

Effect of gravitational lensing on the population inference of binary black holes using gravitational-wave observations

A Thesis

submitted to

Indian Institute of Science Education and Research Pune

in partial fulfillment of the requirements for the

BS-MS Dual Degree Programme

by

Neev Shah



Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

May, 2024

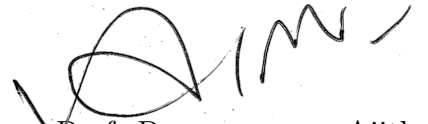
Supervisor: Parameswaran Ajith

© Neev Shah 2024

All rights reserved

Certificate

This is to certify that this dissertation entitled Effect of gravitational lensing on the population inference of binary black holes using gravitational-wave observations towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents work carried out by Neev Shah at the Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Parameswaran Ajith, Professor, International Centre for Theoretical Sciences (ICTS-TIFR), Bengaluru, during the academic year 2023-2024.



Prof. Parameswaran Ajith

Committee:

Prof. Parameswaran Ajith

Prof. Susmita Adhikari

Declaration

I hereby declare that the matter embodied in the report entitled Effect of gravitational lensing on the population inference of binary black holes using gravitational-wave observations, are the results of the work carried out by me at the Department of Physics, Indian Institute of Science Education and Research, Pune, and the International Centre for Theoretical Sciences (ICTS-TIFR), Bengaluru, under the supervision of Prof. Parameswaran Ajith, International Centre for Theoretical Sciences (ICTS-TIFR), and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.



Neev Shah

IISER Roll Number: 20191011

Date: 4th March, 2024

This thesis is dedicated to *my parents*

Acknowledgements

Firstly, I would like to thank Dr. Ajith for providing me the opportunity to work on this project, and for all the valuable inputs, discussions and suggestions throughout. I am also grateful to Mukesh for the useful discussions and helping me at numerous stages with the various tools. I also thank the Astrophysical Relativity group at ICTS for the engaging and lively atmosphere conducive for research, and the Long Term Visiting Student program at ICTS for financial support. I would also like to express my thanks to my thesis expert Dr. Adhikari for her constructive feedback and suggestions at various points over the course of my thesis.

I am extremely grateful to Akanksha, my friends Aniketh, Varun, Amogh Ranade, Divyansh, Garvit, Amogh Rakesh, Vasudha, Aditya Pujari, Vatsal, Pallav and Jezer for their moral support and presence throughout the challenging times. This would be incomplete without thanking my family, and in particular my parents. Thank you for trusting and encouraging me to pursue my interests, even if it was an uncommon path and perhaps difficult to understand, and providing me the freedom and support to figure out my own path.

Lastly, I would like to acknowledge my privilege that has allowed me to reach where I am, and I thank everyone trying to make this world a just, inclusive, and more equitable place.

Abstract

Gravitational lensing due to intervening matter distributions such as galaxies or clusters can (de)–magnify a gravitational-wave (GW) event, which can introduce a bias in the measurement of the astrophysical source’s properties. Hierarchical Bayesian inference on the catalog of detected GW events is performed to study the population properties of compact binaries, such as their mass and redshift distributions. Currently, the lensing probability is low and it is assumed that the events are not significantly (de)–magnified. A higher lensing probability (as expected for the next-generation detectors), can lead to biases in our estimation of the population hyper–parameters. In this work, we investigate the biases in population inference due to lensing, and develop a Bayesian hierarchical inference formalism for correct estimation of both the GW source population hyper–parameters, and the lenses.

Contents

Front Matter	i
Abstract	xii
Contents	xv
List of Figures	xix
I Preliminaries	1
1 Introduction	3
1.1 A very brief outline of the rest of this thesis	4
2 What are Gravitational Waves	5
2.1 Too weak to detect?	6
2.2 Gravitational Wave Astronomy: Past, Present and Future	8
2.3 One to Many	8

2.4	What have we learned?	10
2.5	The Future	13
3	Gravitational Lensing	15
3.1	The Lens Equation	15
3.2	Lensing done in 2 ways	21
4	Lensing of Gravitational Waves	23
4.1	Gravitational Waves: Meet Gravitational Lensing	23
4.2	An Aside on what does a GW signal really tell us?	25
4.3	Oh, but lensing?	26
4.4	Current Searches for Gravitational Wave Lensing	28
4.4.1	How does one know a GW signal is lensed?	28
4.5	Gravitational Wave lensing as a tool	30
4.6	Magnification Probability Distribution	32
II	Developing the tools	35
5	Methods	37
5.1	Parameter Estimation in a Nutshell	37
5.1.1	Bayesian Inference	38
5.1.2	Evaluating the posterior	40
5.2	Population Inference in a nutshell	42
5.3	Common Population Models	44
5.3.1	Redshift Model	44

Power Law Model	44
5.3.2 Mass Model	45
5.4 An Aside on Selection Effects	47
5.4.1 The Detection Fraction	47
5.5 Population Inference in the presence of Lensing	52
III Major Takeaways & Discussion	57
6 Results and Discussion	59
6.1 Toy Model	59
6.1.1 True Population	60
6.2 Toy Model results	62
6.3 Can one learn the lensing magnification distribution?	64
6.3.1 Correlations between the lensing and population parameters	65
6.4 Toy Model for Selection Effects	67
6.5 Towards a realistic scenario	69
7 Conclusion and Future Directions	75
References	79

List of Figures

2.1	PSD of the existing LIGO detectors, and some future planned ground based detectors. Reproduced from Fig 3.3 in [7]	7
2.2	The whitened strain timeseries and the Q-Scan for the first detection, GW150914	9
2.3	<i>Masses in the Stellar Graveyard</i> : The masses of all the events and their merger remnants found in GWTC-3. Figure Credit: Aaron Geller/ Northwestern . .	10
2.4	The primary mass distribution (<i>left</i>) and the mass ratio distribution (<i>right</i>) of black holes as inferred using GWTC-3	11
2.5	The evolution of the merger rate density across redshift of merging black holes as inferred using GWTC-3	12
2.6	The horizon distances for the current and various planned future detectors. Fig 3.3 in [7]	13
3.1	Lensing sketch depicting the geometry and various angular vectors involved for the source and image positions [32].	16
3.2	A Quadruply lensed quasar HE 0435 1223. The foreground galaxy acts as a lens and splits the background quasar into 4 almost evenly spaced images. Image Credit: NASA/ESA/Hubble	19
3.3	Time shifted light curves of the 4 images of a quad quasar, HE 0435 1223 [34]. The background quasar is intrinsically variable, which is seen in the 4 images as well. Importantly, the 4 light curves have a time delay between them due to lensing.	20

3.4	An Einstein Ring, a highly distorted and magnified background galaxy due to very good alignment between us, the lens and the background source. Image Credit: NASA/ESA/Hubble	20
3.5	An explanatory cartoon of the weak lensing phenomenon. We can see that background galaxies undergo small distortions due to an intervening lens. The distortion field is only noticeable for an ensemble of galaxies. Image Credit: Wikipedia	22
4.1	<i>Top:</i> Strong lensing of GWs in the geometrical–optics regime, which leads to multiple images seen at the detector with a time delay, and having identical frequency evolution but with differing amplitudes due to lensing (de)–magnification. <i>Bottom:</i> Lensing of GWs in the wave optics–regime due to a smaller lens. This leads to frequency dependent modulations of the gravitational wave signal, beating patterns etc. Figure Credit: LIGO Scientific Collaboration	24
4.2	An example of an unlensed waveform, and the corresponding lensed waveform with a $\mu = 4$	27
4.3	The weak lensing + strong lensing magnification distribution as it varies with source redshift.	33
5.1	GW150914 corner plot for the 2 component source frame masses	41
5.2	The detection fraction as a function of the population parameters evaluated using an injection set of 100000 events. In each panel, one of the population parameters is varied, while the others are fixed to their values as per the injected population. The bands depict the 2σ uncertainty in the estimate of the detection fraction.	50
5.3	Same as Fig. 5.2, but using an injection set of 1000000 events to show that the uncertainty in the estimate of the detection fraction decreases with more signals.	50
6.1	The blue solid curve in the left and middle panel show the mass and redshift distribution for binaries in our toy population mode, while the orange dashed curve shows the corresponding lensed population. The right panel shows the lensing magnification distribution that we have assumed.	61

6.2	Constraints on the hyper–posteriors of the standard population inference on our unlensed catalog. The diagonal panels represent the marginalized posteriors for the population parameters, while the rest of the panels show the two–dimensional posterior distribution for each pair of hyper–parameters. The orange solid lines mark the true values, and the number quoted on top is for the median values with the the 90% credible interval, which is marked by the black dashed lines.	63
6.3	Constraints on the hyper–posteriors of the standard population inference on our lensed catalog to demonstrate that we recover biased population parameters.	64
6.4	Constraints on the hyper–posteriors of the modified population inference on our lensed catalog to demonstrate that we recover the true population parameters.	65
6.5	Constraints on the hyper–posteriors of the modified population inference on our lensed catalog with the slope of the magnification distribution also as a free hyper parameter. We demonstrate that we can jointly recover the correct slope of the power law, as well as the true population parameters.	66
6.6	<i>Blue</i> : Constraints on the hyper–posteriors of the modified population inference on our lensed catalog with a detection criterion to demonstrate the biases in κ due to not accounting for selection effects. <i>Orange</i> : The same description, except that the population inference procedure accounts for selection effects, and we can see that we recover the correct redshift distribution parameter κ . The dark gray lines represent the true values for the hyper–parameters. . . .	68
6.7	The blue solid curve in the top panels, and the bottom–left panel show the primary mass, redshift and mass–ratio distribution of binaries for our new population model, while the orange dashed curve in the same panels show the corresponding lensed population as obtained using the more realistic lensing magnification model. The bottom–right panel depicts the lensing magnification distribution for the events in our catalog, with $\mu = 1$ marked for reference. .	70
6.8	Constraints on the hyper–posteriors of standard population inference on our unlensed catalog to demonstrate that we recover the true population parameters.	71
6.9	Constraints on the hyper–posteriors of standard population inference on our lensed catalog to demonstrate the biases in the recovered population–parameters.	73

Part I

Preliminaries

Chapter 1

Introduction

Since the first direct detection of gravitational waves (GWs) in 2015, the LIGO-Virgo-KAGRA network of interferometers have detected many tens of binary blackhole mergers over three observing runs, including a few binary neutron star (BNS) and neutron star – black hole (NSBH) coalescences. With the 4th observing run that started in May 2023, it is expected that hundreds of more events will be detected. With such a large number of mergers, one can now start asking questions about the underlying population properties of binary black holes in the universe, like what is the mass distribution of black holes in the universe, merger rates and their evolution with redshift etc. An interesting feature about GWs from CBCs is that they are a standard siren without requiring any distance ladder calibration. This means that the luminosity distance to the merger can directly be estimated from the data along with the redshifted mass. Thus knowing the luminosity distance and the redshifted mass, we can also calculate the source frame masses by assuming a cosmology. One can then study the intrinsic population properties of the binary black holes. However, this assumes that the luminosity distance is unbiased. Gravitational waves, like electromagnetic waves, can be gravitationally lensed by intervening matter like galaxies and clusters. This can provide a (de)–magnification to the amplitude of the gravitational waves, which leads to the luminosity distance to be underestimated, and the source masses to be overestimated, thus leading to a bias in the measurements. Thus, lensing of events can bias the source mass measurements and thus also bias the population properties of the black holes. In order to accurately study the intrinsic properties of the merging binary black hole population, one needs to account for the fact that the detected GW events may be gravitationally lensed. Currently, the

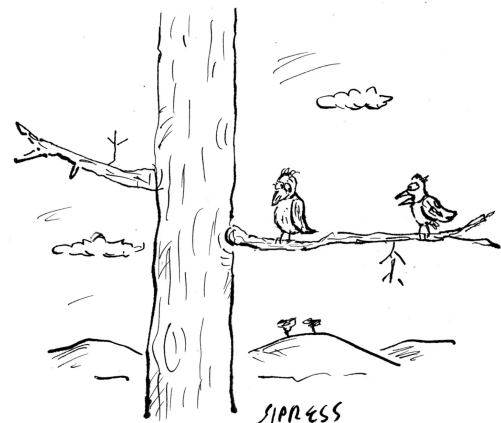
lensing probabilities calculated from the merger rates and known lens population models are expected to be quite small and none of the currently confirmed events are thought to be significantly lensed. However, as the reach of the current GW detectors increases in the future observing runs, the volume of universe probed also increases as well as the chances of detecting a significant number of lensed events. Thus, it is essential to study in detail what features can GW lensing leave on the source population distributions, and how to account for them to understand our true source population. In this project we will use the framework of hierarchical bayesian inference to study the population distributions of the merging BBH's and study the consequences of including lensing in such models.

1.1 A very brief outline of the rest of this thesis

This thesis is divided into three parts. Part I provides some background and overview into what are gravitational waves and what have we learned from them. It also provides a brief introduction to gravitational lensing, before delving deeper into lensing of gravitational waves, their current status and future prospects. Part II then describes some of the bayesian methods used to study individual gravitational wave events, as well as populations from a catalog of detections. We also develop a bayesian hierarchical formalism on how to include lensing in population studies of gravitational waves. In Part III, we apply the methods developed in the previous sections to explore the effects of lensing in population studies and demonstrate that our modified inference procedure can mitigate the resulting biases. We then conclude with a summary of our work and its key points, along with some future directions to extend this study.

Chapter 2

What are Gravitational Waves



"Was that you I heard just now, or was it two black holes colliding?"

Credit: New Yorker comic for the announcement of the first detection

From electromagnetism, we know that radiation can be generated by accelerating charges. This can make one ponder, can accelerating masses do the same? The answer is yes, as was predicted by Einstein more than a century ago using his General Theory of Relativity (GR). Such radiation is called gravitational waves (GWs), which are literally ripples in the fabric of *spacetime*. GWs can be generated by any source of matter or energy that has a time varying quadrupole moment, which implies that spherically symmetric motions do not generate GWs.

Additionally, in GR, these waves propagate at the speed of light, and have 2 polarisation states, called the "+" and "×" polarisations. However, in alternate theories of gravity, GWs may have different propagation effects and there may also be up to 6 polarisation states. Now that we have briefly described what GWs are, how does one actually produce them, *astrophysically speaking*, and can we detect them? In this section, our focus will be on the ground-based GW detectors, which operate in the few Hz to kilo-Hz frequency regime. This is well suited to study transient sources of GWs [1], and the most common among them, and the sole ones detected thus far, are the compact binary coalescences (CBC). The inspiral and merger of 2 compact objects orbiting each other generates gravitational waves that propagate towards us at the speed of light. The amplitude and frequency of the waves increases with time till the merger, giving them a characteristic "chirp" like shape. Other plausible sources, that the ground-based detectors may be able to detect, but haven't been seen yet are gravitational waves generated by nearby core-collapse supernovae, long-duration continuous gravitational waves from spinning neutron stars that have some ellipticity or asymmetry, and the stochastic gravitational wave background from a collection of gravitational wave sources that may not be individually resolvable.

2.1 Too weak to detect?

Thankfully(?), the CBC's that we have detected and expect to detect in the future will be very far from us. This results in the amplitude of the GWs at the detector to be very small, as the amplitude diminishes inversely with the distance to the source¹. The strain amplitude, h for an inspiral can be approximately written as [2] –

$$h \sim 1.4 \times 10^{-22} \left(\frac{f}{100\text{Hz}} \right)^{2/3} \left(\frac{\mathcal{M}}{1.22M_{\odot}} \right)^{5/3} \left(\frac{10\text{Mpc}}{r} \right) \quad (2.1)$$

where f denotes the gravitational wave frequency, \mathcal{M} represents the chirp mass, which is a combination of the 2 component masses, and r is the luminosity distance to the source. As can be seen by the constant pre-factor, GWs for the usual astrophysical sources will always be very weak, with strains of the order of $\mathcal{O}(10^{-22})$. Thus, directly detecting them has

¹This is unlike electromagnetic radiation where the flux/energy falls off as the square of the distance from the source

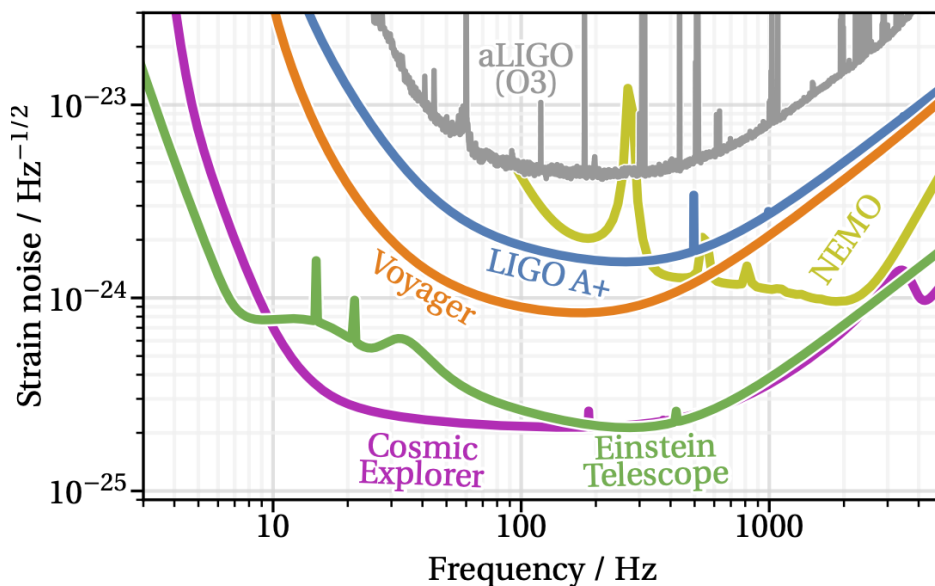


Figure 2.1: PSD of the existing LIGO detectors, and some future planned ground based detectors. Reproduced from Fig 3.3 in [7]

been a challenging effort over multiple decades. We now have kilometer-long ground-based laser interferometers for direct detection, with the current network consisting of the 2 LIGO detectors in Hanford and Livingston in the USA [3], the Virgo detector in Italy [4], and the underground KAGRA detector in Japan [5], with more observatories joining soon in the future [6]. Fig. 2.1 shows the power spectral density (PSD) of the current and future ground-based detectors. The PSD provides an estimate of the frequency-dependent noise in the detectors, and includes contributions from most possible terrestrial noise sources. They have a characteristic shape, which is due to poor sensitivity at the 2 extremes in the frequency space. On the low-frequency side, the ground based detectors encounter the seismic noise wall, which makes them insensitive to the very early parts of the inspiral, as well as the merger of extremely massive sources which merge at low frequencies. On the other side, quantum noise limits the detector performance at high frequencies.

Even with the most sensitive instruments ever constructed, the detection of GWs is still a challenging task as they are typically buried in noise from terrestrial sources. One of the methods to search for GWs in the data is called *Matched Filtering* [8, 9], where we cross-correlate the data with waveform models, that are created based on our understanding of what the signal from CBCs looks like. The cross-correlation is done with a large template bank, covering a range of different possible values of the source parameters like their masses

to find the best possible set of parameters that match with the data. GW interferometers have poor sky resolution, hence it is beneficial to have multiple detectors spread across Earth for better sky localization of the sources, especially to aid in electromagnetic followup with ground and space-based telescopes. Having multiple detectors also increases the significance of candidate events, as often, transient noise sources can mimic as a GW event, but having the same transient noise in all the detectors at roughly the same time is much less likely.

2.2 Gravitational Wave Astronomy: Past, Present and Future

Around 1.4 billion years ago, in a galaxy far, far away, a pair of black holes, each weighing approximately $30M_{\odot}$ collided with each other. Their inspiral dance and subsequent merger generated gravitational waves. On the 14th of September in 2015, the Laser Interferometer Gravitational Wave Observatory (LIGO) at Hanford and Livingston directly detected these waves, thereby opening a new window to *listen* to the universe [10]. This was the culmination of decades of effort, by a worldwide collaboration of scientists building the most sensitive detectors ever made. This was a watershed moment in understanding our universe, akin to when Galileo pointed his telescope in the night sky 400 years back, allowing us to *look* at the heavens in a completely new way. The subsequent birth of modern electromagnetic astronomy has completely revolutionized our understanding of the universe in the past few centuries, occurring hand-in-hand with the development of ever more sophisticated instruments to study and explore various astrophysical phenomena.

Similar is the importance of the first direct detection of gravitational waves. This is just the start, a glimpse of more to come with this new probe of the dark universe in our hands. Since the first detection in 2015, the LVK network of interferometers has detected the inspiral and merger of more than 50 binary black holes (BBH), a couple of binary neutron stars (BNS), and a couple of neutron star–black hole (NSBH) mergers [11]. The ongoing fourth observing run has also found many tens of signals from this final dance and collision of compact objects.

