

Determining Hubble's Constant from Gravitational Wave Sirens

*A Project Report for partial fulfillment of the requirements for
the award of degree of*

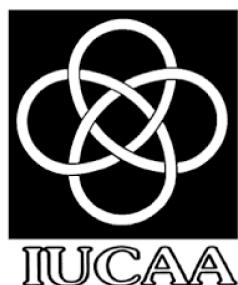
Master of Science
in
Physics with Astrophysics

by

Devesh Yadav

Under the supervision of
Prof. Surhud More

Submitted to



**Inter-University Center for Astronomy
and Astrophysics**

and



**Savitribai Phule Pune University
Pune, India**

Abstract

The Hubble constant quantifies the current rate of expansion of the Universe and is an important parameter in cosmology that sets the scale for distance and time measurements in the Universe. A significant discrepancy in the value of the Hubble constant measurement from different techniques has arisen in the last few years with increased precision. While the local measurements of the Hubble constant based on the cosmic distance ladder provide a value of nearly $72.0 \text{ km}^{-1}\text{Mpc}^{-1}$ ($73.24 \pm 1.74 \text{ km s}^{-1}\text{Mpc}^{-1}$ from Hubble Telescope Measurement), the early-universe measurements based on the cosmic microwave background measurements generally favor a slightly lower value of H_0 (the Plank mission measurements give a value of $H_0 = 67.3 \pm 1.2 \text{ km s}^{-1}\text{Mpc}^{-1}$). Gravitational wave events allow us to infer the luminosity distance to the compact binary coalescence events without having to rely on the cosmic distance ladder and can provide a one-step measurement of the Hubble constant. We present a method to constrain Hubble's constant using binary black hole gravitational wave events by aiding in the inference of their redshift. Since galaxy observations are resolution-limited, very faint and high-redshifted galaxies may not be present in the catalog, so many GW events cannot be associated with the host galaxy to get the redshifts. We try to address this problem by populating galaxies below the resolution limit (some threshold) such that they follow the same large-scale structure as above. We will then look at how the galaxies are clustered in this regime and then use this clustering information to determine the value of different cosmological parameters.

Acknowledgement

I want to express my deepest gratitude to **Prof. Surhud More** for his invaluable guidance, support, and mentorship throughout the course of the project. His expertise in the field and insightful feedback have been instrumental in shaping this report. I sincerely thank **Dr. Sukanata Bose** for his valuable insights and guidance as an external guide.

I am also immensely grateful to **Tathagata Ghosh** for his dedicated assistance, valuable input, and constant encouragement. His enthusiasm and expertise have greatly contributed to the success of this project.

I would also like to acknowledge the **Inter-University Center for Astronomy and Astrophysics** for providing the necessary resources and a conducive environment to complete this project. I also acknowledge using the **IUCAA LDG Cluster - Sarathi** for the computational/numerical work.

Lastly, I extend my heartfelt thanks to all those who have supported and encouraged me throughout this endeavor.

DEVESH YADAV

CERTIFICATE

PHY – IP495

This is to certify that the work incorporated in the Project report entitled "**Determining Hubble's Constant from Gravitational Wave Sirens**" is completed by **Devesh Yadav** (**Seat No: 22024003**) for the partial fulfillment of the requirement of the degree of Master of Science in Physics (Astrophysics) at Savitribai Phule Pune University and Inter-University Center for Astronomy and Astrophysics (IUCAA). This work is original and has not been submitted before for any other degree at this or any other university.

Surhud More

Professor

Inter-University Center for Astronomy and Astrophysics
(Project Guide)

Head

Department of Physics
Savitribai Phule Pune University

Place: Pune, India

Date: 18 May 2024

DECLARATION

PHY – IP495

I hereby declare that the project work entitled "**Determining Hubble's Constant from Gravitational Wave Sirens**" submitted for the partial fulfillment of the degree of Master of Science in Physics (Astrophysics) at Department of Physics, Savitribai Phule Pune University (SPPU) and Inter-University Center for Astronomy and Astrophysics (IUCAA), Pune during the year 2023-24 is carried out by me and has not been previously formed on the basis for the award of any degree or diploma or another similar title of this or any other university.

Place: Pune

Date: 18 May 2024

Name: Devesh Yadav

PRN: 22022002509

Seat No.: 22024003

Contents

Abstract	i
Acknowledgement	i
Certificate	i
Declaration	ii
1 Introduction	4
1.1 The Standard Model of Cosmology	5
1.1.1 Cosmological Principle	5
1.1.2 The Expanding Universe	5
1.1.3 The Scale Factor	6
1.1.4 Friedmann Equations	6
1.1.5 Galaxy Luminosity Function	10
1.2 Objective	10
2 Methodology	13
2.1 Catalog Preparation and Simulating GW Events	13
2.1.1 Associating GW Events with Hosts:	14
2.1.2 Some Bayesian Statistics	15
2.1.3 bilby - The Bayesian Inference Library and finding Posterior . .	15
2.2 Parameter Estimation	17
2.2.1 Posterior	20
3 Galaxy Population	21
3.1 The Idea	21
3.2 The Mass Cut	22
3.3 Density Field	22
3.4 Overdensity	23
3.5 Bias	23
3.6 Some Statistics	25
3.7 Plots	26
3.8 Populating the Galaxies	29
4 Future Work	30

List of Figures

1.1	The plots show how distances vary in an expanding universe for different values of H_0 ($\text{km s}^{-1} \text{Mpc}^{-1}$)	10
1.2	The figure shows luminosity function of galaxies for a particular band taken from SDSS data. The thick line is the Schechter fit and the other one is the data points. The histogram below shows the number distribution of galaxies in each magnitude bin. [1]	11
2.1	The distribution of various parameters for the halo catalog. We associate GW events with these selected halos.	14
2.2	Power Spectral Density	16
2.3	Distribution of Masses of objects for events with $SNR > 12$. Here, 'object' refers to the individual celestial bodies of the binary system undergoing merger events.	17
2.4	Distribution of inclination angle of events with $SNR > 12$	18
2.5	Distribution of redshift and luminosity distance for events with $SNR > 12$	18
2.6	Distribution of dimensionless spins of objects for events with $SNR > 12$	19
2.7	The histograms show the distribution of single parameters, and the scatter plots show the pairwise relationship between them (i.e., the joint distribution and how they are correlated). From the above plot, it can be clearly seen that the joint distribution between luminosity distance and inclination angle is narrow and elongated. This signifies high correlation between these two parameters.	20
3.1	A simulated 3D image of the Cosmic Web, the vast distribution of matter and energy across the universe. (<i>P. Volker Springel/Max Planck Institute for Astrophysics</i>)	21
3.2	Defining the mass cut	22
3.3	The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 0 to 200 $\text{Mpc } h^{-1}$ along the z-direction. The x and y axis spans the entire width of the box i.e. 2500 $\text{Mpc } h^{-1}$. The second plot is just a small patch from the above plot extending from 0 $\text{Mpc } h^{-1}$ to 400 $\text{Mpc } h^{-1}$ along both x and y directions in the box	26
3.4	The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 1000 to 1200 $\text{Mpc } h^{-1}$ along the z-direction. The x and y axis spans the entire width of the box i.e. 2500 $\text{Mpc } h^{-1}$. The second plot is just a small patch from the above plot extending from 0 $\text{Mpc } h^{-1}$ to 400 $\text{Mpc } h^{-1}$ along both x and y directions in the box	27

3.5	The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 2000 to 2200 $\text{Mpc } h^{-1}$ along the z-direction. The x and y axis spans the entire width of the box i.e. 2500 $\text{Mpc } h^{-1}$.The second plot is just a small patch from the above plot extending from 0 $\text{Mpc } h^{-1}$ to 400 $\text{Mpc } h^{-1}$ along both x and y directions in the box.	28
3.6	Distribution of various parameters for the populated galaxies.	29

Chapter 1

Introduction

In the 1920s, Edwin Hubble confirmed that the Universe was expanding by showing that the recession velocity of galaxies increased with their distance from the observer. While this fact has long been established, determining the speed at which the Universe is expanding (known as the Hubble constant H_0) has been a subject of considerable research and debate in astrophysics.

