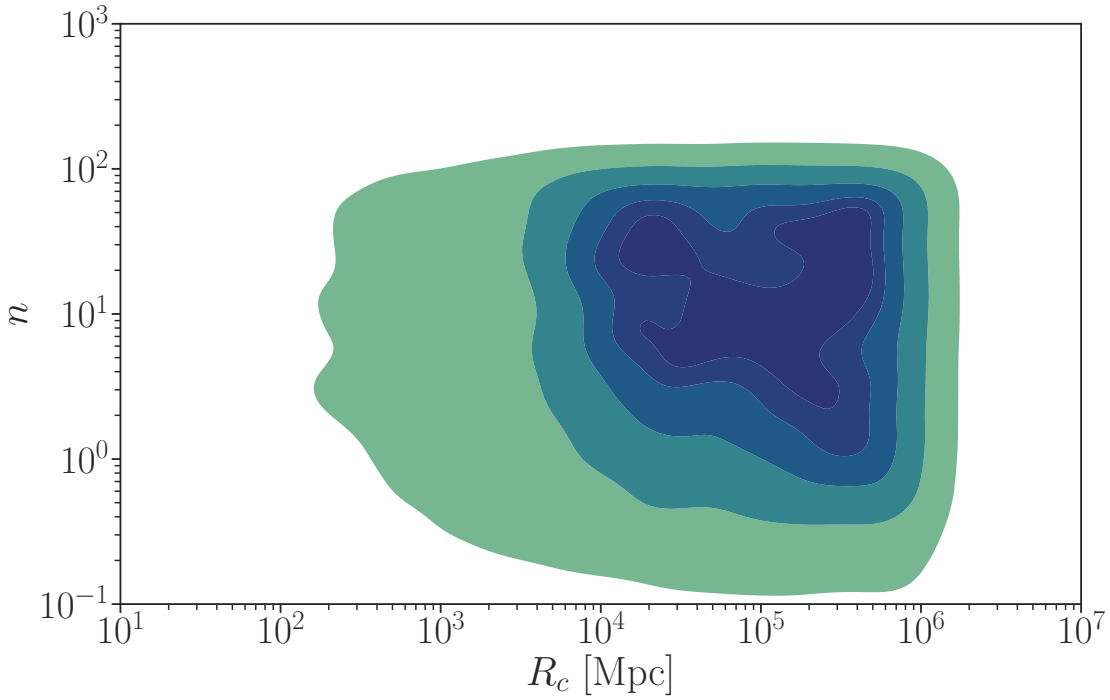


Figure 7.4: Constraints on the screening scale R_c and transition steepness n from GWTC-3 BBH observations under the gravitational leakage model defined in Equation 7.3. We find that $\log_{10} R_c/\mathrm{Mpc} = 4.75^{+0.86}_{-0.99}$ with a transition steepness $\log_{10} n = 0.85^{+0.76}_{-0.86}$ C.L., consistent with previous upper limits placed by GW170817 and GW190521 (Pardo et al., 2018; Corman et al., 2022). We find that these parameters are consistent with GR, e.g. large screening scale.

7.5 DISCUSSION

We placed constraints on the number of spacetime dimensions from binary black hole mergers using the recently released GWTC-3 catalog. We find that our constraints on D agree with general relativity for both gravitational leakage models considered in this work. We also find that the constraints placed on the screening scale and transition steepness for gravitational leakage models are consistent with General Relativity. We note that, further, and more distant observations should allow us to place tighter constraints on these scales.

The constraints presented in this paper are dependent on the chosen binary black hole mass and redshift distribution models. However, the phenomenological models used in this work should be flexible enough to capture the astrophysics that drives different BBH formation channels at current detector sensitivities. The BBH population modeling can be improved to include subpopulations such as hierarchical mergers or even mergers above the PISNe mass gap. Another source of potential systematics with respect to the BBH population modeling is if the BBH mass distribution is not universal, that is, it has redshift dependence. Recent work, by Ezquiaga & Holz (2022) has demonstrated that even if the mass distribution evolves with redshift, such effects can be modeled and consistently taken into account within this framework. However, a precision measurement of cosmology with the approach discussed in this paper will likely require a nonparametric approach and a data-driven approach to model the observed BBH mass distribution.

With respect to gravity leakage models, one can use the Bayesian framework presented here to test different parametrizations for the modified luminosity distance $d_L^{\rm GW}$ under other modified gravitational wave propagation models, for example, alternative phenomenological models such as the graviton leakage model of Pardo et al. (2018) or theory-specific models from specific parametrizations of modified gravity theories.

The Bayesian analysis used in this work can also be extended to work with the dark siren formalism that has been used mainly to perform cosmological measurements on H_0, since it provides

a discrete prior on the allowed locations (assuming they are hosted by a galaxy) for GW sources.

Binary black hole mergers can probe cosmology and extensions of the ΛCDM cosmological model under modified gravity at cosmological scales without the need for an independent redshift measurement of the source or other electromagnetic information. With future, more distant GW catalogs and an improved understanding of the BBH population we can only expect to improve these measurements and better understand the underlying cosmological model.

Part IV

Conclusions

Chapter 8

Summary and Future Work

In this dissertation, we have explored three different topics involving gravitational waves, namely, cosmology, lensing, and modified gravity.

8.1 COSMOLOGY

The cosmological results presented in this thesis involved the development of a robust statistical framework to perform cosmological inference with gravitational wave data and electromagnetic data in the form of associated electromagnetic counterparts or galaxy catalog data. We formulated and validated the statistical framework by fully taking into account both GW and EM selections for the first time. We constructed simulations to validate our formalism in the form of mock galaxy catalogs where we placed approximately 250 GW events. These catalogs allowed us to test our formalism in order of increasing complexity. First, knowing exactly every host galaxy to every GW event allowed us to test the robustness of our inference to GW selection effects. Second, by considering the incompleteness of galaxy catalogs in the form of limited magnitude surveys with varying completeness levels, ranging from 100%, 75%, 50%, down to 25%, we were able to assess the robustness of our methodology in testing for EM selection effects and compensating for the missing galaxies in the constructed catalogs. Finally, we implemented a simple model where we expanded the catalog so that the GW events are most likely to reside on the most luminous galaxies. This is in essence more realistic as different models for GW formation channels predict that GW mergers trace either the stellar mass or star formation rate histories of galaxies. These EM observables can be related to the intrinsic luminosities of galaxies in different magnitude filters. The latter mock data challenge then allowed for testing of the robustness of our algorithm to see

how much better the measurement can be (or if any biases are introduced) if we take the luminosity weighting into account (or not). In summary, the mock data challenge enabled us to establish a robust inference framework, which was later developed into the gwcosmo package. This package has been utilized for GW inference within the LVK Collaboration for cosmological inference with GW observations and is presently in use.

In order to test the gwcosmo formalism with real GW data we provided a summary of the LVC measurement of the Hubble constant with events from the GWTC-1 catalog. We performed a measurement of cosmology using the five best localized BBH events and reanalyzed the BNS merger GW170817 with its host galaxy GW170817 to provide for the first time a joint measurement of H_0 with gravitational wave standard sirens. We used the GLADE galaxy catalog, a compilation catalog comprised of various surveys for most BBH events, and for the GW170814 event, we used the DES survey as it provided a deeper and thus more complete catalog in order to obtain better constraints.

The challenge of real GW data and real galaxy catalog data is the fact that survey incompleteness will always be present. In order to model the incompleteness we had to rely on the simple approaches developed in Chapter 1. However, this was sufficient so long as we considered only the best localized BBH mergers so that their contribution to the H_0 measurement was driven by the galaxy catalog data and not by the assumed BBH population or the galaxy survey assumptions that modeled the incompleteness.

We investigated for the first time how the BBH population can indeed lead to a measurement of H_0 if it has any features that allow for the mass redshift degeneracy to break. We also performed a systematic study of luminosity weighting by both B-band and K-band luminosities, proxies for star formation rates, and stellar masses of galaxies respectively, in order to see if the tracer luminosity had any effect on the inference. Turns out that at the level of five events we are dominated by the statistical uncertainty of the measurement. However, when fixing the BBH population to a fiducial

shape we learned that significant biases can be introduced on the measurement, hence leading us to exclude the worse localized (and typically the farthest) events in the GWTC-1 catalog.

There are obvious future directions to improve the dark siren methodology described in this work, some of these improvements have been implemented but many remain to be explored.

The incompleteness of galaxy catalogs is not uniform across the entire sky. This was an assumption that was made in both the mock data challenge presented in Chapter 2 and subsequently employed in Chapter 3 with the cosmological analysis of GWTC-1. It turns out that galaxy surveys are patchy due to the simple nature of observatories being located either in the north or the south. To make things worse, there will always be regions with almost no galaxies, namely along the galactic plane, where the bright Milky Way sits along with plenty of interstellar dust which makes this region hard to observe. Consequently, this region is colloquially referred to as the zone of avoidance for a reason. Follow-up work by Gray et al. managed to circumvent this issue by analyzing the incompleteness of the EM survey data as a function of sky direction. Using a pixelated approach to both the GW sky localization regions and the EM survey footprint the limiting magnitude maps can be completed for each pixel (sky direction). This allows for the incompleteness variations to be taken into account in a self-consistent way. However, much work remains to be done to fully implement the survey selection function to determine the incompleteness of the catalog in a realistic way. The simple limiting magnitude approach, although working for now, will have to be extended once the number of GW events analyzed increases and this potential systematic begins to be noticeable.

Continuing with EM galaxy survey systematics, much work needs to be done to understand the luminosity function and galaxy distribution models that are assumed a priori to take into account the incompleteness, i.e. the models for the galaxy population that tell you how many galaxies you are really missing. As of writing, simple analytical models for the luminosity function are employed such as Schechter functions, and galaxies are assumed to be distributed uniformly in

comoving volume. The luminosity function for galaxies has been shown to vary with redshift and to depend on the galaxy population such as galaxy type. The assumed spatial distribution of galaxies to be uniform in comoving volume is only realistic at the largest of scales due to isotropy and homogeneity arguments. This assumption breaks down at small scales where the large-scale structure of the universe is more prominent. Further work needs to explore systematic effects due to the assumed luminosity function as well as the lack of galaxy clustering in the assumed spatial distribution affect the inferred results.

The BBH population features can notably affect the inferred cosmological constraints as was shown in Chapter 3. In order to circumvent this and be able to analyze all BBH events without worrying that the constraints are driven by the fixed BBH population one must jointly fit the underlying cosmology and the BBH population model. A parametric approach to the BBH population along with the dark siren formalism is a natural first step, however, a non-parametric approach is likely necessary in order to perform cosmological inference due to the complicated expected shape of the BBH population due to the many predictions of binary merger formation channels. This work remains to be seen in the literature, and to me remains an important step to unify both approaches for cosmological inference with gravitational wave standard sirens and large-scale galaxy surveys.

Finally, the number of GW events in future GW catalogs is expected to grow dramatically. Next-generation galaxy surveys will grow as well and some of these are currently undertaking observations. In order to handle both large datasets, one needs to build robust and scalable codebases in order to future-proof the implementations. This task is important as the gwcosmo code has already shown scalability issues when handling O3-like datasets. The use of modern tools such as hardware accelerated frameworks such as JAX and modern machine learning algorithms might be necessary to perform the work outlined above while keeping the codes scalable, robust, and in the shape required to analyze the latest GW data.

To summarize the gravitational wave cosmology work presented in this thesis paved the way

for future analysis with gwcosmo and subsequent GW catalogs.

