

# **BRIGHT GALAXIES AS REDSHIFT TRACERS FOR DARK SIREN COSMOLOGY**

**Muhammad Khazaifa Naveed**

Student number: 02004453

Supervisor: Prof. Dr. Archisman Ghosh

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of  
Master of Science in Physics and Astronomy

Academic year: 2024 - 2025





Department of Physics and Astronomy  
Ghent Gravity Group

## Bright Galaxies as Redshift Tracers for Dark Siren Cosmology

Muhammad Khuzaifa Naveed

*Supervisor* Prof. Dr. Archisman Ghosh

AY 2024-2025

**Muhammad Khuzaifa Naveed**

*Bright Galaxies as Redshift Tracers for Dark Siren Cosmology*

Master Thesis, AY 2024-2025

Supervisors: Prof. Dr. Archisman Ghosh

**Ghent University**

*Ghent Gravity Group*

Department of Physics and Astronomy

Proeftuinstraat 86

9000 Ghent

# **Declaration**

You can put your declaration here, to declare that you have completed your work solely and only with the help of the references you mentioned.

*Ghent, June, 2025*

---

Muhammad Khuzaifa Naveed



# Acknowledgement

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



# **Abstract**

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# **Abstract (different language)**

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



# Contents

<b>List of Figures</b>	xiii
<b>List of Tables</b>	xv
<b>List of Abbreviations and Symbols</b>	xvii
<b>1 Introduction</b>	1
1.1 The Hubble Tension . . . . .	1
1.2 Standard Siren Cosmology . . . . .	1
1.3 Thesis Objectives . . . . .	1
<b>2 Gravitational-Wave Cosmology</b>	3
2.1 Standard Siren Cosmology . . . . .	3
2.2 Bright vs. Dark Sirens . . . . .	5
2.3 State of the Art Dark Siren Methods . . . . .	6
2.3.1 Redshift Inference Methods . . . . .	7
2.3.2 Galaxy Weighting Choices . . . . .	8
2.4 Refining the Redshift Prior: Brightest Galaxies as Tracers . . . . .	9
<b>3 Data</b>	13
3.1 Gravitaional-Wave (GW) Event Data . . . . .	13
3.1.1 The GWTC-3 Catalog . . . . .	13
3.1.2 Event Selection Criteria . . . . .	13
3.1.3 Distance Posteriors and Sky Localization . . . . .	14
3.1.4 Assumptions About the Source Population . . . . .	14
3.2 electromagnetic (EM) Galaxy Catalogs . . . . .	15
3.2.1 The GLADE+ Catalog . . . . .	15
3.2.2 Catalog Completeness . . . . .	15
3.2.3 Redshift Uncertainty Modeling . . . . .	16
3.2.4 <i>K</i> -band Luminosity as Host Probability Proxy . . . . .	16
<b>4 GLADE+ Bright Galaxies as Redshift Tracers</b>	17
4.1 Bright Galaxy Subsets . . . . .	17
4.1.1 Trade-Offs and Limitations . . . . .	18
4.2 The <code>gwcosmo</code> Pipeline . . . . .	20

4.2.1	Bayesian Framework: <b>TO BE UPDATED</b>	20
4.2.2	Key Features of <code>gwcosmo</code>	21
4.2.3	Implementation and Applications	21
4.3	line-of-sight (LOS) Redshift Prior	22
4.3.1	LOS Redshift Prior Construction	22
4.4	Results	25
4.4.1	LOS Redshift Prior	25
4.4.2	Hubble Constant Posterior	25
4.4.3	Cost–benefit trade-off	27
<b>5</b>	<b>Mock Data Challenge</b>	<b>29</b>
5.1	BUZZARD Mock Catalog	29
5.2	Results	29
<b>6</b>	<b>Discussion</b>	<b>31</b>
<b>7</b>	<b>Conclusion</b>	<b>33</b>
<b>Bibliography</b>		<b>35</b>

# List of Figures

2.1	The $B_J$ -band luminosity function and the $K$ -band luminosity function for the GLADE+ catalog. . . . .	9
2.2	Spatial distribution of galaxies from the GLADE+ galaxy catalog and its subsets. . . . .	11
4.1	Normalized redshift distributions $p(z)$ for the GLADE+ galaxy catalogue and the brightest percentiles. . . . .	18
4.2	LOS redshift prior for GW170809 for different brightness-ranked GLADE subsets . . . . .	26
4.3	$H_0$ posterior distributions for GW170809 using full and brightness-cut galaxy catalogs. . . . .	27
4.4	Cumulative $H_0$ posterior from the selected dark siren events using full GLADE+ and the different subsets. . . . .	28



# List of Tables

3.1	Priors on merger rate shape parameters. . . . .	14
3.2	Priors on the mass model parameters. . . . .	15
4.1	$H_0$ MAP values with 68% confidence ranges, alongside the maximum magnitude limits for GLADE+ and the different subsets. . . . .	28



# List of Abbreviations and Symbols

<b>GW</b>	Gravitaional-Wave . . . . .	xi
<b>CBC</b>	compact binary oalescence . . . . .	3
<b>LVK</b>	LIGO, Virgo, KAGRA . . . . .	4
<b>BNS</b>	binary neutron star . . . . .	5
<b>BBH</b>	binary black hole . . . . .	5
<b>NSBH</b>	neutron star black hole . . . . .	13
<b>EM</b>	electromagnetic . . . . .	xi
<b>LOS</b>	line-of-sight . . . . .	xii
<b>SNR</b>	signal-to-noise ratio . . . . .	13
<b>SNe</b>	supernovae . . . . .	3
<b>GWTC</b>	Gravitational-Wave Transient Catalog . . . . .	13
<b>BCG</b>	Brightest Cluster Galaxy . . . . .	10
<b>MDC</b>	Mock Data Challenge . . . . .	19
<b>DES</b>	Dark Energy Survey . . . . .	19
<b>KDE</b>	kernel density estimation . . . . .	21



# Introduction

## 1.1 The Hubble Tension

- Explain traditional methods of measuring  $H_0$  (e.g. SNe, CMB, BAO)
- Discuss the discrepancy between local measurements (e.g. SNe) and CMB measurements
- Introduce the concept of the Hubble tension and its significance in cosmology

## 1.2 Standard Siren Cosmology

- Explain the concept of standard sirens and their role in cosmology
- Discuss the use of gravitational waves (GWs) as standard sirens
- Introduce the idea of using electromagnetic (EM) counterparts to improve distance measurements
- Discuss the advantages and limitations of standard sirens compared to traditional methods
- Shorter explanation compared to chapter 2

## 1.3 Thesis Objectives

- Explain the core idea of the thesis: using brightness-ranked galaxy catalogs to improve the precision of  $H_0$  measurements
- Discuss the motivation behind using brightness-ranked catalogs
- BUZZARD



# Gravitational-Wave Cosmology

## 2.1 Standard Siren Cosmology

The detection of GW from compact binary coalescences (CBCs) have opened a transformative avenue in modern cosmology. These events act as *standard sirens*, the gravitational analog of standard candles, as the luminosity distance ( $d_L$ ) to a GW source is directly encoded in the strain amplitude and frequency evolution of the GW. The amplitude, corrected for antenna pattern and inclination, provides a direct, calibration-free distance measurement under the assumption of general relativity.

From Maggiore (2007), the strain measured by a GW detector can be expressed as:

$$h(t) = \frac{4}{d_L} \left( \frac{G\mathcal{M}_z}{c^2} \right)^{5/3} (\pi f(t))^{2/3} F(\iota, \psi, \theta, \phi) \quad (2.1)$$

where:

- $h(t)$  is the strain measured at time  $t$ ,
- $d_L$  is the luminosity distance to the source,
- $\mathcal{M}_z = (1+z)\mathcal{M}$  is the redshifted chirp mass,
- $f(t)$  is the instantaneous GW frequency,
- $F(\iota, \psi, \theta, \phi)$  is a function that captures the detector response depending on inclination angle  $\iota$ , polarization angle  $\psi$ , and sky position  $(\theta, \phi)$ .

The amplitude scaling with  $1/d_L$  makes it possible to extract the luminosity distance directly from the signal, once the source's intrinsic properties and orientation are marginalized over. Furthermore, the  $1/d_L$  scaling, in comparison to the  $1/d_L^2$  scaling of EM signals, means that GW signals allow observations of sources at much larger distances, making them ideal for cosmological applications.

The direct self-calibrated measurement of the luminosity distance from the GW signal is a key advantage of standard sirens over traditional EM standard candles. This is in stark contrast to traditional EM standard candles, such as supernovae (SNe), where the distance is inferred from the observed flux and requires a calibration step to

account for the intrinsic brightness of the source. The GW signal is also less affected by the intergalactic medium, as it is not subject to scattering or absorption like EM signals. This allows for a more direct measurement of the distance to the source not affected by dust extinction or other astrophysical uncertainties that plague EM observations. This makes GW standard sirens a powerful tool for cosmology.