Initially, Hubble found the value of H_0 to be about 500 km/s/Mpc, which was highly inaccurate as Cepheid variables used to determine distances exhibit diversity and the Type II Cepheids that Hubble used had a different period-luminosity relation compared to the Type I Cepheids in the Milky way for which this relation was known. As technology as well as new methods have been developed, the uncertainty in the measurements has been reduced to a few percent, which has enabled precision studies of the expansion of the Universe. A discrepancy as to the exact value of the Hubble constant has manifested itself as different measurements produce results that disagree with a large statistical significance. The early-universe measurements (using Cosmic Microwave Background, etc.) favor a slightly smaller value of the Hubble constant than the late-time measurements based on the distance ladder (also called local measurements). This discrepancy is known as Hubble tension, and there are two explanations: Either our measurement techniques aren't precise enough, or our knowledge of the Universe's and its large-scale structure is incomplete.

Two things are required to measure the value of H_0 : the distance to an object (e.g., a galaxy) and the speed at which it moves away from us due to the universe's expansion (called the recession velocity). The recession velocity can be determined by measuring the objects' redshift using spectroscopic techniques. However, measurements of distances pose another problem since there is no single way of measuring them at all scales. So, for this purpose, the *cosmic distance ladder*, a combination of techniques that measure distances at different scales in the Universe, has to be used. Luminosity distances can be measured using *standard candles*, objects of known brightness, such as Cepheids or Type Ia supernovae. By measuring the flux of the source and knowing its intrinsic brightness, its luminosity distance can be calculated.

$$Flux \propto \frac{luminosity}{(distance)^2} \quad (1.1)$$

However, this requires us to determine the object's intrinsic brightness (luminosity).

Geometric measurements of distances to known objects using methods such a parallax can be used to calibrate the intrinsic brightness of objects.

Gravitational Wave events do not have to rely on such a distance ladder. Theorized back in 1915 through Einstein's General Relativity and detected 100 years later in 2015, Gravitational waves are ripples in space-time caused by extreme events like binary mergers in the Universe. Compact Binary Coalescence (CBCs), where two compact objects spiral into each other and merge (like two black holes or neutron stars), emit gravitational waves such that the amplitude of the signal as measured on Earth is inversely proportional to their luminosity distance from Earth. These GW sources are called standard sirens because the shape of the waveform is dictated by the intrinsic parameters of the merging black holes, thereby providing an absolute calibration of the expected amplitude as a function of the distance. GW170817, a gravitational wave event that happened due to the merger of a binary neutron star system, was observed to have an electromagnetic counterpart. The event thus yielded a measurement of the luminosity distance as well as the redshift, which resulted in the first direct measurement of the Hubble constant using gravitational wave events.

1.1 The Standard Model of Cosmology

1.1.1 Cosmological Principle

The cosmological principle states that the Universe is homogeneous and isotropic on large scales. Although there is clumpiness on small scales but, when averaged over sufficiently large scales, every point and every direction is identical.

1.1.2 The Expanding Universe

The Hubble-Lemaître law states that the recessional velocity of objects within the Universe is directly proportional to their distance from Earth, the proportionality constant being the Hubble Parameter.

$$v = H_0 d \quad (1.2)$$

where d is the proper distance between the observer and the object. The distance is usually measured in Mpc , and H_0 is the present-day value of the Hubble Parameter called the Hubble constant and has the units of $kms^{-1}Mpc^{-1}$.

The expansion of the Universe suggests that in the past, it was much smaller than it is today. This leads to the Big Bang Theory, which says that the Universe started with a violent and energetic explosion from a singularity and has been expanding ever since. The Λ CDM (Lambda Cold Dark Matter) is the most widely accepted cosmological model, which begins with the Big Bang and eventually evolves into a universe filled with cold, dark matter, and ordinary matter. Λ , the cosmological constant, indicates the presence of dark energy, which works against gravity's pull to accelerate the Universe's expansion.

1.1.3 The Scale Factor

Since in an expanding universe, every point moves away from each other (neglecting the local effects of gravity), we can describe the Universe by using comoving coordinates, which expand with the Universe. In this coordinate system, the objects have a fixed coordinate irrespective of the expansion. The scale factor $a(t)$ defines the translation between the comoving coordinates and the proper distance. Proper distance (R_{phy}) is the actual distance to an object and changes with time. It is given by:

$$R_{phy}(t) = a(t)R_{com} \quad (1.3)$$

where R_{com} is the comoving distance and $a(t)$ is the value of the scale factor at time t . The Hubble-Lemaître law can be derived from equation (1.2) by taking its time derivative.

$$\frac{d}{dt}(R_{phy}) = \frac{d}{dt}[a(t)R_{com}] \quad (1.4)$$

But

$$\begin{aligned} \frac{d}{dt}(R_{phy}) &= v \\ \implies v &= a(t)\dot{R}_{com} + \dot{a}(t)R_{com} \\ \implies v &= a(t)\dot{R}_{com} + \left[\frac{\dot{a}(t)}{a(t)}\right]R_{phy} \quad (\text{from (1.3)}) \end{aligned}$$

Since for any fundamental observer moving with the Hubble flow, $\dot{R}_{com} = 0$, so:

$$\begin{aligned} v &= \left[\frac{\dot{a}(t)}{a(t)}\right]R_{phy} \\ \implies v &\propto R_{phy} \end{aligned}$$

The proportionality constant is the Hubble Parameter, given by:

$$H = \frac{\dot{a}(t)}{a(t)} \quad (1.5)$$

The value of the Hubble Parameter today is called the Hubble Constant.

1.1.4 Friedmann Equations

The appearance of objects at cosmological distances is affected by the curvature of space-time through which light travels on its way to Earth. Einstein's general theory of relativity provides the most complete description of the geometrical properties of the Universe. In GR, the fundamental quantity is the metric that describes space-time geometry.

Howard P. Robertson and Arthur Geoffrey Walker independently demonstrated in the mid-1930s that the most general metric possible for describing an expanding, homogeneous, and isotropic universe is:

$$ds^2 = -c^2 dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right] \quad (1.6)$$

where k is a constant representing the curvature of space-time.

$$k = \begin{cases} 1 & \text{spherical geometry,} \\ 0 & \text{flat geometry} \\ -1 & \text{hyperbolic geometry (open Universe)} \end{cases}$$

The above metric is also known as the Friedmann–Lemaître–Robertson–Walker metric (FLRW). Solving Einstein's equations assuming a FLRW metric gives rise to the Friedmann equations.

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \quad (1.7)$$

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \quad (1.8)$$

where a is the scale factor, ρ is the density of the universe, p is the pressure, k is the curvature, and Λ is the cosmological constant.

Considering a flat universe with $\Lambda = 0$ to make the cosmological inferences, we can define the density parameter Ω_m , which is the ratio of the density of the Universe (ρ) to its critical density (ρ_c).

$$\Omega_m = \frac{\rho}{\rho_c} \quad (1.9)$$

The critical density is the density for which the Universe is flat today. Considering $a = 1$ today and $\frac{\dot{a}}{a} = H$, and solving (1.7), we get:

$$\rho_c = \frac{3H^2}{8\pi G} \quad (1.10)$$

Rearranging (1.7), using (1.10), and substituting in the expression of Ω_m , we get:

$$\Omega_m = 1 - \frac{\Lambda c^2}{3H^2} + \frac{kc^2}{a^2 H^2} \quad (1.11)$$

Solving the above equation further, we get the curvature density parameter and the dark-energy density parameter, defined as:

$$\Omega_k = -\frac{kc^2}{a^2 H^2} \quad \text{and} \quad \Omega_\Lambda = \frac{\Lambda c^2}{3H^2} \quad (1.12)$$

It also gives:

$$\Omega_m + \Omega_k + \Omega_\Lambda = 1 \quad (1.13)$$

Due to the dependence of these density parameters on the Hubble Parameter, they will change with time as the Universe expands. The present-day values of these parameters are: ($a_0 = 1$ today)

$$\Omega_{m,0} = \frac{\rho_0}{\rho_{c,0}} = \frac{8\pi G}{3H_0^2} \rho_0, \quad \Omega_{k,0} = \frac{-kc^2}{a_0^2 H_0^2} = \frac{kc^2}{H_0^2} \quad \text{and} \quad \Omega_{\Lambda,0} = \frac{\Lambda c^2}{3H_0^2} \quad (1.14)$$