8.2 LENSING

The work presented in Chapter 3 developed a Bayesian framework to statistically identify and characterize strongly lensed GW events by analyzing them jointly in a self-consistent way. Strongly lensed GW events are interesting since the frequency evolution of the lensed images remains unchanged. The strong lensing merely changes the amplitudes, and times of arrival and might add a constant phase shift to each image. Consequently, by taking into account these induced changes the GW signals for each lensed image in principle look the same. Effectively allowing us to coherently fit all images using the joint Bayesian parameter estimation framework introduced in this thesis, so long as we also fit for the lensing-induced changes to the GW waveforms.

The strong lensing joint parameter estimation framework was then used for the first time on real GW event data, namely pairs of events from GWTC-1. We learned that most pairs had very low likelihood under the lensing hypothesis. However, the GW170104 and GW170814 events favored the lensing hypothesis when we jointly analyzed them. This meant that the pairs could either be lensed images of each other or simply events that had similar parameters and by mere chance looked lensed.

To further investigate the potential association we considered the time delay between the two events and the prior probability for lensing at the distances probed by the end of O2. Turns out that if we considered galaxy lenses, the eight-month time delay would be inconsistent with the expected time delays for galaxy-galaxy lensing. Simply speaking, there isn't enough mass in your typical galaxy to create a time delay of this duration. Therefore, under the lensing hypothesis, a galaxy lens would be disfavored. In order to circumvent that we would require at least a galaxy cluster lens to accommodate for the time delay. However, galaxy clusters are less abundant and might require magnifications that would be incompatible with the pair. However, both types of lenses

remained a possibility albeit unlikely and thus reduced the odds for lensing quite dramatically but insufficiently to conclude the pair was not lensed.

When considering the prior probability of lensing at the distance probed in O2, we used simulations that used measured optical depths and monte carlo simulations of GW events to understand the chances of strong lensing as a function of distance. Given the horizon distance for O2, this highly penalized the odds for lensing. Lastly, when considering finer details such as the constant phase shift that we measured between the events, implying one of the events in the pair was a type-III lensed image also penalized the lensing hypothesis as these are quite rare from an EM perspective.

In short, we analyzed pairs of events from GWTC-1 under the lensing hypothesis for the first time and found potential evidence for this pair to be lensed or merely a pair of events that mimicked lensing by having similar system parameters. After taking a closer look at the measured lensing observables and expectations from strong lensing using EM observations we concluded that the pair was unlikely to be a pair of lensed images.

The work of Chapter 3 was used as a foundation for the development of the statistical methods of Chapter 4. Here, we extended the joint parameter estimation formalism to take into account some of the issues that we discovered in Chapter 3. Namely, GW selections are important and need to be fully taken into account when analyzing pairs of GW events. We also self-consistently analyzed all potential constant phase shifts at once and provide a framework that instead gives you posterior distributions for the type of lensed images in the analyzed set. Finally, we incorporated the effects of the BBH population, as it turns out, these can break the mass redshift degeneracy and thus allow for indirect constraints on the redshift to the source affecting the conclusions of our analysis. The statistical framework culminated in the development of the Hanabi library which is the defacto strong lensing analysis library used by the LVK Collaboration.

The joint parameter estimation work presented in Chapters 4 and 5 although very sensitive is

quite computationally expensive and the computational cost grows rapidly when more than two lensed images are analyzed. Therefore, much work is required to speed up the inference and this is critical as the number of events that we expect in O4 can be as many as O(500). This means that it is computationally intractable to analyze every single pair of events in O4. However, this is not what one does in practice, as one can approximately tell if a given pair of events is lensed by say, looking at the sky localization consistency between pairs (from independent analyses).

The odds ratios that hanabi provides need to be interpretable. It is hard to tell if an odds ratio for a given pair is significant enough to warrant a lensing claim. A way to improve the interpretation of these statistics and to understand the false alarms induced by similar-looking events is to construct a realistic simulation of lensed and unleased BBH events and run hanabi on identified pairs from each. This builds two distributions, a background distribution of odds rations for unlensed signals and a foreground distribution for lensed signals. Once these distributions are understood, if a pair of events looks lensed then the odds ratio for the pair can be compared against both the background and foreground distributions and effectively assign a significance compared to a potential false alarm. However, as mentioned above, hanabi is computationally expensive to run so the scope of this work is limited by the computational resources available to construct faithful background and foreground distributions.

The framework could also use more aggressive priors on strong lensing physics in order to be able to best differentiate lensing pairs. If one implements a model for the lens in terms of the absolute magnification parameters as well as the mass of the lens one in principle could break some of the degeneracies that we have with the current implementation of hanabi. The formalism works with relative quantities in order to remain model-independent, yet it remains for models to be implemented at the inference level and a full study is yet to be seen.

In summary, the joint parameter estimation work presented in Chapters 4 and 5 led to the first attempt at realistically identifying pairs of lensed GW events but most importantly to the

development of the Hanabi code which has been subsequently used in LVK Collaboration papers and is still in active use at the time of writing.

8.3 MODIFIED GRAVITY

In the modified gravity section, we explore two vastly different topics. First, we consider how once we have identified a pair of lensed images we could in principle make a measurement of additional GW polarization modes. Second, we also explore constraints on the number of spacetime dimensions predicted by gravitational leakage models using the latest GWTC-3 BBH observations.

The joint parameter estimation framework constructed in Chapters 4 and 5 have been used mainly to identify and perform Bayesian model selection for potential pairs of lensed GW events. However, as was emphasized in Chapter 3, the strong lensing, at least of BBH mergers is imminent. The work of Chapter 6 presents a proof of principle approach to better test General Relativity with lensed GW images compared to constraints obtainable by analyzing events that are not lensed.

It turns out that one can use the joint parameter estimation framework to test for the presence of additional GW polarization modes with lensed images. The effect of strong lensing means that you would get multiple images from the same astrophysical source arriving at different times. However, due to Earth's rotation, the lensed images will be detected by detectors that have rotated by an angle proportional to the relative time delay of the pair effectively increasing the number of detectors that see the system. This is the most important aspect when it comes to probing addition GW polarizations as their measurement is effectively a projection of the GW perturbation on the detector geometry. The additional constraints introduced by the strongly lensed images arriving at different times effectively double the number of GW detectors for a lensed pair allowing you to measure the amplitudes of each mode in some cases.

It was shown that one can effectively measure the amplitudes for each GW polarization mode with a detector network (HLVK) and two lensed images much better than if the images were

analyzed independently. The work presented, although a proof of concept is important as we are at a time where the detection of strongly lensed GW events is a matter of waiting. Thus, the science case presented in Chapter 6 is important, and the constraints there although pessimistic (due to the simple models used) will only improve.

In Chapter 7, we considered modifications to cosmological gravitational wave propagation using the latest LVK catalog of BBH events. We worked with gravitational leakage models and their prediction that additional large-scale spacetime dimensions (other than the four predicted by GR) might exist. Previous studies used GW170817 and its associated host galaxy (as redshift information) to place constraints on such models and found very good agreement with GR. However, we noted that this is expected since if GR were to be incorrect it would happen at cosmological distances. Thus, for the first time, we placed constraints on the number of spacetime dimensions using the roughly 70 BBH mergers detected to date, allowing us to make a measurement across distances much farther than what is possible with BNS events.

We employed two models, a simple model where the number of dimensions changes on the expected $1/D_L$ behavior of GW standard sirens and a model that allows for a transition scale with a corresponding transition strength. To make this measurement we require redshift information for each BBH that we consider. Although we could have used the galaxy catalog approaches of Chapters 2 and 3 we used the BBH population features to statistically break the mass redshift degeneracy and provide the necessary redshift information. We jointly fitted the BBH population and the gravitational leakage model parameter and unsurprisingly found constraints that are compatible with GR and as informative as those placed with GW170817 and NGC4993. Future and much farther observations will allow for stringent tests on extra dimensions.

The strong lensing polarizations constraints with strong lensing although a proof of principle study was performed in order to motivate the science case for strong lensing of GWs. Some studies suggest that the better sky localizations provided by hanabi for a potential pair of lensed events

might be sufficiently small for deep follow-up electromagnetic observations. Such campaigns will map the sky localization region to look for galaxy lenses and their lensed galaxy counterparts. Given that we measure the time delays for lensed GW events very precisely, many of these potential galaxy lenses along their lensed galaxy counterparts and in turn potential hosts for the BBH merger could be ruled out. The potential for discovering the host galaxy for a lensed BBH system is great and the polarization study could only benefit from this additional information.

The gravitational wave propagation study could benefit from the dark siren formalism to provide better constraints. As was noted above, if one wanted to use both approaches self consistently to measure cosmological parameters (in this case modified gravity models) one needs to jointly fit the BBH population with such models. Another avenue for future work is to come up with a unified parameterization for modified gravitational wave propagation and cosmological expansion. Both of these effects are highly degenerate with each other and in the work presented in this thesis, the cosmological model was fixed. Also, gravitational leakage models are only one class of models that affect gravitational propagation, there are other parameterizations that could be used to measure similar quantities. A unified approach remains to be seen.

LIST OF JOURNALS

ApJ	The Astrophysical Journal
ApJS	The Astrophysical Journal Supplement
ARAA	Annual Review of Astronomy and Astrophysics
A&A	Astronomy & Astrophysics
A&ARv	The Astronomy and Astrophysics Review
MNRAS	Monthly Notices of the Royal Astronomical Society

BIBLIOGRAPHY

Aasi, J., et al. (2015). Advanced LIGO. Class. Quant. Grav., 32, 074001.

Abbott, B., et al. (2017a). A gravitational-wave standard siren measurement of the Hubble constant. Nature, 551(7678), 85–88.

Abbott, B., et al. (2017b). GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. Phys. Rev. Lett., 119(16), 161101.

Abbott, B. P., Abbott, R., Abbott, T. D., Abraham, S., & Virgo Collaboration, L. (2019). 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.

Abbott, B. P., et al. (2016a). Binary Black Hole Mergers in the first Advanced LIGO Observing Run. Phys. Rev. X, 6(4), 041015. [Erratum: Phys.Rev.X 8, 039903 (2018)].

Abbott, B. P., et al. (2016b). The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. Astrophys. J., 833(1), L1.

Abbott, B. P., et al. (2017c). GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. Phys. Rev. Lett., 119(16), 161101.

Abbott, B. P., et al. (2019a). Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. Astrophys. J. Lett., 882(2), L24.

Abbott, B. P., et al. (2019b). GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. Phys. Rev. X, 9(3), 031040.

Abbott, B. P., et al. (2019c). Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1. Phys. Rev. D, 100(10), 104036.

Abbott, B. P., et al. (2021a). A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo. Astrophys. J., 909(2), 218.

Abbott, R., et al. (2018a). Advanced LIGO anticipated sensitivity curves. https://dcc.ligo.org/LIGO-T1800044/public.

Abbott, R., et al. (2021b). Constraints on the cosmic expansion history from GWTC-3.

Abbott, R., et al. (2021c). GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. Phys. Rev. X, 11, 021053.

Abbott, R., et al. (2021d). GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run.

Abbott, R., et al. (2021e). GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run.

Abbott, R., et al. (2021f). Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. Astrophys. J. Lett., 913(1), L7.

Abbott, R., et al. (2021g). Search for Lensing Signatures in the Gravitational-Wave Observations from the First Half of LIGO–Virgo's Third Observing Run. Astrophys. J., 923(1), 14.

Abbott, R., et al. (2021h). Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog. Phys. Rev. D, 103(12), 122002.

Abbott, R., et al. (2021i). Tests of General Relativity with GWTC-3.

Abbott, R., et al. (2021j). The population of merging compact binaries inferred using gravitational waves through GWTC-3.

Abbott, T. M. C., et al. (2018b). The Dark Energy Survey Data Release 1. Astrophys. J. Suppl., 239(2), 18.

Acernese, F., et al. (2015a). Advanced Virgo: a second-generation interferometric gravitational wave detector. Class. Quant. Grav., 32(2), 024001.