The concept of standard sirens was first introduced by Bernard F. Schutz, who noted that if the redshift of a GW source can be measured, one could use GW events to trace the expansion history of the universe, in particular an independent measurement of the Hubble constant ( $H_0$ ) (Schutz, 1986). This is due to the relation between the luminosity distance, redshift and the Hubble constant, in a flat  $\Lambda$ CDM cosmology, being:

$$d_L = \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{(1+z')^3 \Omega_m + \Omega_\Lambda}} \quad (2.2)$$

Unlike traditional EM methods, which require cross-calibration across multiple rungs of the distance ladder (parallax, Cepheids, SNe Ia) with each step introducing uncertainties that compound, the standard siren approach provides a direct cosmological probe. This reduces systematic uncertainties and provides an independent check on other  $H_0$  measurement techniques, possibly providing a solution to the Hubble tension, the current discrepancy between local and global measurements of the Hubble constant (Riess et al., 2019; Aghanim et al., 2020).

However, a key limitation is that while GW detectors provide a precise measurement of  $d_L$ , they do not directly measure the redshift. This necessitates an independent redshift measurement to place the source on the Hubble diagram. For *bright sirens*, this redshift is obtained from the host galaxy identified via the electromagnetic counterpart. In contrast, for *dark sirens*, the redshift is inferred statistically by cross-referencing the GW localization volume with galaxy catalogs or by leveraging population-based methods, as discussed in the following sections.

This forms the basis for *standard siren cosmology*, where GW sources are used as cosmic rulers. Over the past decade, this idea has transitioned from theoretical speculation to experimental reality, primarily through observations made by the LIGO, Virgo, KAGRA (LVK) collaboration. The first GW event, **GW150914**, was detected in 2015, marking the beginning of a new era in astrophysics and cosmology (Abbott et al., 2016). Since then, the LVK collaboration has detected numerous GW events, including binary black hole mergers, binary neutron star mergers, and neutron star-black hole mergers. These observations have provided valuable insights into the nature of gravity, the formation of compact objects, and the expansion history of the universe.

## 2.2 Bright vs. Dark Sirens

Standard sirens fall into two broad categories: *bright sirens* and *dark sirens*, depending on whether an electromagnetic counterpart is detected.

Bright sirens are rare but powerful. A prime example is **GW170817**, a binary neutron star (BNS) merger detected by LIGO-Virgo in 2017, which was accompanied by a short gamma-ray burst and subsequent kilonova (Abbott et al., 2017). This multi-messenger event enabled the unambiguous identification of its *host galaxy*, NGC 4993, providing both distance (from GW) and redshift (from optical spectroscopy) measurements. The resulting Hubble constant measurement was a major milestone demonstrating that GW observations could offer an independent and competitive cosmological probe (Abbott et al., 2017).

Such bright sirens provide a straightforward route to cosmological inference, using independent distance and redshift measurements. However, bright sirens require specific astrophysical conditions: the emission of EM signals strong enough to be detected, accurate sky localization, and timely follow-up by optical telescopes. Such conditions are only met for a small fraction of CBC events, especially those involving neutron stars. Majority of the observed mergers, especially binary black holes (BBHs) do not produce detectable EM counterparts, and are thus classified as dark sirens.

In the absence of a direct redshift measurement, dark sirens require a *statistical approach*. This involves cross-matching the sky localization and distance posterior from the GW event with a galaxy catalog covering the relevant region. The redshift information of a number of candidate galaxies is then used to statistically infer the likely redshift distribution of the source. This is done by constructing a *LOS redshift prior* for each GW event, which is then used to infer the Hubble constant. The LOS redshift prior is constructed by taking into account the galaxy number density and redshift distribution within the localization volume of the GW event. The GW localization volume is typically much larger than the volume of a galaxy, leading to a large number of galaxies that could potentially host the GW event. This results in a large number of potential redshift measurements, which can be used to construct a more accurate LOS redshift prior.

The statistical nature of dark sirens allows for the inclusion of a larger number of events, as it does not rely on the detection of an EM counterpart. This is particularly important for BBH events, which are more common and have a higher detection rate than BNS events. The LVK collaboration has detected numerous dark siren events, which have been used to constrain the Hubble constant and test cosmological models.

As the number of BBH detections increases with each observing run, dark sirens will dominate the future of standard siren cosmology. However, this statistical method introduces new complexities and depends heavily on the quality and completeness of the galaxy catalog used, which plays a pivotal role in the reliability of the inferred Hubble constant. The incompleteness of the galaxy catalog can lead to biases in the inferred redshift distribution, which can affect the accuracy of the Hubble constant measurement. This is particularly important for dark sirens, as they rely on the statistical association of GW events with galaxies in the catalog. State of the art dark siren methods have been developed to mitigate these issues, but they still rely on the quality and completeness of the galaxy catalog used. The next section discusses the current state of the art in dark siren cosmology, including the challenges and limitations of existing methods.

## 2.3 State of the Art Dark Siren Methods

The statistical framework for dark siren cosmology has evolved rapidly. Early implementations relied on basic overlap between GW localization volumes and precompiled galaxy catalogs. The first generation of tools, including `gwcosmo` (Gray et al., 2020; Gray et al., 2022; Gray et al., 2023) and `icarogw` (Mastrogiovanni et al., 2021; Mastrogiovanni et al., 2024), introduced a more rigorous Bayesian framework: for each GW event, a *LOS redshift prior* is constructed using galaxy number counts and redshifts, weighted by host probability, typically modeled using stellar mass proxies such as  $K$ -band luminosity.

One major challenge in this process is that current galaxy catalogs, such as **GLADE+** (Dálya et al., 2022), are incomplete beyond low redshift (typically  $z > 0.2 - 0.3$ ). This incompleteness leads to a nontrivial *out-of-catalog* term in the redshift prior, which must be modeled analytically using a *Schechter luminosity function*. This function estimates the number density of missing faint galaxies, often assuming fixed parameters (e.g.,  $M_*$ ,  $\alpha$ ) calibrated from deep surveys. If this modeling is inaccurate, the resulting posterior for  $H_0$  can be biased or artificially broadened.

Despite these challenges, dark siren methods have been successfully applied to increasingly large datasets. The LVK collaboration has published joint analyses using CBC events from the **O1**, **O2**, and **O3** observing runs. For instance, Abbott et al., 2021a; Abbott et al., 2023a presented cumulative constraints on  $H_0$  using numerous CBC events, demonstrating convergence toward a 10–15% precision regime, albeit still limited by catalog systematics.

These analyses highlight the potential of dark sirens to provide competitive cosmological constraints, especially as galaxy catalogs improve and detection rates increase. Future observing runs, such as O4 and O5, are expected to significantly enhance the statistical power of dark siren cosmology, potentially reducing uncertainties in  $H_0$ , providing better constraints on cosmological model, and testing the validity of general relativity on cosmological scales. The next generation of GW detectors, such as the **Einstein Telescope** and **Cosmic Explorer**, will further enhance the sensitivity and detection rates of GW events, opening up new avenues for cosmological exploration.

### 2.3.1 Redshift Inference Methods

In dark siren cosmology, two principal strategies exist for inferring the redshift of gravitational-wave sources: the *catalog-based method*(CITATION NEEDED) and the *population-based method* (Ezquiaga and Holz, 2022). The catalog method, discussed earlier, statistically associates GW sources with galaxies from a survey by constructing a redshift prior along each line of sight, derived from galaxy positions and redshifts within the GW localization volume. This approach captures the clustering of galaxies and allows for redshift inference when direct EM counterparts are absent, but suffers from incompleteness at higher redshifts and spatial variation in survey depth(CITATION NEEDED). Conversely, the population method, often referred to as the *spectral siren approach*, leverages features in the observed distribution of source-frame binary parameters, especially masses, which are redshifted in the detector frame. If the intrinsic mass distribution contains recognizable structure (e.g., a cutoff or peak), its displacement due to cosmic redshift can break the mass-redshift degeneracy, enabling cosmological inference independent of galaxy surveys (Ezquiaga and Holz, 2022). However, this approach is sensitive to modeling assumptions about the underlying binary population.

Recent work has sought to improve robustness by combining dark siren methods with *spectral sirens*, extracting redshift information from the observed mass distribution of GW sources under population synthesis assumptions (Ezquiaga and Holz, 2022). These hybrid approaches aim to reduce dependence on incomplete catalogs and mitigate selection biases while extracting maximal cosmological information from current GW detections. Recently, both `gwcsmo` and `icarogw` have been updated to support such joint inference schemes (Gray et al., 2023; Mastrogiovanni et al., 2024).