Again, rewriting equation (1.7), dividing by H_0^2 , and substituting the present-day density parameters, we get:

$$\left(\frac{H}{H_0}\right)^2 = \frac{\rho}{\rho_0} \Omega_{m,0} + \frac{1}{a^2} \Omega_{k,0} + \Omega_{\Lambda,0} \quad (1.15)$$

We must rephrase our equations regarding the parameters we need to infer. Since the Universe is expanding, the objects' spectra are redshifted, which is related only to the scale factor of the Universe.

$$1 + z = \frac{a(t_0)}{a(t_e)} \quad (1.16)$$

where $a(t_e)$ is the scale factor of the Universe when light from the object was emitted, and $a(t_0)$ is the scale when the light was observed. Since $a(t_0) = 0$ at present, hence (1.16) can be written as:

$$a(t) = \frac{1}{1+z} \quad (1.17)$$

The continuity equation is given by :

$$\frac{d}{dt}(\rho c^2 a^3) + P \frac{d}{dt}(a^3) = 0 \quad (1.18)$$

For a matter-dominated universe: $P \ll \rho c^2$. Using this in equation (1.18), we get $\rho_m \propto a^{-3}$. Hence, the present density of the Universe is related to the density at any time t as:

$$\frac{\rho}{\rho_0} = \frac{1}{a(t)^3} \quad (1.19)$$

Using (1.17) and (1.19) and solving (1.15), we get the value of the dimensionless Hubble Parameter, given by $E(z)$ and expressed as:

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_{\Lambda,0}} \quad (1.20)$$

Comoving Distance (D_C)

It is defined as the distance between two objects that remain constant with epoch if they move with the Hubble flow. The total line-of-sight comoving distance D_C from us to a distant object is computed by integrating the infinitesimal δD_C contributions between nearby events along the radial ray from $z = 0$ to the object. It is given by:

$$D_C = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} \quad (1.21)$$

Luminosity Distance (D_L)

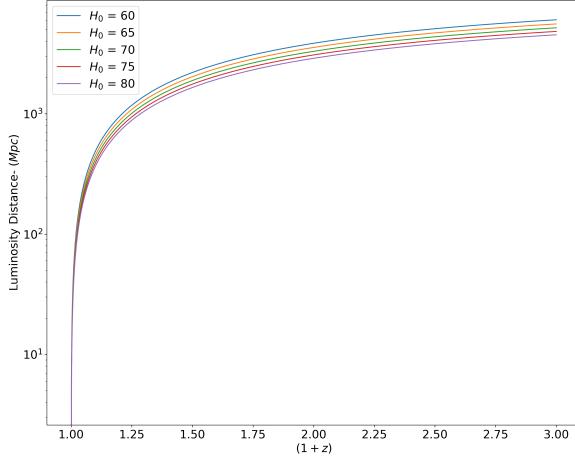
The *luminosity distance* D_L is defined by the relationship between bolometric (i.e., integrated over all frequencies) flux F and bolometric luminosity L , given by:

$$D_L = \sqrt{\frac{L}{4\pi F}} \quad (1.22)$$

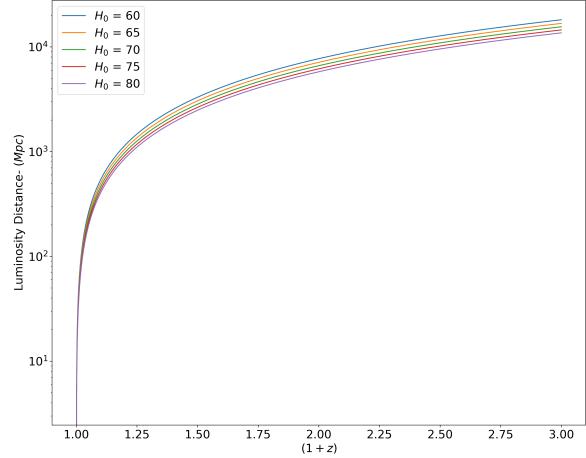
It is also related to the comoving distance by the relation:

$$D_L = (1+z)D_C \quad (1.23)$$

$$\implies D_L = \frac{(1+z)c}{H_0} \int_0^z \frac{dz'}{E(z')} \quad (1.24)$$



(a) Variation of Comoving Distance with changing Redshift



(b) Variation of Luminosity Distance with changing Redshift

Figure 1.1: The plots show how distances vary in an expanding universe for different values of H_0 ($\text{km s}^{-1} \text{Mpc}^{-1}$)

1.1.5 Galaxy Luminosity Function

It is a statistical measurement to describe the distribution of galaxies in terms of their luminosities within a particular volume of space. The Luminosity Function, $\Phi(L)$ describes the relative number of galaxies of different luminosities. Counting the number of galaxies in a unit comoving volume of the universe, $\Phi(L)dL$ is the number of galaxies between L and $L + dL$. An approximation to the luminosity function was suggested by Paul Schechter in 1976 [3]:

$$\Phi(L)dL = n_\star \left(\frac{L}{L_\star} \right)^\alpha \exp\left(-\frac{L}{L_\star}\right) \quad (1.25)$$

where n_\star is the number density of galaxies (specifically the number of galaxies per Mpc), L_\star is a characteristic galaxy luminosity. Any galaxy with luminosity L_\star is bright. The number of galaxies falls sharply for $L > L_\star$. Galaxies with $L < 0.1L_\star$ is considered as a dwarf. The last parameter α defines the "faint end slope" of the luminosity function.

Equation (1.25) can also be written in terms of absolute magnitude M as:

$$\Phi(M) = 0.4 \ln(10) \Phi_\star 10^{-0.4(M-M_\star)(\alpha+1)} \exp[-10^{-0.4(M-M_\star)}] \quad (1.26)$$

1.2 Objective

The advanced LIGO and VIRGO detectors have been detecting GW events since 2015, which mainly results from CBCs. The intrinsic property of a GW signal from a CBC to encode within it the luminosity distance between the source and the observer makes GWs hugely interesting for astrophysical and cosmological reasons. Since most conventional methods to measure cosmic distances rely on the cosmic distance ladder,

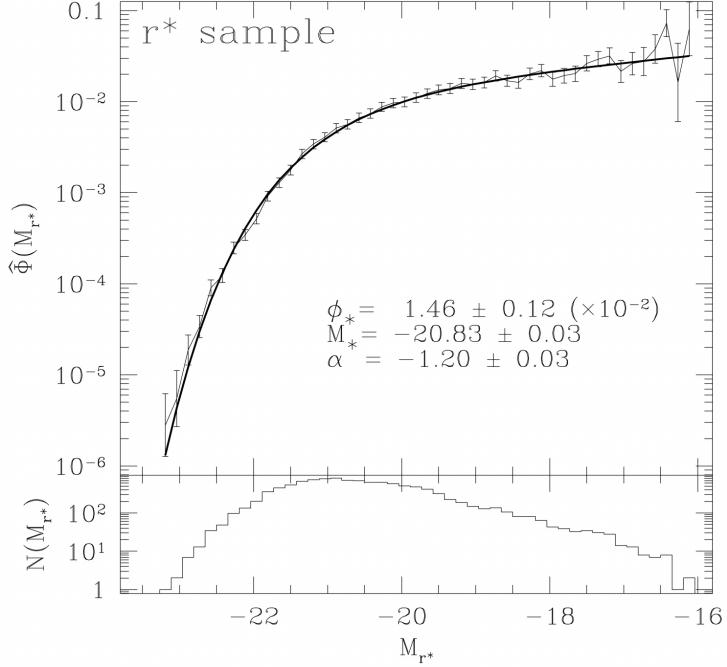


Figure 1.2: The figure shows luminosity function of galaxies for a particular band taken from SDSS data. The thick line is the Schechter fit and the other one is the data points. The histogram below shows the number distribution of galaxies in each magnitude bin. [1]

which poses several problems, the GWs provide a unique way to make a direct measurement completely independent of the cosmic distance ladder or any EM distance measurement.

Binary merges are of two types, one with an electromagnetic counterpart (Binary Neutron Star Mergers, BNS) and the other completely dark in the entire EM spectrum (also called dark sirens, Binary Black Hole Mergers, BBH, for example). The one with an EM counterpart can be well localized in the sky and linked to its host galaxy. The host galaxy's redshift can then be combined with the distance estimate from the BNS to measure H_0 .