Acernese, F., et al. (2015b). Advanced Virgo: a second-generation interferometric gravitational wave detector. Class. Quant. Grav., 32(2), 024001.

Ade, P. A. R., et al. (2016a). Planck 2015 results. XIII. Cosmological parameters. Astron. Astrophys., 594, A13.

Ade, P. A. R., et al. (2016b). Planck 2015 results. XIII. Cosmological parameters. Astron. Astrophys., 594, A13.

Ashton, G., et al. (2019). BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy. Astrophys. J. Suppl., 241(2), 27.

Aso, Y., Michimura, Y., Somiya, K., Ando, M., Miyakawa, O., Sekiguchi, T., Tatsumi, D., & Yamamoto, H. (2013). Interferometer design of the KAGRA gravitational wave detector. Phys. Rev. D, 88(4), 043007.

Belczynski, K., Holz, D. E., Bulik, T., & O'Shaughnessy, R. (2016). The first gravitational-wave source from the isolated evolution of two 40-100 Msun stars. Nature, 534, 512.

Belczynski, K., Ryu, T., Perna, R., Berti, E., Tanaka, T. L., & Bulik, T. (2017). On the likelihood of detecting gravitational waves from Population III compact object binaries. Mon. Not. Roy. Astron. Soc., 471(4), 4702–4721.

Berry, C. P. L., et al. (2015). Parameter estimation for binary neutron-star coalescences with realistic noise during the Advanced LIGO era. Astrophys. J., 804(2), 114.

Bilicki, M., Jarrett, T. H., Peacock, J. A., Cluver, M. E., & Steward, L. (2014). 2MASS Photometric Redshift catalog: a comprehensive three-dimensional census of the whole sky. Astrophys. J. Suppl., 210, 9.

Binney, J., & Tremaine, S. (1987). Galactic Dynamics, chap. 1, (pp. 21–22). Princeton Series in Astrophysics. Princeton University Press.

Broadhurst, T., Diego, J. M., & Smoot, I., George F. (2019). Twin LIGO/Virgo Detections of a Viable Gravitationally-Lensed Black Hole Merger. arXiv e-prints, (p. arXiv:1901.03190).

Buscicchio, R., Moore, C. J., Pratten, G., Schmidt, P., Bianconi, M., & Vecchio, A. (2020). Constraining the lensing of binary black holes from their stochastic background. Phys. Rev. Lett., 125(14), 141102.

Caditz, D., & Petrosian, V. (1989). Cosmological Parameters and Evolution of the Galaxy Luminosity Function. ApJL, 337, L65.

Callister, T., Fishbach, M., Holz, D., & Farr, W. (2020). Shouts and Murmurs: Combining Individual Gravitational-Wave Sources with the Stochastic Background to Measure the History of Binary Black Hole Mergers. Astrophys. J. Lett., 896(2), L32.

Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. (2015). Cosmological parameters from the comparison of peculiar velocities with predictions from the 2M++ density field. Mon. Not. Roy. Astron. Soc., 450(1), 317–332.

Çalışkan, M., Ezquiaga, J. M., Hannuksela, O. A., & Holz, D. E. (2022). Lensing or luck? False alarm probabilities for gravitational lensing of gravitational waves.

Chatziioannou, K., Isi, M., Haster, C.-J., & Littenberg, T. B. (2021). Morphology-independent test of the mixed polarization content of transient gravitational wave signals. Phys. Rev. D, 104(4), 044005.

Chatziioannou, K., Yunes, N., & Cornish, N. (2012). Model-Independent Test of General Relativity: An Extended post-Einsteinian Framework with Complete Polarization Content. Phys. Rev. D, 86, 022004. [Erratum: Phys.Rev.D 95, 129901 (2017)].

Chen, H.-Y., Fishbach, M., & Holz, D. E. (2018). A two per cent Hubble constant measurement from standard sirens within five years. Nature, 562(7728), 545–547.

Corman, M., Ghosh, A., Escamilla-Rivera, C., Hendry, M. A., Marsat, S., & Tamanini, N. (2022). Constraining cosmological extra dimensions with gravitational wave standard sirens: From theory to current and future multimessenger observations. Phys. Rev. D, 105(6), 064061.

Creighton, J. D. E., & Anderson, W. G. (2011). Gravitational-wave physics and astronomy: An introduction to the

Dai, L., & Venumadhav, T. (2017). On the waveforms of gravitationally lensed gravitational waves.

Dai, L., Venumadhav, T., & Sigurdson, K. (2017). Effect of lensing magnification on the apparent distribution of black hole mergers. Phys. Rev. D, 95(4), 044011.

Dai, L., Zackay, B., Venumadhav, T., Roulet, J., & Zaldarriaga, M. (2020). Search for Lensed Gravitational Waves Including Morse Phase Information: An Intriguing Candidate in O2.

Damour, T., Iyer, B. R., & Sathyaprakash, B. S. (2001). A Comparison of search templates for gravitational waves from binary inspiral. Phys. Rev. D, 63, 044023. [Erratum: Phys.Rev.D 72, 029902 (2005)].

Deffayet, C., & Menou, K. (2007). Probing Gravity with Spacetime Sirens. Astrophys. J. Lett., 668, L143–L146.

Del Pozzo, W. (2012). Inference of the cosmological parameters from gravitational waves: application to second generation interferometers. Phys. Rev. D, 86, 043011.

Diaz, C. C., & Mukherjee, S. (2022). Mapping the cosmic expansion history from LIGO-Virgo-KAGRA in synergy with DESI and SPHEREx. Mon. Not. Roy. Astron. Soc., 511(2), 2782–2795.

Drlica-Wagner, A., et al. (2018). Dark Energy Survey Year 1 Results: Photometric Data Set for Cosmology. Astrophys. J. Suppl., 235(2), 33.

Dálya, G., Galgóczi, G., Dobos, L., Frei, Z., Heng, I. S., Macas, R., Messenger, C., Raffai, P., & de Souza, R. S. (2018a). GLADE: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era. Mon. Not. Roy. Astron. Soc., 479(2), 2374–2381.

Dálya, G., Galgóczi, G., Dobos, L., Frei, Z., Heng, I. S., Macas, R., Messenger, C., Raffai, P., & de Souza, R. S. (2018b). GLADE: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era. Mon. Not. Roy. Astron. Soc., 479(2), 2374–2381.

Ezquiaga, J. M. (2021). Hearing gravity from the cosmos: GWTC-2 probes general relativity at cosmological scales. Phys. Lett. B, 822, 136665.

Ezquiaga, J. M., & Holz, D. E. (2021). Jumping the Gap: Searching for LIGO's Biggest Black Holes. Astrophys. J. Lett., 909(2), L23.

Ezquiaga, J. M., & Holz, D. E. (2022). Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries. Phys. Rev. Lett., 129(6), 061102.

Ezquiaga, J. M., Holz, D. E., Hu, W., Lagos, M., & Wald, R. M. (2020). Phase effects from strong gravitational lensing of gravitational waves.

Farr, W. M. (2019). Accuracy Requirements for Empirically-Measured Selection Functions. Research Notes of the AAS, 3(5), 66.

Farr, W. M., Fishbach, M., Ye, J., & Holz, D. (2019). A Future Percent-Level Measurement of the Hubble Expansion at Redshift 0.8 With Advanced LIGO. Astrophys. J., 883(2), L42.

Finke, A., Foffa, S., Iacovelli, F., Maggiore, M., & Mancarella, M. (2021). Cosmology with LIGO/Virgo dark sirens: Hubble parameter and modified gravitational wave propagation. JCAP, 08, 026.

Fishbach, M., Gray, R., Magaña Hernandez, I., Qi, H., & Sur, A. (2019a). A standard siren measurement of the Hubble constant from GW170817 without the electromagnetic counterpart. Astrophys. J., 871(1), L13.

Fishbach, M., Holz, D. E., & Farr, W. M. (2018). Does the Black Hole Merger Rate Evolve with Redshift? Astrophys. J., 863(2), L41. [Astrophys. J. Lett.863,L41(2018)].

Fishbach, M., et al. (2019b). A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart. Astrophys. J. Lett., 871(1), L13.

Fowler, W. A., & Hoyle, F. (1964). Neutrino Processes and Pair Formation in Massive Stars and Supernovae. Astrophys. J. Suppl., 9, 201–319.

Gerosa, D., Pratten, G., & Vecchio, A. (2020). Gravitational-wave selection effects using neural-network classifiers. Phys. Rev. D, 102(10), 103020.

Ghosh, A., Del Pozzo, W., & Ajith, P. (2016). Estimating parameters of binary black holes from gravitational-wave observations of their inspiral, merger and ringdown. Phys. Rev. D, 94(10), 104070.

Goyal, S., Haris, K., Mehta, A. K., & Ajith, P. (2021). Testing the nature of gravitational-wave polarizations using strongly lensed signals. Phys. Rev. D, 103(2), 024038.

Gray, R., Magaña Hernandez, I., Qi, H., Sur, A., et al. (2019). Cosmological Inference using Gravitational Wave Standard Sirens: A Mock Data Challenge. Phys. Rev. D, 101(12), 122001.

Hannuksela, O. A., Collett, T. E., Çalışkan, M., & Li, T. G. F. (2020). Localizing merging black holes with sub-arcsecond precision using gravitational-wave lensing. Mon. Not. Roy. Astron. Soc., 498(3), 3395–3402.

Hannuksela, O. A., Haris, K., Ng, K. K. Y., Kumar, S., Mehta, A. K., Keitel, D., Li, T. G. F., & Ajith, P. (2019). Search for Gravitational Lensing Signatures in LIGO-Virgo Binary Black Hole Events. ApJL, 874(1), L2.

Haris, K., Mehta, A. K., Kumar, S., Venumadhav, T., & Ajith, P. (2018). Identifying strongly lensed gravitational wave signals from binary black hole mergers. arXiv e-prints, (p. arXiv:1807.07062).

Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. (2003). How massive single stars end their life. Astrophys. J., 591, 288–300.

Heger, A., & Woosley, S. E. (2002). The nucleosynthetic signature of population III. Astrophys. J., 567, 532–543.

Holz, D. E., & Hughes, S. A. (2005). Using gravitational-wave standard sirens. Astrophys. J., 629, 15–22.

Hou, S., Fan, X.-L., & Zhu, Z.-H. (2019). Gravitational Lensing of Gravitational Waves: Rotation of Polarization Plane. Phys. Rev. D, 100(6), 064028.

Hoyle, B., et al. (2018). Dark Energy Survey Year 1 Results: Redshift distributions of the weak lensing source galaxies. Mon. Not. Roy. Astron. Soc., 478(1), 592–610.

Husa, S., Khan, S., Hannam, M., Pürrer, M., Ohme, F., Forteza, X. J., & Bohé, A. (2016). Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal. Phys. Rev. D, 93(4), 044006.

Isi, M. (2022). Parametrizing gravitational-wave polarizations.

Isi, M., Pitkin, M., & Weinstein, A. J. (2017). Probing Dynamical Gravity with the Polarization of Continuous Gravitational Waves. Phys. Rev. D, 96(4), 042001.

Khan, S., Husa, S., Hannam, M., Ohme, F., Pürrer, M., Forteza, X. J., & Bohé, A. (2016). Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. Phys. Rev. D, 93(4), 044007.

Kiziltan, B., Kottas, A., De Yoreo, M., & Thorsett, S. E. (2013). The Neutron Star Mass Distribution. Astrophys. J., 778, 66.

Lagos, M., Fishbach, M., Landry, P., & Holz, D. E. (2019). Standard sirens with a running Planck mass. Phys. Rev. D, 99(8), 083504.