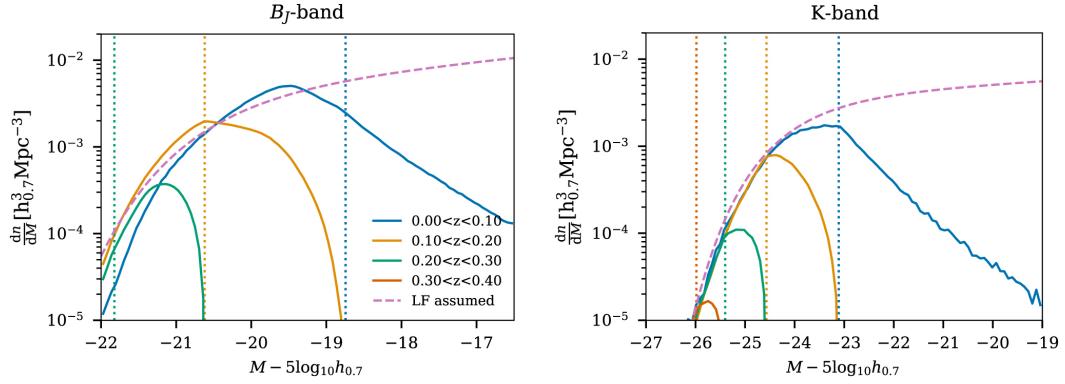
### 2.3.2 Galaxy Weighting Choices

In the construction of the LOS redshift prior, each host galaxy is weighted by the probability of hosting a GW event. This probability is typically modeled as a function of the galaxy's luminosity, and relies on the assumption that the merger host probability of a galaxy is related to its luminosity in a specific band. The choice of band remains uncertain, with different bands correlating with different aspects. The blue band, for instance, correlates with star formation rate and is therefore appropriate for mergers with short time delays (coalescence shortly after formation), while the near-infrared band correlates with the total luminous mass and is thus appropriate for mergers with longer time delays (mergers in older, more evolved galaxies). It is not known how exactly galaxy color is correlated with the merger rate.

For modelling the out-of-catalog contribution, a Schechter luminosity function is assumed, which describes the distribution of galaxy luminosities in a given band. Wrong assumptions on the Schechter parameters, or incorrect description of selection biases can lead to biases in the inferred Hubble constant. The key assumption in constructing the out-of-catalog contribution is that the galaxy catalog is magnitude limited, i.e. galaxies are not detected only because they are too faint. If other selection biases (based on e.g., colors or spectral features) were present, the out-of-catalog contribution would be biased, leading to an incorrect estimate of the redshift prior (Abbott et al., 2023a).

If the assumption of a magnitude-limited catalog is valid, then the luminosity function of the galaxy catalog would match the assumed Schechter function at its bright end, and then start to decrease as it reaches the corresponding absolute magnitude threshold. As can be seen in Figure 2.1, this is indeed the case for the  $K$ -band luminosity function, which is well described by the assumed Schechter function. The  $B_J$ -band luminosity function, on the other hand, is not well described by the assumed luminosity function, and the assumption of a magnitude-limited catalog is not valid as there seem to be some additional missing galaxies at low redshift. This could lead to an incorrect estimate of the out-of-catalog contribution, and therefore an incorrect estimate of the redshift prior (Abbott et al., 2023a).

For this reason, traditionally, the  $K$ -band luminosity has been used for galaxy weighting in GW cosmology inference pipelines. This is also due to the fact that the  $K$ -band luminosity is better correlated with the stellar mass of the galaxy, making it a more reliable tracer of the underlying matter distribution (Strazzullo et al., 2006; Sureshkumar et al., 2021; Abbott et al., 2023a).



**Fig. 2.1:** The  $B_J$ -band luminosity function (left) and the  $K$ -band luminosity function (right) for the GLADE+ catalog. The solid lines show the luminosity function of the catalog, while the dashed lines shows the assumed Schechter function. The vertical dashed line indicates the median absolute magnitude threshold for galaxy detection. The bright end of the  $K$ -band luminosity seems to match the assumed Schechter function, while the  $B_J$ -band luminosity function does not. This indicates that the assumption of a magnitude-limited catalog is not valid for the  $B_J$ -band luminosity function (Abbott et al., 2023a).

However, this choice is not without its limitations. The  $K$ -band luminosity is not always available for all galaxies in the catalog, leading to potential biases in the selection of galaxies used to construct the redshift prior. For example, the  $K$ -band luminosity is only available for about 1 million galaxies in the GLADE+ catalog, which is a small fraction of the total number of galaxies in the catalog. This can lead to uncertainties in the redshift prior, which can in turn affect the inferred value of the Hubble constant. However, this effect should be relatively small, as we are interested in the overall matter distribution, and the  $K$ -band luminosity is a good tracer of the underlying matter distribution. This also forms the basis for this thesis, as we will be using the brightest galaxies in the  $K$ -band to trace the underlying matter distribution for improved  $H_0$  inference as discussed in the next section.

## 2.4 Refining the Redshift Prior: Brightest Galaxies as Tracers

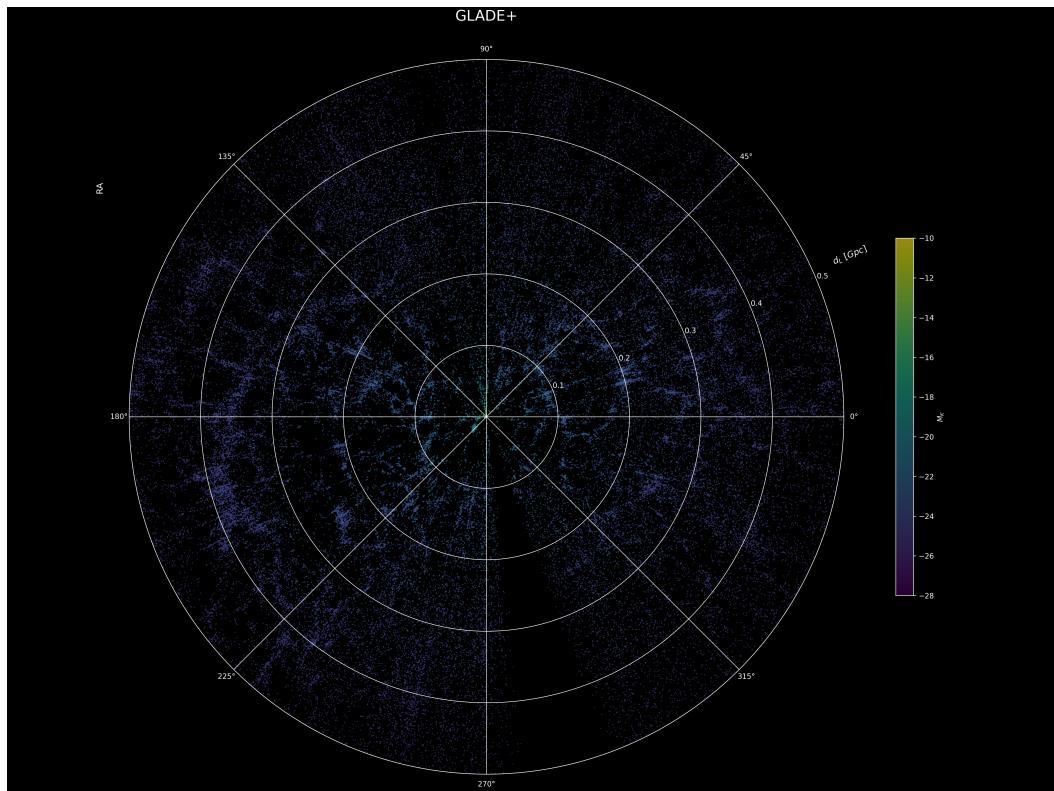
The current state of the art in dark siren cosmology relies heavily on the quality and completeness of the galaxy catalog used to construct the LOS redshift prior. The incompleteness of the galaxy catalog can lead to biases in the inferred redshift distribution, which can affect the accuracy of the Hubble constant measurement. State of the art dark siren methods have been developed to mitigate these issues, but they still rely on the quality and completeness of the galaxy catalog used.

One strategy to mitigate this limitation is by refining the construction of the redshift prior by focusing on the *brightest galaxies*, which are more likely to be catalogued and potentially better tracers of the underlying matter distribution. This would effectively make the catalog more complete and increase precision without introducing significant bias. The rationale is that the brightest galaxies are more likely to be associated with massive halos, which are more likely to host detectable GW events. The brightest galaxies are also more likely to be catalogued, as they are easier to detect and measure. These galaxies would also crudely trace the underlying matter distribution, as they are more likely to be associated with massive halos. This would be particularly useful for dark sirens, as they rely on the statistical association of GW events with galaxies in the catalog. By replacing the redshift of the true host galaxy with the redshift of a nearby bright galaxy in the catalog, the small error incurred would be insignificant compared to the overall uncertainty in the redshift prior and the luminosity distance measurement.