2.3 One to Many

Each detection tells us a story, something unique about the sources that gave rise to the detection. This could be the masses or spins of the compact objects that collided, the

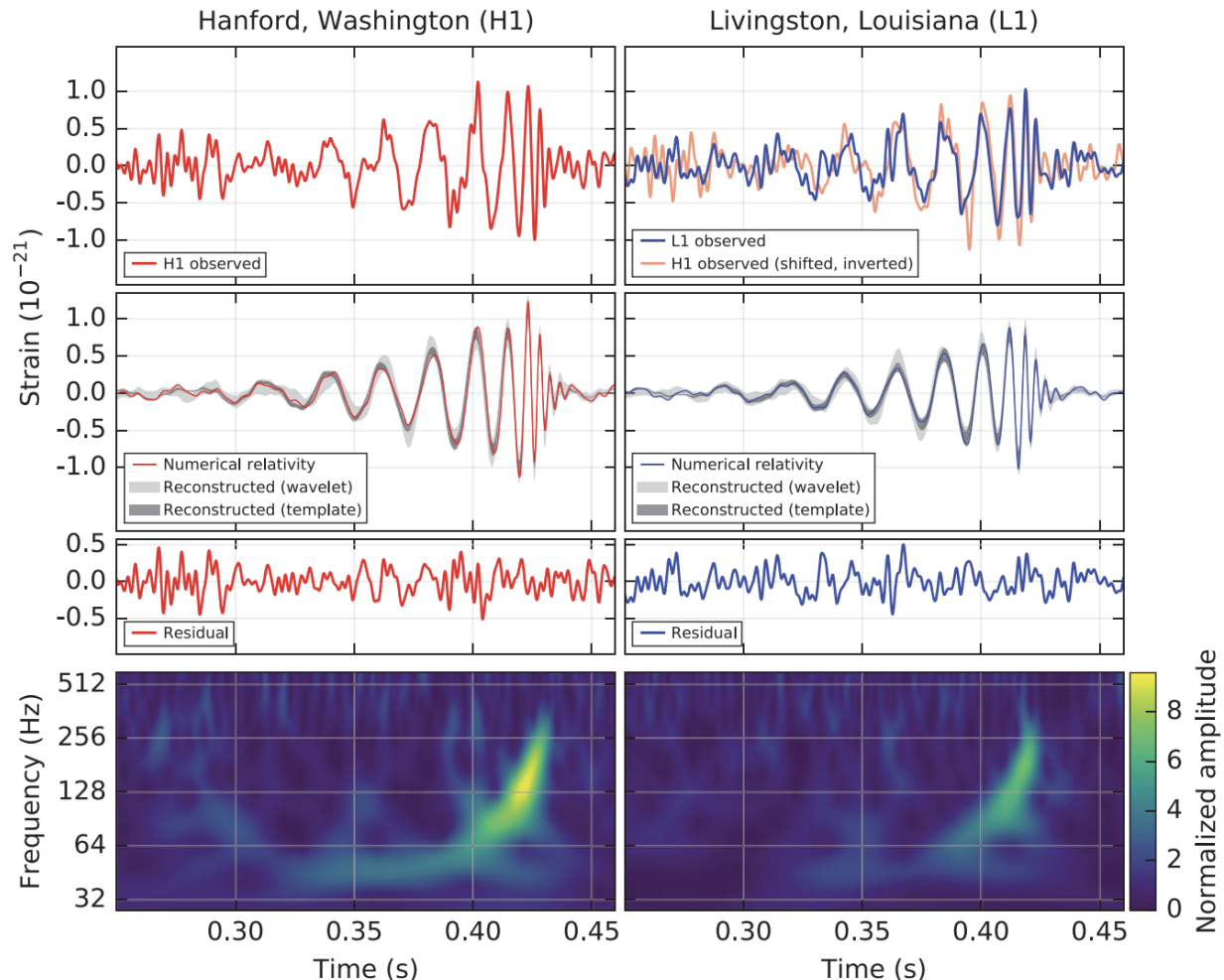


Figure 2.2: The whitened strain timeseries and the Q-Scan for the first detection, GW150914

direction in the sky in which they came from, or how far away did the merger occur. These, and a few other observables can be directly inferred from the data of each detection. However, that is just the tip of the iceberg of what we can achieve with GWs. We can also stitch together the individual short stories from each detection, into a coherent picture, a common thread that is shared by these fascinating events. This is the realm of population analyses of gravitational wave events, where we use our catalog of detections to understand the astrophysical population properties of black holes and neutron stars [12], as well as questions in cosmology [13]. For example, we can ask whether all the GW events that we observe arise from a common population, or are there multiple distinct sub-populations, each forming compact binaries in different environments. What is the mass function of black holes and neutron stars, how are they distributed in redshift, and what are their different formation and

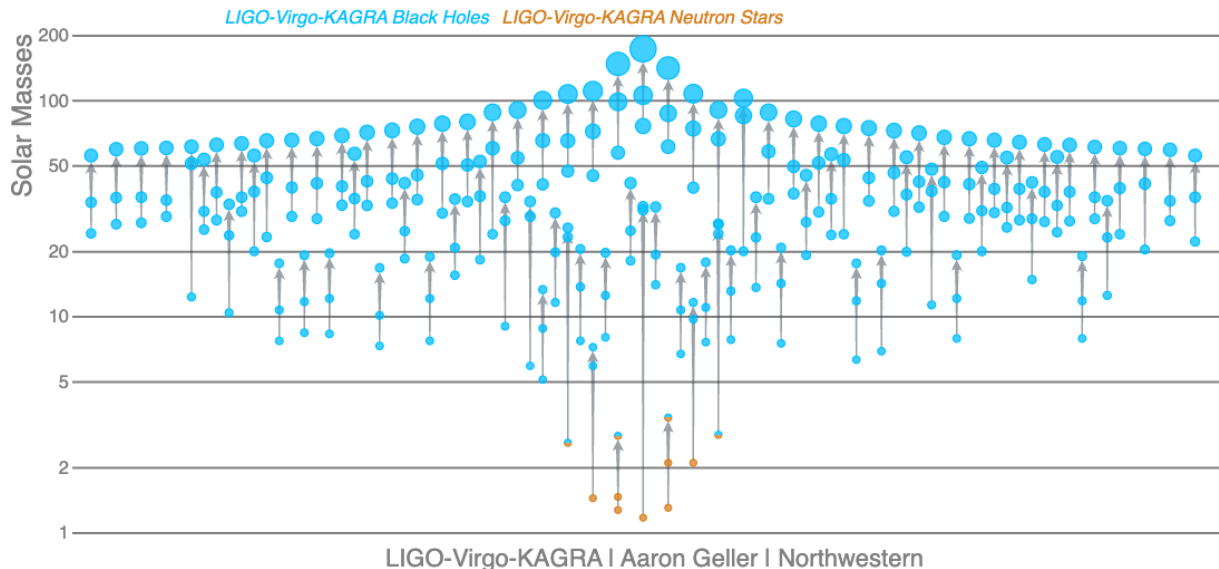


Figure 2.3: *Masses in the Stellar Graveyard:* The masses of all the events and their merger remnants found in GWTC-3. Figure Credit: Aaron Geller/ Northwestern

evolution channels. These can help us understand properties about the stellar progenitors that formed the compact binaries. Since massive stars are the ones that collapse to form neutron stars and black holes, we can also learn about massive star evolution, and how it interplays with binary evolution to produce the gravitational waves that we detect at the end of their life-cycle. We can also use mergers involving a neutron star to infer its interior properties, such as its equation of state, which has yet remained elusive [14]. Population analysis of GW events have also been used to measure the Hubble constant (H_0) [13], and study the implications of the stochastic gravitational-wave background (SGWB) [15].

2.4 What have we learned?

It has been less than 10 years since we have opened the window to the gravitational wave universe. Nevertheless, with the current catalog of fewer than hundred events, we have learned a lot about the astrophysical properties of black holes and neutron stars. Population studies usually focus on the mass, spin and redshift distribution of compact objects, as different formation channels of how these compact binaries form and evolve often leave unique fingerprints on the mass and spin distributions of the merging binaries. Some of the

interesting questions pertaining to the mass distribution are –

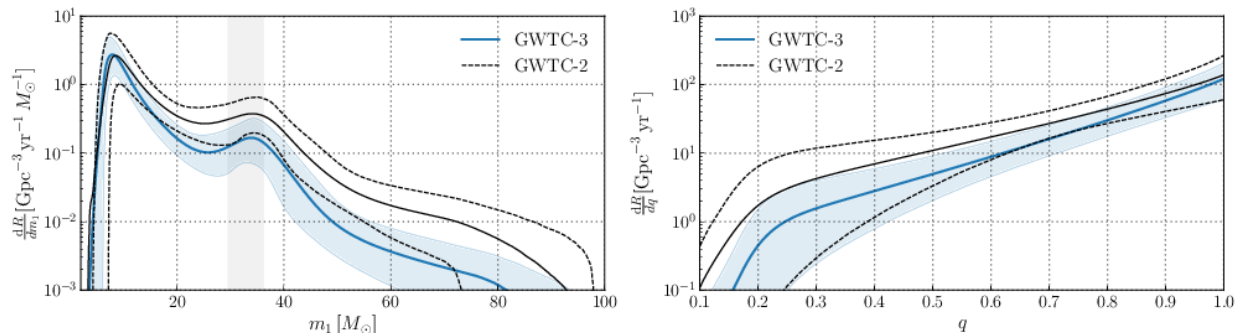


Figure 2.4: The primary mass distribution (*left*) and the mass ratio distribution (*right*) of black holes as inferred using GWTC-3

1. **Features in the mass distribution?** This is a particularly interesting question as observations of black holes through both X-Ray binaries and gravitational waves have hinted at a possible dearth of black holes between $3 - 5M_{\odot}$, i.e there seems to be a *lower mass gap* [12, 16] between the most massive neutron stars and the lightest black holes. This has important implications for understanding how massive stars die, and gravitational waves can help identify this feature in the mass distribution if it exists. An upper mass gap in the black hole mass function is also postulated to exist due to the (Pulsational) Pair-instability supernovae. This roughly predicts an absence of black holes formed due to stellar collapse in the range of $\sim 50 - 120M_{\odot}$ [17]. The exact location of the edges of the mass gap are dependent on several unknown factors, such as uncertain physics of massive star evolution, nuclear reaction rates etc. Using GWTC-3 [11], the LVK has inferred the mass distribution of the heavier (primary) black hole to be well fit by a POWERLAW + PEAK model [12]. There is no definite evidence of an upper mass gap in the data yet, but there is a possible pileup of black holes around $\sim 35M_{\odot}$, for which the source of origin is unknown. More detections will provide crucial information about the presence and exact location of possible mass gaps, bumps and valleys in the mass distribution, which will help better understand the formation of black holes. If we find black holes in the upper mass gap that we do not expect from isolated binary evolution, they may hint towards the presence of additional formation channels, such as dynamical formation in dense star clusters [18], where hierarchical mergers can populate the upper mass gap.

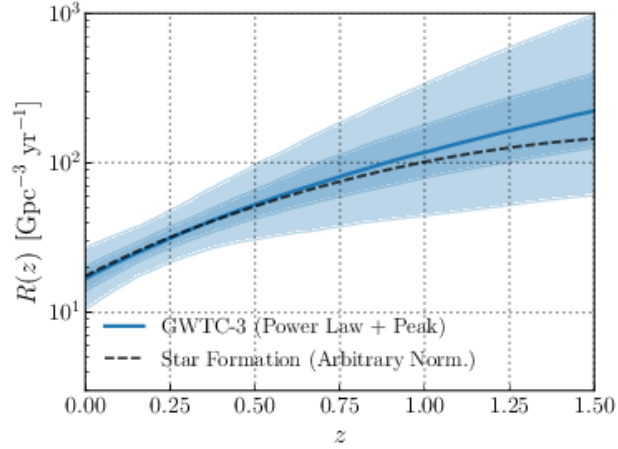


Figure 2.5: The evolution of the merger rate density across redshift of merging black holes as inferred using GWTC-3

2. **Redshift Evolution of Merger Rate:** Typically, black holes and neutron stars form when massive stars collapse at the end of their lives. The expected merger rate of compact objects would then certainly depend on the the number of such massive stars present at any given time, which is traced by the star formation history, probability of occurring in a binary, and also how much time does it take for a pair of compact binaries to merge after formation, also called the "delay-time distribution" [19–21]. The delay-time distribution is unknown, and there are several other uncertainties in the formation and evolution of compact binaries, which results in their expected merger rate to have large uncertainties with many competing models. This is why it is essential to empirically measure the merger rate of different types of compact binaries (BBH, BNS, NSBH) directly with observations [12], as this can help better understand the number of events that we expect to see with our detectors. Measuring the evolution of the merger rate across redshift [12] would also help in understanding how compact binaries have formed and evolved across cosmic time, and its relation with the star formation history of the universe. This can also have important implications for the strength of the SGWB, which strongly depends on the number of merging binaries present at high redshifts [22].

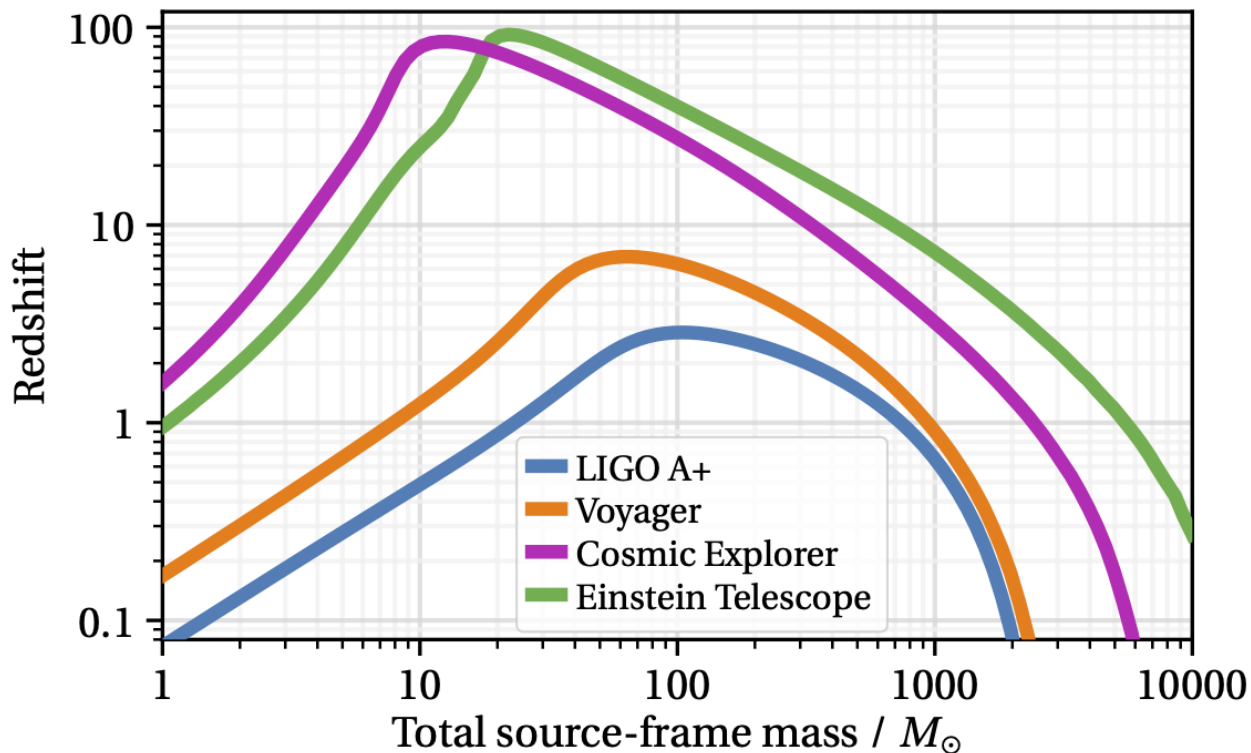


Figure 2.6: The horizon distances for the current and various planned future detectors. Fig 3.3 in [7]

2.5 The Future

In the coming observing runs, and with the addition of a new LIGO detector in Aundha [23], thousands of detections are expected in this decade, and with possible future upgrades to existing detectors, this number could rise significantly. The 2030s and 40s will likely see the start of the next generation of ground-based gravitational wave observatories, such as US-based Cosmic Explorer [7] and the Einstein Telescope [24] in Europe. These will see *millions* of mergers *every year*, and will be sensitive to mergers till very high redshifts. There will be an enormous explosion in the rich data from these detections, which will help probe several questions in astrophysics and cosmology [25], some of which we have mentioned above.

In fact, that is not all that we can see in gravitational waves. We know that we can observe electromagnetic radiation at different frequencies with different instruments, and they often probe different astrophysical sources and emission mechanisms. Similarly, the ground-based detectors are sensitive to a very narrow part of the GW frequency spectrum, i.e the high

frequency range. However, there are a lot of interesting astrophysical and cosmic sources at lower frequencies as well. Recently, Pulsar Timing Arrays across the globe announced evidence for the detection of the nanohertz gravitational wave background [26–29]. These extremely low frequency GWs could have been produced by an ensemble of merging supermassive black holes that reside in the centers of galaxies, or from other exotic sources. The 2030s will also see the launch of LISA, the Laser-Interferometer Space Based Antenna [30], which will be sensitive to sources in the milli-Hz regime. This window includes a wide variety of sources such as double white dwarf binaries in the Milky Way, extreme mass ratio inspirals, and massive black hole mergers at high redshifts.

Chapter 3

Gravitational Lensing

When Einstein formulated GR, one of its first successes was in the correct prediction of the anomaly in the precession of Mercury's perihelion. Another major prediction was the bending of distant starlight as it passes near the limb of the Sun, due to the Sun's gravity¹. This was experimentally verified during the 1919 Total Solar Eclipse by Arthur Eddington and others. In GR, light follows null geodesics, which are straight lines in flat spacetime. However, if there is a massive body present in a photon's path, it bends the spacetime around it. Now, due to a change in the metric due to the gravitational potential of a massive object, null geodesics are no longer straight lines between the source and the observer. Light rays now follow a curved path near the massive object. This phenomenon is known as gravitational lensing, and it has led to enormous discoveries in astronomy in the past few decades, and has also become the bedrock for performing several studies in cosmology.

3.1 The Lens Equation

Before we describe some of the the interesting phenomena caused due to lensing, in this section, we will briefly describe the mathematical framework to calculate the change in source positions and magnifications due to a lensing potential, which could be any massive source such as stars, galaxies or clusters of galaxies etc. We will roughly follow the references [31, 32].

¹Famously, the GR deflection is twice the deflection predicted in Newtonian gravity.

Firstly, we will assume that we are working in the geometrical optics limit. This is when the wavelength λ of the propagating radiation is much smaller than the length scale associated with the lens mass, i.e $\lambda \ll \frac{GM_{\text{lens}}}{c^2}$. This is practically always true for lensing of electromagnetic radiation, but as we will see later, it need not always hold true for lensing of GWs due to their large wavelength. However we will work in this limit, which is valid for lensing due to large lens masses, such as galaxies or clusters, which is the focus of our study.

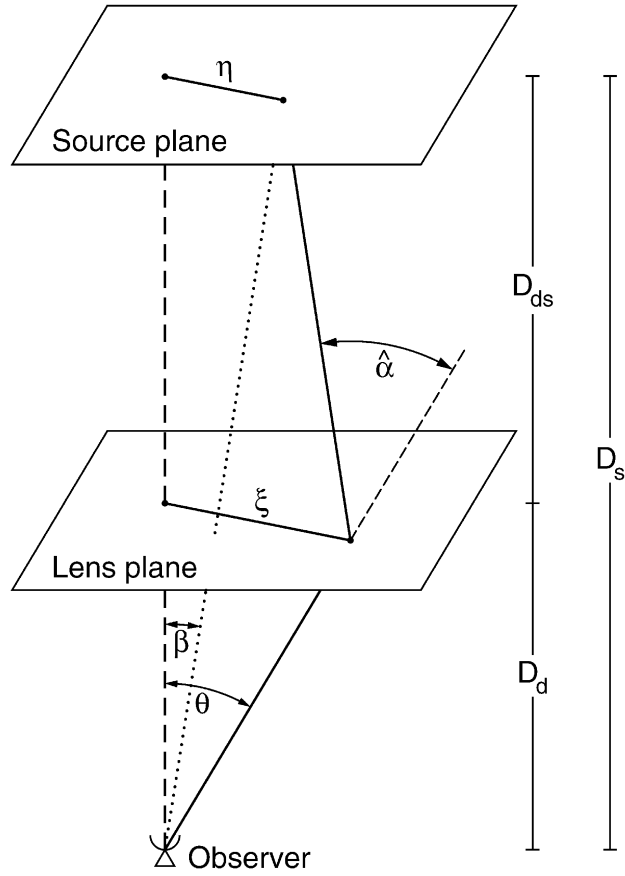


Figure 3.1: Lensing sketch depicting the geometry and various angular vectors involved for the source and image positions [32].

To proceed further, we will make another assumption (valid for the usual astronomical lens), which is called the *thin-lens approximation*. We assume that the size of the lens is much smaller than the distances involved between the lens, source or observer.

In this framework, we can now calculate the deflection angle –

$$\boxed{\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})} \quad (3.1)$$

This equation is widely known as the *lens equation*. In this context, $\vec{\beta}$ represents the actual angular position of the source, $\vec{\theta}$ denotes its apparent angular position, and $\vec{\alpha}$ signifies the deflection angle of a light ray originating from the source, which bends due to the presence of the intervening lens. This equation looks relatively simple, but all the rich physics about the natures of the lens and the distances involved are included in the term $\vec{\alpha}(\vec{\theta})$. To make things clearer, we introduce the effective lensing potential $\hat{\Psi}$ below, which is obtained by projecting the 3D lens potential Φ with some appropriate scaling factors. It is given by –

$$\Psi(\vec{\theta}) = \frac{D_{\text{ds}}}{D_{\text{d}}D_{\text{s}}} \frac{2}{c^2} \int \Phi(D_{\text{d}}\vec{\theta}, z) dz = \frac{1}{\pi\Sigma_{\text{cr}}} \int d\theta' \Sigma(\theta') \ln|\theta - \theta'| \quad (3.2)$$

where $\Sigma_{\text{cr}} = \frac{c^2}{4\pi G} \frac{D_{\text{s}}}{D_{\text{d}}D_{\text{ds}}}$ is called the critical surface density of the lens which depends on the angular diameter distances involved, and $\Sigma(\theta')$ is the surface mass density of the lens.

As it turns out (we will skip the derivation), the deflection angle and the effective potential are related by –

$$\nabla\vec{\Psi}(\vec{\theta}) = \vec{\alpha}(\vec{\theta}) \quad (3.3)$$

We can now write our lens equation as –

$$\vec{\beta} = \vec{\theta} - \nabla\vec{\Psi}(\vec{\theta}) \quad (3.4)$$

This equation explicitly shows us how to calculate the true source position given the apparent source positions and the geometry of the lens, which includes the distances between

the lens, source and observer, and the mass profile of the lens. Common mass profiles used to treat lens systems are the SIS model which stands for Singular Isothermal Sphere, and on relaxing the spherical symmetry assumption, we get the Singular Isothermal Ellipsoid (SIE) model.

Note that, if we are given the true source position and the lens geometry, the lens equation can have *multiple solutions* for $\vec{\theta}$, i.e the apparent source positions in the sky. This corresponds to the phenomena of multiple images that may be seen during strong lensing of a source by a lens.

Interestingly, lensing can also cause magnification and distortion of images, because of differential deflection of different regions of an extended source. We can study these effects by assuming the source is much smaller than the angular size at which the lens' properties change. This allows us to linearize the lens equations, and we can write down the jacobian matrix A between the lensed and unlensed coordinates as –

$$A_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = \delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j} \quad (3.5)$$

We can use the relation between the deflection angle and the effective potential to rewrite the above equation as –

$$A_{ij} = \delta_{ij} - \frac{\partial^2 \Psi}{\partial \theta_i \partial \theta_j} \quad (3.6)$$

We can now decompose the jacobian into 2 parts, one that is trace-free and anti-symmetric, quantifying "shear", which describes the source getting distorted tangentially around the lens. The remaining part of the matrix is symmetric, and quantifies "convergence", which simply describes the (de)-magnification of the source. We can exactly evaluate the magnification factor μ of the source by calculating the determinant of the inverse of the jacobian–

$$\mu = \det A^{-1} = \frac{1}{\det A} \quad (3.7)$$

Having briefly described a simplified mathematical framework to understand gravitational lensing, we will now show a few examples about how all these interesting phenomena show up in the universe. Gravitational lensing can affect background sources of light in multiple ways such as –



Figure 3.2: A Quadruply lensed quasar HE 0435 1223. The foreground galaxy acts as a lens and splits the background quasar into 4 almost evenly spaced images. Image Credit: NASA/ESA/Hubble

1. Multiple Images: As we have studied, in certain situations, there can be multiple solutions for the projected images of a background source of light (such as a galaxy) due to the presence of a massive lens in between (could be another galaxy or a cluster of galaxies). Fig. 3.2 depicts one such example of a quadruply lensed quasar.
2. Time Delay: If there are multiple images, the light rays travel around different paths around the lens exploring different areas of the lens potential, and this could lead to a gap between the time of arrival of photons that were emitted by the source at the same time. Fig. 3.3 depicts the time-shifted light curve of the same quadruply lensed quasar, where the original quasar undergoes intrinsic variability which can be well measured.

In fact, The time delay distribution is particularly sensitive to the value of the Hubble constant, H_0 [33].