Li, A. K. Y., Lo, R. K. L., Sachdev, S., Chan, C. L., Lin, E. T., Li, T. G. F., & Weinstein, A. J. (2019). Finding diamonds in the rough: Targeted Sub-threshold Search for Strongly-lensed Gravitational-wave Events.

Li, S.-S., Mao, S., Zhao, Y., & Lu, Y. (2018). Gravitational lensing of gravitational waves: A statistical perspective. Mon. Not. Roy. Astron. Soc., 476(2), 2220–2229.

LIGO Scientific Collaboration (2018). LIGO Algorithm Library.

Liu, X., Hernandez, I. M., & Creighton, J. (2021). Identifying strong gravitational-wave lensing during the second observing run of Advanced LIGO and Advanced Virgo. Astrophys. J., 908(1), 97.

Lo, R. K. L., & Magaña Hernandez, I. (2021). A Bayesian statistical framework for identifying strongly-lensed gravitational-wave signals.

LVC (2019). A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO-Virgo. https://dcc.ligo.org/LIGO-P1900015, TBD.

Madgwick, D. S., Lahav, O., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lewis, I., Lumsden, S., Maddox, S., Norberg, P., Peacock, J. A., Peterson, B. A., Sutherland, W., & Taylor, K. (2002). The 2dF Galaxy Redshift Survey: galaxy luminosity functions per spectral type. MNRAS, 333(1), 133–144.

Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. (2014). HyperLEDA. III. The catalogue of extragalactic distances. Astron. Astrophys., 570, A13.

Mancarella, M., Genoud-Prachex, E., & Maggiore, M. (2022). Cosmology and modified gravitational wave propagation from binary black hole population models. Phys. Rev. D, 105(6), 064030.

Mandel, I., & de Mink, S. E. (2016). Merging binary black holes formed through chemically homogeneous evolution in short-period stellar binaries. Mon. Not. Roy. Astron. Soc., 458(3), 2634–2647.

Mandel, I., Farr, W. M., & Gair, J. R. (2019). Extracting distribution parameters from multiple uncertain observations with selection biases. Mon. Not. Roy. Astron. Soc., 486(1), 1086–1093.

Mandel, I., Farr, W. M., & Gair, J. R. (2019). Extracting distribution parameters from multiple uncertain observations with selection biases. Mon. Not. Roy. Astron. Soc., 486(1), 1086–1093.

McIsaac, C., Keitel, D., Collett, T., Harry, I., Mozzon, S., Edy, O., & Bacon, D. (2020). Search for strongly lensed counterpart images of binary black hole mergers in the first two LIGO observing runs. Phys. Rev. D, 102(8), 084031.

Messick, C., Blackburn, K., Brady, P., Brockill, P., Cannon, K., Cariou, R., Caudill, S., Chamberlin, S. J., Creighton, J. D. E., Everett, R., Hanna, C., Keppel, D., Lang, R. N., Li, T. G. F., Meacher, D., Nielsen, A., Pankow, C., Privitera, S., Qi, H., Sachdev, S., Sadeghian, L., Singer, L., Thomas, E. G., Wade, L., Wade, M., Weinstein, A., & Wiesner, K. (2017). Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data. Phys. Rev. D, 95, 042001.

Mukherjee, S., Broadhurst, T., Diego, J. M., Silk, J., & Smoot, G. F. (2021a). Inferring the lensing rate of LIGO-Virgo sources from the stochastic gravitational wave background. Mon. Not. Roy. Astron. Soc., 501(2), 2451–2466.

Mukherjee, S., Wandelt, B. D., Nissanke, S. M., & Silvestri, A. (2021b). Accurate precision Cosmology with redshift unknown gravitational wave sources. Phys. Rev. D, 103(4), 043520.

Mukherjee, S., Wandelt, B. D., & Silk, J. (2021c). Testing the general theory of relativity using gravitational wave propagation from dark standard sirens. Mon. Not. Roy. Astron. Soc., 502(1), 1136–1144.

Nair, R., Bose, S., & Saini, T. D. (2018). Measuring the Hubble constant: Gravitational wave observations meet galaxy clustering. Phys. Rev. D, 98(2), 023502.

Ng, K. K. Y., Wong, K. W. K., Broadhurst, T., & Li, T. G. F. (2018). Precise LIGO Lensing Rate Predictions for Binary Black Holes. Phys. Rev. D, 97(2), 023012.

Ng, K. K. Y., Wong, K. W. K., Broadhurst, T., & Li, T. G. F. (2018). Precise LIGO lensing rate predictions for binary black holes. Phys. Rev. D, 97(2), 023012.

Oguri, M. (2018). Effect of gravitational lensing on the distribution of gravitational waves from distant binary black hole mergers. Mon. Not. Roy. Astron. Soc., 480, 3842–3855.

Oguri, M. (2018). Effect of gravitational lensing on the distribution of gravitational waves from distant binary black hole mergers. Mon. Not. Roy. Astron. Soc., 480(3), 3842–3855.

Oguri, M. (2019). Strong gravitational lensing of explosive transients. Rept. Prog. Phys., 82(12), 126901.

Palmese, A., Bom, C. R., Mucesh, S., & Hartley, W. G. (2023). A Standard Siren Measurement of the Hubble Constant Using Gravitational-wave Events from the First Three LIGO/Virgo Observing Runs and the DESI Legacy Survey. Astrophys. J., 943(1), 56.

Palmese, A., et al. (2020). A statistical standard siren measurement of the Hubble constant from the LIGO/Virgo gravitational wave compact object merger GW190814 and Dark Energy Survey galaxies. Astrophys. J. Lett., 900(2), L33.

Pardo, K., Fishbach, M., Holz, D. E., & Spergel, D. N. (2018). Limits on the number of spacetime dimensions from GW170817. JCAP, 07, 048.

Planck Collaboration (2018). Planck 2018 results. VI. Cosmological parameters. Astron. Astrophys., 641, A6.

Pratten, G., et al. (2020). Let's twist again: computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes.

Pâris, I., et al. (2017). The Sloan Digital Sky Survey Quasar Catalog: twelfth data release. Astron. Astrophys., 597, A79.

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. (2019). Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM. Astrophys. J., 876(1), 85.

Robertson, A., Smith, G. P., Massey, R., Eke, V., Jauzac, M., Bianconi, M., & Ryczanowski, D. (2020). What does strong gravitational lensing? The mass and redshift distribution of high-magnification lenses.

Romero-Shaw, I. M., et al. (2020). Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue. Mon. Not. Roy. Astron. Soc., 499(3), 3295–3319.

Sadeh, I., Abdalla, F. B., & Lahav, O. (2016). ANNz2 - photometric redshift and probability distribution function estimation using machine learning. Publ. Astron. Soc. Pac., 128(968), 104502.

Sathyaprakash, B. S., & Schutz, B. F. (2009). Physics, Astrophysics and Cosmology with Gravitational Waves. Living Rev. Rel., 12, 2.

Saunders, W., Rowan-Robinson, M., Lawrence, A., Efstathiou, G., Kaiser, N., Ellis, R. S., & Frenk, C. S. (1990). The 60- and far-infrared luminosity functions of IRAS galaxies. Monthly Notices of the Royal Astronomical Society, 242(3), 318–337.

Schneider, P., Ehlers, J., & Falco, E. E. (1992). Gravitational Lenses.

Schutz, B. F. (1986). Determining the Hubble constant from gravitational wave observations. Naturet, 323, 310.

Schutz, B. F. (1986). Determining the Hubble Constant from Gravitational Wave Observations. Nature, 323, 310–311.

Seto, N. (2004). Strong gravitational lensing and localization of merging massive black hole binaries with LISA. Phys. Rev. D, 69, 022002.

Singer, L. P., Goldstein, D. A., & Bloom, J. S. (2019). The Two LIGO/Virgo Binary Black Hole Mergers on 2019 August 28 Were Not Strongly Lensed.

Singer, L. P., Price, L. R., Farr, B., Urban, A. L., Pankow, C., Vitale, S., Veitch, J., Farr, W. M., Hanna, C., Cannon, K., Downes, T., Graff, P., Haster, C.-J., Mandel, I., Sidery, T., & Vecchio, A. (2014). The First Two Years of Electromagnetic Follow-up with Advanced LIGO and Virgo. ApJ, 795, 105.

Skilling, J. (2004). Nested Sampling. In R. Fischer, R. Preuss, & U. V. Toussaint (Eds.) Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, vol. 735 of American Institute of Physics Conference Series, (pp. 395–405).

Skrutskie, M. F., et al. (2006). The Two Micron All Sky Survey (2MASS). Astron. J., 131, 1163–1183.

Smith, G., Bianconi, M., Jauzac, M., Richard, J., Robertson, A., Berry, C. L., Massey, R., Sharon, K., Farr, W., & Veitch, J. (2019). Deep and rapid observations of strong-lensing galaxy clusters within the sky localization of GW170814. Mon. Not. Roy. Astron. Soc., 485(4), 5180–5191.

Smith, G., et al. (2017). Strong-lensing of Gravitational Waves by Galaxy Clusters. IAU Symp., 338, 98–102.

Smith, G. P., Jauzac, M., Veitch, J., Farr, W. M., Massey, R., & Richard, J. (2018). What if LIGO's gravitational wave detections are strongly lensed by massive galaxy clusters? Mon. Not. Roy. Astron. Soc., 475(3), 3823–3828.

Smith, R. J. E., Ashton, G., Vajpeyi, A., & Talbot, C. (2020). Massively parallel Bayesian inference for transient gravitational-wave astronomy. Mon. Not. Roy. Astron. Soc., 498(3), 4492–4502.

Soares-Santos, M., Palmese, A., et al. (2019a). First measurement of the Hubble constant from a dark standard siren using the Dark Energy Survey galaxies and the LIGO/Virgo binary-black-hole merger GW170814. Astrophys. J., 876(1), L7.

Soares-Santos, M., et al. (2019b). First measurement of the Hubble constant from a dark standard siren using the Dark Energy Survey galaxies and the LIGO/Virgo binary-black-hole merger GW170814. Astrophys. J., 876(1), L7.

Soares-Santos, M., et al. (2019c). First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary–Black-hole Merger GW170814. Astrophys. J. Lett., 876(1), L7.

Speagle, J. S. (2020). dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. Mon. Not. Roy. Astron. Soc., 493(3), 3132–3158.

Takahashi, R., & Nakamura, T. (2003). Wave effects in gravitational lensing of gravitational waves from chirping binaries. Astrophys. J., 595, 1039–1051.

Takeda, H., Morisaki, S., & Nishizawa, A. (2021). Pure polarization test of GW170814 and GW170817 using waveforms consistent with modified theories of gravity. Phys. Rev. D, 103(6), 064037.

Talbot, C., & Thrane, E. (2018). Measuring the binary black hole mass spectrum with an astrophysically motivated parameterization. Astrophys. J., 856(2), 173.

Taylor, S. R., & Gair, J. R. (2012). Cosmology with the lights off: standard sirens in the Einstein Telescope era. Phys. Rev., D86, 023502.

Taylor, S. R., Gair, J. R., & Mandel, I. (2012). Hubble without the Hubble: Cosmology using advanced gravitational-wave detectors alone. Phys. Rev. D, 85, 023535.

Thrane, E., & Talbot, C. (2019). An introduction to Bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models. PASA, 36, e010.

Turner, E. L., Ostriker, J. P., & Gott, I., J. R. (1984). The statistics of gravitational lenses : the distributions of image angular separations and lens redshifts. ApJ, 284, 1–22.

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

Wang, Y., Lo, R. K. L., Li, A. K. Y., & Chen, Y. (2021). Identifying Type-II Strongly-Lensed Gravitational-Wave Images in Third-Generation Gravitational-Wave Detectors.