This approach is similar to the *Brightest Cluster Galaxy (BCG)* method used in traditional cosmology, where the brightest galaxy in a cluster is used as a standard candle. The BCG method has been shown to be effective in reducing the scatter in distance measurements and improving the precision of cosmological constraints (Lauer et al., 2014). By applying a similar approach to dark sirens, one can potentially improve the precision of the inferred Hubble constant without introducing significant bias.

To implement this approach, we define a subset of the galaxy catalog, GLADEPXX, which includes only the top  $XX\%$  of galaxies ranked by  $K$ -band luminosity, as the  $K$ -band luminosity is better associated with the mass of the galaxies (Strazzullo et al., 2006; Sureshkumar et al., 2021). The construction of the LOS redshift prior is then modified to account for this restricted catalog. Specifically, the Schechter luminosity function is adjusted to reflect the brighter subset, making the out-of-catalog contribution smaller, effectively making the catalog more complete.

This modified prior is then used in the Bayesian framework to infer the Hubble constant. The next chapter provides a detailed description of the methodology, including the selection criteria for GLADEPXX, the modeling of the out-of-catalog part and the validation of this approach using mock data.



**Fig. 2.2:** Spatial distribution of galaxies from the GLADE+ catalog, employed for dark standard siren cosmology. The plot shows an 8-degree slice in declination, centered at  $0^\circ$ , with the radial coordinate representing luminosity distance  $d_L$  (in Gpc) and the angular coordinates being right ascension (RA). This shows how the bright galaxies trace the large-scale structure of the universe. [PLACEHOLDER FIGURE: place figures from P100-P10]



# Data

The foundation of any standard siren cosmological analysis lies in the quality, completeness, and coverage of the data. In this thesis, we rely on two primary data sources: GW observations from the LVK detectors, and galaxy catalogs providing EM redshift information. These datasets are used in tandem to statistically construct a redshift prior for each GW event and ultimately to constrain the Hubble constant  $H_0$ .

## 3.1 GW Event Data

### 3.1.1 The GWTC-3 Catalog

The GW data used in this study are obtained from the *Gravitational-Wave Transient Catalog (GWTC)* (Abbott et al., 2019; Abbott et al., 2021b; Abbott et al., 2024; Abbott et al., 2023b). This catalog comprises all CBC events detected during the three observing runs (O1, O2, O3) of the Advanced LIGO and Virgo detectors. GWTC includes over 93 confident detections, primarily BBH mergers, with a few neutron star black hole (NSBH) and BNS systems.

### 3.1.2 Event Selection Criteria

To focus our analysis on dark siren cosmology, we restrict our sample to events that do not have any confirmed electromagnetic counterpart (i.e., no associated kilonova or GRB), have a network signal-to-noise ratio (SNR) exceeding 11, have somewhat good sky localization, and are accompanied by publicly available posterior samples for distance and sky position.

These criteria are designed to ensure both the statistical robustness of the inference and compatibility with existing galaxy catalogs. After applying these cuts, a subset of 47 events, with 42 BBHs, 3 NSBHs, and 2 BNSs, is retained for cosmological analysis.

### 3.1.3 Distance Posteriors and Sky Localization

Each event is characterized by a posterior distribution over the three-dimensional localization volume, typically encoded using the HEALPix format. The posterior provides the probability density  $p(d_L, \alpha, \delta)$ , where  $d_L$  is the luminosity distance and  $(\alpha, \delta)$  denote right ascension and declination.

These distance posteriors are crucial for constructing the LOS redshift prior when cross-matched with galaxy catalogs. Events with poor localization or multimodal distance distributions contribute more uncertainty to the inferred  $H_0$ , emphasizing the importance of event quality.

### 3.1.4 Assumptions About the Source Population

Although this thesis primarily focuses on statistical redshift modeling, it is important to acknowledge that the cosmological inference also depends weakly on assumptions about the source population. For instance, `gwcosmo` allows one to specify priors on:

- Merger rate evolution with redshift, typically modeled as  $R(z) \propto (1+z)^\kappa$
- Mass distributions of the binaries

For our analysis we assume that CBCs follow a Madau-Dickinson merger redshift evolution model (Madau and Dickinson, 2014):

$$R(z) = R_0(1+z)^\gamma \frac{1 + (1+z_p)^{-(\gamma+\kappa)}}{1 + \left(\frac{1+z}{1+z_p}\right)^{\gamma+\kappa}} \quad (3.1)$$

Here  $R_0$  is the local merger rate,  $\kappa$  the high-z slope,  $\gamma$  the low-z slope and  $z_p$  the break-point. All the assumed priors are given in Table 3.1. **TODO: add sources for the parameters.**

**Tab. 3.1:** Priors on merger rate shape parameters.

Parameter	Prior
$R_0$	$1/R_0$ (implicit)
$\kappa$	2.86
$\gamma$	4.59
$z_p$	2.47

For BBH we use power-law with a Gaussian peak mass model, with powerlaw slope  $\alpha$ , the mean of the Gaussian  $\sigma_g$ , the width of the peak  $\mu_g$ , and the relative weight between power-law and the peak  $\lambda_g$ . We also assume the minimum and maximum masses for the black hole distribution. For neutron stars we assume that mass is

uniformly distributed between  $M_{min,NS}$  and  $M_{max,NS}$ . All the assumed priors are given in Table 3.2. **TODO: add sources for the parameters.**

**Tab. 3.2:** Priors on the mass model parameters.

Parameter	Prior
$\alpha$	3.78
$\sigma_g$	3.88
$\mu_g$	32.27
$\lambda_g$	0.03
$M_{min,BH}$	$4.98M_\odot$
$M_{max,BH}$	$112.5M_\odot$
$M_{min,NS}$	$1.0M_\odot$
$M_{max,NS}$	$3.0M_\odot$

## 3.2 EM Galaxy Catalogs

### 3.2.1 The GLADE+ Catalog

For redshift information, we use the **GLADE+** galaxy catalog (Dálya et al., 2022), an extended version of the original GLADE catalog designed for GW follow-up. GLADE+ is a composite catalog that combines several large surveys to achieve all-sky coverage:

- 2MASS XSC and 2MPZ (infrared-based all-sky surveys)
- WISExSCOS (photometric redshifts)
- HyperLEDA (spectroscopic redshifts)
- SDSS DR16 (optical photometric and spectroscopic data)

As of the latest release, GLADE+ includes over 22 million objects with available positions, redshifts (spectroscopic or photometric), and photometry in multiple bands, most critically the  $K$ -band.

### 3.2.2 Catalog Completeness

A significant limitation of GLADE+ is its incompleteness beyond redshift  $z \sim 0.3$ . While the catalog is approximately complete for bright galaxies at low redshift, its coverage of fainter galaxies or more distant regions is limited. This incompleteness introduces a bias when performing statistical redshift inference, as missing galaxies contribute to the out-of-catalog part of the redshift prior.

To address this, the `gwcosmo` pipeline models the galaxy population using a truncated Schechter luminosity function:

$$\phi(M) \propto 10^{0.4(\alpha+1)(M-M^*)} \exp[-10^{0.4(M-M^*)}] \quad (3.2)$$

where  $M$  is the absolute magnitude in the  $K$ -band,  $M^*$  is the characteristic magnitude, and  $\alpha$  is the faint-end slope. For our analysis, we adopt the standard parameters:  $\alpha = -1.09$ ,  $M_K^* = -23.39 + 5 \log h$ , consistent with Kochanek et al., 2001, and truncate at  $M_{K,min}^* = -27.00 + 5 \log h$  and  $M_{K,max}^* = -19.00 + 5 \log h$ .

### 3.2.3 Redshift Uncertainty Modeling

Redshift measurements in GLADE+ are heterogeneous. Spectroscopic redshifts are generally precise, but photometric redshifts can have uncertainties on the order of  $\sigma_z \sim 0.01 - 0.05$ . These uncertainties are modeled by convolving the redshift of each galaxy with a Gaussian of fixed width, an assumption used in constructing the redshift prior.

Additionally, to reduce systematic error, galaxies with unphysical redshifts or unreliable photometry are filtered out during preprocessing. The net result is a cleaned, sky-localized, and redshift-tagged galaxy distribution that can be used for LOS prior construction.

Furthermore,  $K$ -corrections are applied in `gwcosmo` following Kochanek et al., 2001, and the apparent magnitude thresholds  $m_{thr}$  are computed per HEALPix pixel as a median apparent magnitude in a given pixel, giving a  $K$ -band threshold of 13.5 on average.