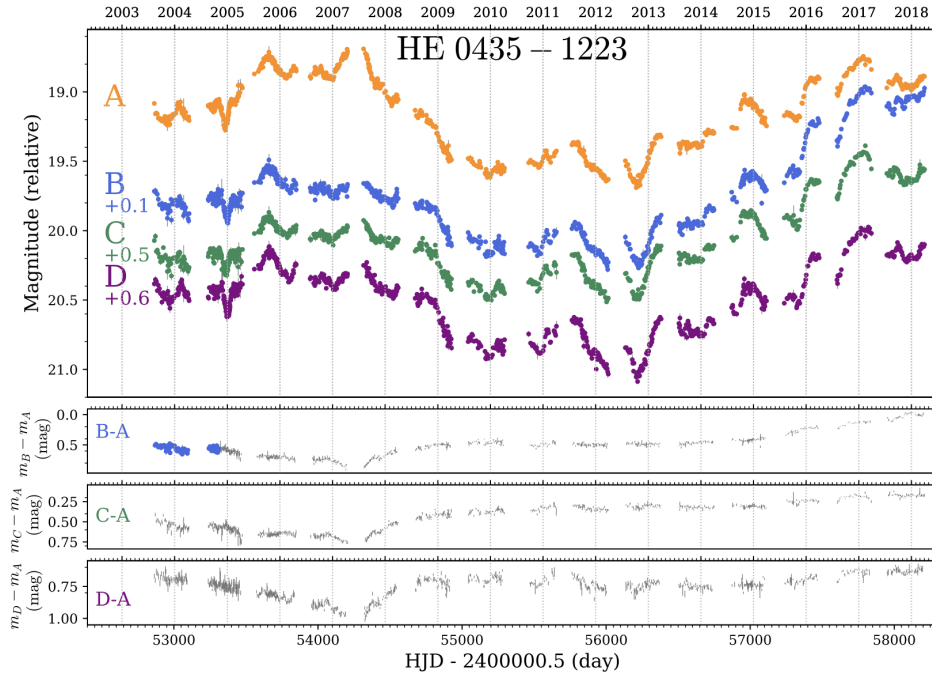


Figure 3.3: Time shifted light curves of the 4 images of a quad quasar, HE 0435 1223 [34]. The background quasar is intrinsically variable, which is seen in the 4 images as well. Importantly, the 4 light curves have a time delay between them due to lensing.

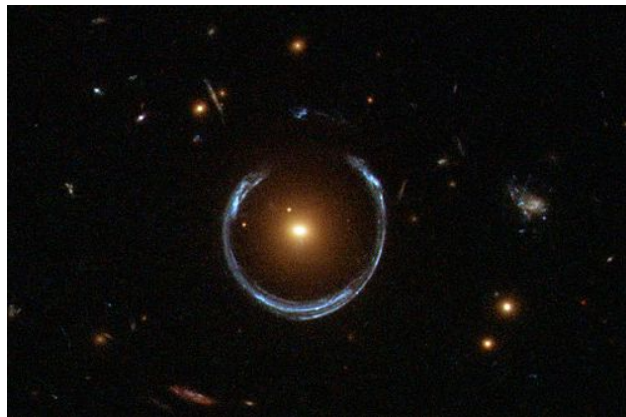


Figure 3.4: An Einstein Ring, a highly distorted and magnified background galaxy due to very good alignment between us, the lens and the background source. Image Credit: NASA/ESA/Hubble

3. Magnification: Lensing does not create or destroy photons, it only changes their path

as we have studied, and can concentrate them towards our line of sight. This can lead to (de)–magnification of the sources, making them appear (dimmer)brighter than they actually are.

4. Distortion: If the background source is an extended object, different regions within the source would get bent along different paths due to the massive lens, This could lead to an apparent distortion of its original shape. If the distortion increases its apparent size, it would appear brighter due to the conservation of surface brightness in lensing as explained above. Fig. 3.4 depicts one such extreme example, where the background galaxy has been distorted and appears as almost a ring like image due to lensing.

3.2 Lensing done in 2 ways

Depending on how the source, lens and observer are arranged geometrically, lensing broadly occurs in 2 different regimes – strong lensing [35, 36], and weak lensing [37]. To first order, these depend on the angular separation between the lens and the source, and also on the critical surface density.

1. If the angular separation between them is sufficiently small, i.e they are aligned close to our line of sight, we observe strong lensing which may lead to a variety of rich phenomena that we described above, such as multiple images, and these may be highly distorted as well if the background source is an extended object.
2. Conversely, if the angular separation between them is large, the effects of lensing are small and subtle, i.e we observe weak lensing. This only leads to minute changes in the shape of the background galaxies. These changes are invisible for a single background source, since we do not always know their intrinsic shape, but the effects of weak lensing can be extracted by averaging over an ensemble of galaxies in a region, since they encounter a similar gravitational potential due to the lens, and there is a coherent distortion in their shapes. Weak lensing has become a particularly powerful tool for precision cosmology in recent years in the era of big–data, especially with the start of large–sky–surveys such as DES, and soon the LSST etc.

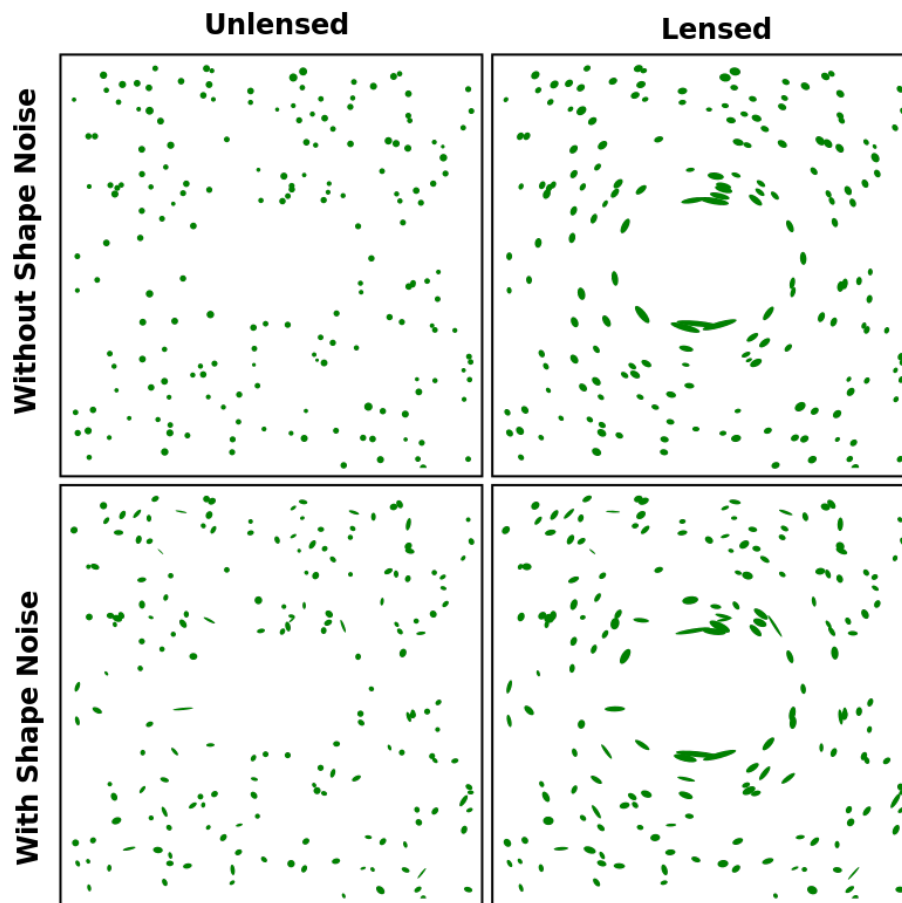


Figure 3.5: An explanatory cartoon of the weak lensing phenomenon. We can see that background galaxies undergo small distortions due to an intervening lens. The distortion field is only noticeable for an ensemble of galaxies. Image Credit: Wikipedia

Chapter 4

Lensing of Gravitational Waves

GWs, like electromagnetic radiation can also be lensed by intervening matter between the source and the observer, such as galaxies or clusters [38]. A confident detection of lensing of gravitational waves has remained elusive, however with more observing runs and the third generation of detectors detecting millions of events every year up to very high redshifts, the possibilities of finding lensed GWs is not a question of if, but when?

Lensing of gravitational wave signals may provide important information about fundamental physics through tests of General Relativity, and also in probing cosmology as it may enable the measurement of the Hubble constant as well as constrain properties of dark matter.

But before we proceed further, it is worthwhile to describe what exactly does lensing do to a gravitational wave?

4.1 Gravitational Waves: Meet Gravitational Lensing

In general, lensing distorts the gravitational wave signal, but the exact nature of the distortion depends on the size of the lens. In the geometrical optics regime, which is relevant for lensing due to massive lenses such as galaxies or clusters of galaxies, lensing *does not affect the frequency evolution* of the signal, but it can *(de)magnify its amplitude*¹ [38]. It can also lead to other interesting effects, such as multiple signals that may be separated by a time

¹There can be additional phase effects in some situations which we will describe later.

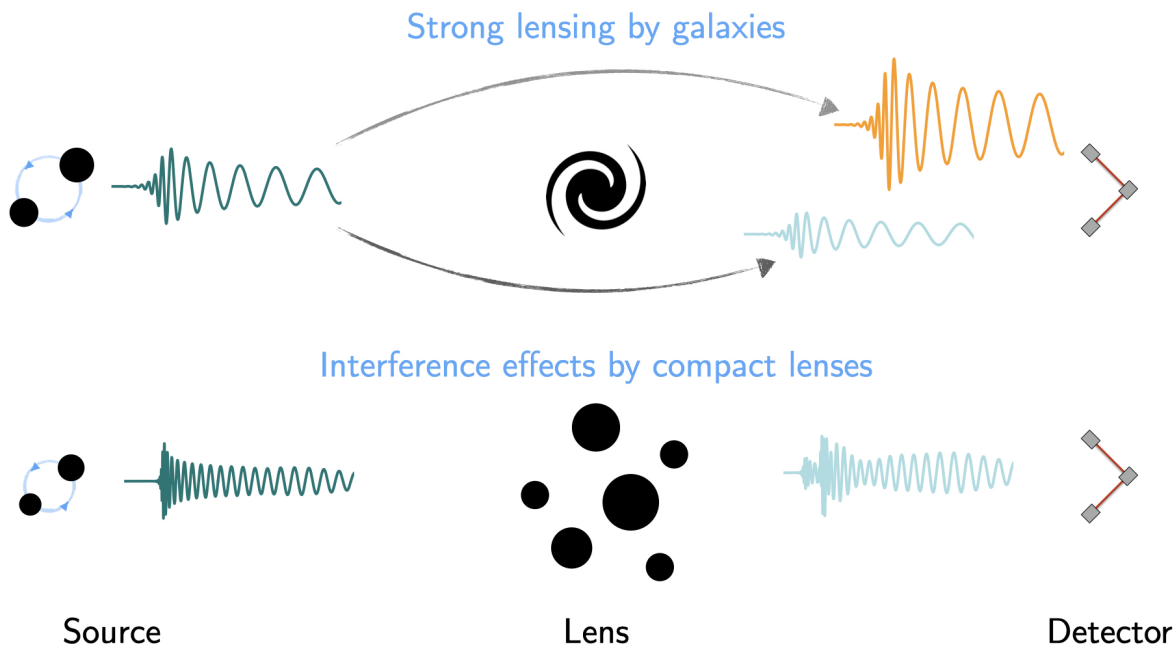


Figure 4.1: *Top:* Strong lensing of GWs in the geometrical–optics regime, which leads to multiple images seen at the detector with a time delay, and having identical frequency evolution but with differing amplitudes due to lensing (de)–magnification. *Bottom:* Lensing of GWs in the wave optics–regime due to a smaller lens. This leads to frequency dependent modulations of the gravitational wave signal, beating patterns etc. Figure Credit: LIGO Scientific Collaboration

delay of minutes to months for galaxies, and up to years for clusters. For low mass lenses, such as stars and compact objects like black holes or small dark matter halos, lensing occurs in the wave–optics regime. This leads to *frequency–dependent modulations to the signal* [39], such as beating patterns in the signal due to the superposition of multiple images. These are extremely fascinating sources and one can learn a lot from the (non–)detection of such signals. Fig. 4.1 depicts the regimes of GW lensing. However, in this work we will only focus on lensing in the geometric optics regime, where only the amplitude is modified. We will also ignore the possibilities of multiple images, which is relevant in the regime of strong–lensing, as a majority of the sources may only be weakly lensed.

4.2 An Aside on what does a GW signal really tell us?

In astronomy, Cepheid variables were one of the first sources that could be used to measure how far away they are from us. They were known to have their observed brightness change periodically over time, and Henrietta Leavitt noticed that this oscillation cycle was related to the luminosity of the star, i.e its intrinsic brightness. Then by measuring the observed brightness, we could measure how far away the star is. Such objects are called "standard candles", where we can use some quantity that can we directly measure to infer an intrinsic property about the source. GWs from CBC's are not too different², and are called "standard sirens" (since we are *listening not seeing*). Essentially, a GW signal is a waveform whose frequency and amplitude changes with time. The frequency evolution is set by General Relativity, and tells us about the masses of the merging objects. To first order, the frequency evolution is given by [40] –

$$\frac{df}{dt} = \frac{96}{5}\pi^{8/3} \left(\frac{G\mathcal{M}}{c^3}\right)^{5/3} f^{11/3} \quad (4.1)$$

where \mathcal{M} is the "chirp mass", which is a combination of the component masses is defined as –

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \quad (4.2)$$

The chirp mass is the best measured parameter of a GW signal, since it directly enters the first order evolution of the signal, which is often well measured. We can obtain the component masses from the chirp mass and another parameter, called the symmetric mass ratio η , which is given by –

$$\eta = \frac{m_1 m_2}{(m_1 + m_2)^2} \quad (4.3)$$

²I thank Prof. BS Sathyaprakash for the analogy between Cepheids and GWs from CBC's.

Meanwhile, from the amplitude of the signal, we can measure how far away the source is, i.e its luminosity distance d_L , as the amplitude falls over as the inverse of the luminosity distance. However, as the universe is expanding, the GWs also suffer from cosmological redshift and are stretched. What this means with respect to parameter estimation is that we are sensitive to not the source chirp mass \mathcal{M}_s , but the redshifted chirp mass \mathcal{M}_d , which is also called the detector-frame chirp mass. They are related by –

$$\mathcal{M}_d = \mathcal{M}_s(1 + z) \tag{4.4}$$

This is known as the "mass-redshift degeneracy" which basically says that the frequency evolution of a nearby, massive source is identical to that of a distant, lighter source. In normal scenarios, this degeneracy can be broken as we have an additional observable, the amplitude, which directly gives us the luminosity distance d_L to the source. By assuming a cosmology, we have a $z(d_L)$ relation, and we can infer the source chirp mass using Eqn 4.4 –

$$\mathcal{M}_s = \frac{\mathcal{M}_d}{1 + z(d_L)} \tag{4.5}$$

4.3 Oh, but lensing?

The breaking of the mass-redshift degeneracy in Eqn 4.5 relied on the correct measurement of the luminosity distance to the source. However, as we learned in Section 4.1, lensing modifies the amplitude of the GW, which restores the degeneracy. To be more precise, lensing of a GW (in the geometric optics regime) can be described by a magnification factor μ . The new luminosity distance due to this magnification factor is given by [38] –

$$\tilde{d}_L = \frac{d_L}{\sqrt{\mu}} \tag{4.6}$$

If one directly uses this new luminosity distance to infer the source mass, we get –

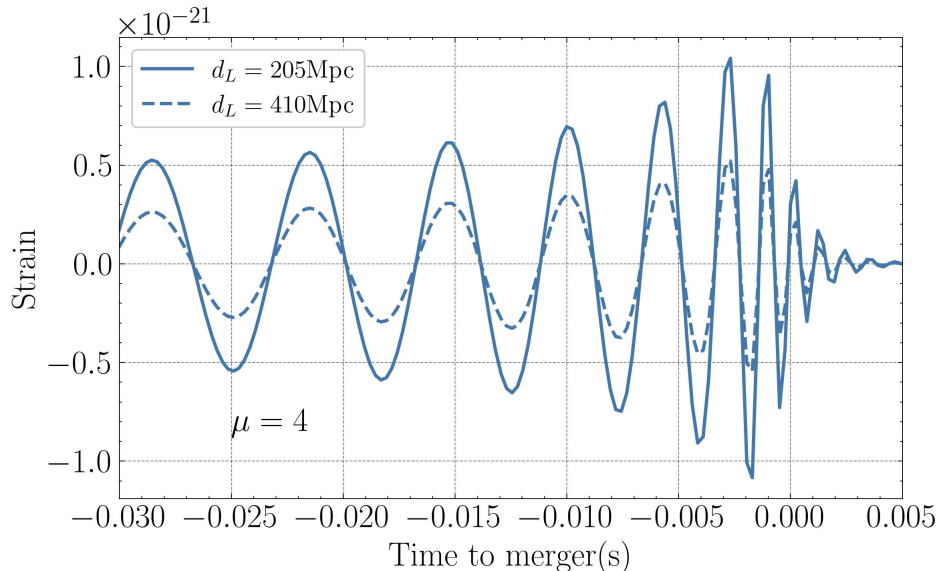


Figure 4.2: An example of an unlensed waveform, and the corresponding lensed waveform with a $\mu = 4$

$$\tilde{z} = z(\tilde{d}_L) = z\left(\frac{d_L}{\sqrt{\mu}}\right) \quad (4.7)$$

$$\tilde{m}_s(1 + \tilde{z}) = m_s(1 + z) \quad (4.8)$$

$$\boxed{\tilde{m}_s = m_s \left(\frac{1 + z}{1 + \tilde{z}} \right)} \quad (4.9)$$

Eqn 4.9 shows us that the source mass that we infer of a lensed signal is biased and not equal to the true source mass. Note that to derive this, we used the fact that the detector frame masses are unchanged due to lensing, as they solely depend on the frequency evolution of the signal.

In general, we would never be able to identify the individual magnification factors of weakly lensed events, as they would just appear as a perfectly normal GW signal at our detector. If we measure the source parameters using the standard procedure, we would make

biased estimates of its source masses and redshift as demonstrated above.

4.4 Current Searches for Gravitational Wave Lensing

GW150914, the first direct detection of GWs came from a binary black hole merger, with each black hole weighing an enormous $30M_{\odot}$. This was sufficiently unexpected, as the stellar mass black holes observed through electromagnetic observations (mainly through X-Ray Binaries) are much lower in mass. With further detections, it was soon realised that there is indeed a population of stellar mass black holes of such heavy masses that can be formed through stellar collapse etc, and aided by LIGO/Virgo’s better sensitivity for heavier mergers, it is not surprising that these are the sources that are often detected. However, some groups still claim that the heavy black holes that the LVK has detected do not exist, and that almost *all* the mergers are strongly lensed events [41]. This requires making a different set of assumptions, such as a merger rate that is high at high redshift, and yet still explain the non-observation of the SGWB. Such models may get quite convoluted and require making untested assumptions, however more events detected in future observing runs will be significantly helpful in better understanding the lensing of GW events.

But this raised an interesting question about the possibility of observing gravitationally lensed GWs, and what could be their implications. But the first question that arises is how would one confidently claim a detected signal to be lensed? and how many such lensed signals should one expect? This has led to an entire effort both within and outside the LVK collaboration to develop methods to search for GW lensing, and the intriguing astrophysical and cosmological questions one could probe with them. In this section, we give a brief overview about such searches.

4.4.1 How does one know a GW signal is lensed?

In electromagnetic astronomy, lensing can be relatively easily identified by just directly observing multiple images, distorted shapes, Einstein rings in our telescope images. However, the complication for GWs is that they have really poor spatial resolution, i.e we cannot localize their sky location accurately. Conversely, they have excellent temporal resolution, i.e we can accurately detect when a GW signal arrives at the detector, and infer the source parameters. Let us assume for now that we are in the strong lensing and geometrical optics regime. Multiple images in this scenario would correspond to repeated GW events at the detector, coming from the same location in the sky, but with a time delay due to lensing. The

events would also undergo different magnifications, and therefore have different amplitudes, but the frequency evolution would remain unchanged. The time delays between the events can be of the order of minutes to days for lensing due to galaxies, and up to months or years for large galaxy clusters. How would one then identify a pair (or more) of events to be lensed among hundreds to thousands of unlensed events? To do this, we can take advantage of the fact that the different images will have common parameters, except for their luminosity distance, since it is degenerate with the lensing magnification of the source. We can define an a Posterior Overlap statistic called $\mathcal{B}_{\mathcal{U}}^{\mathcal{L}}$ that computes the ratio of the overlap of the posteriors for the common set of parameters (except luminosity distance) for the 2 images under 2 hypotheses, i) the 2 events are a lensed pair, ii) the 2 events are not lensed. It is given by [42] –

$$\mathcal{B}_{\mathcal{U}}^{\mathcal{L}} = \frac{p(\{d_1, d_2\}|\mathcal{H}_L)}{p(\{d_1, d_2\}|\mathcal{H}_U)} = \int \frac{p(\vec{\theta}|d_1)p(\vec{\theta}|d_2)}{p(\vec{\theta})} d\vec{\theta} \quad (4.10)$$

We can also account for the time–delay in the arrival of 2 events, as the time–delay distribution for lensed pairs is different from unlensed GW events. For unlensed GW events, the rate of observing GW events follows a Poisson process, while for lensed pairs, the time delay distribution is qualitatively different, depending on the lens model and geometry. This is called the \mathcal{R}_{gal} statistic [42] and is given by –

$$\mathcal{R}_{\text{gal}} = \frac{p(t_0|\mathcal{H}_L)}{p(t_0|\mathcal{H}_U)} \quad (4.11)$$

Earlier, we mentioned that strong lensing in the geometric optics limit only affects the amplitudes. But it turns out that it can also introduce a phase–shift (also called the *Morse phase*) in the waveform which can be identified for a particular class of images called type II images [43, 44]. In the geometrical optics regime, we can write down a lensed waveform in terms of the unlensed waveform in full generality as –

$$h_L(f; \Lambda, \mu, \Delta t, \Delta \phi) = \sqrt{|\mu|} h(f; \Lambda, \Delta t) e^{i\Delta \phi \text{sign}(f)} \quad (4.12)$$

where $\Delta \phi$ is a constant difference that takes the values $0, \pi/2, \pi$ depending on whether it is a type I, II, III image.

For low mass lenses, such as stars or compact objects, GW lensing occurs in the wave-optics regime, and this leads to frequency-dependent modulations in the waveform, which can lead to beating patterns. This is called microlensing, and with some lens modelling assumptions, methods exist to find such microlensing signatures in a signal [39, 45].

Current searches have looked for signatures of lensing in the data up to the third observing run, and there has not been any confident detection of gravitational wave lensing yet [45, 46]. However, with thousands of events expected to be detected in the coming observing runs, it is not a matter a question of if, but when? Current lens models predict 0.1 – 1% of events to be strongly lensed [47, 48].

Nevertheless, the absence of a detection also tells us something! Current detections help us constrain the merger rate of BBH at low redshifts. But due to the limited detector sensitivity, the merger rate is poorly understood at high redshifts. The only information we have is an upper limit from the non-observation of the stochastic GW background. Similarly, the non-observation of lensing signatures in the detected events can help in constraining the merger rate at high redshifts [45]. In fact, the non-observation of microlensing signatures till now can also help constrain the number of lenses that can cause microlensing. In particular, objects of such masses could be dark matter candidates, for which there are several mass windows where their abundance is unconstrained. Non-observation of microlensing limits the number of such objects, and this can put a limit on their contribution to being a constituent of dark matter [45].