White, D. J., Daw, E. J., & Dhillon, V. S. (2011). A List of Galaxies for Gravitational Wave Searches. Class. Quant. Grav., 28, 085016.

Woosley, S. E. (2017). Pulsational Pair-Instability Supernovae. Astrophys. J., 836(2), 244.

Woosley, S. E. (2019). The Evolution of Massive Helium Stars, Including Mass Loss. ApJ, 878(1), 49.

Woosley, S. E., Heger, A., & Weaver, T. A. (2002). The evolution and explosion of massive stars. Rev. Mod. Phys., 74, 1015–1071.

Xu, F., Ezquiaga, J. M., & Holz, D. E. (2022). Please Repeat: Strong Lensing of Gravitational Waves as a Probe of Compact Binary and Galaxy Populations. Astrophys. J., 929(1), 9.

Appendix A

Single Event and Combined MDC Results

In this section we provide detailed Figures for the Mock Data Challenge results of Chapter 2.

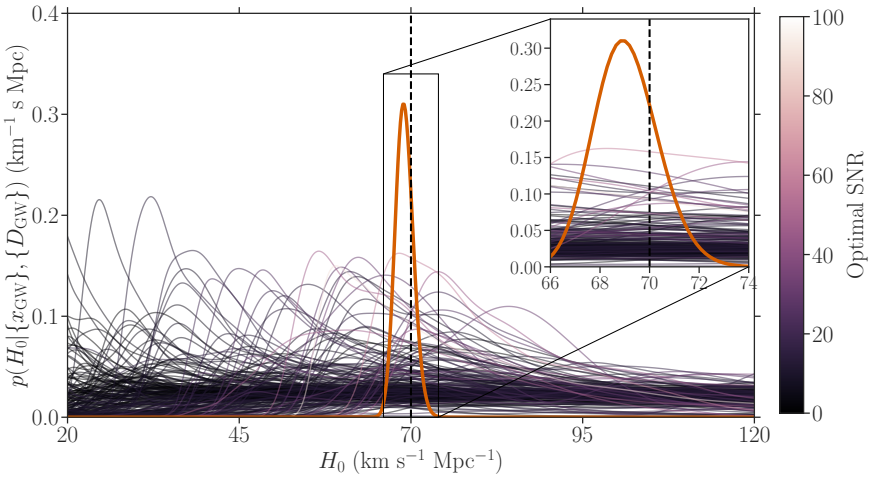


Figure A.1: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

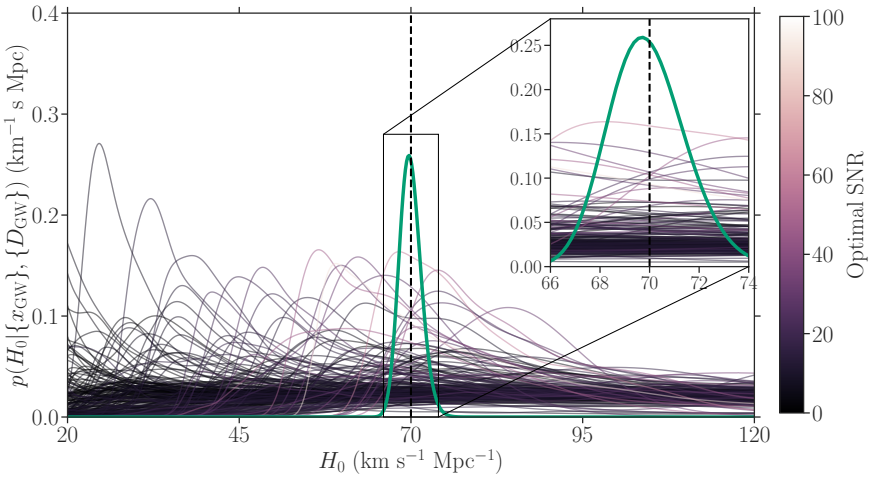


Figure A.2: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

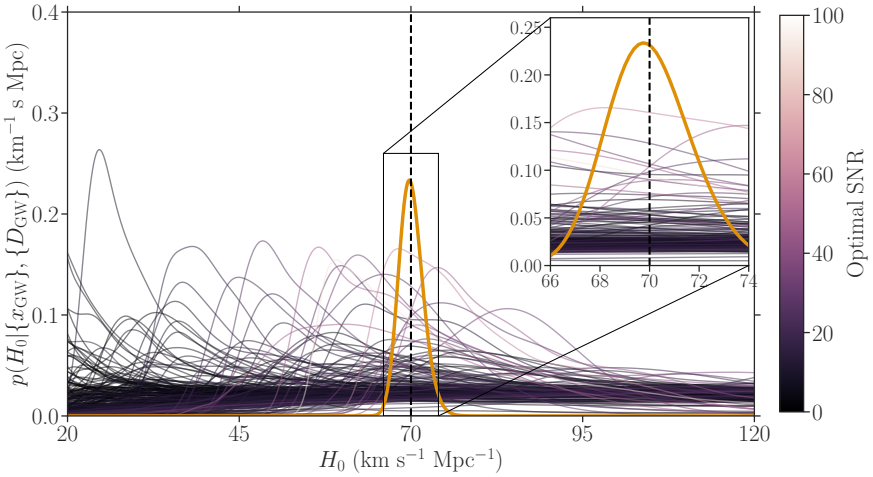


Figure A.3: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

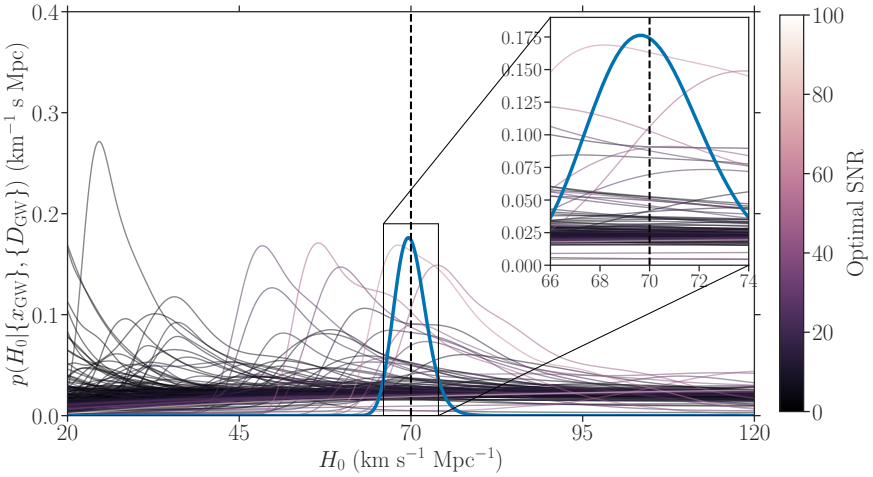


Figure A.4: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

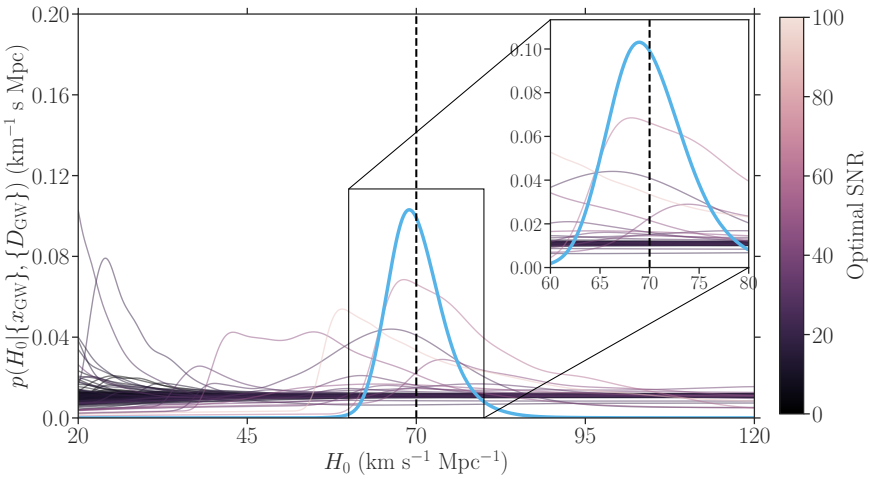


Figure A.5: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

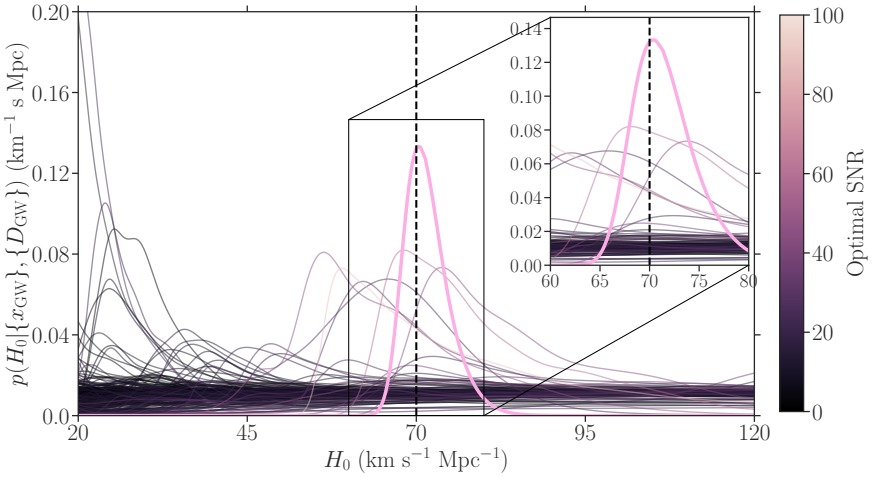


Figure A.6: Likelihoods (purple shades representing optimal SNR) and combined posterior (orange) for a 25% complete galaxy catalog (see **??**).

Appendix B

Detailed analysis of DES-Y1

The high completeness fraction of DES-Y1 within the GW170814 sky localization is apparent from Fig. 3.1. The catalog is expected to be more complete than GLADE since it has a limiting magnitude of approximately 23.8 for DES-Y1. We analyze the EM information coming from this catalog in greater detail. It is helpful to have an assessment of the contribution from potential host galaxies as a function of redshift for these events. In order to quantify this contribution, we perform a treatment analogous to Fishbach et al. (2019a) and compute the ratio $p_{\rm cat}(z)/p_{\rm vol}(z)$ between the probability distribution for the redshifts of potential host galaxies $p_{\rm cat}(z)$ and of a uniform in comoving volume distribution of galaxies $p_{\rm vol}(z)$. When computing $p_{\rm cat}(z)$ we include all galaxies brighter than $0.05L_g^*$ within the corresponding event's 99% sky localization region defined as,

$$p_{\rm cat}(z) \equiv \int p(x_{\rm GW}|\Omega)\, p_0(z,\Omega)\, d\Omega\,, \tag{B.1}$$

where $p(x_{\rm GW}|\Omega)$ is the GW likelihood as a function of the sky position Ω (this effectively weights each galaxy with the 2D skymap probability), and $p_0(z,\Omega)$ represents the galaxy catalog contribution, obtained from the distribution of galaxies in the catalog, marginalized over their redshift uncertainties also obtained from the catalog, and weighted by their probability of hosting a GW source (assuming a Planck 2015 cosmology for the required magnitude conversion). We consider weights for each galaxy proportional to their g-band luminosity as well as uniform weights to explore the effects due to this choice.