### 3.2.4 $K$ -band Luminosity as Host Probability Proxy

In dark siren analysis, we must assign each galaxy a probability of being the true host. Following common practice, this probability is assumed to scale with stellar mass, for which  $K$ -band luminosity serves as a good proxy CITATION NEEDED. The host probability  $p_i$  for each galaxy is given by:

$$p_i \propto L_{K,i} \quad (3.3)$$

This choice is motivated by the assumption that more massive galaxies are more likely to host CBCs events. While simplistic, it provides a reasonable baseline.

# GLADE+ Bright Galaxies as Redshift Tracers

The core idea investigated in this thesis is the use of **bright galaxies** as redshift tracers for GW events. The GLADE+ galaxy catalog is a comprehensive database of galaxies in the local universe, but it suffers from incompleteness at high redshifts. This incompleteness arises from the limited depth of the catalog, which is primarily designed to cover the local universe, owing to the sensitivity constraints of the current EM telescope. We overcome this limitation by focusing on the brightest galaxies in the catalog, which are more likely to be well-represented and can provide a more accurate estimate of the redshift of GW events. This essentially makes the catalog deeper and more complete at high redshifts, while allowing us to use the GLADE+ catalog as a proxy for the underlying matter distribution in the universe.

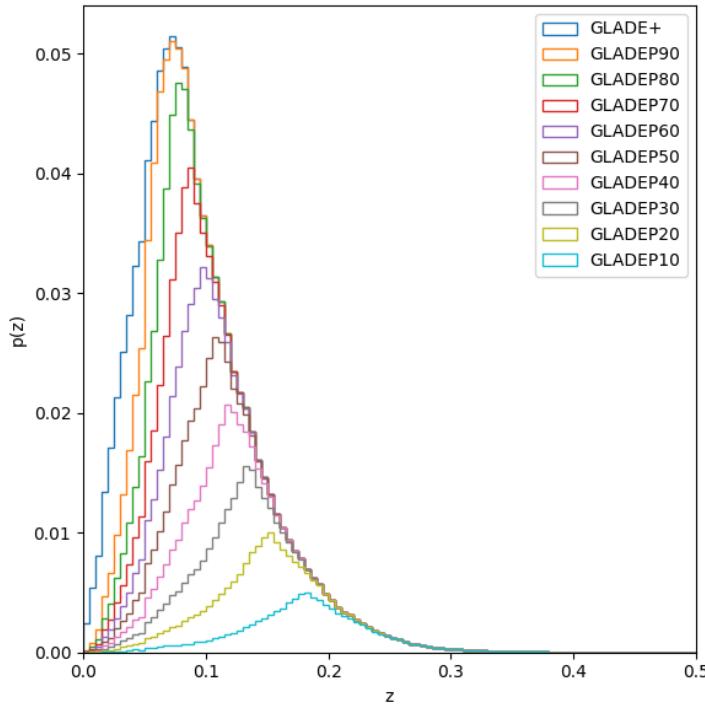
## 4.1 Bright Galaxy Subsets

The analysis begins with the construction of **bright galaxy subsets**, denoted GLADEPXX, where only the top  $XX\%$  of galaxies by cumulative  $K$ -band luminosity are included. The  $K$ -band luminosity is chosen as it is better associated with the mass of the galaxies (Strazzullo et al., 2006; Sureshkumar et al., 2021). Limiting the analysis to these bright subsets allows us to:

- Reduce the out-of-catalog correction by focusing on galaxies likely to be well-represented.
- Improve the signal-to-noise ratio of the statistical redshift prior.
- Minimize the inclusion of poorly characterized or faint galaxies.

The bright galaxy subsets are defined by selecting galaxies based on their absolute magnitude, which is inferred from their redshift and apparent magnitude. For a given subset, the dimmest galaxy sets the absolute magnitude limit,  $M_{max}$ , for that subset. This limit is then used in the Schechter luminosity function to adjust the out-of-catalog contribution, effectively reducing its impact. By focusing on the brightest galaxies, we assume that they trace the mass distribution in the Universe more effectively, and are more likely to host CBC events.

This approach improves the completeness of the galaxy catalog in the high-redshift regime, as the out-of-catalog contribution becomes smaller. Figure 4.1 illustrates how the redshift distribution shift towards higher values making the catalog more complete, highlighting the effectiveness of this method in probing deeper  $z$ -ranges.



**Fig. 4.1:** Normalized redshift distributions  $p(z)$  for the full GLADE+ galaxy catalogue (blue) and the brightest percentiles (other curves). Each subset, labeled “GLADEPXX,” includes only the top XX% brightest galaxies in the  $K$ -band. Focusing on increasingly brighter subsets shifts and sharpens the redshift distribution, effectively probing deeper ranges in  $z$ .

#### 4.1.1 Trade-Offs and Limitations

While bright subsets reduce catalog incompleteness, they may inadvertently exclude genuine host galaxies, especially if those lie in less luminous systems or in under-represented regions of the catalog. The effectiveness of these cuts depends on the intrinsic distribution of host galaxies, the depth and completeness of the original catalog, and the accuracy of  $K$ -band photometry.

One major limitation of this approach is that brightest galaxies may not be representative of the overall galaxy population, as they may be biased towards certain types of galaxies or regions of the sky. This effect is however mitigated by the fact that we are using the  $K$ -band luminosities, which are good tracer for the stellar mass (Strazzullo et al., 2006; Sureshkumar et al., 2021), and in turn the overall matter distribution. Furthermore, the bright galaxy subsets are constructed from the

GLADE+ catalog, which is designed to be a complete and representative sample of galaxies in the local universe. This means that the bright galaxy subsets are likely to be representative of the overall galaxy population, and give a crude estimate of the matter distribution, even if they are not a complete sample.

Estimating the true redshift of a GW event by the nearest brightest galaxies is supposed to have a negligible effect on the results. This is because the bright galaxies are more likely to be located in regions of high density, such as galaxy clusters or groups, which are also likely to host GW events. This means that the bright galaxies are likely to be located in the same large-scale structure as the GW event, and thus provide a good estimate of the redshift of the event.

One important thing to note is that while the use of a bright subset enhances the completeness of the catalogue at high redshifts, it also introduces a trade-off. If the brightness cut is too restrictive, the exclusion of fainter galaxies may lead to an underrepresentation of the overall matter distribution, potentially biasing the results. One needs to carefully balance these considerations by comparing different brightness thresholds. A moderate brightness cut will maximize the benefit in terms of depth without incurring significant bias.

The appropriate brightness cut can be established only via a set of astrophysics-motivated large-scale simulations. A *GW Mock Data Challenge (MDC)* using the simulated BUZZARD galaxy catalog from the Dark Energy Survey (DES) Collaboration (DeRose et al., 2019; DeRose et al., 2022). It is designed to determine the optimal brightness threshold which maximizes the measurement precision with bright subsets while keeping the biases in control. Tests performed in the process will help refine the methodology and ensure that the improved constraints on the Hubble constant are robust against additional systematic uncertainties.

While the bright galaxy subsets may not be a perfect representation of the overall galaxy population, they are still a useful tool for improving the completeness of the galaxy catalog and reducing the out-of-catalog correction. The error incurred by using the bright galaxies as redshift tracers maybe small compared to the current errors in luminosity distance measurements from the GW events, this may become a problem in the future as the GW measurements become more precise. In this case, one may need to use more sophisticated methods to estimate the redshift of the GW event, such as more complex models of the galaxy population, but for now this is a good first step towards improving the completeness of the galaxy catalog and reducing the out-of-catalog correction, by leveraging the currently available EM data.

## 4.2 The `gwcosmo` Pipeline

Based on the work presented in Gray et al., 2020; Gray et al., 2022; Gray et al., 2023, `gwcosmo` is a Python package specifically designed for the joint inference of cosmological parameters using data from standard sirens and galaxy catalogs. The pipeline employs Bayesian inference techniques to simultaneously constrain cosmological parameters, such as the Hubble constant and dark energy equation of state parameters. It does this using various data inputs, including the gravitational wave strain data from gravitational wave observatories such as LIGO, Virgo, and KAGRA, entries from galaxy catalogs containing photometric or spectroscopic redshift information, as well as other relevant galaxy properties, and CBC source population models.

`gwcosmo` represents a significant advancement in extracting cosmological information especially from CBC events that lack unique EM counterparts. It addresses the challenge of redshift inference in dark siren cosmology by statistically marginalizing over potential host galaxies from a flux-limited catalog. This is achieved through hierarchical Bayesian modeling that incorporates GW strain data, the spatial and redshift distribution of galaxies, and astrophysical population models.

The pipeline is designed to be flexible and extensible, allowing users to customize the models and priors used in the analysis. It also provides tools for visualizing the results, including posterior distributions and confidence intervals for the inferred parameters. The package is built on top of established libraries such as `numpy`, `scipy`, and `matplotlib`, making it accessible to researchers familiar with these tools. The `gwcosmo` pipeline is a particularly useful tool in the intersection of gravitational wave astronomy and cosmology, as it enables the extraction of valuable cosmological information from gravitational wave events, which can complement traditional methods of measuring cosmological parameters. The package is actively maintained and updated to incorporate new developments in both gravitational wave detection and cosmological modeling, ensuring that it remains a relevant tool for researchers in the field.