4.5 Gravitational Wave lensing as a tool

In the previous sections, we have described how lensing could affect the original gravitational wave signal from a source, how it modifies the waveform and can lead to biases in parameter estimation. But, similar to lensing of electromagnetic radiation, lensing of

gravitational waves can probe numerous questions in astrophysics, cosmology and the nature of gravity. This active area of current research has been explored in several studies, some of which we highlight below.

that time delays between arrival of different images in strong lensing is particularly sensitive to the value of the H_0 , and the Hubble tension is a particularly challenging problem that remains unsolved. This method of using time delays to measure H_0 has been implemented by observing lensed supernovae where the time differences of the different lensed copies of the same supernova can be measured at the order of a day. Observing multiple copies of the same GW signal will be extremely useful in this regard, as their time of arrival can be measured at the precision of a *few milliseconds*. With the third generation of detectors (3G), [49] showed that the number and delay–time distribution of lensed events could be used for estimating cosmological parameters in a redshift range not explored by other cosmological probes, thereby aiding in resolving the Hubble tension and also measuring the evolution of the Hubble parameter across redshift.

As has been mentioned earlier, the GW signal is sensitive to the detector frame mass, and it would be extremely useful to obtain a signal’s redshift estimate, which would enable extracting the source–frame masses and also for measuring the Hubble constant. However, GW signals typically have large sky localizations, with many possible host galaxies. In some cases, such as BNS or NSBH mergers, it is possible to directly obtain a redshift estimate if an electromagnetic (EM) counterpart is present and detected. But for BBH mergers, which typically do not have EM counterparts, [50] showed that one can use multiple copies of a lensed GW signals to improve sky-localization, and in some cases achieve sub-arcsecond regions which would aid in reducing the number of possible host–galaxies to a few or even just one.

Additionally, in theories with modifications to GR, different polarisations of the GW (such as $+$ or \times) can propagate at different speeds in the vicinity of a lens. This is called lens-induced birefringence (LIB) which essentially introduces a time delay between the arrival of different polarisations of the signal. [51] searched for LIB signatures using GWTC-3. They did not find evidence for LIB, which helped in constraining the modified gravity theories. Observing multiple copies of GW signals can also help in better extracting the different polarisations in the signal, as each copy measures a different linear combination of the polarisation states. This can be used to test GR, which predicts only 2 polarisation modes,

but modified theories can have up to 6 of them.

Gravitational lensing has also been used to predict the approximate time of arrival of a lensed SNe signal, using previously identified images and lens modelling. [52] showed that for lensed BNS or NSBH with an identified electromagnetic counterpart, one can predict the time of arrival of the next image hours to days before the merger. This can help in pointing telescopes at the source to study any EM emission prior to or just after the merger, which has never been observed before.

In recent years, we have detected numerous Fast radio Bursts (FRB's), however their source of origin and emission mechanism is still unknown, with several different proposed models, including compact binary coalescences. [53] showed that just from the associated time delays between different FRB images and CBC images (which can be measured up to a precision of nanoseconds and milliseconds correspondingly), we can make a unambiguous $> 5\sigma$ association between them, which would significantly help in understanding where and how FRB's are generated.

4.6 Magnification Probability Distribution

Previously, we showed that the GW amplitude is affected by an unknown magnification factor, μ . In this section, we describe the probability distribution function of lensing magnifications, as expected in our universe.

Let us say we have a pair of black holes merging at a source redshift z . What is the probability that such a merger is lensed by a magnification factor μ . If we assume that mergers are uniformly distributed on the source plane at redshift z , a random merger has a probability

$$dP = \frac{dP(\mu; z)}{d\ln\mu} d\ln\mu \quad (4.13)$$

of being magnified by μ , where $\frac{dP(\mu; z)}{d\ln\mu}$ is the lensing magnification probability density function (PDF), which would obviously depend on redshift, as one would expect. Since the

total solid angle is approximately conserved, the mean magnification is ≈ 1 [54, 55], i.e –

$$\langle \mu \rangle = \int_0^\infty \frac{dP(\mu; z)}{d \ln \mu} d \ln \mu \approx 1 \quad (4.14)$$

In general, this PDF has a contribution from all intervening matter that is present between the source and the observer, and therefore includes both the weak lensing regime and the strong lensing regime at high magnifications. These functions have been well studied with the help of ray–tracing through large–scale–structure in simulations, and have a characteristic shape. At high magnifications, they have a universal power-law of $dP/d\mu \sim \mu^{-3}$, and its amplitude increases rapidly with redshift, because of the larger lensing optical depth. Such a power law behaviour is also theoretically predicted, and is a generic feature that arises when a point background source is located near the caustics in the lens plane [56]. Fig. 4.3 shows an analytical fit for the lensing magnification PDF as given in [38], and it is important to note that lensing can cause de–magnifications as well, as can be seen by the non zero probabilities of $\mu < 1$, particularly at high redshift.

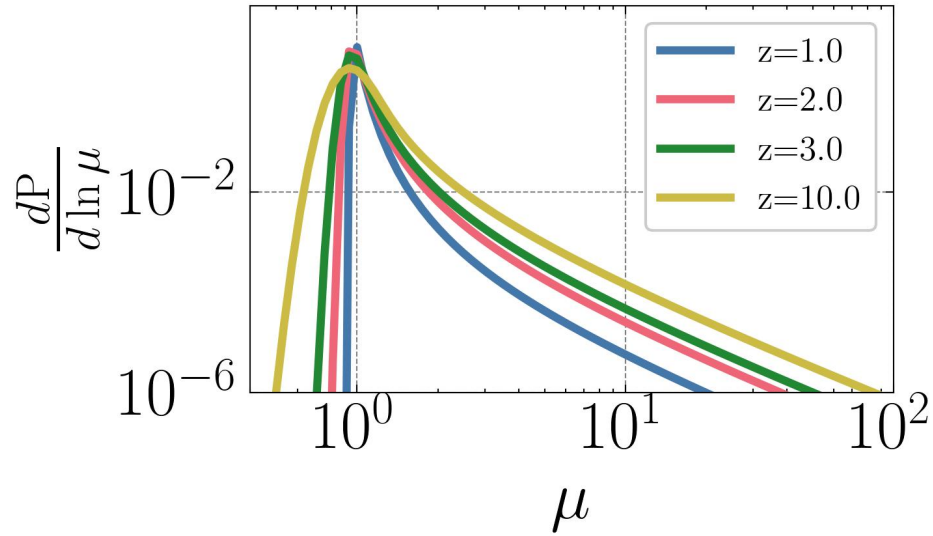


Figure 4.3: The weak lensing + strong lensing magnification distribution as it varies with source redshift.

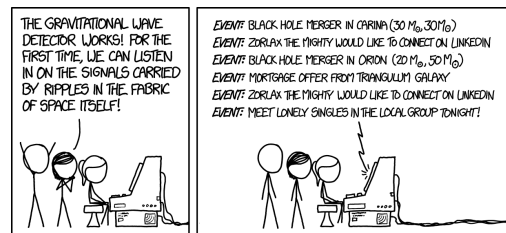
Caveat: There are some uncertainties in the lensing magnification distribution as well, as different groups use different simulations to derive it, and some may or may not include the contributions of baryons in the central galaxies of the lens, which can affect the behaviour of the tail at high magnifications.

Part II

Developing the tools

Chapter 5

Methods



Credit: xkcd comic for the announcement of the first detection

5.1 Parameter Estimation in a Nutshell

Having found a GW signal in the noisy data, what can one do with it? One might want to know what kind of merger produced such a signal? Is it a BBH, a BNS, a NSBH merger or some other exotic source? This often requires one to know the masses of the merging compact object. One might also want to know where in the sky did the merger occur, and how far is it from us. This is important for following up on the detection with electromagnetic telescopes in the space and on ground, to look for a electromagnetic counterpart. To answer all these questions, and many more, we need to measure the parameters of the source that created the GW signal. Since the detector data is noisy, we can never measure the parameters accurately with perfect precision, but with some uncertainty in them. Typically, parameter estimation

methods estimate the source properties using the framework of bayesian inference. In this section, we will give a very short overview about how this is done. For more details, we refer the reader to [57] for a brief introduction to parameter estimation of GWs, and [58, 59] for a more exhaustive review.

5.1.1 Bayesian Inference

In the past few years, bayesian inference has become the bedrock of modern statistical analysis, especially in astronomy, where we are using observations, that are often noisy, to learn about some hidden underlying physics. These methods often require us to assume a model that can directly allow us to compare between theory and data. In gravitational wave astronomy, this theory is General Relativity, which gives us an excellent prediction of the gravitational waves that are generated from the inspiral, merger and ringdown of compact objects, such as black holes or neutron stars. We detect these signals at our interferometer, but along with the GW signal, there is also noise, often from a range of terrestrial disturbances. Bayesian analysis are extremely useful in such situations, to measure the parameters of merging binaries using noisy data. It provides us the full probability distributions of these parameters, which allows us to statistically quantify the uncertainties in our measurements as well. Here, we briefly describe Bayesian inference, in a context specific to gravitational wave astronomy, as we will see similar techniques being used later as well. The main ingredient in Bayesian Inference is Bayes' theorem, which is essentially a statement about conditional probabilities. Specifically, suppose we have 2 events, even A, and event B. Using Bayes' theorem, we can write:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \quad (5.1)$$

But what do A and B correspond to in GW astronomy, particularly in the context of parameter estimation, where we are trying to infer the parameters of the CBC's that produced the GW signal that we have detected. Note that here, we will always assume that we know our underlying model, GR, i.e we know what the signal would look like given its parameters. Let θ (could be one or multiple parameters) refer to the parameters of the binary, and let d refer to the data segment that contains the signal which we have detected¹ (like the strain

¹We will ignore the definition of what is considered a detection. Different search methods have separate

timeseries measured at the detector). We can rewrite Bayes' theorem by using $A = \theta$ and $B = d$. We obtain –

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)} \quad (5.2)$$

where $p(\theta)$ denotes the prior distribution (henceforth called $\pi(\theta)$), which encapsulates our prior belief regarding the parameters before detecting the signal. Consequently, $p(\theta|d)$ is called the posterior probability distribution. The term $p(d|\theta)$ is usually called the likelihood function $\mathcal{L}(d|\theta)$. It encodes our knowledge about both the detector noise and the underlying data generation model. It tells us how *likely* are we to have observed the data d if the binary parameters that produced the signal are θ . The denominator, $p(d)$ is a normalizing factor, referred to as the marginalized likelihood or the evidence \mathcal{Z} .

$$\mathcal{Z} = \int \mathcal{L}(d|\theta)d\theta \quad (5.3)$$

The evidence is useful when we want to compare between different models, for example between GR and other alternative theories of gravity. We can ignore it for now, as our sole interest is in estimating the posterior probability distribution, having already fixed our underlying model.

Lets say we found a GW signal in a data-segment d . Assuming that the noise n is additive, we can write –

$$d = h(\theta) + n \quad (5.4)$$

Now, the main goal of parameter estimation algorithms based on bayesian inference is to evaluate the *posterior* probability distribution $p(\theta|d)$.

detection criteria for example, the matched filter SNR. If we somehow perform parameter estimation on a data segment that does not contain a signal, we would likely recover broad posteriors that resemble more to our initial prior.

$$p(\theta|d) \propto \mathcal{L}(d|\theta)\pi(\theta) \quad (5.5)$$

Since the detector noise is gaussian to an excellent approximation, the likelihood is called the *Whittle likelihood* and it looks like –

$$\mathcal{L}(d|\theta) \propto \exp\left(-\frac{1}{2}(d - h(\theta)|d - h(\theta))\right) \quad (5.6)$$

$(a|b)$ denotes the inner product, which is mathematically written as –

$$(a|b) = 2 \int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)}{S_n(f)} df \quad (5.7)$$

where $S_n(f)$ denotes the PSD of the noise in the detectors. \tilde{a}, \tilde{b} are the fourier transforms of the data or signal, and * represents their complex conjugate.

The prior function $\pi(\theta)$ is also chosen by us. When we do not know anything about the parameters before analyzing the signal, we can use uniform/isotropic prior distributions during parameter estimation. However, in some situations, other prior choices may be more beneficial. We refer the reader to [22] for a collection of the common priors used in gravitational wave astronomy.

5.1.2 Evaluating the posterior

Typically, GW signals are completely described by at least 15 parameters. These are the 8 intrinsic parameters, 2 for the component masses (m_1, m_2), and 6 for their corresponding spin vectors ($\vec{\chi}_1, \vec{\chi}_2$). The usual 7 extrinsic parameters are the luminosity distance to the source (d_L), its inclination/ orientation (θ_{JN}), location in the sky (α, δ), the polarization angle of the detector ψ , and the coalescence time and phase (t_c, ϕ_c). For mergers involving neutron stars, there can be up to 2 additional parameters (Λ_1, Λ_2) for the tidal deformation of the neutron star which is an important probe of the neutron star interior.

The posterior distributions in most cases are not simple distributions that we can directly sample from. One can also not calculate the posterior by simply evaluating it on a grid, as

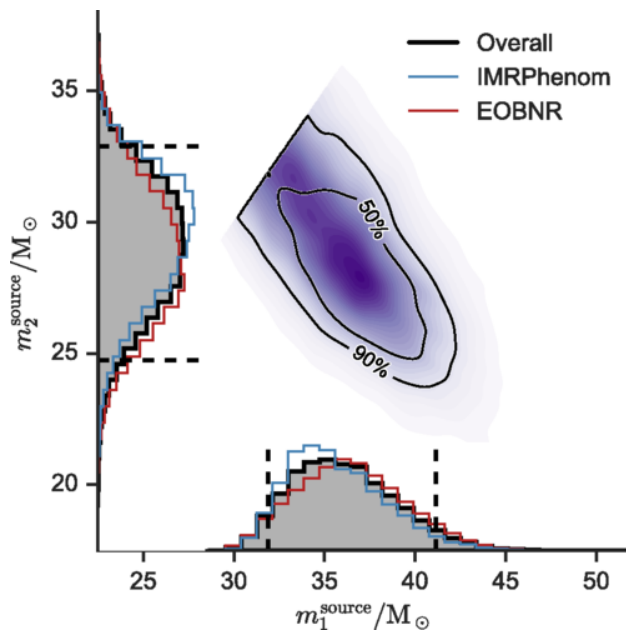


Figure 5.1: GW150914 corner plot for the 2 component source frame masses

the grid is at least 15 dimensional, and we suffer from the curse of dimensionality. Instead, the common approach to solve such inverse problems with many parameters is to use stochastic samplers such as MCMC [60], nested sampling [61] and other similar methods to explore this multi-dimensional parameter space. `PyCBC Inference` [62], `Bilby` [63, 64], and `LALInference` [65] are packages specifically developed for various inference tasks in gravitational wave analyses, and provide the implementation of numerous stochastic samplers. The end goal of such sampling algorithms is to generate a set of N multi-dimensional posterior samples $\vec{\theta}$, such that they approximately resemble the distribution $p(\vec{\theta}|d)$, i.e –

$$\{\vec{\theta}_j\}_{j=1}^N \sim p(\vec{\theta}|d) \quad (5.8)$$

Fig. 5.1 demonstrates one such example. It shows the posterior distribution of the analyses results for the first detection, GW150914 [10], as evaluated using 2 different waveform models. The corner plot shows the 2D probability density for the component source frame masses, along with their individual marginalized posterior distributions on the 2 axes.

5.2 Population Inference in a nutshell

Lets assume we have a catalog of N confirmed GW detections, and we have performed parameter estimation on each of them to obtain the posterior distributions of the event-level parameters. Let us focus on only one of those parameters, which we refer to as θ , and our goal is to infer the underlying population distribution of θ , which could be for example the mass function of the black hole and neutron stars. We parameterize the underlying probability distribution of θ with hyper-parameters Λ , i.e $\pi(\theta|\Lambda)$, and our aim is to infer the posterior distribution for Λ . This is done via bayesian hierarchical inference, and in this section we provide a very brief overview about the derivation of the population likelihood following [57]. For a more exhaustive treatment about hierarchical inference and the many intricacies involved in it, we refer the reader to [66].

We want to evaluate $\mathcal{L}(\vec{d}|\Lambda)$, where \vec{d} is the data of the catalog of N detections. Assuming that the N events are independently drawn from the population, we can decompose the likelihood and write it as –

$$\mathcal{L}_{\text{pop}}(\vec{d}|\Lambda) = \prod_{i=1}^N \mathcal{L}_{\text{pop}}(d_i|\Lambda) \quad (5.9)$$

To evaluate $\mathcal{L}_{\text{pop}}(d_i|\Lambda)$, we can marginalize over θ as follows –

$$\mathcal{L}_{\text{pop}}(d_i|\Lambda) = \int d\theta \mathcal{L}(d_i|\theta) \pi(\theta|\Lambda) \quad (5.10)$$

where $\mathcal{L}(d_i|\theta)$ is the likelihood used for analysis of individual events. Now, using Eqn 5.10 in Eqn 5.9, we get –

$$\mathcal{L}_{\text{pop}}(\vec{d}|\Lambda) = \prod_{i=1}^N \int d\theta \mathcal{L}(d_i|\theta) \pi(\theta|\Lambda) \quad (5.11)$$

Given the population likelihood, we can write down the posterior for Λ using our friendly neighbourhood Bayes' theorem –

$$p(\Lambda|\vec{d}) = \frac{\mathcal{L}_{\text{pop}}(\vec{d}|\Lambda)\pi(\Lambda)}{\int d\Lambda \mathcal{L}_{\text{pop}}(\vec{d}|\Lambda)\pi(\Lambda)} \quad (5.12)$$

where the denominator is now called the "hyper-evidence", and like the evidence in individual event analyses, it is used to compare between different population models.

Evaluating the hyper-posterior involves a lot of multi-dimensional integrals, but we can simplify this by estimating the individual integrals for each event with the help of the event level posteriors, that were obtained during their parameter estimation. We can rewrite Eqn 5.11 as –

$$\mathcal{L}_{\text{pop}}(\vec{d}|\Lambda) \propto \prod_{i=1}^N \int d\theta p(\theta|d_i) \frac{\pi(\theta|\Lambda)}{\pi(\theta|\text{PE})} \quad (5.13)$$

where $\pi(\theta|\text{PE})$ is the prior used for parameter estimation of an individual event, and we used Bayes' theorem to rewrite $\mathcal{L}(d_i|\theta)$. We can now estimate the integral as a weighted Monte Carlo sum, using the posterior samples $p(\theta|d_i)$ from the parameter estimation analyses of the i 'th event.

$$\int d\theta p(\theta|d_i) \frac{\pi(\theta|\Lambda)}{\pi(\theta|\text{PE})} \sim \frac{1}{n_i} \sum_{j=0}^{n_i} \frac{\pi(\theta_i^j|\Lambda)}{\pi(\theta_i^j|\text{PE})} \Big|_{\theta_i \sim p(\theta|d_i, \text{PE})} \quad (5.14)$$

where n_i is the number of posterior samples for the i 'th event. Using Eqn 5.14, we can now write the posterior for the hyper-parameters as –

$$p(\Lambda|\vec{d}) \propto \pi(\Lambda) \prod_{i=1}^N \frac{1}{n_i} \sum_{j=0}^{n_i} \frac{\pi(\theta_i^j|\Lambda)}{\pi(\theta_i^j|\text{PE})} \Big|_{\theta_i \sim p(\theta|d_i, \text{PE})} \quad (5.15)$$

This is much simpler to evaluate, since essentially all we need are the posterior samples from parameter estimation analysis of individual events in our catalog (which we do have!).

In this derivation, we made an important assumption that does not hold true in practice. We have assumed that our catalog of detections represents an unbiased subset of our underlying population. But GW detectors suffer from selection effects ("Malmquist Bias"), and we will delve into this in more detail soon, but before that we provide some of the frequently used models in GW population inference to fit the mass and redshift distribution of BBH.

5.3 Common Population Models

5.3.1 Redshift Model

Power Law Model

This simple redshift evolution model has only one parameter, λ and is defined as follows [12] –

$$p(z|\lambda) \propto \frac{dV_c}{dz} \frac{1}{1+z} (1+z)^\lambda \quad (5.16)$$

where $\frac{dV_c}{dz}$ is the differential comoving volume. In this model, $\lambda = 0$ refers to a merger rate density that does not intrinsically evolve across redshift, i.e it is uniform in comoving volume and source–frame time. The factor $(1+z)^{-1}$ is required for conversion from detector–frame time to source–frame time. However, we can have more complicated redshift evolution models as well, for example if the BBH merger rate follows the cosmic star formation rate (SFR) of the universe, we can use a redshift evolution model given by –

$$p(z|\{\lambda\}) \propto \frac{dV_c}{dz} \frac{1}{1+z} \psi(z) \quad (5.17)$$

with $\psi(z)$ denoting the specific SFR. One particular parameterization for $\psi(z)$ is –

$$\psi(z) = \frac{(1+z)^\gamma}{1 + \left(\frac{1+z}{1+z_p}\right)^\kappa} \quad (5.18)$$

where $\psi(z)$ roughly behaves as a power law with slope γ at lower redshifts, z_p is roughly where the function peaks, after which there is a power law drop-off at high redshifts with slope κ . This is similar to the parameterization used for the cosmic SFR, which is well fit by the Madau-Dickinson model given by [67] –

$$\psi(z) = 0.015 \frac{(1+z)^{2.7}}{1 + [(1+z)/2.9]^{5.6}} M_\odot \text{yr}^{-1} \text{Mpc}^{-3} \quad (5.19)$$

5.3.2 Mass Model

Different formation channels of compact binary mergers predict different mass distributions for the merging binaries. Some of them may predict a mass gap in the distribution as well. A simple model describing the BBH primary mass distribution is that of a power law model between a minimum and maximum mass. Using the catalog of detections, it is possible to infer the lower and upper edges of the mass function (m_{\min}, m_{\max}), as well as the slope of the power law, α . The distribution is given by [12] –

$$p_{\text{pow}}(m) \propto m^{-\alpha} : m_{\min} \leq m \leq m_{\max} \quad (5.20)$$

However, such a model is quite simple, and may not be able to fit the rich diversity of features that might be present in the true mass function. In fact, it has been shown that

such a model is disfavoured by the data [12, 68], as it cannot explain the current catalog of detections. To account for this, we try to make our models more complicated, with more parameters to fit for the extra features seen in addition to a simple power law. Commonly used models in LVK analyses is the `POWERLAW + PEAK model`, which has a gaussian feature and a smoothing at lower masses. This model searches for bumps in the mass distribution, and has been used to identify the tentative peak seen at $35M_\odot$ in the data. It has 7 free hyper-parameters and is give by [12] –

$$p(m_1|\alpha, m_{\min}, \delta_m, m_{\max}, \lambda_{\text{peak}}, \mu_m, \sigma_m) = [(1 - \lambda_{\text{peak}})B(m_1| - \alpha, m_{\max}) + \lambda_{\text{peak}}G(m_1|\mu_m, \sigma_m)] S(m_1|m_{\min}, \delta_m) \quad (5.21)$$

where $B(m_1| - \alpha, m_{\max})$ is a power law with slope $-\alpha$ and a upper cutoff at m_{\max} . $G(m_1|\mu_m, \sigma_m)$ is the gaussian component for modelling any peak present in the mass distribution, with λ_{peak} fixing the fraction of mergers in the power law component, B , and the gaussian component, G . As mentioned above, there is a tapering at the lower end, which is modeled by the smoothing function $S(m_1|m_{\min}, \delta_m)$ which increases from 0 to 1 in $(m_{\min}, m_{\min} + \delta_m)$. It is given by –

$$S(m|m_{\min}, \delta_m) = \begin{cases} 0 & (m < m_{\min}) \\ \frac{1}{f(m-m_{\min}, \delta_m)+1} & (m_{\min} \leq m < m_{\min} + \delta_m) \\ 1 & (m \geq m_{\min} + \delta_m) \end{cases} \quad (5.22)$$

with

$$f(m', \delta_m) = \exp\left(\frac{\delta_m}{m'} + \frac{\delta_m}{m' - \delta_m}\right) \quad (5.23)$$

The mass-ratio distribution is usually modelled as a conditional power-law with spectral index β_q and is give by [12] –

$$p(q|\beta_q, m_{\min}, m_1) \propto \begin{cases} q^{\beta_q} & \text{if } m_{\min} < m_2 < m_1 \\ 0 & \text{otherwise} \end{cases} \quad (5.24)$$

where $q = m_2/m_1$ is the mass ratio of a binary.