In Fig. B.1 we show the distributions $p_{\rm cat}(z)/p_{\rm vol}(z)$ for the DES-Y1 galaxies selected within the GW170814 sky localization region, for the redshift range $0 < z < 0.5$. The unweighted curve traces the over/under-density of galaxies, and then falls off at larger redshift due to incompleteness in the catalog. The luminosity-weighted redshift distribution is driven partially by the overdensity of galaxies at $z \approx 0.4$, and partially by bright high-redshift galaxies. The host galaxies for GW170814 are more likely to be located near the higher galaxy density regions in the DES-Y1 catalog – these features in the redshift prior are expected to drive the inferred H_0 posteriors for the corresponding events. Features we see in the DES-Y1 catalog are not as pronounced as the overdensity in the DES-Y3 data seen in Soares-Santos et al. (2019a). While the DES-Y3 survey is deeper, and may reveal finer features, a part of the above difference is likely also driven by the difference in the photometric redshift estimation algorithms, namely, template fitting methods such as BPZ (Hoyle et al., 2018) and machine learning based methods such as the `ANNz2` algorithm (Sadeh et al., 2016), with the latter used for GW170814 in (Soares-Santos et al., 2019a).The different selection criteria for choosing galaxies from the two catalogs, such as the stringent redshift cut placed in Soares-Santos et al. (2019a) versus a more relaxed redshift prior used in this work, is another potential source of difference between the corresponding redshift distributions.

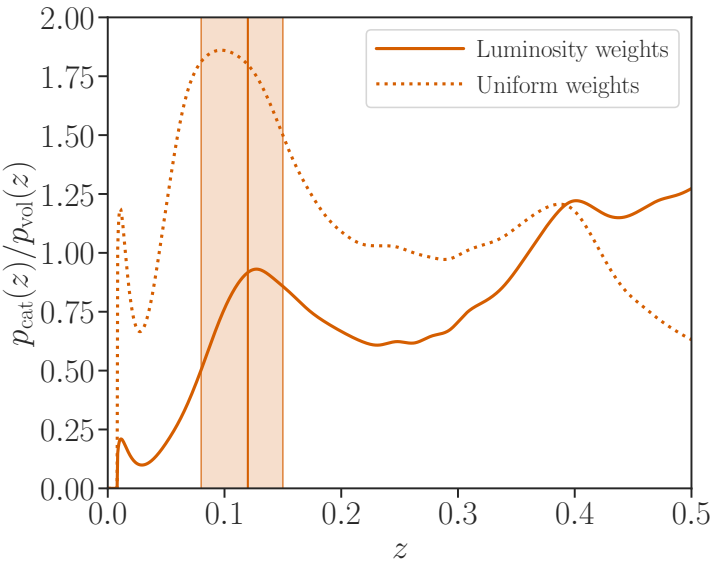


Figure B.1: The probability distribution for the redshifts of potential host galaxies $p_{\rm cat}(z)$, with redshift uncertainties taken into account, divided out by a uniform in comoving volume distribution $p_{\rm vol}(z)$ of galaxies. When computing $p_{\rm cat}(z)$ we include all galaxies brighter than $0.05L_g^*$ within the corresponding event's 99% sky localization region and weight each galaxy by weights proportional to their g-band luminosity (solid lines) as well as with uniform weights (dotted lines). We show these distributions for the DES-Y1 galaxies within the GW170814 sky localization region. We also show the 90% median estimated redshift range for GW170814 (calculated assuming a Planck 2015 cosmology) for reference.

Appendix C

Full Posterior Distributions on Polarization mode amplitudes

We provide the full posterior distributions for the cases investigated in Section 6.5 for completeness. In Table C.1 we show the priors that we used to infer the BBH parameters as well as the polarization mode amplitudes and other lensing observables.

Parameter	Shape	Limits	Boundary
d_L/Mpc	Powerlaw in d_L^2	$[1, 3000]$	-
μ	Uniform in Log	$[0.1, 10]$	-
ϵ_A	Uniform	$[0, 1]$	-
$\mathcal{M}_c/M_\odot$	Uniform	$[25, 100]$	-
q	Uniform	$[0.125, 1]$	-
$\chi_{1,2}$	Uniform	$[0, 0.99]$	-
ι	Sinusoidal	$[0, \pi]$	-
ψ	Uniform	$[0, \pi]$	Periodic
ϕ_c	Uniform	$[0, 2\pi]$	Periodic
α	Uniform	$[0, 2\pi]$	Periodic
δ	Cosinusoidal	$[-\pi/2, \pi/2]$	-
t_c	Uniform	$[t_c - 0.1, t_c + 0.1]$	-

Table C.1: We show the priors that we used to infer the BBH parameters as well as the polarization mode amplitudes and other lensing observables for the cases that we explored in Section 6.5. For the polarization mode amplitudes, we use a Dirichlet prior such that $\sum_A \epsilon_p = 1$ for $A \in \{+, \times, x, y, s\}$. For each lensed image that we analyze jointly, we take independent priors on t_c centered around the corresponding injected values.

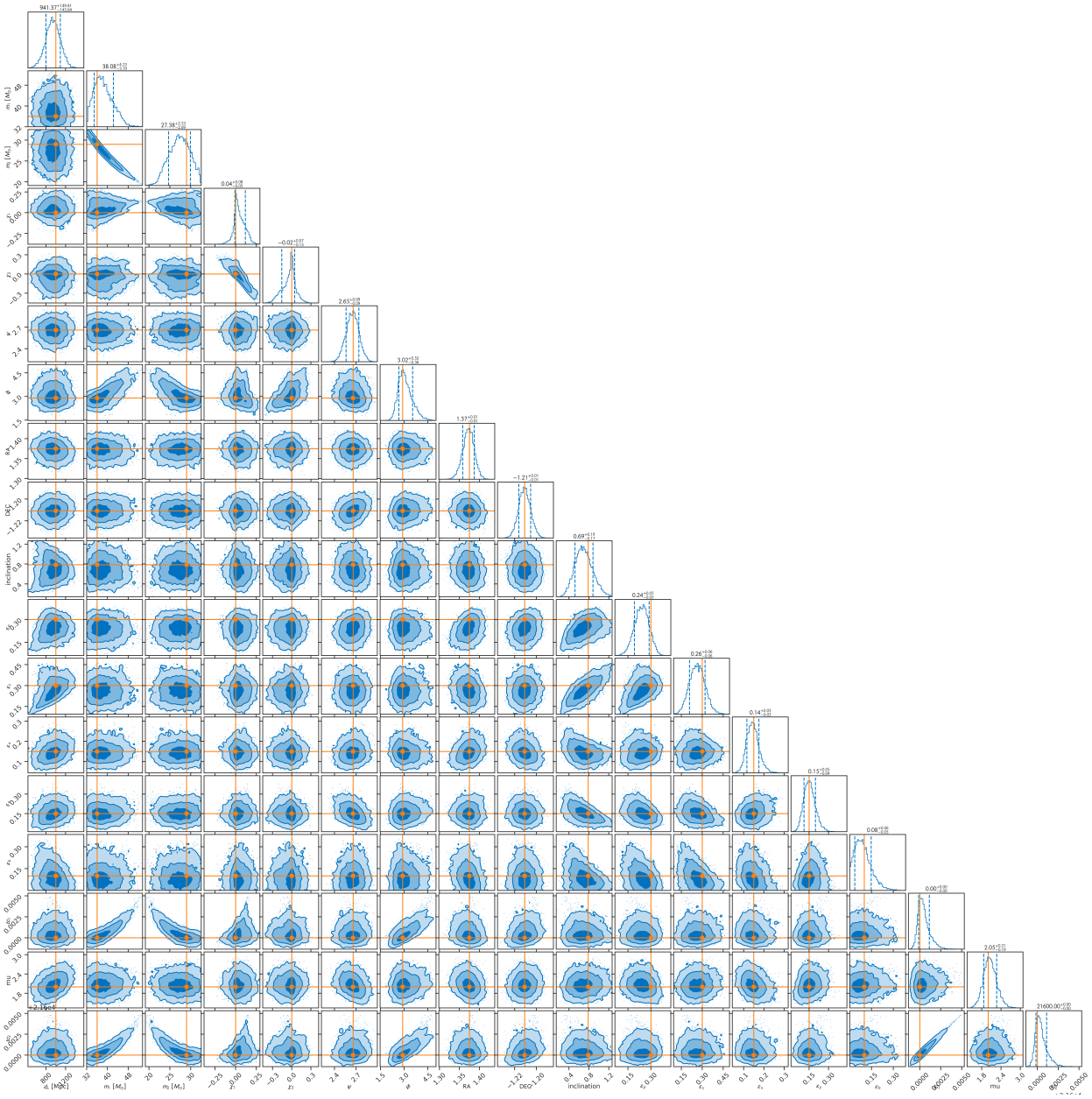


Figure C.1: Posterior distribution on the pair of lensed images as described in Section 6.5 observed by four detectors (HLVK) with a relative time delay of six hours with polarization mode amplitude of $\epsilon_+ = \epsilon_\times = 0.35$, $\epsilon_x = \epsilon_y = 0.15$ and $\epsilon_s = 0.05$.

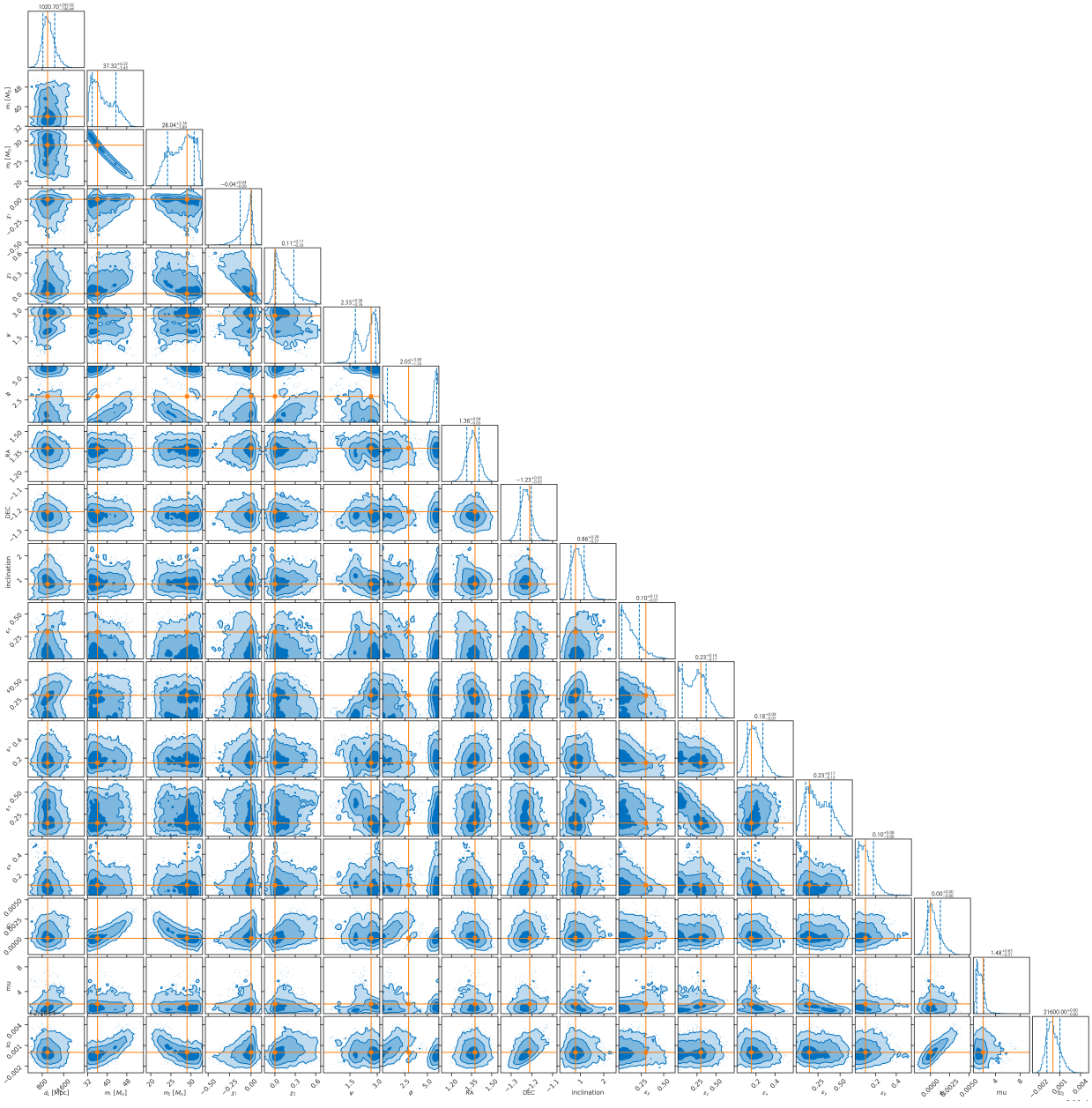


Figure C.2: Posterior distribution on the pair of lensed images as described in Section 6.5 observed by two detectors (HL).