### 4.2.1 Bayesian Framework: TO BE UPDATED

The goal is to evaluate the posterior distribution:

$$p(\Lambda \mid \{d_{\text{GW}}\}, \{d_{\text{EM}}\}) \propto p(\Lambda) p(\{d_{\text{GW}}\}, \{d_{\text{EM}}\} \mid \Lambda), \quad (4.1)$$

where:

- $\Lambda$  are the hyperparameters of interest (e.g., cosmological parameters, population hyperparameters),
- $d_{\text{GW}}$  are the GW observables (e.g., luminosity distance posteriors),
- $d_{\text{EM}}$  are the galaxy catalogue data (e.g., redshifts, magnitudes, positions).

The likelihood is approximated by marginalizing over all galaxies  $k$  within the GW localization volume:

$$p(\Lambda \mid \{d_{\text{GW}}\}, \{d_{\text{EM}}\}) \propto \prod_i \sum_k w_k p(d_{\text{GW},i} \mid z_k, \Lambda), \quad (4.2)$$

where  $w_k$  is the weight assigned to each galaxy, typically modeled based on a proxy for host probability (e.g., luminosity in a selected band, merger rate evolution).

### 4.2.2 Key Features of `gwcosmo`

Some of the key features of the `gwcosmo` pipeline include:

- **Joint Population + Cosmology Inference:** Simultaneously samples over cosmological parameters and population hyperparameters, avoiding bias from fixed population assumptions.
- **Galaxy Catalogue Weighting:** Utilizes LOS redshift priors constructed from galaxy distributions, incorporating completeness corrections based on flux thresholds.
- **Selection Effects:** Corrects for GW detector sensitivity and incompleteness in galaxy catalogues by modeling detection probabilities and magnitude limits.

### 4.2.3 Implementation and Applications

The pipeline is implemented in Python and uses kernel density estimation (KDE) of GW posteriors and efficient catalog summation to handle discrete and incomplete galaxy data. The most recent version was applied by Gray et al., 2023 to reanalyze 47 events from the GWTC-3 catalog using the GLADE+ galaxy catalog, yielding an updated measurement of the Hubble constant:

$$H_0 = 69^{+12}_{-7} \text{ km s}^{-1} \text{ Mpc}^{-1}.$$

`gwcosmo` is publicly available and is expected to be a useful tool for future gravitational-wave cosmology with upcoming detector runs and deeper galaxy surveys.

## 4.3 LOS Redshift Prior

We base our analysis on the `gwcosmo` pipeline (Gray et al., 2020; Gray et al., 2022; Gray et al., 2023), which relies on a precomputed LOS redshift prior, a prior on the GW signal’s redshift and direction. This LOS redshift prior is used in tandem with the luminosity distance posterior from the GW signal, to get a measurement for the Hubble constant  $H_0$ .

### 4.3.1 LOS Redshift Prior Construction

The LOS redshift prior is constructed from the GLADE+ galaxy catalog, which contains a wealth of information about the galaxies in the local universe. The redshift prior is constructed by taking into account the distribution of galaxies along the line of sight to the GW event, as well as their luminosity and redshift. This allows us to obtain a more accurate estimate for the redshift of the GW signal, which is crucial for cosmological measurements. The prior is constructed by dividing the sky into HEALPix pixels, and computing the redshift distribution of galaxies in each pixel. The redshift prior is then weighted using the luminosity of the potential host galaxies, allowing us to obtain a more accurate estimate for the Hubble constant.

Furthermore, the prior accounts for the incompleteness of the galaxy catalog, using source population models and the magnitude threshold calculated per pixel, which is particularly important for high-redshift events where the number of galaxies is significantly reduced. This allows us to separate the LOS redshift prior into an in-catalog and out-of-catalog contribution.

Taking into account, the fact that the host galaxy can be present, marked as  $G$ , or not, marked as  $\bar{G}$ , inside the catalog, one can write the LOS redshift prior as:

$$p(z|\Omega_i, \Lambda, s, I) = \iint \sum_{g=G, \bar{G}} p(z, M, m, g|\Omega_i, \Lambda, s, I) dM dm \quad (4.3)$$

$$= p(G|\Omega_i, \Lambda, s, I) \iint p(z, M, m|G, \Omega_i, \Lambda, s, I) dM dm + p(\bar{G}|\Omega_i, \Lambda, s, I) \iint p(z, M, m|\bar{G}, \Omega_i, \Lambda, s, I) dM dm \quad (4.4)$$

The first term in the equation represents the contribution from galaxies that are present in the catalog, while the second term represents the contribution from galaxies not present in the catalog. The two terms are weighted by their respective probabilities of being present or not in the catalog. The terms inside the integral are the priors on the redshift  $z$ , absolute magnitude  $M$ , and apparent magnitude of the galaxies  $m$ , informed by the galaxy catalog, within the sky area covered by pixel  $i$ . Here the parameters  $G/\bar{G}$  give the presence or absence of the galaxy in the catalog,  $\Omega_i$  the sky location of the GW event,  $\Lambda$  the cosmological and population hyperparameters of interest,  $s$  the presence of a GW source, and  $I$  the additional assumptions which are not explicitly expressed. One also needs to marginalize over the absolute magnitude  $M$  and the apparent magnitude  $m$  of the galaxy, as these determine, to the leading order, which galaxies are present in a flux-limited EM survey (Gray et al., 2023).

The integral in the in-catalog term can be expressed as the sum over the possible host galaxies in the catalog, weighted by their respective probabilities of being the host galaxy. These galaxies are weighted by their luminosity, which is a function of the absolute magnitude and redshift. The rationale being that the more luminous, and thus heavier galaxies are more likely to host CBC events, and therefore contribute more to the LOS redshift prior. This reduces the in-catalog part to a weighted sum over the galaxies in the catalog, where the galaxies are treated as point sources modeled by a Gaussian. This term can thus be expressed as:

$$\begin{aligned} \iint p(z, M, m | G, \Omega_i, \Lambda, s, I) dM dm &= \frac{1}{p(s|G, \Omega_i, \Lambda, I) N_{gal}(\Omega_i)} \\ &\times \sum_k^{N_{gal}(\Omega_i)} p(z|\hat{z}_k) p(s|z, M(z, \hat{m}_k, \Lambda), \Lambda, I) \end{aligned} \quad (4.5)$$

where the term  $p(z|\hat{z}_k)$  represents the probability of a galaxy being at redshift  $z$ , given its observed redshift  $\hat{z}_k$ . This term is used to weight the contribution from each galaxy in the catalog, based on its observed redshift.

The integral in the out-of-catalog term marginalizes over the possible host galaxies not present in the catalog. This term is more complex, as it requires a model for the distribution of galaxies in the universe, which is not directly available from the catalog. We use a Schechter luminosity function to model the distribution of galaxies

in the universe, which allows us to estimate the contribution from galaxies outside the catalog. The out-of-catalog term can be expressed as:

$$\begin{aligned}
& \iint p(z, M, m | \bar{G}, \Omega_i, \Lambda, s, I) dM dm \\
&= \frac{1}{p(s | \bar{G}, \Omega_i, \Lambda, I) p(\bar{G} | \Omega_i, \Lambda, I)} \\
&\times \left[ \Theta[z_{cut} - z] \int_{M(z, m_{th}(\Omega_i), \Lambda)}^{M_{max}(H_0)} p(z, M | \Lambda, I) p(s | z, M, \Lambda, I) dM \right. \\
&\quad \left. + \Theta[z - z_{cut}] \int_{M_{min}(H_0)}^{M_{max}(H_0)} p(z, M | \Lambda, I) p(s | z, M, \Lambda, I) dM \right] \quad (4.6)
\end{aligned}$$

Here we also account for the EM selection effects of the catalog. Due to the flux limited nature of the galaxy catalog, the probability of a galaxy being present in the catalog depends on the galaxy's apparent magnitude, and whether it is greater or smaller than apparent magnitude threshold of the catalog along the same line of sight,  $m_{th}(\Omega_i)$ . The Heaviside function  $\Theta$  is used to separate the two cases of the out-of-catalog contribution, depending on whether the redshift  $z$  is below or above a certain threshold  $z_{cut}$ . This is due to the exclusion of galaxies with redshift  $z$  greater than  $z_{cut}$  from the catalog, which is a result of unreliable redshift or color information at these higher redshifts.