5.4 An Aside on Selection Effects

Gravitational waves are rather weak by the time of arrival at Earth, which is why it took decades of effort to build the current interferometers sensitive enough to detect them. But still, the signals are buried in the detector noise, and search algorithms need to parse through this data to look for signals, which are literally needles in the haystack. These algorithms can only identify events if they cross some threshold, such as their SNR should be high enough for an event to be considered real. What this also means is that the current interferometers are not sensitive to all the compact binary coalescences occurring in the universe, but only a subset of them which cross the detection threshold. Only the loudest signals would cross these thresholds. The strength of a GW when it reaches Earth depends on several factors, including properties of the source and the noise sensitivity of the ground-based interferometers. For example, the amplitude of the GW falls as we increase the luminosity distance. This implies that if the other properties remain fixed, nearby mergers would have a higher SNR than distant ones. Similarly, heavier mass mergers are louder than lighter ones, but this is only true as long as the merger remains in the frequency band where the detectors are sensitive. Essentially, we observe a biased subset of the true population of mergers occurring in the universe. This is extremely important to consider when we are trying to use detected signals to infer properties about the underlying population. Not including this information would lead to biases in population inference, and we would not infer the correct underlying population in the universe. In this section we briefly describe about how selection effects are handled in population analyses. We refer the reader to [66], which provides an exhaustive treatment of how to handle selection effects in population inference.

5.4.1 The Detection Fraction

Essentially, our goal is to calculate the detection fraction, $\alpha(\Lambda)$ where Λ are the parameters describing the underlying population. The detection fraction represents the proportion of

events that would be detected from a population of mergers that occur in the universe described by the population parameters. One can write it down explicitly as –

$$\alpha(\Lambda) = \int d\theta \pi(\theta|\Lambda) p(\mathcal{D}|\theta) \quad (5.25)$$

where $\pi(\theta|\Lambda)$ is the population distribution, $p(\mathcal{D}|\theta)$ denotes the probability of an event characterized by parameters θ crossing the set threshold of detection. Mathematically, it is written as –

$$p(\mathcal{D}|\theta) = \int_{\rho(d)_{\uparrow}} dd p(d|\theta) \quad (5.26)$$

where d are all possible data realizations, and $\rho(d)_{\uparrow}$ refers to when the data realization crosses the detection threshold. Eqn 5.25 is complicated to evaluate, and in most cases, does not have an analytical form. The usual technique to get around this is with a monte-carlo approach. Essentially, we draw N_{draw} samples from a reference population with parameters Λ_{ref} , inject the corresponding signals into real detector data that includes noise, run the detection algorithms to look for signals in the data, and record the observations which are detected. Using the detected samples, we can estimate the detection fraction of any general population with parameters Λ using a weighted monte-carlo integral which is given by [69] –

$$\alpha(\Lambda) \approx \frac{1}{N_{\text{draw}}} \sum_{j=1}^{N_{\text{det}}} \frac{\pi(\theta_j|\Lambda)}{\pi(\theta_j|\Lambda_{\text{ref}})} \quad (5.27)$$

There are a few caveats to this, since this is just a point estimate of the integral which has an uncertainty. The uncertainty can be estimated by –

$$\sigma^2(\Lambda) \equiv \frac{\alpha^2(\Lambda)}{N_{\text{eff}}} \approx \frac{1}{N_{\text{draw}}^2} \sum_{j=1}^{N_{\text{det}}} \left[\frac{\pi(\theta_j|\Lambda)}{\pi(\theta_j|\Lambda_{\text{ref}})} \right]^2 - \frac{\alpha^2(\Lambda)}{N_{\text{draw}}} \quad (5.28)$$

where we have introduced N_{eff} that approximately denotes the effective number of samples used for calculating the detection fraction. The accuracy of such an estimate of the detection fraction increases as we increase N_{draw} . As a rule of thumb, if we want to infer the population using N_{obs} detections, we must have at least a sufficient number of N_{draw} such that $N_{\text{eff}} > 4N_{\text{obs}}$ [69].

We demonstrate the above procedure with a mock injection population. We simulate $N = 100000$ non-spinning GW signals from BBH that come from a population where the primary black hole comes from a power-law distribution, the mass-ratio distribution is also a conditional power-law distribution and they are distributed in the universe such that the merger rate density does not intrinsically evolve across redshift ($\kappa = 0$), with no events above a maximum redshift of 2. The parameters for the model are –

1. $\alpha = 1.6$
2. $m_{\text{min}} = 10$
3. $m_{\text{max}} = 300$
4. $\beta_q = 1.7$
5. $\kappa = 0$

We randomize rest of the extrinsic parameters, such that the binaries are randomly oriented and isotropically distributed in the sky. We evaluate their SNR's using only a single detector network, i.e LIGO–Livingston at O3 sensitivity. Setting the detection threshold to be an SNR=8, we find that out of 100000 signals, only 2904 are detected, which corresponds to a detection fraction of 0.02904. Now using the draw probabilities of these 2904 signals, we can estimate the detection fraction for any other simulated population as well as the uncertainty in it using the procedure described above. We show how the detection fraction changes as we change the population parameters in Fig. 5.2, where we vary each parameters individually, and fix the remaining to the corresponding reference population values. The shaded band around the point estimates of the detection fraction show the 2σ uncertainty evaluated using Eqn. 5.28. We can clearly see that as we go away from the parameters that were used to generated the injected reference population, the uncertainty in our estimate increases. In Fig. 5.3, we repeat the same process with $N = 1000000$ events. The factor of 10

increases clearly decreases the uncertainty in our estimates, as we can see the narrower bands in all the panels. This illustrates the significance of having a huge injection population, so as to have an accurate estimate of the detection fraction throughout the space of parameters that we are interested in.

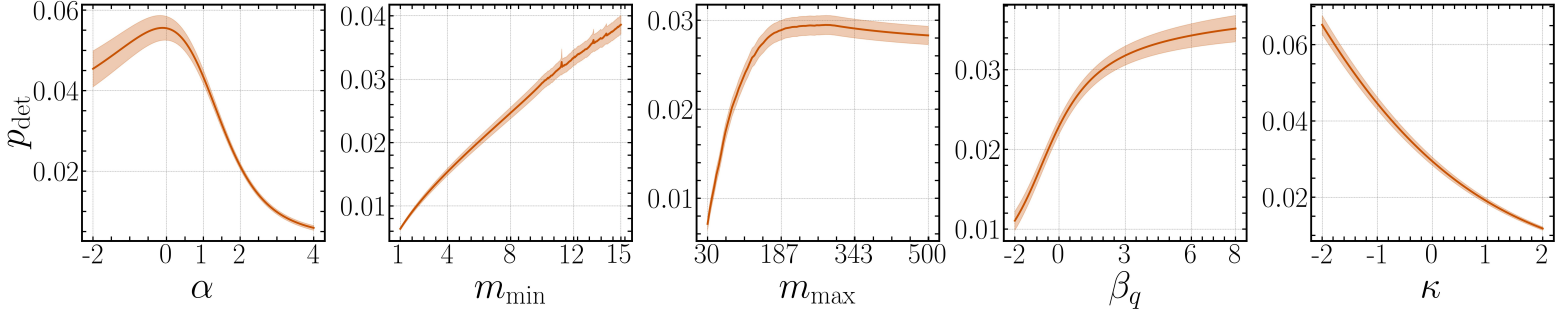


Figure 5.2: The detection fraction as a function of the population parameters evaluated using an injection set of 100000 events. In each panel, one of the population parameters is varied, while the others are fixed to their values as per the injected population. The bands depict the 2σ uncertainty in the estimate of the detection fraction.

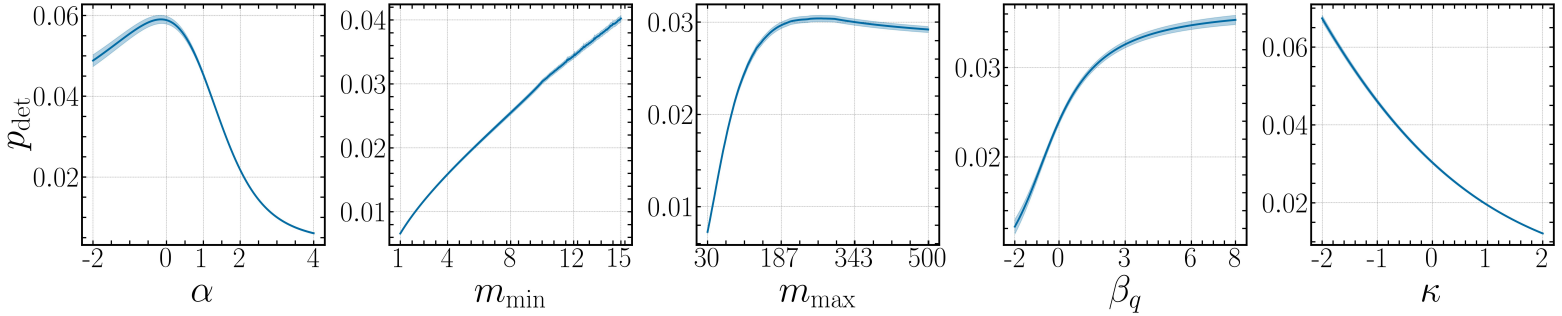


Figure 5.3: Same as Fig. 5.2, but using an injection set of 1000000 events to show that the uncertainty in the estimate of the detection fraction decreases with more signals.

We can also notice some interesting trends in how the detection fraction p_{det} depends on the population parameters, and it is worthwhile to explore why that is the case.

1. p_{det} vs α . This trend can be understood by 2 competing factors, i) heavier mass events are easier to detect than lower mass events, but this is only true up to a certain limit, since very high mass events also merge at lower frequencies, and can get redshifted out of the detectors sensitivity bands. Since α is the negative slope spectral index, a larger alpha corresponds to more lower mass mergers and less heavier mass mergers. This explains why the detection fraction increases as we slowly decrease α , but after a

certain point, when there are a lot of extremely heavy mergers, our sensitivity towards them decreases, and we see that as a decrease in p_{det} .

2. p_{det} vs m_{min} . This trend is relatively easier to understand. As we increase the minimum mass in our population, we have a larger chance of detecting more signals, as we do not have the lower mass signals in our population anymore which have a lower SNR in general.
3. p_{det} vs m_{max} . This trend is also intuitive, as increasing the maximum mass in our population allows us to have more heavier mass signals and we have a larger chance of detecting more signals. But the flattening and subsequent shallow decrease in p_{det} is because of the reason we mentioned earlier, i.e extremely heavy masses are not easy to detect as they merge outside the detectors sensitivity band.
4. p_{det} vs β . This trend is not directly obvious. It relies on the fact that for GWs from CBC's, symmetric mass mergers have a higher SNR than asymmetric mergers, everything else remaining equal. This is why p_{det} increases as β_q increases, as it allows for more mergers in the population with a mass ratio closer to 1.
5. p_{det} vs κ . This trend is fairly intuitive. Since the GW amplitude falls off inversely with distance, nearby mergers are easier to detect than more distant ones. A larger κ denotes a merger rate that increases with redshift, and we have more events at a higher redshift than at $\kappa = 0$. This explains why we see that p_{det} decreases as κ increases.

Having discussed how to evaluate the selection function, we can now include selection effects in bayesian hierarchical inference. We will ignore the derivation, which is given in detail in [66], and will only show the final result.

In the presence of selection effects, the only difference that arises is that Eqn 5.10 includes the detection fraction in the denominator, i.e

$$\mathcal{L}_{\text{pop}}(d_i|\Lambda) = \frac{1}{\alpha(\Lambda)} \int d\theta \mathcal{L}(d_i|\theta) \pi(\theta|\Lambda) \quad (5.29)$$

Using this, Eqn 5.15 now looks like –

$$p(\Lambda|\vec{d}) \propto \pi(\Lambda) \prod_{i=1}^N \frac{1}{\alpha(\Lambda)n_i} \sum_{j=0}^{n_i} \frac{\pi(\theta_i^j|\Lambda)}{\pi(\theta_i^j|\text{PE})} \Big|_{\theta_i \sim p(\theta|d_i, \text{PE})} \quad (5.30)$$

5.5 Population Inference in the presence of Lensing

Let's assume that we want to learn the underlying mass and redshift distribution of binaries in the universe from a catalog of detected events. We parameterize the mass function with Λ_m and the redshift function with Λ_z . The likelihood now reads as –

$$\mathcal{L}_{\text{pop}}(\{d_i\}|\Lambda_m, \Lambda_z) \propto \prod_i^N \frac{1}{\alpha(\Lambda_m, \Lambda_z)} \int d\theta \mathcal{L}(d_i|\theta) p(\theta|\Lambda_m, \Lambda_z) \quad (5.31)$$

For now, we will assume that we can de-couple the remaining event-level parameters from the mass and redshift parameters which we call m, z . We now have –

$$\mathcal{L}_{\text{pop}}(\{d_i\}|\Lambda_m, \Lambda_z) \propto \prod_i^N \frac{1}{\alpha(\Lambda_m, \Lambda_z)} \int dm dz \mathcal{L}(d_i|m, z) p(m, z|\Lambda_m, \Lambda_z) \quad (5.32)$$

where $p(m, z|\Lambda_m, \Lambda_z)$ is the population model for the mass and redshift function that is assumed beforehand. It important thing to notice here is that if we want to study the true underlying population, we need to make sure that we are using the correct and unbiased mass and redshift measurements. But as we learned in Section 4.3, we measure biased mass and redshifts for events that are lensed. The strength of the bias depends on the magnification factor μ . If we knew the μ values for all the events in our catalog, we could in principle just "de-lens" them on a event-by-event basis prior to performing population studies. However, the crucial issue here is that we *cannot measure the magnification factors of individual events*, since lensing in the geometrical optics regime only appears as a (de)-magnification of the amplitude of the signal. So how do we get around this? Will we never be able to understand the underlying population in the presence of lensing? Luckily, our universe does not assign

magnification factors to events randomly. Events that are further away are more likely to have μ 's that are non-unity than events that are nearby, due to the fact that the lensing optical depth increases with redshift due to more intervening matter. Let us assume that along with parameters representing the population of binaries, there is also an underlying lensing magnification distribution in the universe, which we parameterize with Λ_{lens} . Essentially, we want to evaluate $p(m_b, z_b | \Lambda_m, \Lambda_z, \Lambda_{\text{lens}})$ where m_b, z_b are the inferred source mass and redshifts that could be biased due to lensing. Λ_{lens} does not tell us about the individual magnifications of the events, but only the overall probability distribution. So what can we do, if we know the lensing magnification distribution, but not the exact μ ? We can *marginalize over all possible magnification factors!* Condensing $\{\Lambda_m, \Lambda_z\}$ together as Λ_{pop} , we can write down $p(m_b, z_b | \Lambda_m, \Lambda_z, \Lambda_{\text{lens}})$ as –

$$p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{lens}}) = \int p(m_b, z_b, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{lens}}) d\mu \quad (5.33)$$

If we know the μ for a given m_b, z_b , we can de-lens them and obtain the true masses and redshifts, which we refer to as m_s, z_s . They are given by –

$$z_s = z(\sqrt{\mu}d_L(z_b)) \quad (5.34)$$

$$m_s = m_b \left(\frac{1 + z_b}{1 + z_s} \right) \quad (5.35)$$

This implies that there is a 1–1 transformation between $\{m_s, z_s, \mu\}$ and $\{m_b, z_b, \mu\}$. Taking advantage of this relation, we can rewrite the integrand in Eqn. 5.33 as –

$$p(m_b, z_b, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{lens}}) = p(m_s, z_s, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{lens}}) \frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} \quad (5.36)$$

where we have introduced the jacobian to account for the transformation of variables.

Below, we list a few equations that will be needed to calculate the jacobian –

$$\frac{\partial m_s}{\partial m_b} = \frac{1 + z_b}{1 + z_s} \quad (5.37)$$

$$\frac{\partial z_s}{\partial z_b} = \sqrt{\mu} \frac{dz}{dd_L} \Big|_{d_{L_s}} \frac{dd_L}{dz} \Big|_{z_b} \quad (5.38)$$

Now, the jacobian is given by –

$$\frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} = \begin{bmatrix} \frac{\partial m_s}{\partial m_b} & \frac{\partial m_s}{\partial z_b} & \frac{\partial m_s}{\partial \mu} \\ \frac{\partial z_s}{\partial m_b} & \frac{\partial z_s}{\partial z_b} & \frac{\partial z_s}{\partial \mu} \\ \frac{\partial \mu}{\partial m_b} & \frac{\partial \mu}{\partial z_b} & \frac{\partial \mu}{\partial \mu} \end{bmatrix} \quad (5.39)$$

$$\frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} = \begin{bmatrix} \frac{\partial m_s}{\partial m_b} & \frac{\partial m_s}{\partial z_b} & \frac{\partial m_s}{\partial \mu} \\ 0 & \frac{\partial z_s}{\partial z_b} & \frac{\partial z_s}{\partial \mu} \\ 0 & 0 & 1 \end{bmatrix} \quad (5.40)$$

$$\frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} = \left(\frac{\partial m_s}{\partial m_b} \right) \left(\frac{\partial z_s}{\partial z_b} \right) \quad (5.41)$$

$$\frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} = \left(\frac{1 + z_b}{1 + z_s} \right) \sqrt{\mu} \frac{dz}{dd_L} \Big|_{d_{L_s}} \frac{dd_L}{dz} \Big|_{z_b} \quad (5.42)$$

Now, using Eqn 5.36 and Eqn 5.42, we can rewrite Eqn 5.33 as –

$$p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_b, z_b, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) d\mu \quad (5.43)$$

$$p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_s, z_s, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) \frac{\partial(m_s, z_s, \mu)}{\partial(m_b, z_b, \mu)} d\mu \quad (5.44)$$

$$p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_s, z_s, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) \left(\frac{1 + z_b}{1 + z_s} \right) \sqrt{\mu} \frac{dz}{dd_L} \Big|_{d_{L_s}} \frac{dd_L}{dz} \Big|_{z_b} d\mu \quad (5.45)$$

The closed–form expression of the derivative of luminosity distance with respect to redshift in flat spacetime is given by [70] –

$$\frac{dd_L}{dz} = \frac{d_L}{1 + z} + (1 + z) \frac{d_H}{E(z)} \quad (5.46)$$

where d_H denotes the Hubble distance, and $E(z)$ is the ratio of the Hubble expansion rate at redshift z with respect to the current rate –

$$E(z) = \sqrt{\Omega_m(1 + z)^3 + \Omega_\Lambda} \quad (5.47)$$

To simplify things further, we define a dummy variable $A(z)$ as –

$$A(z) = d_L(z) + \frac{(1 + z)^2 d_H}{E(z)} \quad (5.48)$$

This helps us rewrite Eqn 5.45 as –

$$p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_s, z_s, \mu | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) \frac{A(z_b)}{A(z_s)} \sqrt{\mu} d\mu \quad (5.49)$$

$$\boxed{p(m_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_s | \Lambda_{\text{pop}}) p(z_s | \Lambda_{\text{pop}}) p(\mu | z_s, \Lambda_{\text{len}}) \frac{A(z_b)}{A(z_s)} \sqrt{\mu} d\mu} \quad (5.50)$$

Eqn. 5.50 shows us how to correctly incorporate lensing in our framework, so that we can jointly estimate the population and lensing parameters in an unbiased manner, even if our individual measurements are biased.

This result can be easily accommodated for additional population parameters. For example, if we also wish to infer the mass-ratio distribution, the final result can be written as –

$$\boxed{p(m_b, q_b, z_b | \Lambda_{\text{pop}}, \Lambda_{\text{len}}) = \int p(m_s, q_s | \Lambda_{\text{pop}}) p(z_s | \Lambda_{\text{pop}}) p(\mu | z_s, \Lambda_{\text{len}}) \frac{A(z_b)}{A(z_s)} \sqrt{\mu} d\mu} \quad (5.51)$$

where it is important to note that $q_b = q_s$, since both the component masses are affected in the same manner by lensing. Hence the jacobian transformation remains the same, and there are no additional modifications to our result.

Part III

Major Takeaways & Discussion

Chapter 6

Results and Discussion

We have now explored what are GWs, what is gravitational lensing, and how does lensing affects GWs. We have also developed all the tools necessary for performing analyses of individual events, using a catalog of detections to understand the underlying population, and how to account for lensing in population analyses. We can now explore how gravitational lensing really affects population studies of GWs, and whether the tools that we have developed can handle lensing.

6.1 Toy Model

As a proof-of-principle, we first demonstrate the biases caused by lensing under a simplified framework, and later on mitigate the resulting biases using our modified population inference procedure that we described in Section 5.5. We list down certain assumptions that we make below for ease in computational expense. These do not hold in reality, however they provide a good starting point to understand, test, and interpret the implications of our method. The assumptions that we make are as follows:

1. All mergers in the universe are detected (*no selection effects*), and their masses and redshifts are measured with perfect accuracy, with no measurement uncertainties or noise¹. We also assume the binaries to be equal-mass as well for simplicity and reduce the number of population parameters.

¹This essentially means that we aren't really performing parameter estimation on an individual event here.

2. We only model 2 properties of the population of merging binaries, i.e their mass and redshift distributions.
3. The true lensing magnification distribution is a power-law in μ , and does not depend on the source redshift.

6.1.1 True Population

We assume the true mass function of merging black holes to be a gaussian distribution, i.e –

$$p(m|\mu_m, \sigma) \sim \mathcal{N}(\mu_m, \sigma) \quad (6.1)$$

This population distribution has 2 hyper-parameters, i.e the gaussian’s mean and standard deviation, for which we set their true values to be $\mu_m = 20M_\odot$ and width to be $\sigma = 2M_\odot$.

Similarly, we assume the redshift distribution is a POWERLAW model, that we described in Section 5.3.1. This model has 1 hyper-parameter, κ which measures the evolution of the merger rate across redshift. We set its true value to be $\kappa = 0$, which means that there is no intrinsic evolution in the merger rate across redshift. The probability distribution is normalized by assuming that there are no mergers in the universe beyond $z_{\max} = 3^2$.

Lastly, our magnification distribution model is also a power-law distribution, i.e –

$$p(\mu|\gamma) \propto \mu^{-\gamma} \quad ; \quad \mu \in [1, 10] \quad (6.2)$$

In general, strongly magnified events are less likely to occur than weakly magnified events, hence the negative factor in the definition. We have one hyper-parameter in this model, for which we set the true value as $\gamma = 3$.

²The z_{\max} is not a free population parameter in our analyses, but in principle we could also try to keep it as a free variable and infer it from the data.