On the other hand, if it is not possible to identify the EM counterpart, then the GW event provides a localization volume, and all the galaxies lying within that volume are considered as potential hosts. Out of all the galaxies in the catalog, the true host galaxy have a redshift that, when combined with the distance from the GW event, gives the actual value of H_0 , while the other galaxies give potential values of H_0 . Using multiple GW events, the contribution from true hosts stack over time, and the contributions from non-host galaxies statistically average out.[2]

Why use dark sirens?

The vast majority of events detected so far correspond to mergers of massive binary black holes (BBH) compared to binary neutron star mergers. The difference is that of an order of magnitude. These BBH mergers do not emit any electromagnetic counterpart; thus, the cosmological inference is only possible statistically.

What if the host galaxy is not in the catalog?

Since galaxy observations are resolution-limited, very faint and high-redshifted galaxies may not be present in the catalog, so many GW events cannot be associated with the host galaxy. In such a case, getting the redshift information is not easy.

This project aims to address this problem and determine the value of Hubble's constant. Since we know the galaxy clustering information and the large-scale structure above the resolution limit, by using this, we will develop a model to statistically populate the fainter galaxies below the resolution limit such that they follow the same large-scale structure. We will then look at how the galaxies are clustered in this regime and then use this clustering information to determine the value of different cosmological parameters using the abovementioned method.

Chapter 2

Methodology

This work can be divided into many different subsections, which include:

- Catalog Preparation and Simulation of GW Events
- Galaxy Population
- Cosmological Inference

This chapter focuses on the first part of this project i.e, how to prepare the catalog, simulate the GW events and estimate posterior distribution of various GW event parameters.

2.1 Catalog Preparation and Simulating GW Events

The Cosmosim Database provides results from cosmological simulations performed with different projects. We use the `bigmdpl.rockstarz_0` catalog of dark matter halos and then associate GW events to it by putting appropriate cuts on the data.

Some details of the simulation are as follows:

Box Size	$2.5 \text{ Gpc}/h$
Number of Particles	3840^3
Hubble Parameter, h	0.6777
Ω_Λ	0.692885
Ω_m	0.307115

Table 2.1: Simulation Details

The initial reference point for the positions of the halos in the simulation box is one of the box's corners. So, to have a uniform distribution of the halos in all directions, we shift this reference point to the center of the box and consider the halos lying within a sphere of $1.25 \text{ Gpc } h^{-1}$. RA and Dec of objects are simply the azimuthal angle (ϕ) and the compliment of the polar angle ($\pi/2 - \theta$) of the spherical coordinate system with its origin at the center of the box. The ' r ' coordinate is the comoving distance of the halos.

To get the redshift, the comoving distances for a range of redshifts have been calculated using the `astropy.cosmology` package. Interpolation of these distances establishes the relationship between redshift and the comoving distance. This relationship is then used to get the redshift of the halos from the catalog.

Since the halos are distributed uniformly in all directions, RA and Dec follow the distribution shown in the plot. We have selected only halos with a maximum velocity greater than a certain limit (180 km/s) due to the finite resolution of the simulation (corresponding to heavy-mass halos). This is because low-mass halos cannot be resolved properly in such simulations. The redshift distribution is uniform in comoving volume.

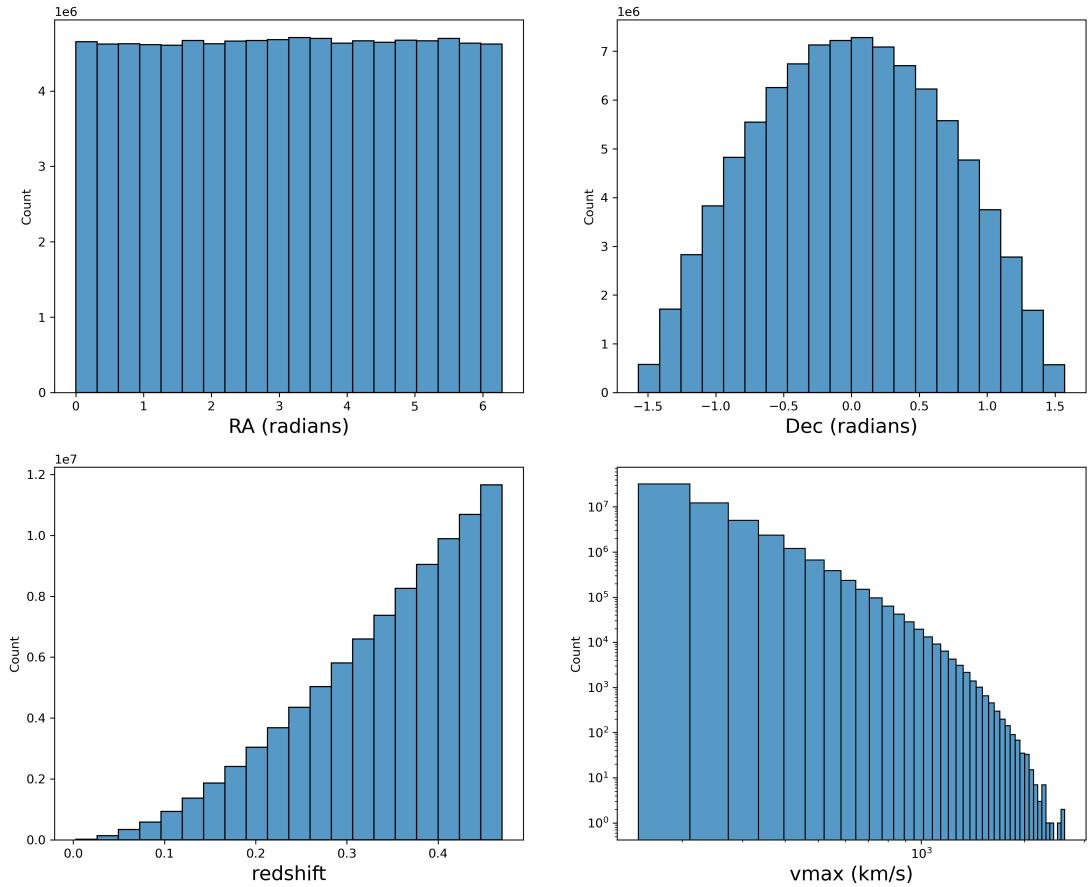


Figure 2.1: The distribution of various parameters for the halo catalog. We associate GW events with these selected halos.

2.1.1 Associating GW Events with Hosts:

After applying the above cuts and preparing the catalog, we randomly select some halos from it and associate GW events (BBH mergers) with each. This is done by associating

a prior distribution of parameters to each randomly selected halo. Luminosity distance for each selected Halo has been calculated using the formula: $cz = H_0 d_L$, considering $H_0 = 67.7 \text{ km/s/Mpc}$, and then only those have been selected for which $d_L < 1000 \text{ Mpc}$. This reduced total number of events to 678. After the catalog preparation, we need to determine the posterior distribution of parameters.

2.1.2 Some Bayesian Statistics

The main aim of Bayesian inference is to obtain the posterior distribution of parameters θ , given the data d associated with the parameter.[4] The posterior distribution is the probability density function for the variable θ given the data d . It is given by:

$$p(\theta|d) \quad \text{also:} \quad \int d\theta p(\theta|d) = 1 \quad (2.1)$$

For a CBC, θ can be considered as the 15 parameters, and d is the strain data obtained from a network of Gravitational Wave Detectors. According to Bayes theorem:

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}} \quad (2.2)$$

where $\mathcal{L}(d|\theta)$ is the likelihood function of the data given the parameters θ , $\pi(\theta)$ is the prior distribution of the parameters and \mathcal{Z} is the evidence given by:

$$\mathcal{Z} \equiv \int d\theta \mathcal{L}(d|\theta)\pi(\theta) \quad (2.3)$$

Here \mathcal{Z} acts as the normalization factor.