The term  $p(s | z, M, \Lambda, I)$  is the weighting factor for the contribution from each galaxy, based on its luminosity and redshift. The galaxies are weighted by their luminosity in the  $K$ -band, which is a good tracer of the mass of the galaxies (Strazzullo et al., 2006; Sureshkumar et al., 2021). Furthermore, the merger host probability is also taken into account, which is a function of the redshift. This is modeled by a Madau-Dickinson merger rate evolution model (Madau and Dickinson, 2014), which describes the evolution of the merger rate with redshift, discussed in Section 3.1.4. The term  $p(s | z, M, \Lambda, I)$ , also incorporates the source population models, used to populate the out-of-catalog contribution. These models are also discussed in Section 3.1.4.

The term  $p(z, M | \Lambda, I)$  represents the luminosity function of the galaxies, taken to be the Schechter luminosity function, discussed in Section 3.2.2. The integration limits are set by the minimum and maximum absolute magnitudes of the galaxies,  $M_{min}(H_0)$  and  $M_{max}(H_0)$ . These are  $H_0$ -dependent, as the parameters of the Schechter luminosity function are also  $H_0$ -dependent, but the final distribution remains insensitive to the exact values of  $H_0$  (Gray et al., 2023).

## 4.4 Results

In this chapter we present the main outcomes of our analysis: how applying brightness cuts to the GLADE+ catalog alters LOS redshift prior for individual events, how these modified priors propagate into the inferred posterior on the Hubble constant  $H_0$ , and where the optimal balance lies between catalog depth and sampling variance.

To test the impact of catalog completeness, we apply successive cuts to the galaxy catalog by selecting only the brightest  $XX\%$  of galaxies in  $K$ -band luminosity, forming subsets labeled GLADEPXX. This process shifts the effective magnitude threshold upward, reducing the number of galaxies but improving the catalog's completeness above the cut.

### 4.4.1 LOS Redshift Prior

Figure 4.2 shows the LOS redshift prior for the event GW170809 under different catalog cuts. Bright galaxy subsets (e.g., GLADEP20) result in sharper redshift distributions and reduce the weight of the out-of-catalog term. The low- $z$  tail contributed by faint galaxies in the nearby universe is also suppressed. Notably, these priors maintain consistency in shape, indicating that bright galaxies are reliable tracers of large-scale structure. The application of a brightness cut significantly modifies the LOS redshift distribution. Specifically, the bright galaxy subsets yield an amplified redshift prior at higher distances, effectively extending the reach of the catalogue, somewhat mitigating the incompleteness issues that arise at deeper redshifts. This behavior is qualitatively consistent across events.

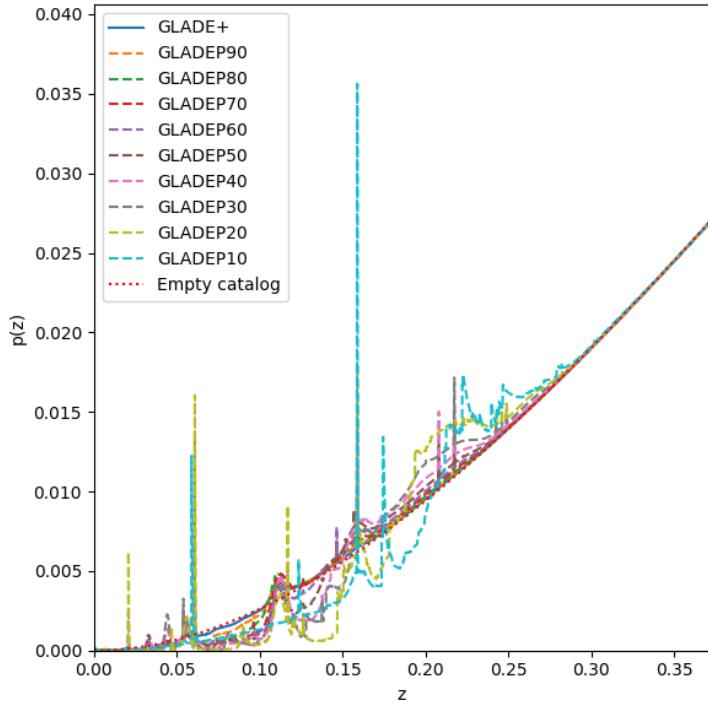
### 4.4.2 Hubble Constant Posterior

Once one has the LOS redshift prior, one can use it to compute the posterior distribution of the Hubble constant  $H_0$  given the GW event data, as the relation between the redshift  $z$ , luminosity distance  $d_L$  and the Hubble constant  $H_0$  is given by:

$$d_L = \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{(1+z')^3 \Omega_m + \Omega_\Lambda}} \quad (4.7)$$

for a flat  $\Lambda$ CDM cosmology. At lower redshifts, this can be simplified to:

$$d_L \approx \frac{cz}{H_0} \quad (4.8)$$



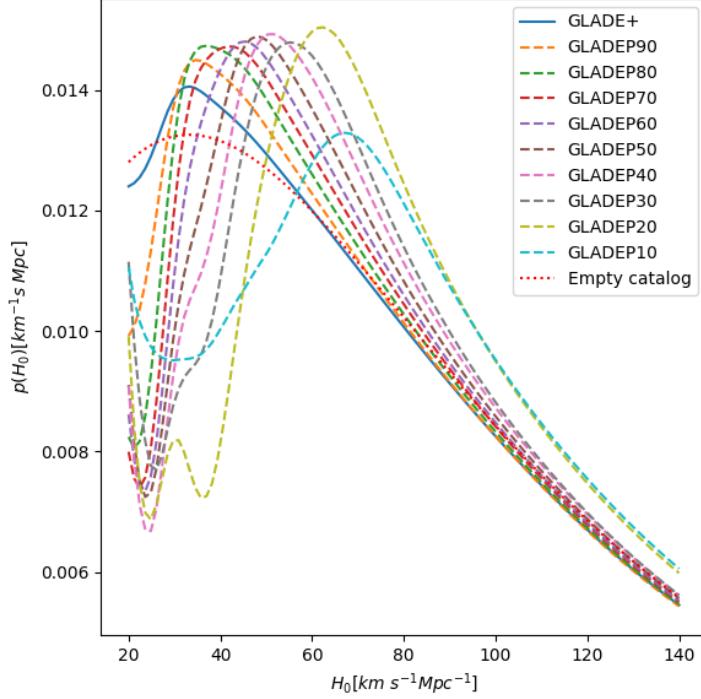
**Fig. 4.2:** LOS redshift prior for GW170809 for different brightness-ranked GLADE subsets. The full GLADE+ catalog (solid blue) is compared to the different subsets. Applying a brightness cut amplifies the prior at higher distances, extending the effective reach of the catalog and partially mitigating incompleteness at greater redshifts.

Thus the LOS redshift prior,  $p(z|\Omega_i, \Lambda, s, I)$ , and the luminosity distance posterior,  $p(d_L, \Omega | \text{GW})$ , from the GW event data can be combined to obtain the posterior distribution of  $H_0$ , which can then be used to obtain constraints on the value of  $H_0$ .

Figure 4.3 shows the resulting  $H_0$  posteriors for GW170809 under various catalog cuts. As the catalog is restricted to brighter galaxies, the  $H_0$  posterior becomes increasingly narrow. For example, GLADEP20 yields a visibly tighter posterior compared to the full GLADE+, with minimal shift in the median. However, the most aggressive cut (GLADEP10) introduces broader tails, likely due to insufficient galaxy sampling in the localization volume. (**LINK TO THE 20% LIMIT IN BUZZARD?**)

We repeat this procedure for a subset of BBH events from GWTC that meet our selection criteria. Figure 4.4 shows the combined  $H_0$  posterior from the selected dark siren events using the full GLADE+ catalog and the different subsets.

Table 4.1 summarizes the  $H_0$  posteriors for all cuts. The uncertainty is minimized around the GLADEP20 subset, showing 30-40% tighter constraints, while median values shift towards a higher value. For less extreme cuts, the median values doesn't



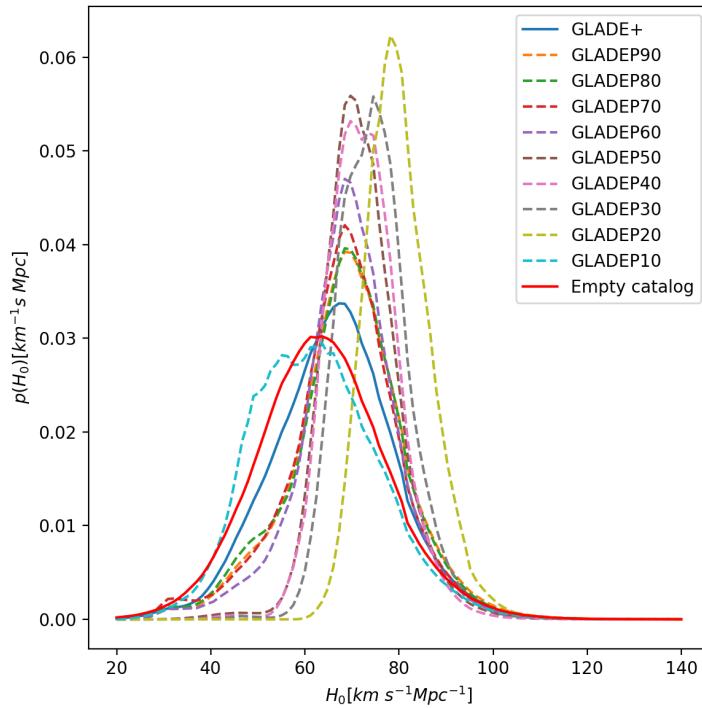
**Fig. 4.3:**  $H_0$  posterior distributions for GW170809 using full and brightness-cut galaxy catalogs. GLADEP20 yields the tightest credible interval. GLADEP10 suffers from sparse sampling.

show a huge shift. This suggests that moderate brightness cuts improve precision without introducing systematic bias.