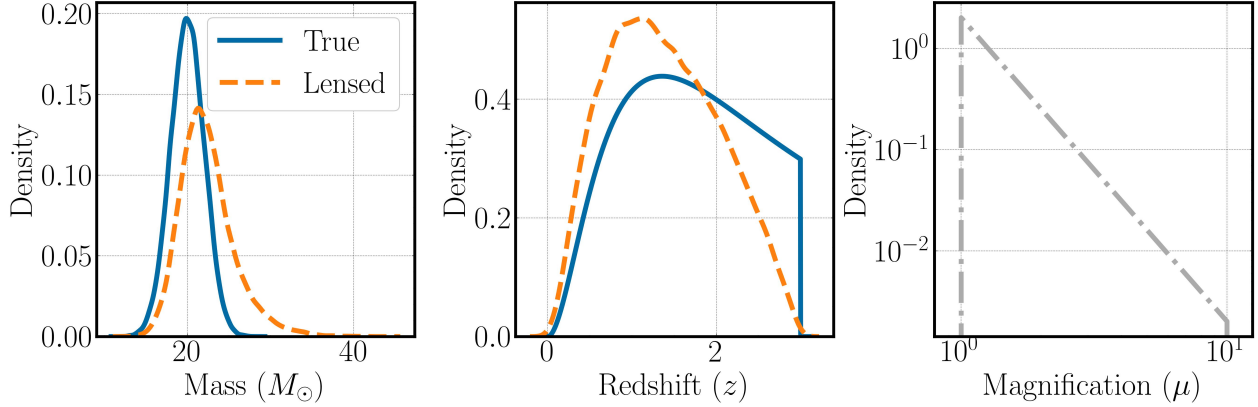


Figure 6.1: The blue solid curve in the left and middle panel show the mass and redshift distribution for binaries in our toy population mode, while the orange dashed curve shows the corresponding lensed population. The right panel shows the lensing magnification distribution that we have assumed.

Caveat: This is partly true, as the lensing magnification distribution in the strong-lensing regime is indeed a power-law of spectral index -3, i.e $p(\mu) \propto \mu^{-3}$. However, the complete magnification distribution includes the contribution of the weak lensing regime also, and the overall distribution depends on the source redshift as well. In fact, most mergers in the universe would only be weakly lensed, with only a small fraction among them being strongly lensed.

Now that we have everything set-up, we generate 2 mock catalogs of $N = 500$ events each

1. Unlensed (True) Population: We draw N samples from the true mass and redshift distributions as defined earlier. Since there are no selection effects, our detected event catalog consists of all the events, with their masses and redshifts perfectly measured. Since there is no lensing, these measurements are also unbiased.
2. Lensed Population: We repeat the same procedure as above, but along with the mass and redshift samples, we correspondingly also draw N samples from the magnification distribution. Now using the mass, redshift and μ samples, we calculate the biased mass and redshifts that we would measure. Although the measurements would be perfect, they would not be the "precise" true values due to lensing biases.

Now that we have 2 catalogs of 500 detected events, we use them to infer the underlying

population. We use our mechanism of bayesian hierarchical inference for this purpose. To demonstrate the biases in population inference due to lensing, and that our modified procedure is successful in mitigating the resulting biases, we perform three studies –

1. Standard population inference on the unlensed population. This can be thought of as our control study because there is no lensing, hence we expect to recover the true population parameters that we had created our catalog from.
2. Standard population inference on the lensed population. This is to demonstrate the biases that lensing might cause in population studies.
3. Modified population inference on the lensed population. This includes the marginalization over μ step as we had described earlier. This is to demonstrate that biases on a population level can be corrected if we include the lensing magnification model in our inference procedure.

6.2 Toy Model results

In our first study, we recover our injected population parameters reasonably well. We have shown the 3D posteriors (corner plot) for the hyper-parameters in Fig. 6.2, where the black dashed lines denote the 90% credible intervals (C.I), with the orange lines marking the true values. The plots along the diagonal are the marginalized posteriors for the 3 population parameters. The width of these posteriors would decrease as the number of events in our catalog increases³.

In our second study, where we perform standard population inference on our lensed population. The masses and redshifts of each event in our catalog are biased from their true values, but we are not correcting for it in our population inference procedure. Hence we see biases in the inferred population–parameter posteriors. For example, we estimate μ_m to be *higher* than injected. This is intuitive because in our model, lensing biases masses to higher values. We also see that the parameter for the redshift distribution κ is *inferred to be lower* than the true value. This is also intuitive because lensing biases redshifts to lower values than

³One might think that since we had perfect measurements for all our detected events, shouldn't the hyper-posteriors be centered exactly on the true values? This is actually not true, as we are working with a finite sized catalog, and the maximum posteriori estimate (maximum likelihood estimate in case of a uniform prior) need not lie on the true value.

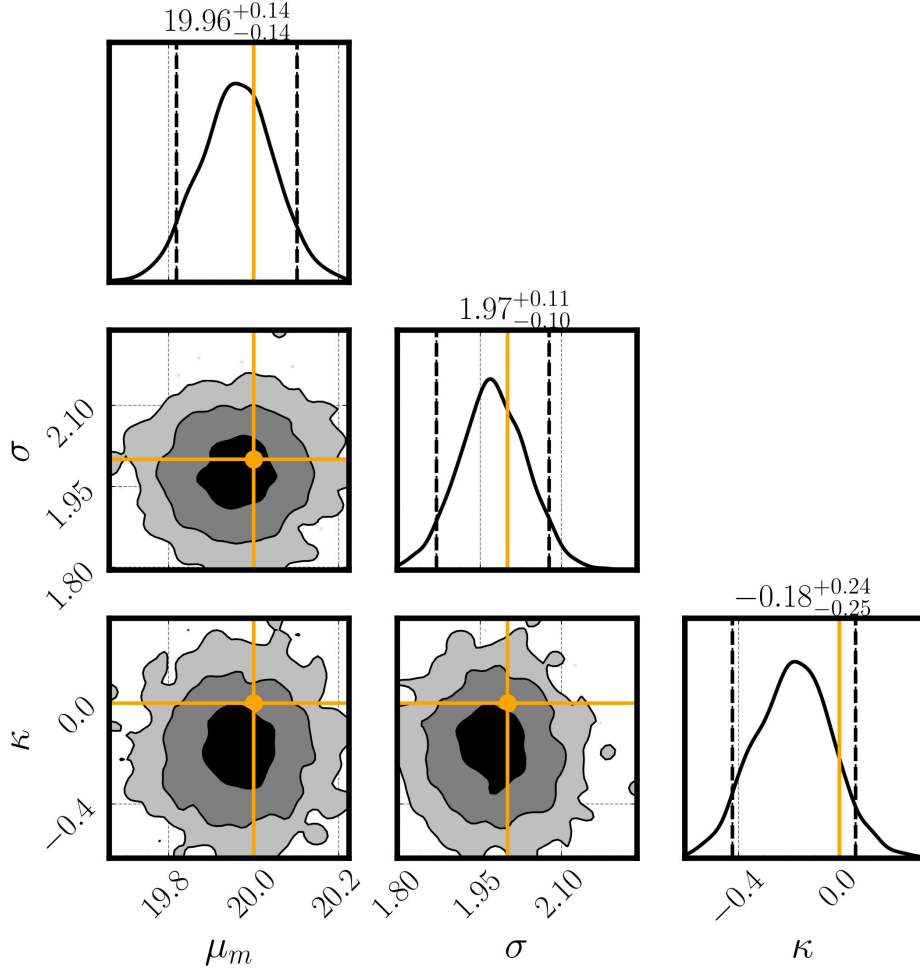


Figure 6.2: Constraints on the hyper-posteriors of the standard population inference on our unlensed catalog. The diagonal panels represent the marginalized posteriors for the population parameters, while the rest of the panels show the two-dimensional posterior distribution for each pair of hyper-parameters. The orange solid lines mark the true values, and the number quoted on top is for the median values with the the 90% credible interval, which is marked by the black dashed lines.

the true values. Hence we see a lot more mergers at lower redshifts than existed in the true population, and to fit for such an observation, the κ is underestimated. The overestimation of σ is also obvious when one looks at the orange dashed mass distribution in Fig. 6.1, which is clearly wider than the blue solid mass distribution.

Finally, in our last study, we show how our modified population inference algorithm can mitigate the biases that we saw in our second study. We again use our lensed catalog, but include the marginalization step in population inference to account for lensing. Assuming

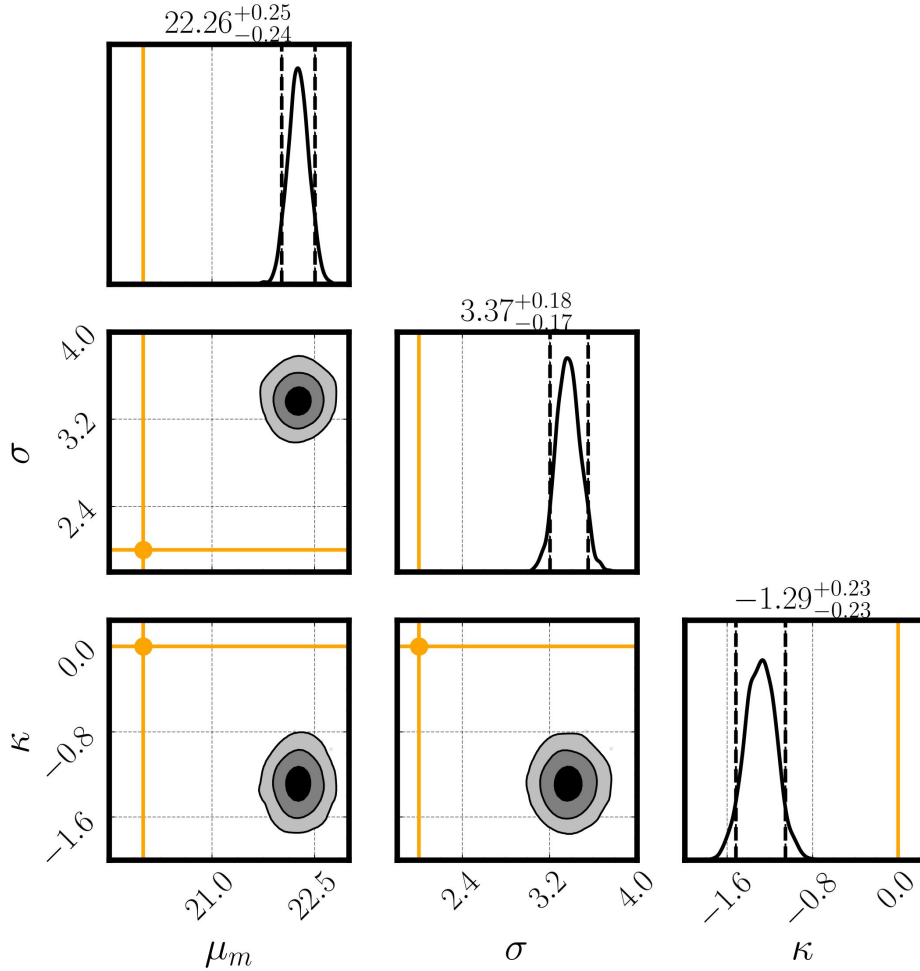


Figure 6.3: Constraints on the hyper-posteriors of the standard population inference on our lensed catalog to demonstrate that we recover biased population parameters.

that we know our lensing magnification distribution, we set $\gamma = 3$, and try to infer only the population parameters μ_m, σ, κ . In Fig. 6.4, we can clearly see that even though our catalog was lensed, we are able to recover the true population parameters reasonably well.

6.3 Can one learn the lensing magnification distribution?

Previously, we showed that we can de-lens our population assuming that we know our magnification distribution. But we can in fact go a step further by keeping the lensing distribution parameters as additional hyper-parameters that we can vary and try to infer from the data. This can help us check if we can recover the correct spectral index of the power law distribution that we have used. Therefore, we repeat the same study once again,

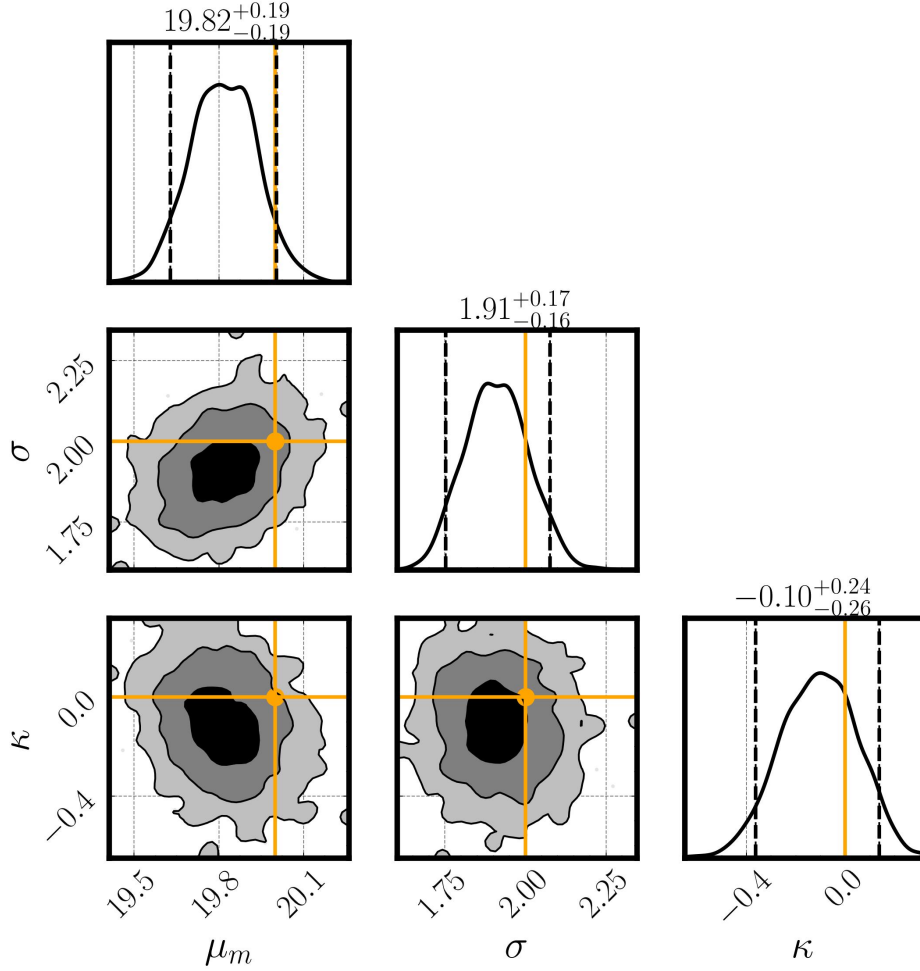


Figure 6.4: Constraints on the hyper-posteriors of the modified population inference on our lensed catalog to demonstrate that we recover the true population parameters.

but with an additional hyper-parameter γ for the spectral index of the lensing magnification distribution. As we can notice in Fig. 6.5, not only do we recover the correct population parameters, but also the correct value for the slope of the power-law which was injected at $\gamma = 3$. Essentially, in the presence of lensing, one needs to jointly infer the parameters for the population and magnification distributions in order to obtain unbiased results.

6.3.1 Correlations between the lensing and population parameters

In the blue panels of Fig. 6.5, we can notice an interesting correlation between the lensing distribution parameter γ , and the population parameters. Firstly we see a positive correlation between γ and μ_m . This can be understood by realizing that a higher γ prohibits large lensing

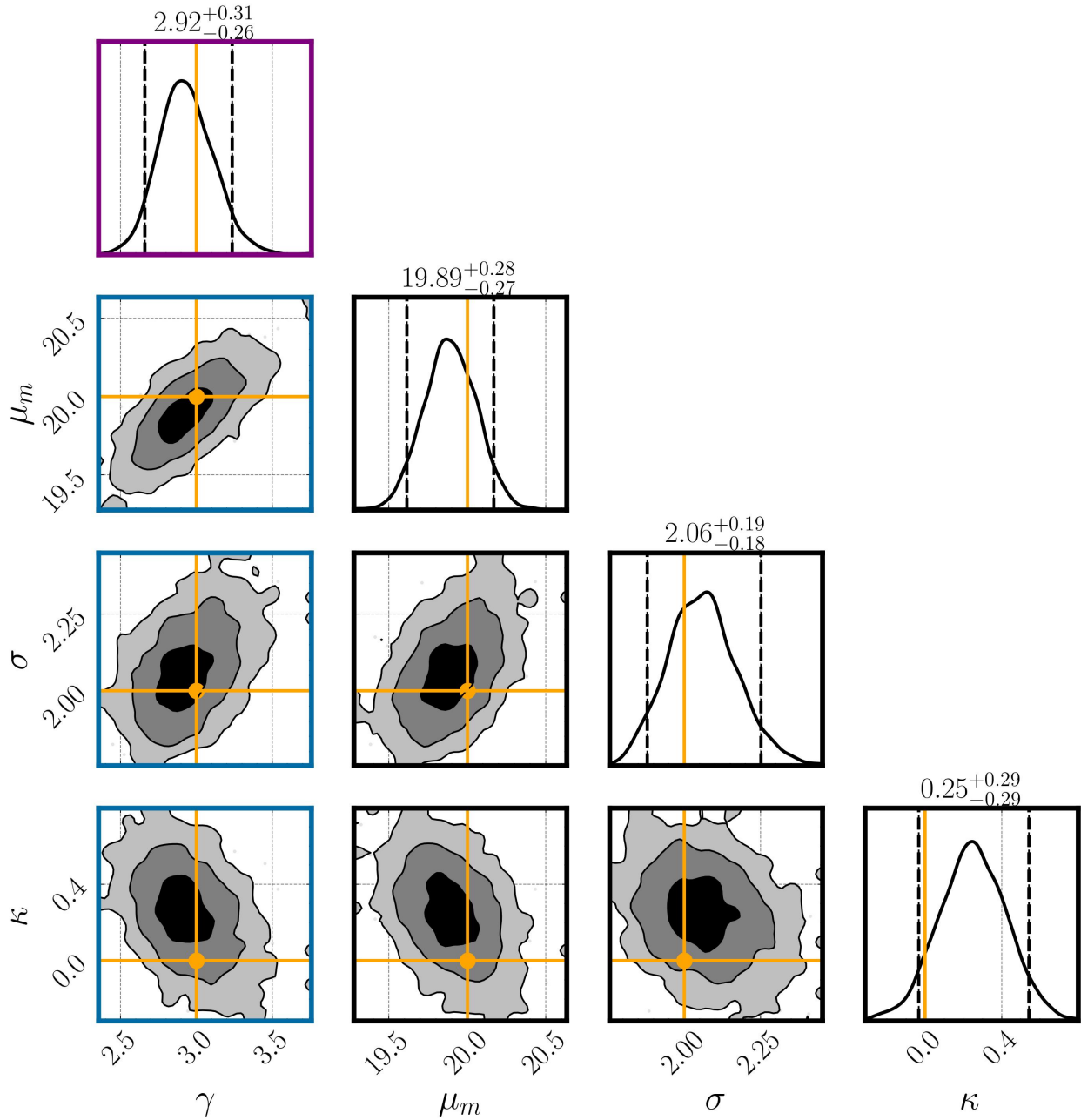


Figure 6.5: Constraints on the hyper-posteriors of the modified population inference on our lensed catalog with the slope of the magnification distribution also as a free hyper parameter. We demonstrate that we can jointly recover the correct slope of the power law, as well as the true population parameters.

magnifications, i.e they have a much lower probability. Hence, to fit for the observed mass distribution, we would need an intrinsic population that itself favours higher masses, which is

realized by a larger μ_m . Another interesting feature that can be noticed is an anti-correlation between γ and κ . A higher γ diminishes the probability of non-unity lensing magnifications, but since our catalog seems to have an apparent excess of events at low redshift, explaining them requires a lower κ as it increases the merger rate at low redshift.

6.4 Toy Model for Selection Effects

In all our previous results, we had assumed that we could detect all the events in our population, irrespective of their source parameters or redshifts. As we have explained in Section 5.4, this is not true in reality, as our detectors are sensitive to only a biased subset of the overall population. In reality, we only consider an event to be of astrophysical origin if it crosses some set threshold, which could be its signal-to-noise-ratio (SNR). An event's SNR depends on the data segment that we collect, which includes information both about the noise realization, and the true signal (which encodes the source parameters). In this section, we provide a crude demonstration about the biases that may arise in population inference due to not accounting for selection effects. To demonstrate this, we construct a toy model for what is considered a detection, while using the same model for the toy population and lensing. We set the detection threshold to be all events that are within a redshift of 1, i.e. an event in our population is only considered in our detection catalog if it has a $z < 1$. In reality, the threshold depends not just on the redshift, but on the source masses, spins, sky location, noise realization etc., but we ignore such dependencies for now.

Caveat: Since we have assumed perfect measurements of source parameters and zero noise, we do not need to worry about noise realizations, which can cause an event to be detectable even if the SNR calculated only from its source parameters is below the detection threshold. But in reality, it is essential to note that detection thresholds are always conditioned on the data that we collect, since that is the only thing we have access to in reality. Hence it is important to make sure that we calculate the detection fraction also based on the data and not on the source parameters. We refer the reader to [71] for more details.

In Fig. 6.6, we can see the difference in results depending on whether we include the detection fraction $\alpha(\Lambda)$ term in our population likelihood or not. The blue corner plot shows our inferred hyper-parameters when we do not include selection effects. We only see a bias in the parameter for the redshift distribution κ in the purple panel. This is not surprising since our simplified detection criteria solely depended on the redshift of an event. We see that κ

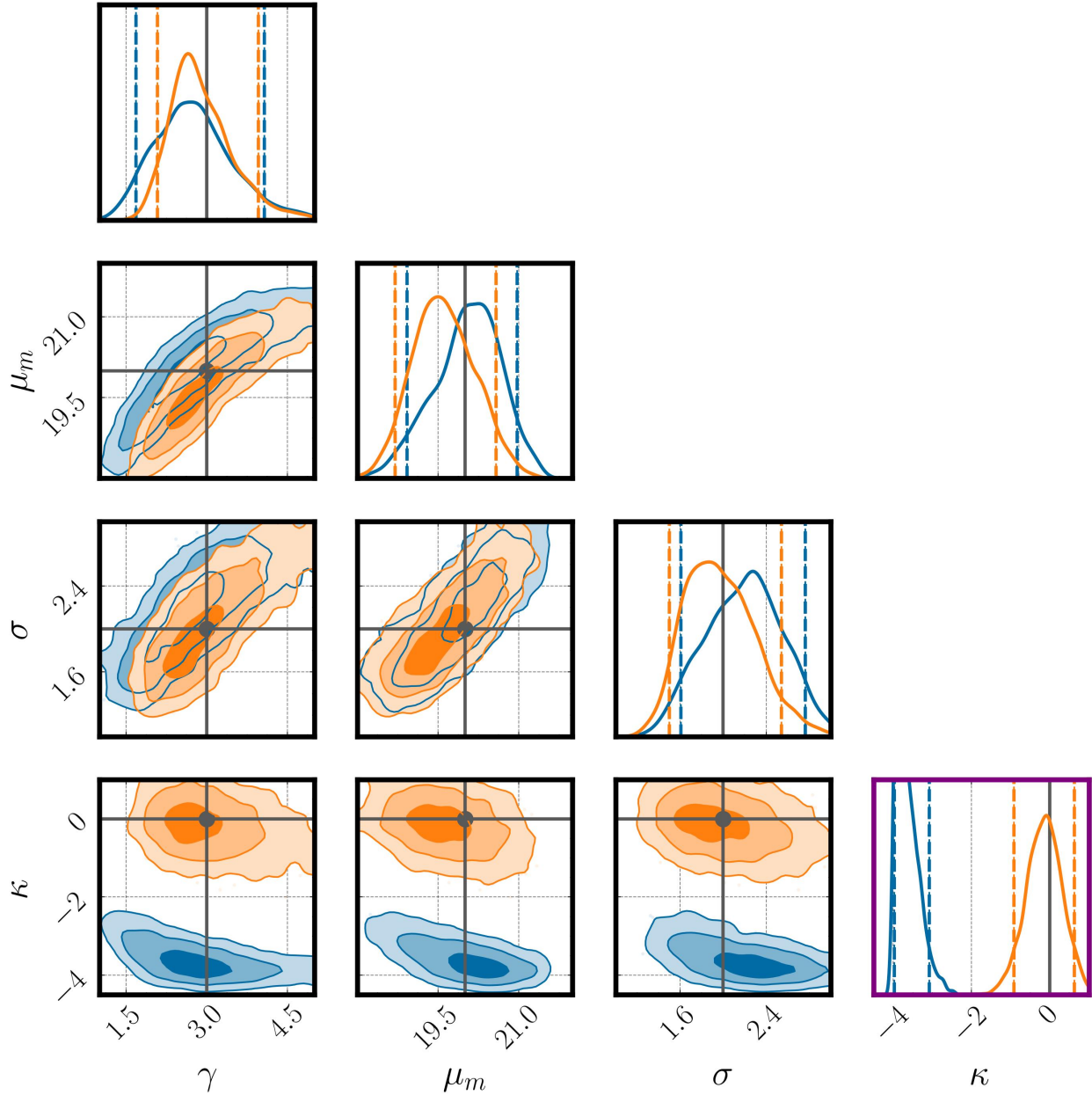


Figure 6.6: *Blue:* Constraints on the hyper-posteriors of the modified population inference on our lensed catalog with a detection criterion to demonstrate the biases in κ due to not accounting for selection effects. *Orange:* The same description, except that the population inference procedure accounts for selection effects, and we can see that we recover the correct redshift distribution parameter κ . The dark gray lines represent the true values for the hyper-parameters.

is underestimate, which means that the model prefers a merger rate that is higher at lower redshifts than in reality. This is intuitive as all our detected events have a $z < 1$, but the

inference procedure does not account for the fact that this is because we are only sensitive to detect such events. Meanwhile, if we do include the $\alpha(\Lambda)$ term, we recover the correct population and lensing distribution parameters as can be seen in the orange corner plot.