Appendix D

Full Posterior Distributions on gravitational leakage models

In Figure D.1 we show the full posterior distribution on the BBH population parameters and on the gravitational leakage model parameters described by Equation 7.2. Similarly, Figure D.2, shows the full posterior distribution on the BBH population parameters but with the gravitational leakage model with screening defined in Equation 7.3.

In Table D.1 we show the prior ranges for the uniform priors that we use for the BBH population and gravitational leakage model parameters.

Parameter	Prior Range
$m_{\rm min}$	$[2, 10]M_\odot$
$m_{\rm max}$	$[50, 100]M_\odot$
α	$[-4, 12]$
β	$[-4, 12]$
$m_{\rm pp}$	$[20, 50]M_\odot$
$\sigma_{\rm pp}$	$[1, 10]M_\odot$
$\log_{10} f_{\rm pp}$	$[-5, 0]$
γ	$[-25, 25]$
β	$[0, 10]$
$z_{\rm p}$	$[0, 4]$
D	$[3, 7]$
$\log_{10} R_c/{\rm Mpc}$	$[1, 6]$
$\log_{10} n$	$[1, 2]$

Table D.1: We use uniform priors for the BBH population parameters and on the gravitational leakage model parameters described by Equation 7.2 and Equation 7.3 with the following ranges.

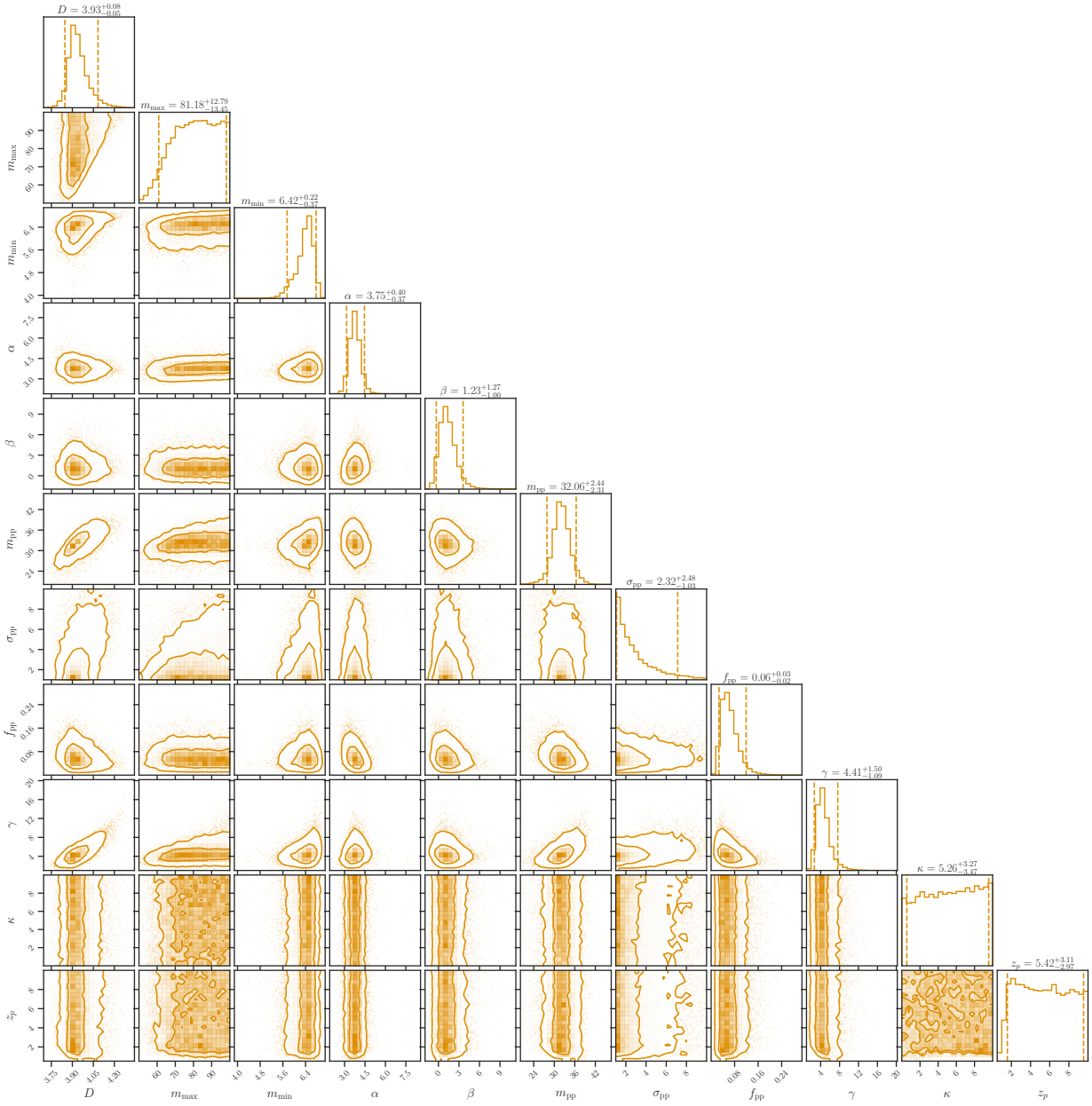


Figure D.1: Posterior distribution on number of spacetime dimensions D using the model in Equation 7.2, as well as the "POWER LAW + PEAK" mass distribution parameters $m_{\rm min}$, $m_{\rm max}$, α, β, $m_{\rm pp}$, $\sigma_{\rm pp}$ and $f_{\rm pp}$ and the SFR-like redshift evolution model with parameters γ, κ and $z_{\rm p}$.

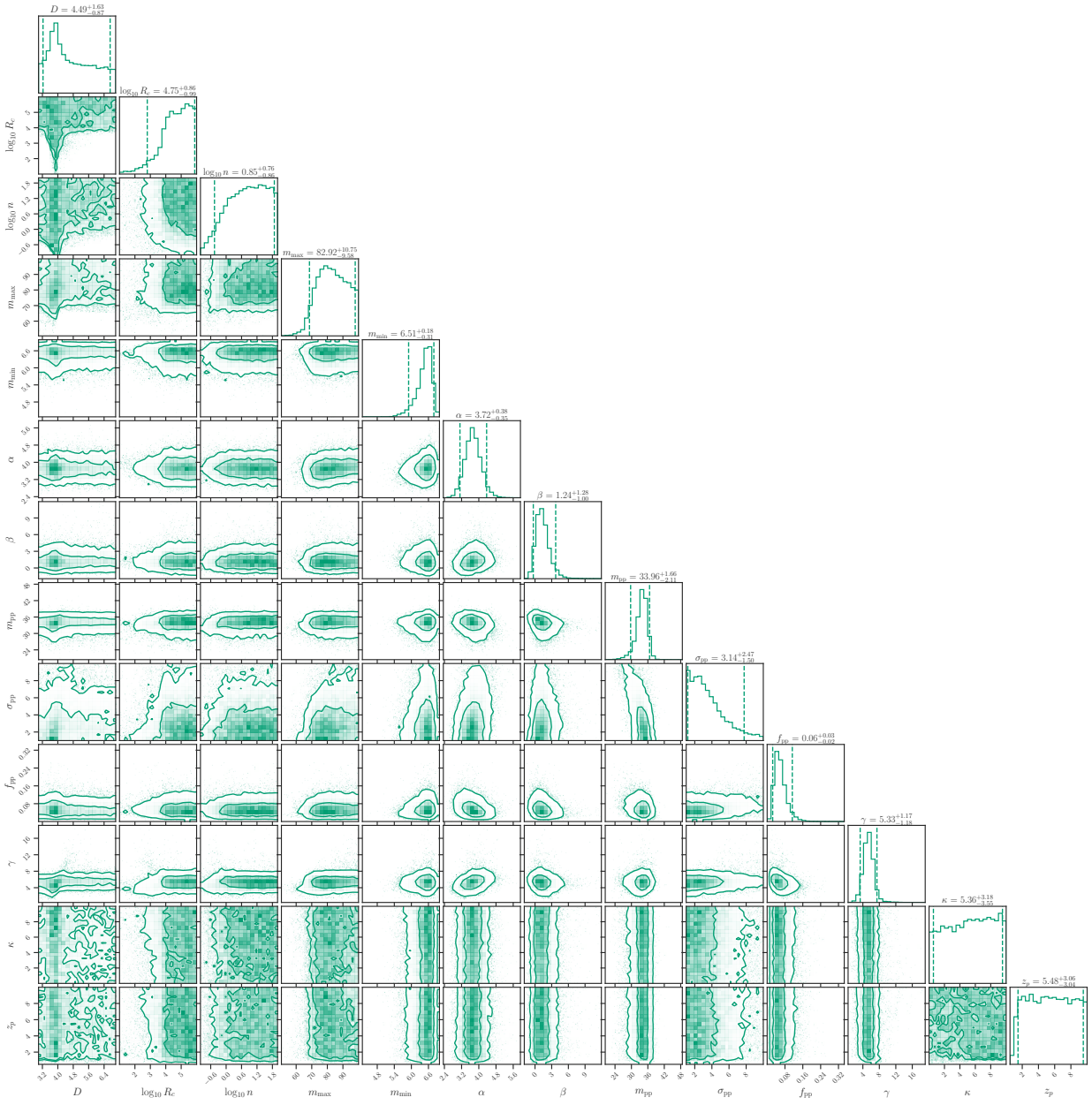


Figure D.2: Posterior distribution on the number of spacetime dimensions D, screening scale R_c and transition steepness n using the model in Equation 7.3, as well as the "POWER LAW + PEAK" mass distribution parameters the SFR-like redshift evolution model with parameters.

Ignacio Magaña Hernandez

Email: maganah2@uwm.edu
Web: http://ignaciomagana.github.io
Citizenship: USA

Address: Department of Physics
University of Wisconsin-Milwaukee
3135 N Maryland Ave
Milwaukee, WI 53211

Education

University of Wisconsin-Milwaukee — Milwaukee, WI
PhD in Physics — *September 2017 - Present*

University of California, Santa Barbara — Santa Barbara, CA
B.S., Physics (Highest Honors) — *September 2014 - September 2017*

Foothill College — Los Altos Hills, CA
Transfer Student, Physics — *September 2012 - June 2014*

Research Interests

Gravitational wave cosmology: measurements of the Hubble constant and other cosmological parameters using gravitational wave standard sirens such as merging binary black holes and neutron stars.

Gravitational wave lensing: its detection, data analysis and cosmological applications.

Modified gravitational wave propagation: testing deviations of ΛCDM and general relativity with gravitational wave standard sirens.