#### 4.4.3 Cost–benefit trade-off

We observe that precision improves with increasing brightness until an optimum is reached (near GLADEP20). Beyond this point, the posterior widens again, moving towards an empty catalog result, due to under-sampling. This defines a practical limit for catalog pruning. In real applications, the exact optimum may depend on event SNR, catalog completeness, and the merger rate model.

Our results suggest that targeting the brightest galaxies in a well-characterized catalog can substantially improve the statistical precision of  $H_0$  constraints from dark sirens. This approach complements other efforts to mitigate catalog incompleteness, including joint inference with galaxy clustering and population-informed redshift estimation.



**Fig. 4.4:** Cumulative  $H_0$  posterior from the selected dark siren events using full GLADE+ (blue) and the different subsets (dashed lines). The brightness-weighted catalog yields a tighter constraint without a significant shift.

**Tab. 4.1:** Maximum *a posteriori* probabilities with 68% confidence ranges of the  $H_0$  posterior distributions alongside the maximum magnitude limits for the different percentiles of the GLADE+ galaxy catalogue.

Catalogue	$M_{K,max}$	$H_0 [km s^{-1} Mpc^{-1}]$
GLADE+	-19.00	$67.87^{+8.97}_{-10.29}$
GLADEP90	-23.07	$68.94^{+9.24}_{-7.55}$
GLADEP80	-23.62	$68.93^{+9.25}_{-7.57}$
GLADEP70	-23.94	$68.63^{+8.61}_{-7.42}$
GLADEP60	-24.19	$68.85^{+7.72}_{-6.43}$
GLADEP50	-24.14	$70.05^{+6.12}_{-5.20}$
GLADEP40	-24.63	$69.94^{+7.46}_{-3.88}$
GLADEP30	-24.87	$74.70^{+4.69}_{-6.58}$
GLADEP20	-25.15	$78.28^{+5.53}_{-4.95}$
GLADEP10	-25.53	$63.46^{+6.39}_{-14.37}$

## Mock Data Challenge

### 5.1 BUZZARD Mock Catalog

### 5.2 Results



# Discussion



# Conclusion



# Bibliography

- Abbott, B. P. et al. (2017). „A gravitational-wave standard siren measurement of the Hubble constant“. In: *Nature* 551.7678, pp. 85–88. arXiv: [1710.05835 \[astro-ph.CO\]](#) (cit. on p. 5).
- Abbott, Benjamin P, R Abbott, TD Abbott, et al. (2016). „GW150914: First results from the search for binary black hole coalescence with Advanced LIGO“. In: *Physical Review D* 93.12, p. 122003 (cit. on p. 4).
- Abbott, Benjamin P et al. (2019). „GWTC-1: a gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs“. In: *Physical Review X* 9.3, p. 031040 (cit. on p. 13).
- Abbott, BP et al. (2021a). „A gravitational-wave measurement of the Hubble constant following the second observing run of advanced LIGO and Virgo“. In: *The Astrophysical Journal* 909.2, p. 218 (cit. on p. 6).
- Abbott, R et al. (2023a). „Constraints on the cosmic expansion history from GWTC-3“. In: *The Astrophysical Journal* (cit. on pp. 6, 8, 9).
- Abbott, R. et al. (2024). „GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run“. In: *Physical Review D* 109.2, p. 022001 (cit. on p. 13).
- Abbott, Richard et al. (2021b). „GWTC-2: compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run“. In: *Physical Review X* 11.2, p. 021053 (cit. on p. 13).
- (2023b). „GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run“. In: *Physical Review X* 13.4, p. 041039 (cit. on p. 13).
- Aghanim, N. et al. (2020). „Planck 2018 results. VI. Cosmological parameters“. In: *Astron. Astrophys.* 641. [Erratum: *Astron.Astrophys.* 652, C4 (2021)], A6. arXiv: [1807.06209 \[astro-ph.CO\]](#) (cit. on p. 4).
- Dálya, Gergely et al. (2022). „GLADE+: an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors“. In: *Monthly Notices of the Royal Astronomical Society* 514.1, pp. 1403–1411 (cit. on pp. 6, 15).
- DeRose, J. et al. (Jan. 2019). „The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs“. In: arXiv: [1901.02401 \[astro-ph.CO\]](#) (cit. on p. 19).

- DeRose, J. et al. (2022). „Dark Energy Survey Year 3 results: Cosmology from combined galaxy clustering and lensing validation on cosmological simulations“. In: *Phys. Rev. D* 105.12, p. 123520. arXiv: [2105.13547 \[astro-ph.CO\]](#) (cit. on p. 19).
- Ezquiaga, Jose María and Daniel E Holz (2022). „Spectral sirens: Cosmology from the full mass distribution of compact binaries“. In: *Physical Review Letters* 129.6, p. 061102 (cit. on p. 7).
- Gray, Rachel et al. (2020). „Cosmological inference using gravitational wave standard sirens: A mock data analysis“. In: *Physical Review D* 101.12, p. 122001 (cit. on pp. 6, 20, 22).
- (2022). „A pixelated approach to galaxy catalogue incompleteness: improving the dark siren measurement of the Hubble constant“. In: *Monthly Notices of the Royal Astronomical Society* 512.1, pp. 1127–1140 (cit. on pp. 6, 20, 22).
  - (2023). „Joint cosmological and gravitational-wave population inference using dark sirens and galaxy catalogues“. In: *Journal of Cosmology and Astroparticle Physics* 2023.12, p. 023 (cit. on pp. 6, 7, 20–24).
- Kochanek, CS et al. (2001). „The k-band galaxy luminosity function“. In: *The Astrophysical Journal* 560.2, p. 566 (cit. on p. 16).
- Lauer, Tod R, Marc Postman, Michael A Strauss, Genevieve J Graves, and Nora E Chisari (2014). „Brightest cluster galaxies at the present epoch“. In: *The Astrophysical Journal* 797.2, p. 82 (cit. on p. 10).
- Madau, Piero and Mark Dickinson (2014). „Cosmic star-formation history“. In: *Annual Review of Astronomy and Astrophysics* 52.1, pp. 415–486 (cit. on pp. 14, 24).
- Maggiore, Michele (Oct. 2007). *Gravitational Waves: Volume 1: Theory and Experiments*. Oxford University Press (cit. on p. 3).
- Mastrogiovanni, S et al. (2021). „On the importance of source population models for gravitational-wave cosmology“. In: *Physical Review D* 104.6, p. 062009 (cit. on p. 6).
- Mastrogiovanni, Simone et al. (2024). „ICAROGW: A python package for inference of astrophysical population properties of noisy, heterogeneous, and incomplete observations“. In: *Astronomy & Astrophysics* 682, A167 (cit. on pp. 6, 7).
- Riess, Adam G., Stefano Casertano, Wenlong Yuan, Lucas M. Macri, and Dan Scolnic (2019). „Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond  $\Lambda$ CDM“. In: *Astrophys. J.* 876.1, p. 85. arXiv: [1903.07603 \[astro-ph.CO\]](#) (cit. on p. 4).
- Schutz, Bernard F (1986). „Determining the Hubble constant from gravitational wave observations“. In: *Nature* 323.6086, pp. 310–311 (cit. on p. 4).
- Strazzullo, Veronica, Piero Rosati, SA Stanford, et al. (2006). „The near-infrared luminosity function of cluster galaxies beyond redshift one“. In: *Astronomy & Astrophysics* 450.3, pp. 909–923 (cit. on pp. 8, 10, 17, 18, 24).
- Sureshkumar, Unnikrishnan, A Durkalec, Agnieszka Pollo, et al. (2021). „Galaxy and Mass Assembly (GAMA)-Tracing galaxy environment using the marked correlation function“. In: *Astronomy & Astrophysics* 653, A35 (cit. on pp. 8, 10, 17, 18, 24).