6.5 Towards a realistic scenario

The mass distribution of black holes that we have found is certainly not a gaussian distribution, and the lensing distribution is also not a power law in reality. Previously, we worked with a toy model for the population and lensing as a proof-of-principle to demonstrate that our approach works. But now, we try to modify our original scenario to step towards a more probable effect of lensing in our universe. We will still ignore selection effects and measurement uncertainties for computational ease, but will use a more realistic population model and lensing magnification distribution.

In reality the BBH mass distribution resembles more like a power law distribution⁴. Now, we also model the mass-ratio distribution now with the conditional power-law model. For the redshift distribution, we still use the power law model, but with a $z_{\max} = 20$ to explore lensing effects at a high redshift. Therefore, we now have 5 population parameters, i.e $\alpha, m_{\min}, m_{\max}, \beta_q$ for the mass distribution, and κ for the redshift distribution.

For our true population, we set the parameters to be $\alpha = 1.6, m_{\min} = 5, m_{\max} = 50, \beta_q = 1.7, \kappa = 0$.

As mentioned earlier, a power law model for lensing magnification is not realistic, as the true distribution not only includes contribution from weak lensing and strong lensing, but also depends on the source redshift. We use the more realistic magnification distribution as described in [38] (reproduced in Fig. 4.3) to explore lensing biases in this section.

For our catalog generation, we repeat the procedure described previously, i.e we first sample component masses and redshifts from the true population. For each event in our catalog, we also sample a μ for the merger *using* the realistic lensing magnification model that also utilizes information about the true source redshift. Using them, we can now generate our lensed catalog. We have shown our lensed and unlensed (true) catalog in Fig. 6.7, each consisting of 5000 events (all of which we assume to be detected as we do not include selection effects here). We also show the mass ratio distribution, which remains the same as lensing

⁴We will ignore the low mass smoothening and the gaussian peak identified in the data for now.

does not affect the mass ratio measurements. The lower right panel shows the magnifications of the 5000 events. We can see that not only do we have magnified events ($\mu > 1$), but also de-magnified events ($\mu < 1$).

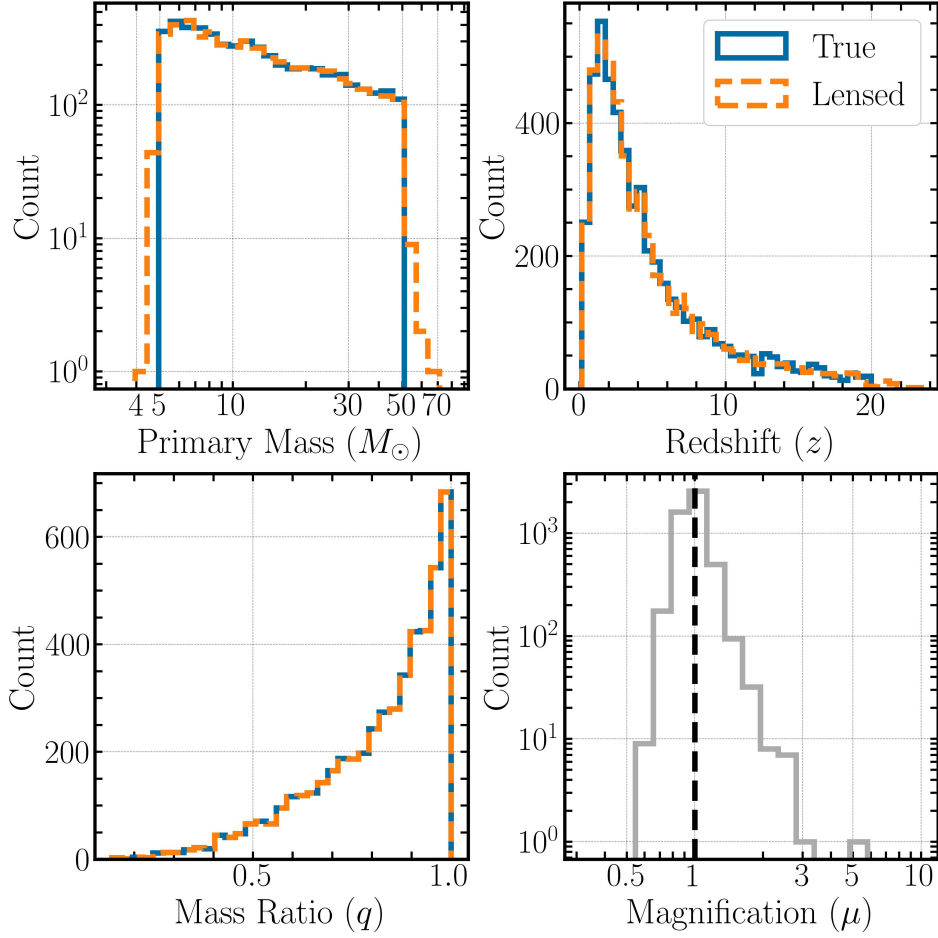


Figure 6.7: The blue solid curve in the top panels, and the bottom-left panel show the primary mass, redshift and mass-ratio distribution of binaries for our new population model, while the orange dashed curve in the same panels show the corresponding lensed population as obtained using the more realistic lensing magnification model. The bottom-right panel depicts the lensing magnification distribution for the events in our catalog, with $\mu = 1$ marked for reference.

We can immediately notice a few important effects of lensing on the population distribution directly from this figure, without performing hierarchical inference. Firstly, the lensed mass distribution does not resemble a power law anymore (especially at the edges), because it has been biased with magnification factors, and the magnifications depend on the true redshift distribution and the lensing magnification distribution. Since we would detect events from the

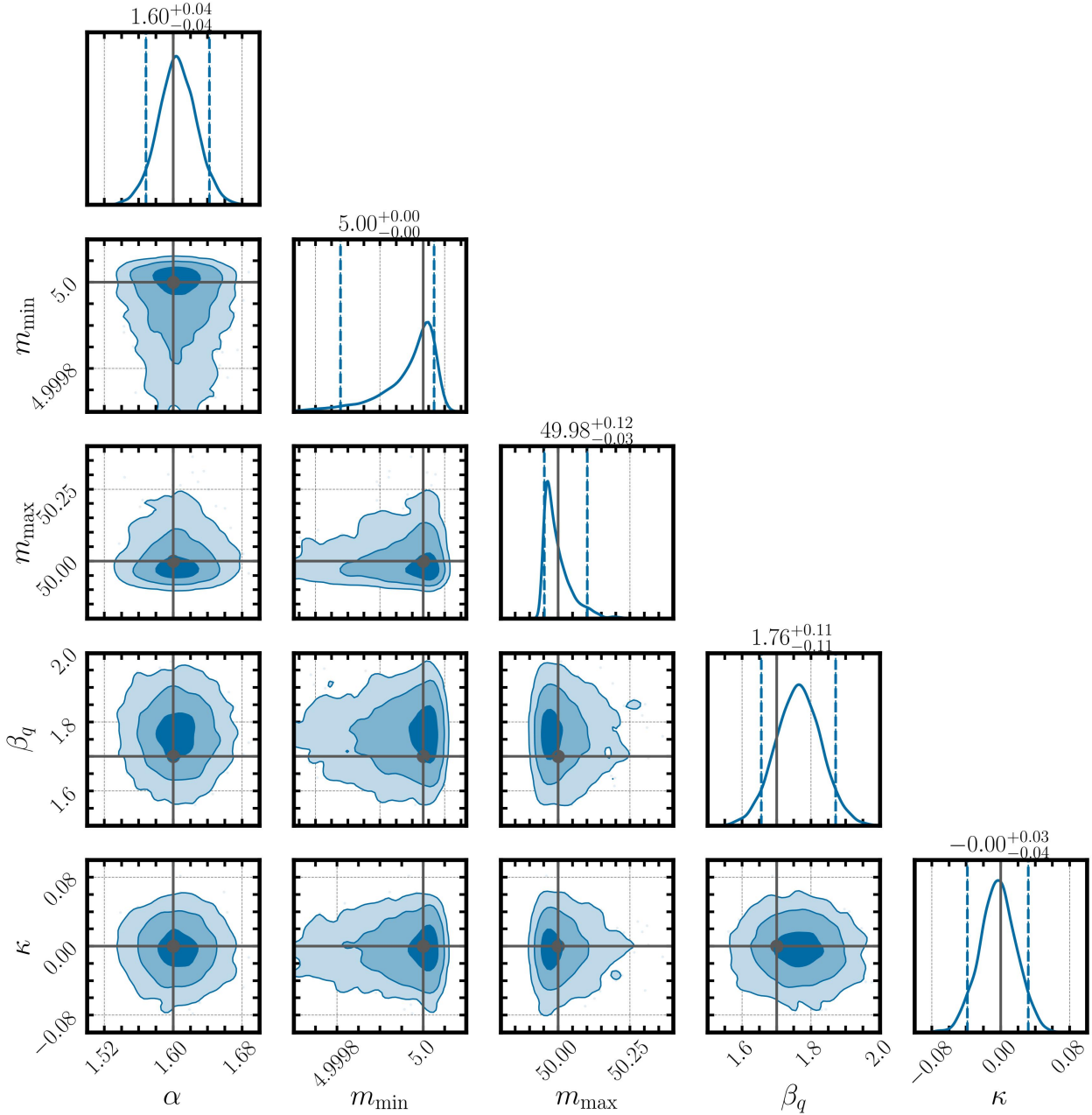


Figure 6.8: Constraints on the hyper-posteriors of standard population inference on our unlensed catalog to demonstrate that we recover the true population parameters.

orange dashed distribution, it would be incorrect to perform standard population inference with a power law mass model and hope to recover the solid blue distribution. Secondly, in the original population, there exist sharp cutoffs in the mass distribution at the upper and lower ends (by construction). Due to lensing, these cutoffs may get smeared, as we can have events

close to (but less than) m_{\max} that get lensed and appear as mergers with masses greater than the maximum allowed mass. Similarly, some events close to (but greater than) m_{\min} can get de-magnified due to lensing, and appear to have masses lower than the minimum mass in the distribution.

Sharp edges in the mass function are particularly useful, since their locations are important probes of the formation of compact binaries. The location of the upper mass gap in particular is related to the Pair-Instability Supernovae (PISN) phenomena, and the lower mass gap has important implications for the maximum possible mass of a neutron star, and the transition from neutron stars to black holes. In fact, different theoretical models make different predictions, with some of them predicting a mass-gap between the most massive neutron stars and the least massive black holes, while others do not predict a gap between them. Features in the mass distribution have also been proven to be useful in measuring cosmological parameters, especially through the *spectral sirens* method [72]. Thus accurately measuring their locations in the data empirically is crucial to understand not only compact object formation, but also in other areas in gravitational wave astronomy that utilize features in the gravitational wave population.

We perform population inference on the unlensed and lensed population using the standard algorithm that does not account for lensing. As we can see in Fig. 6.8, we infer the correct population hyper-parameters for our unlensed population as one would expect. For our lensed population in Fig. 6.9, we can see that we do not recover some of the population parameters correctly. In particular, there is a mild bias in the recovered slope of the mass distribution, α . We also underestimate the m_{\min} parameters, and overestimate the m_{\max} parameter. This was anticipated previously, as we have some events that get magnified and appear to be greater than the true m_{\max} , while some events get de-magnified and appear to be lighter than the true m_{\min} . We do not see any noticeable bias in the redshift distribution parameter κ , but we do see a bias in the mass-ratio spectral index β_q . At first, this may seem counter-intuitive, as we know that the mass-ratio does not change due to lensing, and the marginal mass-ratio distribution for our lensed and unlensed catalogs looks identical in Fig. 6.7 for the same reason. But it is crucial to understand that in population inference, we assume the mass-ratio distribution as a *conditional* power-law distribution, as defined in Eqn. 5.24. This distribution clearly depends on m_1 and m_{\min} , and biases in them have led to biases in the inferred β_q .

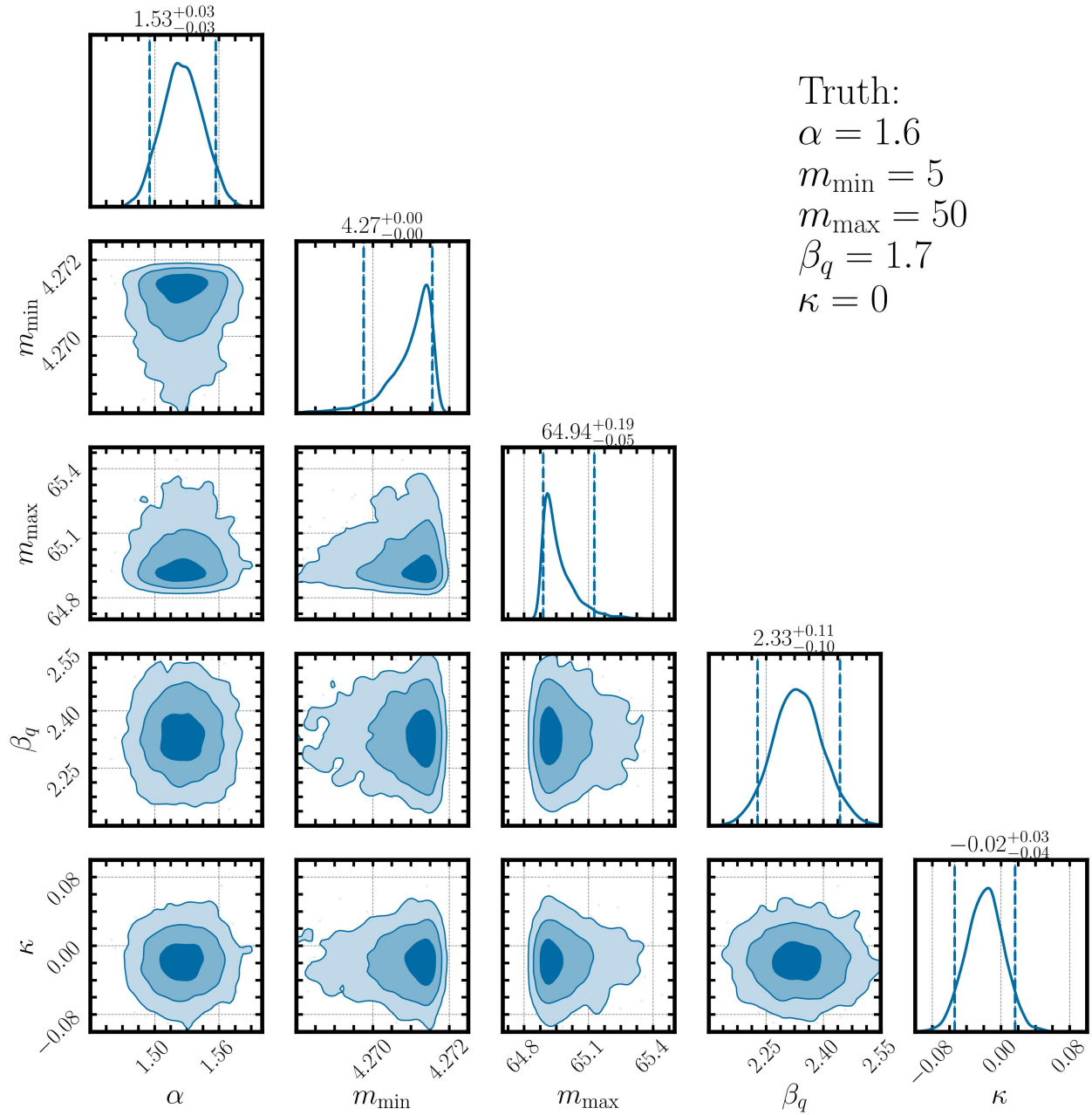


Figure 6.9: Constraints on the hyper-posteriors of standard population inference on our lensed catalog to demonstrate the biases in the recovered population-parameters.

Chapter 7

Conclusion and Future Directions

Gravitational lensing of gravitational waves is an exciting topic that may soon be detected in the near future with the existing detectors, and surely with the next generation of detectors. Like in electromagnetic astronomy, lensing of GWs can be a powerful tool to probe several questions in astrophysics, cosmology and the nature of gravity. However, as we have learned, lensing can also lead to biases in source mass and redshift measurements of individual events.

Population inference is an important probe of the formation and evolution of compact binary mergers in the universe, but it relies on accurate measurement of the source parameters such as masses and redshifts. In this work, we explored the effects of gravitational lensing on a population of GW sources to show the biases that may occur in population inference. We also developed a bayesian hierarchical inference algorithm that can mitigate the resulting biases in the population parameters, by including a step where we marginalize over all possible lensing magnifications, weighted by their corresponding probabilities. We demonstrated our approach using a toy population, and also show the biases with a more realistic population. However, we made several assumptions in our results for ease in computational expense and better interpretability. In the future, we plan on expanding our study to address and relax them as follows –

1. Measurement Uncertainties: We had assumed perfect measurements (just 1 posterior sample at the true value) for each event, but we will next include realistic measurements, where the properties of the CBC are measured with some uncertainty, i.e we have a

posterior distribution, and posteriors for different parameters may be correlated with each other as well (for example source mass and redshifts are often correlated as we measure their combination, the detector frame mass). The posterior widths of different parameters depend on the SNR of the detected event as well. To obtain realistic measurements of the source parameters, we will perform bayesian parameter estimation for each event in our catalog of detections, including the detector sensitivities of our network as well.

2. Selection Effects: In most of our results, we did not include detector selection effects, and only a toy model was implemented in one case where the detection statistic solely depended on the redshift. In the next step, we plan on including realistic selection effects for our detector network by calculating the detection fraction with a large-scale simulation study.
3. De-lensing a realistic population: We demonstrated how we can jointly learn the population parameters and the slope of the lensing magnification distribution for a toy model. We also showed the possible biases that may arise with a realistic mass distribution and a lensing magnification distribution that includes the dominant weak lensing component as well. In the future, we will demonstrate how we can de-lens any possible biases in our population parameters for a realistic mass and redshift distribution that has been lensed by a realistic lensing magnification distribution. We will also explore the possibilities of learning the free parameters of the weak lensing + strong lensing magnification distribution along with the population parameters.

Including measurement uncertainties and selection effects would mean that our constraints on the population-parameters would become poorer, i.e the hyper-posteriors would have larger widths than they have right now. In that case, it may be possible that any systematic bias that may be present due to lensing is sub-dominant to the overall uncertainties in our recovery of the population-parameters. This is the likely situation right now, where we are data limited, and only have $\mathcal{O}(100)$ events in our catalog, with all of them being low redshift events. In these situations, the systemic error (if present) is likely sub-dominant to the overall error in the population-parameters, and lensing may not be a concern right now. However, the uncertainties in recovery will decrease as the number of events in our catalog increases, and as we probe events till high redshifts with the next generation of detectors.

Lensing of gravitational waves will be an important tool in the future to study various

interesting phenomena, but at the same time it is imperative to develop methods to make sure that it does not lead to biases in studies involving GW detections, and in this thesis we explored one such scenario involving population studies of gravitational waves.

References

- [1] B. S. Sathyaprakash and B. F. Schutz. “Physics, Astrophysics and Cosmology with Gravitational Waves”. In: *Living Rev. Rel.* 12 (2009), p. 2. DOI: [10.12942/lrr-2009-2](https://doi.org/10.12942/lrr-2009-2). arXiv: [0903.0338](https://arxiv.org/abs/0903.0338) [gr-qc].
- [2] *Advanced Gravitational Wave Detectors*. Cambridge University Press, 2012.
- [3] J. Aasi et al. “Advanced LIGO”. In: *Class. Quant. Grav.* 32 (2015), p. 074001. DOI: [10.1088/0264-9381/32/7/074001](https://doi.org/10.1088/0264-9381/32/7/074001). arXiv: [1411.4547](https://arxiv.org/abs/1411.4547) [gr-qc].
- [4] F. Acernese et al. “Advanced Virgo: a second-generation interferometric gravitational wave detector”. In: *Class. Quant. Grav.* 32.2 (2015), p. 024001. DOI: [10.1088/0264-9381/32/2/024001](https://doi.org/10.1088/0264-9381/32/2/024001). arXiv: [1408.3978](https://arxiv.org/abs/1408.3978) [gr-qc].
- [5] T. Akutsu et al. “KAGRA: 2.5 Generation Interferometric Gravitational Wave Detector”. In: *Nature Astron.* 3.1 (2019), pp. 35–40. DOI: [10.1038/s41550-018-0658-y](https://doi.org/10.1038/s41550-018-0658-y). arXiv: [1811.08079](https://arxiv.org/abs/1811.08079) [gr-qc].
- [6] M. Bailes et al. “Gravitational-wave physics and astronomy in the 2020s and 2030s”. In: *Nature Rev. Phys.* 3.5 (2021), pp. 344–366. DOI: [10.1038/s42254-021-00303-8](https://doi.org/10.1038/s42254-021-00303-8).
- [7] Matthew Evans et al. “A Horizon Study for Cosmic Explorer: Science, Observatories, and Community”. In: (Sept. 2021). arXiv: [2109.09882](https://arxiv.org/abs/2109.09882) [astro-ph.IM].
- [8] Samantha A. Usman et al. “The PyCBC search for gravitational waves from compact binary coalescence”. In: *Class. Quant. Grav.* 33.21 (2016), p. 215004. DOI: [10.1088/0264-9381/33/21/215004](https://doi.org/10.1088/0264-9381/33/21/215004). arXiv: [1508.02357](https://arxiv.org/abs/1508.02357) [gr-qc].
- [9] Surabhi Sachdev et al. “The GstLAL Search Analysis Methods for Compact Binary Mergers in Advanced LIGO’s Second and Advanced Virgo’s First Observing Runs”. In: (Jan. 2019). arXiv: [1901.08580](https://arxiv.org/abs/1901.08580) [gr-qc].
- [10] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116.6 (2016), p. 061102. DOI: [10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102). arXiv: [1602.03837](https://arxiv.org/abs/1602.03837) [gr-qc].
- [11] R. Abbott et al. “GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run”. In: *Phys. Rev. X* 13.4 (2023), p. 041039. DOI: [10.1103/PhysRevX.13.041039](https://doi.org/10.1103/PhysRevX.13.041039). arXiv: [2111.03606](https://arxiv.org/abs/2111.03606) [gr-qc].