Likelihood function is a description of the measurement, and by writing it down, we implicitly introduce a noise model. For gravitational wave astronomy, it is given as:

$$\mathcal{L}(d|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(d - \mu(\theta))^2}{\sigma^2}\right) \quad (2.4)$$

where $\mu(\theta)$ is the template for the gravitational strain waveform given θ and σ is the detector noise. $\mu(\theta)$, the template, refers to a theoretical model or waveform representing the expected shape of a gravitational wave signal emitted by a specific astrophysical source, such as the merger of compact binary systems. It is a function of the parameters θ .

2.1.3 bilby - The Bayesian Inference Library and finding Posterior

Bilby is used to perform parameter estimation. It is primarily designed and built to infer CBC events in interferometric data. We have used this to estimate the posterior for our model data.

Noise Curves and SNR Calculation for Events

Noise curves, also known as power spectral densities (PSDs), represent the frequency dependence of the noise in gravitational-wave detectors. Understanding noise curves is crucial in detecting gravitational-wave signals because they help characterize the detector's sensitivity across different frequencies. Noise curves aid in computing the SNR of detected signals. They directly affect the shape and scale of the likelihood function in Bayesian analysis.

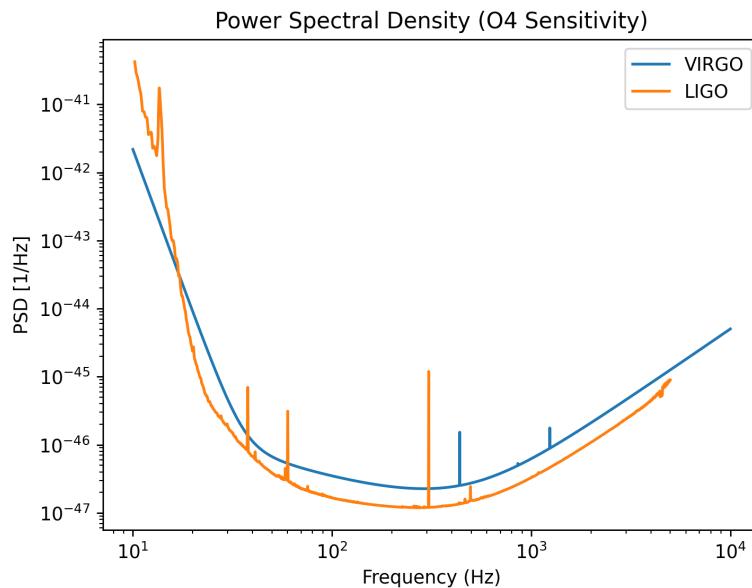


Figure 2.2: Power Spectral Density

Using Bilby, we calculate each event's network SNR from our model data by providing the Noise Curves corresponding to detectors. Only those events are selected for which $\text{SNR} \geq 12$. Then we calculate the posterior for each parameter by injecting our priors into the detectors.

Parameters for a GW Event

- **mass:** The masses of each compact object undergoing merger in solar masses
- **a:** Dimensionless spin
- **tilt:** The angle in radians between spin angular momentum and orbital angular momentum
- **phi_12:** Relative azimuthal angle between the spins of the two objects
- **phi_jl:** Angle between the total angular momentum and the orbital angular momentum

- **luminosity_distance**: Luminosity distance to the source in Mpc
- **theta_jn**: Angle between the line of sight and the total angular momentum vector
- **phase**: GW phase at reference time
- **RA and Dec**: Location of the source on celestial sphere
- **geocent_time**: Time of the event's arrival at the geo-center in GPS seconds
- **psi**: Polarization angle of the gravitational wave

2.2 Parameter Estimation

Total **678** events have been recognized for which $d_L < 1000\text{Mpc}$. Since LIGO detectors are noise sensitive, we have then taken only those events for which **SNR > 12**. After applying this SNR cut, only **347** Binary Black Hole merger events remain. We now run the parameter estimation code on this.

The prior distribution of different parameters for all events are shown in the plots below:

1. Mass Distribution for both binaries

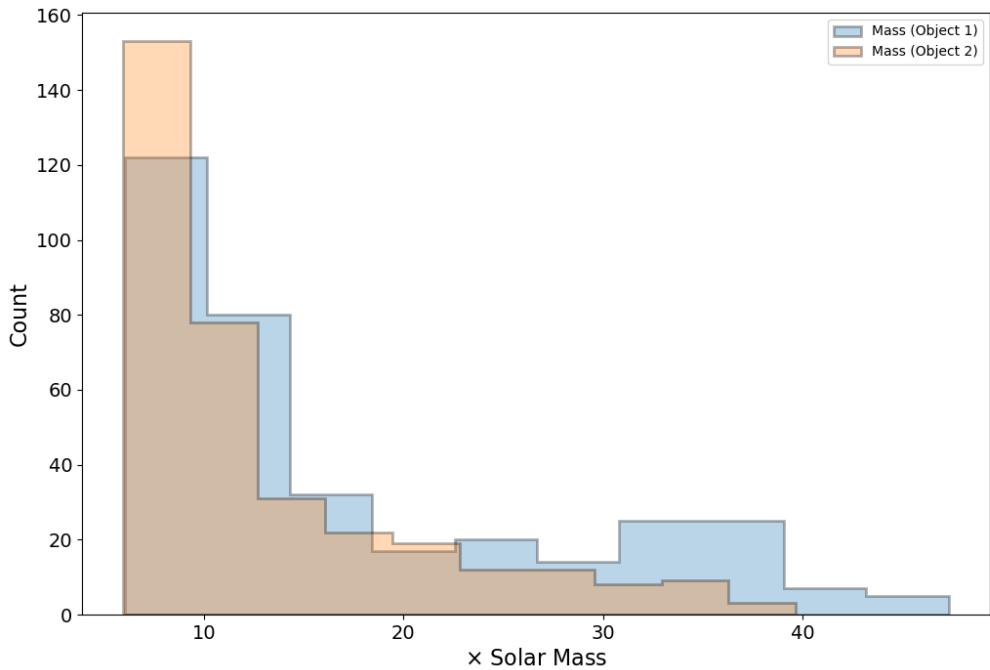


Figure 2.3: Distribution of Masses of objects for events with $\text{SNR} > 12$. Here, 'object' refers to the individual celestial bodies of the binary system undergoing merger events.