Awards & Fellowships

UWM Advanced Opportunity Program Fellowship — 2019 - Present

NSF Graduate Research Fellowship — 2017 - Present

NASA Wisconsin Space Grant Consortium Fellowship — Summer 2020 and Summer 2022

UWM Physics Department Lichtman Fellowship — Spring 2020

UCSB Physics Department Highest Academic Honors — Summer 2017

LIGO SURF Blanco Fellowship — Summer 2015

UCSB Dean's list — Fall 2014 - Spring 2017

NSF Science, Math and Engineering (S-STEM) Scholarship — Fall 2013 - Spring 2014

Research Positions

LIGO, University of Wisconsin-Milwaukee — *Research Assistant, September, 2017 - Present*

- My main work is on gravitational wave cosmology with standard sirens. I am most interested in using binary black hole events for which no electromagnetic counterpart is expected to infer cosmological parameters statistically with large galaxy surveys.
- I am one of the lead developers of the `gwcosmo` package[1] which is used by the LVK Collaboration for cosmological results using GWs and galaxy surveys.
- Additional work on strongly lensed gravitational wave events: detection, model selection and tests of GR through alternative polarization mode measurements. Additionally, one of the lead developers for the LVK Collaboration strong lensing model selection package, `hanabi`[2]

LIGO, The Chinese University of Hong Kong — *Research Assistant, March, 2016 - January 2017*

- Worked with Prof. Tjonnie Li performing GW data analysis for the CUHK LIGO group, work culminated in a successful NSF GRFP fellowship application.

LIGO Lab, Caltech — *Summer Undergraduate Research Fellow, June, 2015 - September, 2015*

- Worked with Prof. Rana Adhikari at the Caltech 40 meter prototype interferometer. I developed feedforward IIR Wiener filters to subtract seismic noise out the interferometer mode cleaner and arm cavities.

[1] https://git.ligo.org/lscsoft/gwcosmo
[2] https://github.com/ricokaloklo/hanabi/

Student Supervision

- **Pratyusava Baral** — University of Wisconsin-Milwaukee
  PhD student (co-advised) — *September 2021 - Present*
- **Amanda Baylor** — University of Wisconsin-Milwaukee
  PhD student (co-advised) — *September 2021 - Present*
- **Brandon Piotrzkowski** — University of Wisconsin-Milwaukee
  PhD student (co-advised) — *September 2019 - February 2020*

Short Author Publications

Citations according to inspirehep.net on October 27, 2022

1. *Dark siren cosmology with binary black hole mergers*
   I. Magaña Hernandez, V. D'Emilio and P. Brady
   (In preparation)
2. *On the association of GW190425 with its potential electromagnetic counterpart FRB 20190425A*
   I. Magaña Hernandez, V. D'Emilio, S. Morisaki, M. Bhardwaj and A. Palmese
   (In preparation)
3. *Localization of binary neutron star mergers with a single Cosmic Explorer*
   P. Baral, S. Morisaki, **I. Magaña Hernandez** and J.D.E. Creighton
   (arXiv:2304.09889)
4. *Non-parametric inference of the population of compact binaries from gravitational wave observations using binned Gaussian processes*
   A Ray, **I. Magaña Hernandez**, S. Mohite, J.D.E. Creighton and S. Kapadia
   (arXiv:2304.08046)
5. *The Hitchhiker's guide to the galaxy catalog approach for gravitational wave cosmology*
   J. Gair et al. (including **I. Magaña Hernandez**)
   (arXiv:2212.08694)
6. *Measuring the polarization content of gravitational waves with strongly lensed binary black hole mergers*
   I. Magaña Hernandez
   (arXiv:2211.01272, 2 citations)
7. *Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers*
   I. Magaña Hernandez
   Phys. Rev. D 107, 084033 (2023) (arXiv:2112.07650, 5 citations)
8. *A Bayesian statistical framework for identifying strongly-lensed gravitational-wave signals*
   R.K.L. Lo and **I. Magaña Hernandez**
   (arXiv:2104.09339, 20 citations)
9. *First measurement of the Hubble parameter from bright binary black hole GW190521*
   S. Mukherjee et al. (including **I. Magaña Hernandez**)
   (arXiv:2009.14199, 34 citations)
10. *A joint ranking statistic for multi-messenger astronomical searches with gravitational waves*
    B. Piotrzkowski, A. Baylor and **I. Magaña Hernandez**
    Class. Quantum Grav. 39 085010 (2022) (arXiv:2111.12814, 1 citations)
11. *Current observations are insufficient to confidently associate the binary black hole merger GW190521 with AGN J124942.3 + 344929*
    G. Ashton, K. Ackley, **I. Magaña Hernandez** et al.
    Class. Quantum Grav. 38 235004 (2021) (arXiv:2009.12346, 41 citations)
12. *Rapid model comparison of equations of state from gravitational wave observation of binary neutron star coalescences*
    S. Ghosh, X. Liu, J.D.E. Creighton, **I. Magaña Hernandez** et al.
    Phys. Rev. D 104, 083003 (2021) (arXiv:2104.08681, 7 citations)
13. *Gravitational wave friction in light of GW170817 and GW190521*
    S. Mastrogiovanni, L. Haegel, K. Karathanasis, **I. Magaña Hernandez** and D.A. Steer
    JCAP 02 043 (2021) (arXiv:2010.04047, 22 citations)

14. *Identifying strong gravitational-wave lensing during the second observing run of Advanced LIGO and Advanced Virgo*
X. Liu, **I. Magaña Hernandez** and J.D.E. Creighton
ApJ 908 97 (2021) (arXiv:2009.06539, 31 citations)

15. *Cosmological inference using gravitational wave standard sirens: A mock data analysis*
R. Gray, **I. Magaña Hernandez** et al.
ApJ 908 97 (2021) (arXiv:1908.06050, 79 citations)

16. *Bayesian inference for compact binary coalescences with BILBY: Validation and application to the first LIGO–Virgo gravitational-wave transient catalogue*
I.M. Romero-Shaw et al. (including **I. Magaña Hernandez**)
MNRAS 499, 3295 (2020) (arXiv:2006.00714, 155 citations)

17. *A statistical standard siren measurement of the Hubble Constant from the LIGO/Virgo gravitational wave compact object merger GW190814 and Dark Energy Survey galaxies*
A. Palmese et al. (including **I. Magaña Hernandez**)
ApJL 900 L33 (2019) (arXiv:2006.14961, 60 citations)

18. *A standard siren measurement of the Hubble Constant from GW170817 without the electromagnetic counterpart*
M. Fishbach, R. Gray, **I. Magaña Hernandez** et al.
ApJL 871 L13 (2019) (arXiv:1807.05667, 137 citations)

LIGO-Virgo Publications to which I highly contributed:

1. *GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object*
LIGO Scientific and Virgo Collaborations (including **I. Magaña Hernandez**)
ApJL 871 L13 (2019) (arXiv:2006.12611, 871 citations) **Led and wrote the Hubble constant analysis**

2. *A Gravitational-wave Measurement of the Hubble Constant Following the Second Observing Run of Advanced LIGO and Virgo*
LIGO Scientific and Virgo Collaborations (including **I. Magaña Hernandez**)
ApJ 909 218 (2021) (arXiv:1908.06060, 185 citations) **Paper writing and analysis team member**

3. *First measurement of the Hubble constant from a dark standard siren using the Dark Energy Survey galaxies and the LIGO/Virgo binary-black-hole merger GW170814*
M. Soares-Santos, A. Palmese et al. and LIGO Scientific and Virgo Collaborations (including **I. Magaña Hernandez**)
Astrophys. J. 875, 161 (2019) (arXiv:1901.01540, 169 citations) **Validated analysis using `gwcosmo`**

Invited Seminars

1. **Cardiff University** September 2022
Cosmology and modified gravity with binary black holes
2. **UC Berkeley** June 2022
Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers
3. **University of Wisconsin-Milwaukee** April 2022
Cosmology and modified gravity with binary black hole mergers
4. **Eotvos Lorand University** September 2019
Gravitational wave cosmology with large galaxy surveys
5. **University of Glasgow** August 2019
Measuring the Hubble constant with gravitational-wave standard sirens
6. **The Chinese University of Hong Kong** February 2019
A measurement of the Hubble constant from the second observation run of Advanced LIGO-Virgo
7. **Nagoya University** February 2019
A measurement of the Hubble constant from the second observation run of Advanced LIGO-Virgo

Talks at Conferences

1. **APS April Meeting** April 2023
Measuring the polarization content of gravitational waves with strongly lensed binary black hole mergers
2. **11th Australasian Conference on General Relativity and Gravitation** February 2022
Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers
3. **Midwest Relativity Meeting** October 2021
Constraining the number of spacetime dimensions from GWTC-3 binary black hole mergers

4. **Midwest Relativity Meeting** October 2020
   On the candidate electromagnetic counterpart of GW190521 and its implications on the Hubble constant
5. **Gravitational Waves Physics and Astronomy Workshop** October 2019
   A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO and Virgo
6. **Multi-Messenger Astrophysics in the Gravitational Wave Era** September 2019
   A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO and Virgo
7. **YITP Asian-Pacific Winter School and Workshop on Gravitation and Cosmology** February 2019
   A standard siren measurement of the Hubble constant from GW170817 without the electromagnetic counterpart
8. **Midwest Relativity Meeting** October 2018
   Cosmological inference using gravitational wave standard sirens: A mock data challenge
9. **19th Annual Physical Society of Hong Kong** June 2016
   Feedforward Seismic Noise Cancellation at the 40m Prototype Interferometer

Posters

1. **LVK Meeting: Cardiff University** September 2022
   Measuring the polarization content of gravitational waves with strongly lensed binary black hole mergers
2. **Merging Visions: Exploring Compact-Object Binaries with Gravity and Light** June 2019
   Identifying Strong Gravitational-wave Lensing during the Second Observing Run of Advanced LIGO and Advanced Virgo
3. **Gravitational Waves Physics and Astronomy Workshop** December 2018
   Cosmological inference using gravitational wave standard sirens: A mock data analysis

Additional Training

1. **Black Holes and Neutron Stars with Gravitational Waves** October 2019
   Yukawa Institute for Theoretical Physics
2. **The New Era of Gravitational-Wave Physics and Astrophysics** May - July 2019
   Kavli Institute for Theoretical Physics
3. **Asian-Pacific Winter School and Workshop on Gravitation and Cosmology** February 2019
   Yukawa Institute for Theoretical Physics
4. **Les Houches Summer School: Gravitational-Wave Astronomy** July 2018
   Ecole de Physics des Houches
5. **Summer School on Gravitational-Wave Astronomy** August 2016
   ICTS-TIFR

Societies/Affiliations

- LIGO Scientific Collaboration member since June 2015
- American Physical Society
- Society for Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS)
- National Society of Hispanic Physicists

Skills

- **Programming skills:**
  - Proficiency with: Python (NumPy, SciPy, Pandas, Scikit-learn), JAX, C, Mathematica, LaTeX.
  - Familiar with: Stan, C++, MATLAB/Simulink.
  - LVK related software: Bilby, LALSuite, PyCBC, ligo.skymap, gwcosmo, hanabi.
- **Data Analysis:**
  - Bayesian model selection and parameter estimation, Monte Carlo methods.
  - Digital signal processing, machine learning and basic numerical relativity experience.
- **Languages:** English (fluent), Spanish (fluent) and Japanese (beginner).

References

Patrick R. Brady (PhD Supervisor)
University of Wisconsin-Milwaukee
prbrady[at]uwm.edu

Jolien Creighton
University of Wisconsin-Milwaukee
jolien[at]uwm.edu

Antonella Palmese
Carnegie Mellon University
apalmese[at]andrew.cmu.edu

Archisman Ghosh
Ghent University
archisman.ghosh[at]ugent.be

Jonathan Gair
Max Planck Institute for Gravitational Physics
jonathan.gair[at]aei.mpg.de

Soichiro Morisaki
The University of Tokyo
soichiro[at]icrr.u-tokyo.ac.jp