- [12] R. Abbott et al. “Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3”. In: *Phys. Rev. X* 13.1 (2023), p. 011048. DOI: [10.1103/PhysRevX.13.011048](https://doi.org/10.1103/PhysRevX.13.011048). arXiv: [2111.03634](https://arxiv.org/abs/2111.03634) [[astro-ph.HE](#)].
- [13] R. Abbott et al. “Constraints on the Cosmic Expansion History from GWTC-3”. In: *Astrophys. J.* 949.2 (2023), p. 76. DOI: [10.3847/1538-4357/ac74bb](https://doi.org/10.3847/1538-4357/ac74bb). arXiv: [2111.03604](https://arxiv.org/abs/2111.03604) [[astro-ph.CO](#)].
- [14] Katerina Chatziioannou. “Neutron star tidal deformability and equation of state constraints”. In: *Gen. Rel. Grav.* 52.11 (2020), p. 109. DOI: [10.1007/s10714-020-02754-3](https://doi.org/10.1007/s10714-020-02754-3). arXiv: [2006.03168](https://arxiv.org/abs/2006.03168) [[gr-qc](#)].
- [15] Rory J. E. Smith et al. “Inferring the population properties of binary black holes from unresolved gravitational waves”. In: *Mon. Not. Roy. Astron. Soc.* 496.3 (2020), pp. 3281–3290. DOI: [10.1093/mnras/staa1642](https://doi.org/10.1093/mnras/staa1642). arXiv: [2004.09700](https://arxiv.org/abs/2004.09700) [[astro-ph.HE](#)].
- [16] Laura Kreidberg et al. “MASS MEASUREMENTS OF BLACK HOLES IN X-RAY TRANSIENTS: IS THERE A MASS GAP?” In: *The Astrophysical Journal* 757.1 (Sept. 2012), p. 36. ISSN: 1538-4357. DOI: [10.1088/0004-637x/757/1/36](https://doi.org/10.1088/0004-637x/757/1/36). URL: <http://dx.doi.org/10.1088/0004-637X/757/1/36>.
- [17] R. Farmer et al. “Mind the gap: The location of the lower edge of the pair instability supernovae black hole mass gap”. In: (Oct. 2019). DOI: [10.3847/1538-4357/ab518b](https://doi.org/10.3847/1538-4357/ab518b). arXiv: [1910.12874](https://arxiv.org/abs/1910.12874) [[astro-ph.SR](#)].
- [18] Carl L. Rodriguez et al. “Dynamical Formation of the GW150914 Binary Black Hole”. In: *Astrophys. J. Lett.* 824.1 (2016), p. L8. DOI: [10.3847/2041-8205/824/1/L8](https://doi.org/10.3847/2041-8205/824/1/L8). arXiv: [1604.04254](https://arxiv.org/abs/1604.04254) [[astro-ph.HE](#)].
- [19] Mohammadtaher Safarzadeh, Sylvia Biscoveanu, and Abraham Loeb. “Constraining the delay time distribution of compact binary objects from the stochastic gravitational wave background searches”. In: *Astrophys. J.* 901.2 (2020), p. 137. DOI: [10.3847/1538-4357/abb1af](https://doi.org/10.3847/1538-4357/abb1af). arXiv: [2004.12999](https://arxiv.org/abs/2004.12999) [[astro-ph.HE](#)].
- [20] Maya Fishbach and Vicky Kalogera. “The Time Delay Distribution and Formation Metallicity of LIGO-Virgo’s Binary Black Holes”. In: *Astrophys. J. Lett.* 914.2 (2021), p. L30. DOI: [10.3847/2041-8213/ac05c4](https://doi.org/10.3847/2041-8213/ac05c4). arXiv: [2105.06491](https://arxiv.org/abs/2105.06491) [[astro-ph.HE](#)].
- [21] Susmita Adhikari et al. “The Binary–Host Connection: Astrophysics of Gravitational-Wave Binaries from Host Galaxy Properties”. In: *Astrophys. J.* 905.1 (2020), p. 21. DOI: [10.3847/1538-4357/abfb7](https://doi.org/10.3847/1538-4357/abfb7). arXiv: [2001.01025](https://arxiv.org/abs/2001.01025) [[astro-ph.GA](#)].
- [22] T. Callister. “A Thesaurus for Common Priors in Gravitational-Wave Astronomy”. In: (Apr. 2021). arXiv: [2104.09508](https://arxiv.org/abs/2104.09508) [[gr-qc](#)].
- [23] M. Saleem et al. “The science case for LIGO-India”. In: *Class. Quant. Grav.* 39.2 (2022), p. 025004. DOI: [10.1088/1361-6382/ac3b99](https://doi.org/10.1088/1361-6382/ac3b99). arXiv: [2105.01716](https://arxiv.org/abs/2105.01716) [[gr-qc](#)].
- [24] Michele Maggiore et al. “Science Case for the Einstein Telescope”. In: *JCAP* 03 (2020), p. 050. DOI: [10.1088/1475-7516/2020/03/050](https://doi.org/10.1088/1475-7516/2020/03/050). arXiv: [1912.02622](https://arxiv.org/abs/1912.02622) [[astro-ph.CO](#)].
- [25] Ish Gupta et al. “Characterizing Gravitational Wave Detector Networks: From A[#] to Cosmic Explorer”. In: (July 2023). arXiv: [2307.10421](https://arxiv.org/abs/2307.10421) [[gr-qc](#)].

- [26] Gabriella Agazie et al. “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background”. In: *Astrophys. J. Lett.* 951.1 (2023), p. L8. DOI: [10.3847/2041-8213/acdac6](https://doi.org/10.3847/2041-8213/acdac6). arXiv: [2306.16213](https://arxiv.org/abs/2306.16213) [[astro-ph.HE](#)].
- [27] J. Antoniadis et al. “The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals”. In: *Astron. Astrophys.* 678 (2023), A50. DOI: [10.1051/0004-6361/202346844](https://doi.org/10.1051/0004-6361/202346844). arXiv: [2306.16214](https://arxiv.org/abs/2306.16214) [[astro-ph.HE](#)].
- [28] Daniel J. Reardon et al. “Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array”. In: *Astrophys. J. Lett.* 951.1 (2023), p. L6. DOI: [10.3847/2041-8213/acdd02](https://doi.org/10.3847/2041-8213/acdd02). arXiv: [2306.16215](https://arxiv.org/abs/2306.16215) [[astro-ph.HE](#)].
- [29] Heng Xu et al. “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I”. In: *Res. Astron. Astrophys.* 23.7 (2023), p. 075024. DOI: [10.1088/1674-4527/acdfa5](https://doi.org/10.1088/1674-4527/acdfa5). arXiv: [2306.16216](https://arxiv.org/abs/2306.16216) [[astro-ph.HE](#)].
- [30] Monica Colpi et al. “LISA Definition Study Report”. In: (Feb. 2024). arXiv: [2402.07571](https://arxiv.org/abs/2402.07571) [[astro-ph.CO](#)].
- [31] Massimo Meneghetti. *Introduction to Gravitational Lensing: With Python Examples*. 2022.
- [32] Matthias Bartelmann and Peter Schneider. “Weak Gravitational Lensing”. In: *Physics Reports* 340.4 (Jan. 2001), pp. 291–472. ISSN: 03701573. DOI: [10.1016/S0370-1573\(00\)00082-X](https://doi.org/10.1016/S0370-1573(00)00082-X). arXiv: [astro-ph/9912508](https://arxiv.org/abs/astro-ph/9912508). URL: <http://arxiv.org/abs/astro-ph/9912508> (visited on 02/21/2024).
- [33] S. Birrer et al. “Time-Delay Cosmography: Measuring the Hubble Constant and other cosmological parameters with strong gravitational lensing”. In: (Oct. 2022). arXiv: [2210.10833](https://arxiv.org/abs/2210.10833) [[astro-ph.CO](#)].
- [34] V. Bonvin et al. “H0LiCOW – V. New COSMOGRAIL time delays of HE 0435–1223: H_0 to 3.8 per cent precision from strong lensing in a flat Λ CDM model”. In: *Mon. Not. Roy. Astron. Soc.* 465.4 (2017), pp. 4914–4930. DOI: [10.1093/mnras/stw3006](https://doi.org/10.1093/mnras/stw3006). arXiv: [1607.01790](https://arxiv.org/abs/1607.01790) [[astro-ph.CO](#)].
- [35] Tommaso Treu. “Strong Lensing by Galaxies”. In: *Annual Review of Astronomy and Astrophysics* 48.1 (Aug. 2010), pp. 87–125. ISSN: 1545-4282. DOI: [10.1146/annurev-astro-081309-130924](https://doi.org/10.1146/annurev-astro-081309-130924). URL: <http://dx.doi.org/10.1146/annurev-astro-081309-130924>.
- [36] P. Natarajan et al. “Strong Lensing by Galaxy Clusters”. In: *Space Science Reviews* 220.2 (Feb. 15, 2024), p. 19. ISSN: 1572-9672. DOI: [10.1007/s11214-024-01051-8](https://doi.org/10.1007/s11214-024-01051-8). URL: <https://doi.org/10.1007/s11214-024-01051-8>.
- [37] Rachel Mandelbaum. “Weak lensing for precision cosmology”. In: *Ann. Rev. Astron. Astrophys.* 56 (2018), pp. 393–433. DOI: [10.1146/annurev-astro-081817-051928](https://doi.org/10.1146/annurev-astro-081817-051928). arXiv: [1710.03235](https://arxiv.org/abs/1710.03235) [[astro-ph.CO](#)].
- [38] Liang Dai, Tejaswi Venumadhav, and Kris Sigurdson. “Effect of lensing magnification on the apparent distribution of black hole mergers”. In: *Physical Review D* 95.4 (Feb. 10, 2017). Publisher: American Physical Society, p. 044011. DOI: [10.1103/PhysRevD.95.044011](https://doi.org/10.1103/PhysRevD.95.044011). URL: <https://link.aps.org/doi/10.1103/PhysRevD.95.044011>.

- [39] Anuj Mishra et al. “Gravitational lensing of gravitational waves: effect of microlens population in lensing galaxies”. In: *Mon. Not. Roy. Astron. Soc.* 508.4 (2021), pp. 4869–4886. DOI: [10.1093/mnras/stab2875](https://doi.org/10.1093/mnras/stab2875). arXiv: [2102.03946](https://arxiv.org/abs/2102.03946) [astro-ph.CO].
- [40] Curt Cutler and Éanna E. Flanagan. “Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform?” In: *Phys. Rev. D* 49 (6 Mar. 1994), pp. 2658–2697. DOI: [10.1103/PhysRevD.49.2658](https://doi.org/10.1103/PhysRevD.49.2658). URL: <https://link.aps.org/doi/10.1103/PhysRevD.49.2658>.
- [41] Jose M. Diego, Tom Broadhurst, and George Smoot. “Evidence for lensing of gravitational waves from LIGO-Virgo data”. In: *Phys. Rev. D* 104.10 (2021), p. 103529. DOI: [10.1103/PhysRevD.104.103529](https://doi.org/10.1103/PhysRevD.104.103529). arXiv: [2106.06545](https://arxiv.org/abs/2106.06545) [gr-qc].
- [42] K. Haris et al. *Identifying strongly lensed gravitational wave signals from binary black hole mergers*. July 18, 2018. DOI: [10.48550/arXiv.1807.07062](https://doi.org/10.48550/arXiv.1807.07062). arXiv: [1807.07062](https://arxiv.org/abs/1807.07062)[gr-qc]. URL: <http://arxiv.org/abs/1807.07062> (visited on 06/21/2023).
- [43] Jose María Ezquiaga et al. “Phase effects from strong gravitational lensing of gravitational waves”. In: *Phys. Rev. D* 103.6 (2021), p. 064047. DOI: [10.1103/PhysRevD.103.064047](https://doi.org/10.1103/PhysRevD.103.064047). arXiv: [2008.12814](https://arxiv.org/abs/2008.12814) [gr-qc].
- [44] Liang Dai et al. “Search for Lensed Gravitational Waves Including Morse Phase Information: An Intriguing Candidate in O2”. In: (July 2020). arXiv: [2007.12709](https://arxiv.org/abs/2007.12709) [astro-ph.HE].
- [45] R. Abbott et al. “Search for gravitational-lensing signatures in the full third observing run of the LIGO-Virgo network”. In: (Apr. 2023). arXiv: [2304.08393](https://arxiv.org/abs/2304.08393) [gr-qc].
- [46] Justin Janquart et al. “Follow-up analyses to the O3 LIGO–Virgo–KAGRA lensing searches”. In: *Mon. Not. Roy. Astron. Soc.* 526.3 (2023), pp. 3832–3860. DOI: [10.1093/mnras/stad2909](https://doi.org/10.1093/mnras/stad2909). arXiv: [2306.03827](https://arxiv.org/abs/2306.03827) [gr-qc].
- [47] A. Renske A. C. Wierda et al. “Beyond the Detector Horizon: Forecasting Gravitational-Wave Strong Lensing”. In: *Astrophys. J.* 921.2 (2021), p. 154. DOI: [10.3847/1538-4357/ac1bb4](https://doi.org/10.3847/1538-4357/ac1bb4). arXiv: [2106.06303](https://arxiv.org/abs/2106.06303) [astro-ph.HE].
- [48] Ken K. Y. Ng et al. “Precise LIGO Lensing Rate Predictions for Binary Black Holes”. In: *Phys. Rev. D* 97.2 (2018), p. 023012. DOI: [10.1103/PhysRevD.97.023012](https://doi.org/10.1103/PhysRevD.97.023012). arXiv: [1703.06319](https://arxiv.org/abs/1703.06319) [astro-ph.CO].
- [49] Souvik Jana et al. “Cosmography Using Strongly Lensed Gravitational Waves from Binary Black Holes”. In: *Phys. Rev. Lett.* 130 (26 June 2023), p. 261401. DOI: [10.1103/PhysRevLett.130.261401](https://doi.org/10.1103/PhysRevLett.130.261401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.130.261401>.
- [50] Otto A. Hannuksela et al. “Localizing merging black holes with sub-arcsecond precision using gravitational-wave lensing”. In: *Mon. Not. Roy. Astron. Soc.* 498.3 (2020), pp. 3395–3402. DOI: [10.1093/mnras/staa2577](https://doi.org/10.1093/mnras/staa2577). arXiv: [2004.13811](https://arxiv.org/abs/2004.13811) [astro-ph.HE].

- [51] Srashti Goyal et al. “Probing lens-induced gravitational-wave birefringence as a test of general relativity”. In: *Phys. Rev. D* 108.2 (2023), p. 024052. DOI: [10.1103/PhysRevD.108.024052](https://doi.org/10.1103/PhysRevD.108.024052). arXiv: [2301.04826](https://arxiv.org/abs/2301.04826) [gr-qc].
- [52] Sourabh Magare et al. “Gear Up for the Action Replay: Leveraging Lensing for Enhanced Gravitational-wave Early Warning”. In: *Astrophys. J. Lett.* 955.2 (2023), p. L31. DOI: [10.3847/2041-8213/acf668](https://doi.org/10.3847/2041-8213/acf668). arXiv: [2302.02916](https://arxiv.org/abs/2302.02916) [astro-ph.HE].
- [53] Mukesh Kumar Singh et al. “Déjà-vu et Déjà-entendu: Associating fast radio bursts with compact binary mergers via gravitational lensing”. In: (Apr. 2023). _eprint: 2304.02879.
- [54] S. Weinberg. “Apparent luminosities in a locally inhomogeneous universe.” In: *The Astrophysical Journal Letters* 208 (Aug. 1976), pp. L1–L3. DOI: [10.1086/182216](https://doi.org/10.1086/182216).
- [55] Nick Kaiser and John A. Peacock. “On the bias of the distance–redshift relation from gravitational lensing”. In: *Mon. Not. Roy. Astron. Soc.* 455.4 (2016), pp. 4518–4547. DOI: [10.1093/mnras/stv2585](https://doi.org/10.1093/mnras/stv2585). arXiv: [1503.08506](https://arxiv.org/abs/1503.08506) [astro-ph.CO].
- [56] Peter Schneider, Jürgen Ehlers, and Emilio E. Falco. *Gravitational Lenses*. 1992. DOI: [10.1007/978-3-662-03758-4](https://doi.org/10.1007/978-3-662-03758-4).
- [57] Eric Thrane and Colm Talbot. “An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models”. In: *Publications of the Astronomical Society of Australia* 36 (2019), e010. ISSN: 1323-3580, 1448-6083. DOI: [10.1017/pasa.2019.2](https://doi.org/10.1017/pasa.2019.2). arXiv: [1809.02293](https://arxiv.org/abs/1809.02293)[astro-ph]. URL: <http://arxiv.org/abs/1809.02293> (visited on 06/17/2023).
- [58] Nelson Christensen and Renate Meyer. “Parameter estimation with gravitational waves”. In: *Rev. Mod. Phys.* 94.2 (2022), p. 025001. DOI: [10.1103/RevModPhys.94.025001](https://doi.org/10.1103/RevModPhys.94.025001). arXiv: [2204.04449](https://arxiv.org/abs/2204.04449) [gr-qc].
- [59] Javier Roulet and Tejaswi Venumadhav. “Inferring Binary Properties from Gravitational Wave Signals”. In: (Feb. 2024). DOI: [10.1146/annurev-nucl-121423-100725](https://doi.org/10.1146/annurev-nucl-121423-100725). arXiv: [2402.11439](https://arxiv.org/abs/2402.11439) [gr-qc].
- [60] Joshua S. Speagle. “A Conceptual Introduction to Markov Chain Monte Carlo Methods”. In: (Sept. 2019). arXiv: [1909.12313](https://arxiv.org/abs/1909.12313) [stat.OT].
- [61] John Skilling. “Nested Sampling”. In: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Ed. by Rainer Fischer, Roland Preuss, and Udo Von Toussaint. Vol. 735. American Institute of Physics Conference Series. Nov. 2004, pp. 395–405. DOI: [10.1063/1.1835238](https://doi.org/10.1063/1.1835238).
- [62] C. M. Biwer et al. “PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals”. In: *Publ. Astron. Soc. Pac.* 131.996 (2019), p. 024503. DOI: [10.1088/1538-3873/aaef0b](https://doi.org/10.1088/1538-3873/aaef0b). arXiv: [1807.10312](https://arxiv.org/abs/1807.10312) [astro-ph.IM].
- [63] Gregory Ashton et al. “BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy”. In: *Astrophys. J. Suppl.* 241.2 (2019), p. 27. DOI: [10.3847/1538-4365/ab06fc](https://doi.org/10.3847/1538-4365/ab06fc). arXiv: [1811.02042](https://arxiv.org/abs/1811.02042) [astro-ph.IM].

- [64] I. M. Romero-Shaw et al. “Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue”. In: *Mon. Not. Roy. Astron. Soc.* 499.3 (2020), pp. 3295–3319. DOI: [10.1093/mnras/staa2850](https://doi.org/10.1093/mnras/staa2850). arXiv: [2006.00714](https://arxiv.org/abs/2006.00714) [[astro-ph.IM](#)].
- [65] J. Veitch et al. “Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library”. In: *Phys. Rev. D* 91.4 (2015), p. 042003. DOI: [10.1103/PhysRevD.91.042003](https://doi.org/10.1103/PhysRevD.91.042003). arXiv: [1409.7215](https://arxiv.org/abs/1409.7215) [[gr-qc](#)].
- [66] Salvatore Vitale et al. “Inferring the properties of a population of compact binaries in presence of selection effects”. In: 2021, pp. 1–60. DOI: [10.1007/978-981-15-4702-7_45-1](https://doi.org/10.1007/978-981-15-4702-7_45-1). arXiv: [2007.05579](https://arxiv.org/abs/2007.05579) [[astro-ph, physics:gr-qc](#)]. URL: <http://arxiv.org/abs/2007.05579> (visited on 06/17/2023).
- [67] Piero Madau and Mark Dickinson. “Cosmic Star Formation History”. In: *Ann. Rev. Astron. Astrophys.* 52 (2014), pp. 415–486. DOI: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615). arXiv: [1403.0007](https://arxiv.org/abs/1403.0007) [[astro-ph.CO](#)].
- [68] The LIGO Scientific Collaboration et al. “Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo”. In: *The Astrophysical Journal* 882.2 (Sept. 11, 2019), p. L24. ISSN: 2041-8213. DOI: [10.3847/2041-8213/ab3800](https://doi.org/10.3847/2041-8213/ab3800). arXiv: [1811.12940](https://arxiv.org/abs/1811.12940) [[astro-ph](#)]. URL: <http://arxiv.org/abs/1811.12940> (visited on 12/20/2023).
- [69] Will M. Farr. “Accuracy Requirements for Empirically Measured Selection Functions”. In: *Research Notes of the AAS* 3.5 (May 2019). Publisher: The American Astronomical Society, p. 66. ISSN: 2515-5172. DOI: [10.3847/2515-5172/ab1d5f](https://doi.org/10.3847/2515-5172/ab1d5f). URL: <https://dx.doi.org/10.3847/2515-5172/ab1d5f> (visited on 07/30/2023).
- [70] David W. Hogg. *Distance measures in cosmology*. Dec. 15, 2000. arXiv: [astro-ph/9905116](https://arxiv.org/abs/astro-ph/9905116). URL: <http://arxiv.org/abs/astro-ph/9905116> (visited on 09/18/2023).
- [71] Reed Essick and Maya Fishbach. *DAGnabbit! Ensuring Consistency between Noise and Detection in Hierarchical Bayesian Inference*. Oct. 3, 2023. arXiv: [2310.02017](https://arxiv.org/abs/2310.02017) [[astro-ph, physics:gr-qc](#)]. URL: <http://arxiv.org/abs/2310.02017> (visited on 01/08/2024).
- [72] Jose María Ezquiaga and Daniel E. Holz. “Spectral sirens: cosmology from the full mass distribution of compact binaries”. In: *Physical Review Letters* 129.6 (Aug. 3, 2022), p. 061102. ISSN: 0031-9007, 1079-7114. DOI: [10.1103/PhysRevLett.129.061102](https://doi.org/10.1103/PhysRevLett.129.061102). arXiv: [2202.08240](https://arxiv.org/abs/2202.08240) [[astro-ph, physics:gr-qc](#)]. URL: <http://arxiv.org/abs/2202.08240> (visited on 08/24/2023).