2. Distribution of Inclination Angle

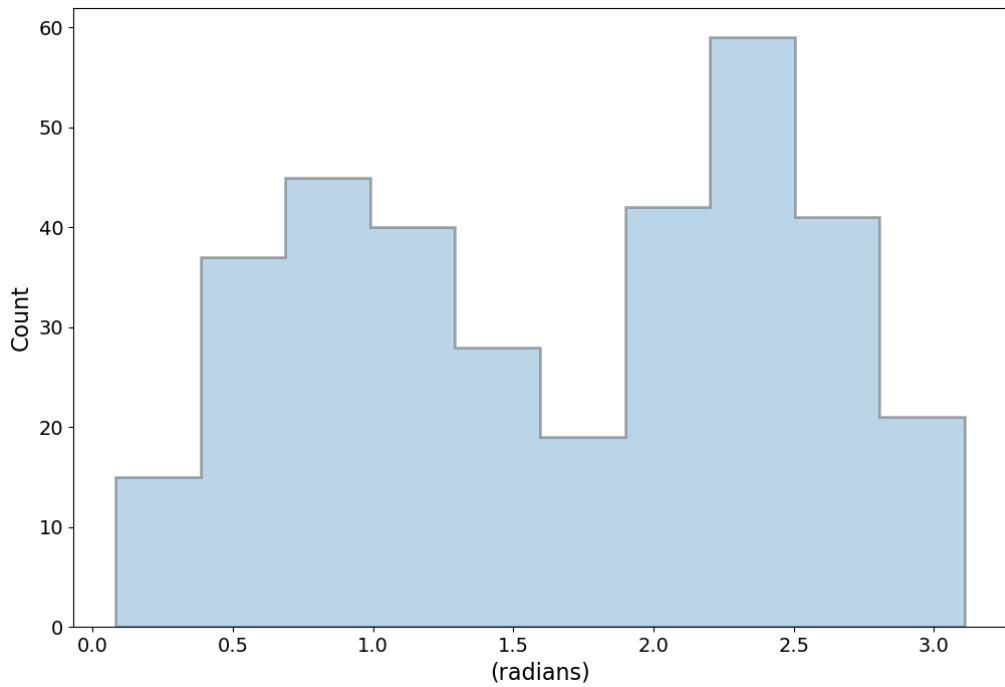


Figure 2.4: Distribution of inclination angle of events with $SNR > 12$

3. Distribution of Redshift and Luminosity Distance

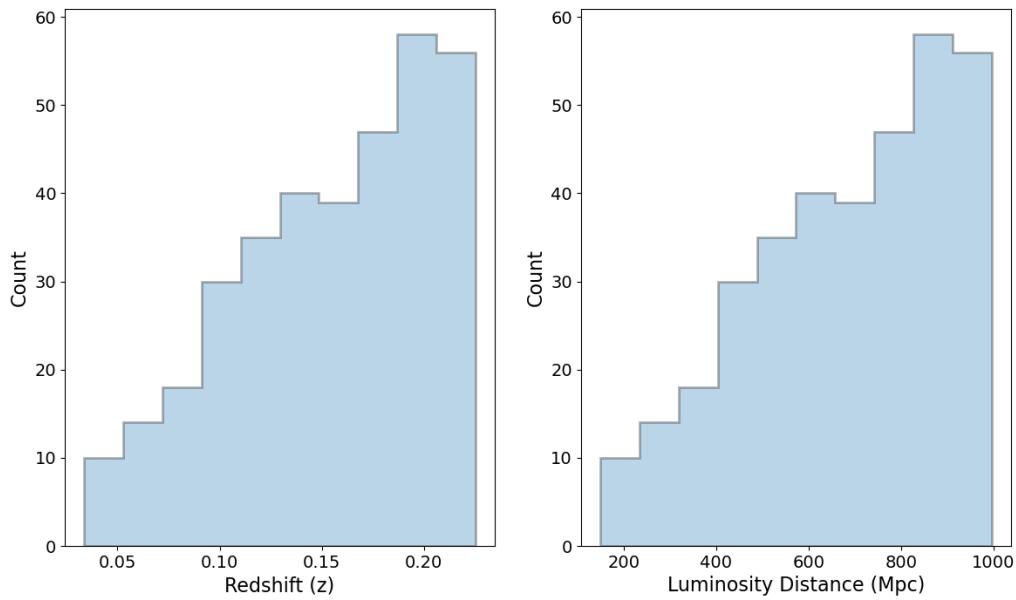


Figure 2.5: Distribution of redshift and luminosity distance for events with $SNR > 12$.

4. Distribution of Dimensionless Spins

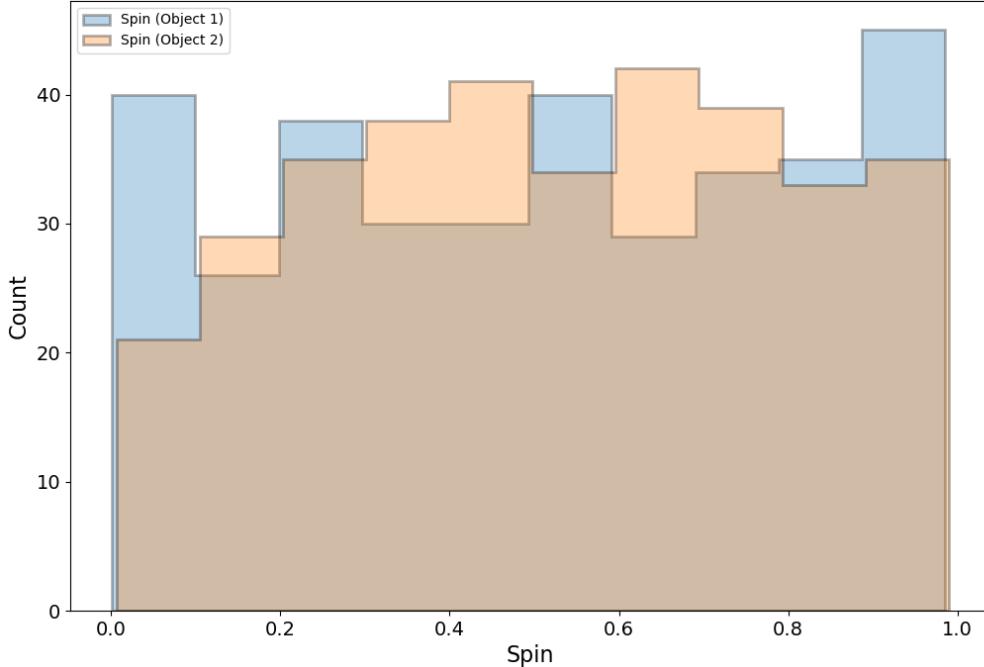


Figure 2.6: Distribution of dimensionless spins of objects for events with $SNR > 12$

Once we have our events satisfying the threshold on SNR, we then run the parameter estimation code of bilby to find how different parameters are distributed in the given volume of space from where the event originated. This will give us the posterior distribution of parameters corresponding to each event. The output generates a JSON file containing inferred values of different parameters as well as injected values. It also generates a plot showing joint distribution between different parameters using those inferred values. Some more plots are also generated, which do not have any significance here.

2.2.1 Posterior

Parameter estimation using Bilby produces corner plots that display the joint distribution of parameters. Here is an example of a corner plot showing the joint distribution of various parameters corresponding to a CBC of two black holes with 26.52 and 15.17 solar masses from the catalog we created.

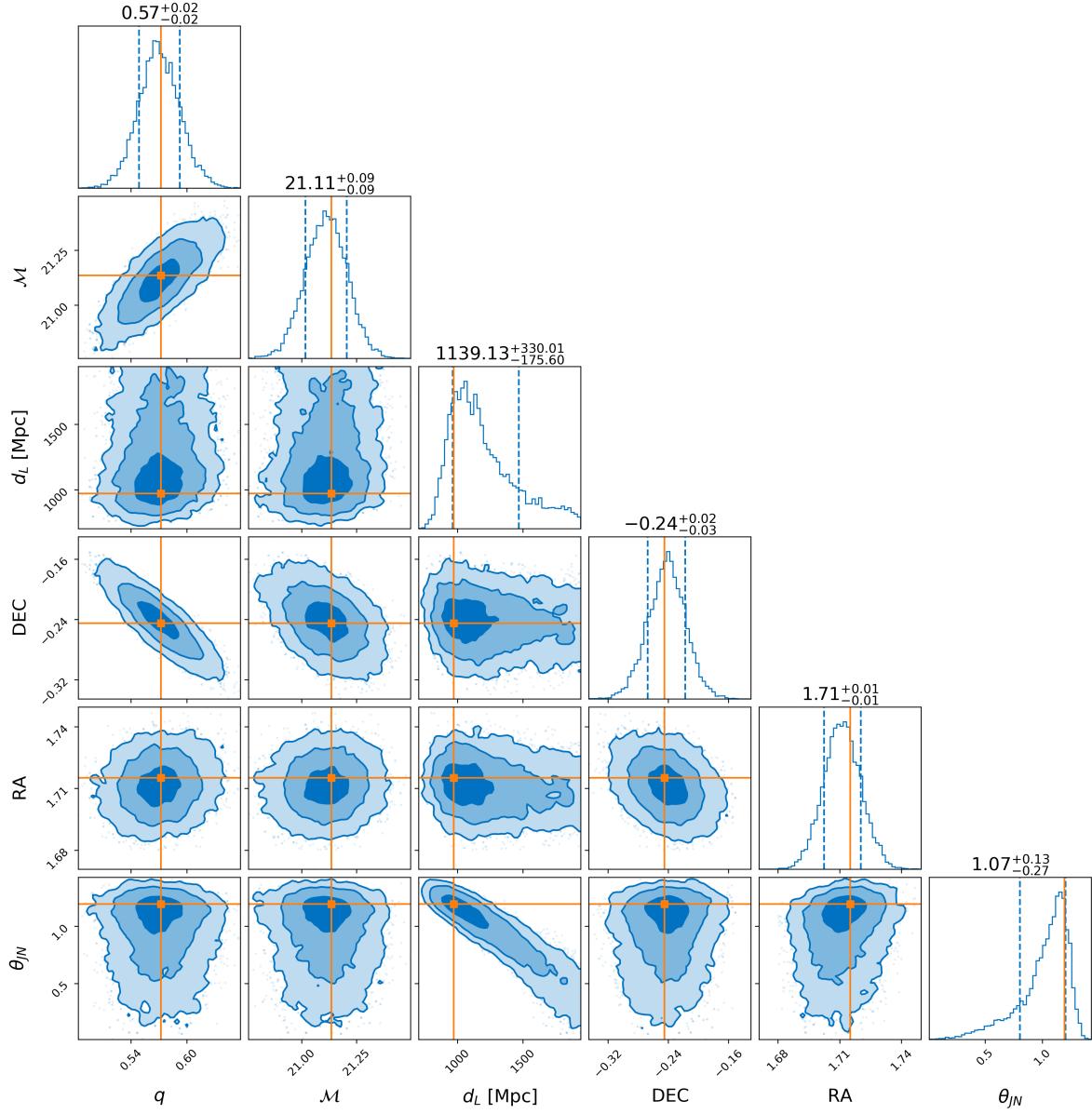


Figure 2.7: The histograms show the distribution of single parameters, and the scatter plots show the pairwise relationship between them (i.e., the joint distribution and how they are correlated). From the above plot, it can be clearly seen that the joint distribution between luminosity distance and inclination angle is narrow and elongated. This signifies high correlation between these two parameters.

Chapter 3

Galaxy Population

In this chapter, we develop a method to populate the galaxies below a certain threshold, given how they form structures above this. The threshold is in the form of a mass cut, which is constant over all redshifts.

3.1 The Idea

The basic idea behind this method is to look at the galaxy/halo clustering above the defined threshold limit. Since the large-scale structure (cosmic web) remains the same, we will use this fact to estimate the galaxy clustering information below the set limit. For this, we will look at how all the galaxies below the limit are biased with respect to those above. While going through this, we will use the fact that the halos and galaxies follow the same distribution and their average number density are equal. We will apply all of this to a simulated catalog.

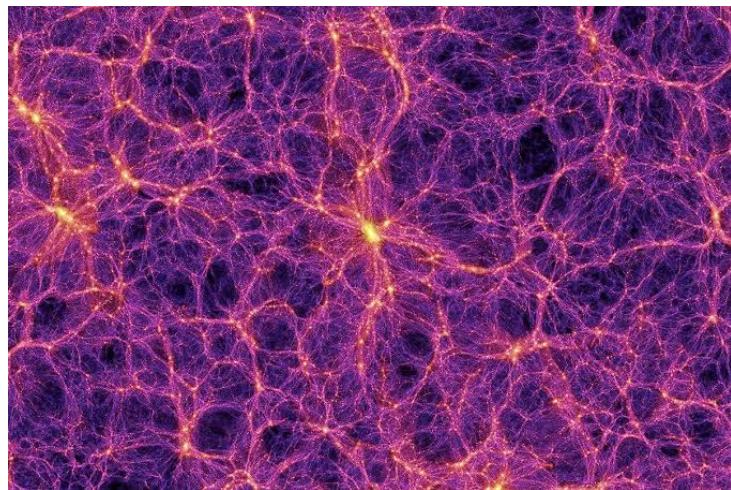


Figure 3.1: A simulated 3D image of the Cosmic Web, the vast distribution of matter and energy across the universe. (P. Volker Springel/Max Planck Institute for Astrophysics)

3.2 The Mass Cut

We have defined a constant mass cut (same for all redshifts) and have considered all the halos below to be absent from the catalog. The catalog spans 5 orders of magnitudes of masses, from $4.7 \cdot 10^{10} M_{\odot}$ to $5.4 \cdot 10^{15} M_{\odot}$. Since low-mass halos cannot be resolved in such simulations, we have only selected halos above $M_{min} = 10^{13} M_{\odot}$, and then defined the mass cut to be at $M_{th} = 10^{14} M_{\odot}$.

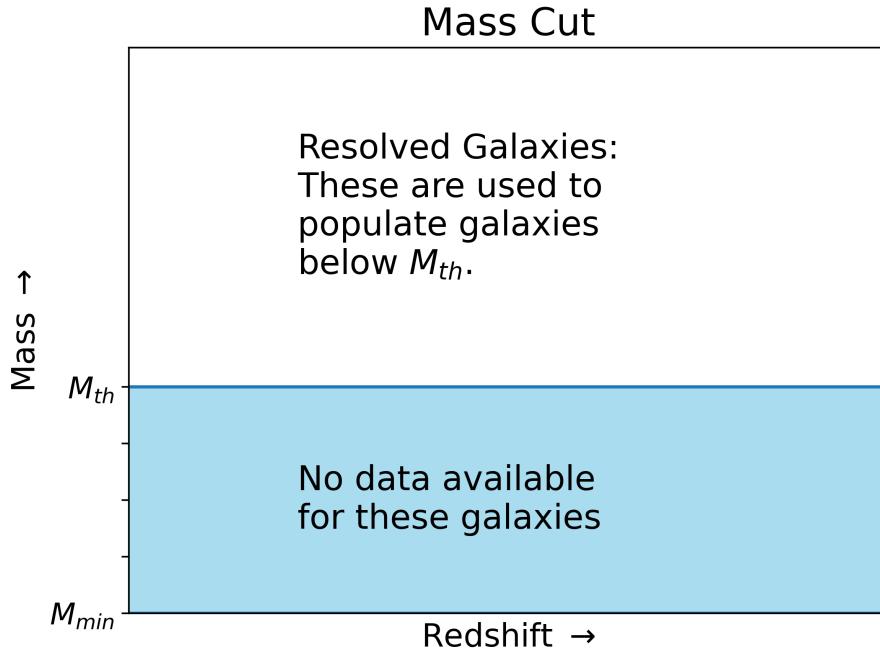


Figure 3.2: Defining the mass cut

3.3 Density Field

The density field of halos refers to their spatial distribution across the universe. In cosmology, halos are not uniformly distributed; they clump together to form structures on various scales, such as clusters, filaments, and voids. The density field characterizes the variations in the number of halos per unit volume of space.

To calculate the density field from our catalog, we first divide the entire box into small cubes such that the number of such cubes along each axis is 128 to get a better resolution. This way, the size of each small cube is $2500/128 \text{ Mpc } h^{-1}$ on each side. Now, we count the number of halos residing in each small cube and divide it by the volume of the respective cube. Now we have a box with 128 cubes on each side, each containing the number density of halos. This gives the density field of the box.

The volume of each small cube, $V = (2500/128)^3 \text{ Mpc}^3 h^{-3}$. Let N_{ijk} be the number of halos in each of the cubes. Then, number density, $n_{ijk} = N_{ijk}/V$, where all the three indices i, j and k run from 1 to 128.

3.4 Overdensity

It is conventional in cosmology to refer to structures by the density contrast they represent with respect to the critical density of the universe. Overdensity refers to the deviation of the density at a point from the average density. It is defined with respect to the mean density \bar{n} . It is given by:

$$\delta = \frac{n - \bar{n}}{\bar{n}} \quad (3.1)$$

where n is the number density of a specific region. If $\delta < 0$, the region is underdense; if $\delta > 0$, it is overdense; and if $\delta = 0$, then it has a density equal to the mean density.

We calculate the overdensity for each of the small cubes of the box by first calculating the mean number density (\bar{n}) of the box. This is done by counting the total number of galaxies divided by the total volume of the entire box, i.e., $V = 2500 \times 2500 \times 2500 \text{ Mpc}^3 h^{-3}$. Then, the overdensity in each box is:

$$\delta_{ijk} = \frac{n_{ijk} - \bar{n}}{\bar{n}} \quad ; \quad i, j, k = 1, 2, 3, \dots, 128 \quad (3.2)$$

Now, we make a heatmap plot of $\log_{10}(1 + \delta)$ to get the structure. Since the length of each of the cubes is only $2500/128 \text{ Mpc } h^{-1}$, there is a chance that one of the cubes does not contain any halo; hence, in this case, the halo overdensity in that particular cube is -1. This will give $\log_{10}0$, which is undefined. So, to address this issue, we stack cubes up to a length of 200 Mpc (10 cubes along a particular direction) and then take their average overdensity to make the plots.

3.5 Bias

Now, to populate the galaxies, we need to find out the the galaxies below the limit are biased with those above it. Galaxy bias, in general, is defined by the relationship between the spatial distribution of galaxies and the underlying dark matter density field. It is given by:

$$b = \frac{\int_{M1}^{M2} n(M, z)b(M, z)dM}{\int_{M1}^{M2} n(M, z)dM} \quad (3.3)$$

where;

$n(M, z)$: number density of halos with a given mass at a redshift.

$b(M, z)$: large-scale linear halo bias at a given redshift and mass.

Linear Bias: It is called "linear" because it's often assumed to hold on large scales where linear perturbation theory applies, meaning that the density fluctuations are small enough that they evolve approximately linearly with cosmic time. If b is greater than 1, galaxies are more clustered than underlying dark matter distribution.

The above integral can also be written in the following form:

$$b = \frac{\int_{\log M_1}^{\log M_2} M \cdot n(M, z) b(M, z) dM}{\int_{\log M_1}^{\log M_2} M \cdot n(M, z) dM} \quad (3.4)$$

where $M = 10^{\log M}$. The overdensity of galaxies is then given by:

$$\delta_g = b \cdot \delta_h \quad (3.5)$$

To populate the galaxies below the limit, we need to know the bias of the unresolved galaxies with respect to that of galaxies below the limit. For this we need to calculate the bias for both regions and then get the relative bias by taking ratios of two. It can be done as:

$$b_{above} = \frac{\int_{\log M_{th}}^{\log M_{max}} M \cdot n(M, z) b(M, z) dM}{\int_{\log M_{th}}^{\log M_{max}} M \cdot n(M, z) dM} \quad (3.6)$$

$$b_{below} = \frac{\int_{\log M_{min}}^{\log M_{th}} M \cdot n(M, z) b(M, z) dM}{\int_{\log M_{min}}^{\log M_{th}} M \cdot n(M, z) dM} \quad (3.7)$$

where $M = 10^{\log M}$.

So, the relative bias is given as:

$$b_{rel} = \frac{b_{below}}{b_{above}} \quad (3.8)$$

Now, we have all the components to evaluate the number density of galaxies below the threshold limit, using the overdensity of halos above it. We assume that the overdensity of galaxies in both regions is same as that of halos (\bar{n}). This can be done as:

$$\begin{aligned} \delta_g &= b_{rel} \cdot \delta_h \\ \implies \frac{n_g - \bar{n}}{\bar{n}} &= b_{rel} \cdot \delta_h \\ \implies n_g &= \bar{n} \cdot (1 + b_{rel} \cdot \delta_h) \end{aligned} \quad (3.9)$$

This will give us the number density of the galaxies inside each box below the set threshold mass. Now, knowing the individual densities of the cube, we calculate their overdensity and then plot it to check how the new structure looks and compares with the original one.

3.6 Some Statistics

The box spans $2500 \text{ Mpc } h^{-1}$ on each side. After applying the cut on vmax and choosing M_{min} to be 10^{13} solar masses, the total number of halos remaining is 7349808. Out of all these halos, 6921639 of them lie below the set threshold limit of 10^{14} solar masses, and the rest ($7349808 - 6921639 = 428169$) lie above the limit, i.e., are the number of resolved halos.

After dividing the box into 128 smaller grids on each side (the grid size is $2500/128$ on each side), we calculate the number of halos in each. This number varies for different scenarios:

1. For halos above the limit, it varies from 0 to 11 in each grid with a mean value of 0.2041.
2. For halos below the limit, it varies from 0 to 41 with a mean value of 3.3004.
3. For the halos below plus above, it varies from 0 to 49 with a mean value of 3.5046.

The average number density of halos in each of the cases above is:

1. For $M_{min} < M < M_{th}$ it is 4.4298×10^{-4} halos/ Mpc^3/h^3 . We have assumed this number density to be equal to the number density of galaxies to be populated.
2. For $M > M_{th}$ it is 2.7402×10^{-5} halos/ Mpc^3/h^3
3. For the entire box, it is 4.7038×10^{-4} halos / Mpc^3/h^3 .

The value of bias as evaluated using the integrals on the previous page is coming out to be: **bias = 0.5040**.

For the populated galaxies using the algorithm discussed above:

1. Total Number of galaxies = 6921392
2. The number of galaxies in each grid ranges from 0 to 92
3. Mean Galaxy Density = 4.4429×10^{-4} galaxies/ Mpc^3/h^3

3.7 Plots

After calculating the overdensity, we now plot $\log(1 + \delta)$ for each of the following:

1. The overall overdensity of the simulation box
2. The overdensity of galaxies below the set limit after calculating the bias
3. The generated overdensity of galaxies by going through the steps mentioned above.

For all the plots, we have taken the average of overdensities up to a distance of 200 Mpc, each at a different distance from the set origin along the z -direction.

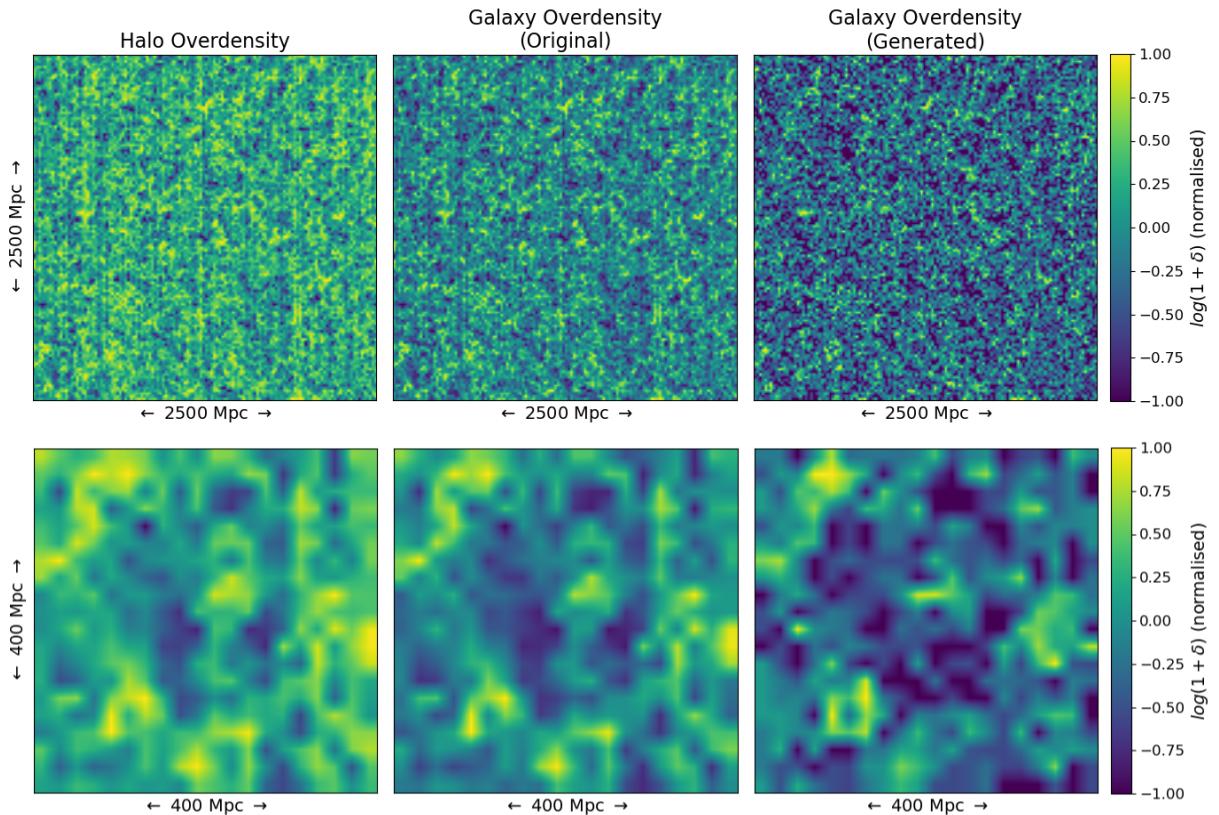


Figure 3.3: The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 0 to 200 Mpc h^{-1} along the z -direction. The x and y axis spans the entire width of the box i.e. 2500 Mpc h^{-1} . The second plot is just a small patch from the above plot extending from 0 Mpc h^{-1} to 400 Mpc h^{-1} along both x and y directions in the box

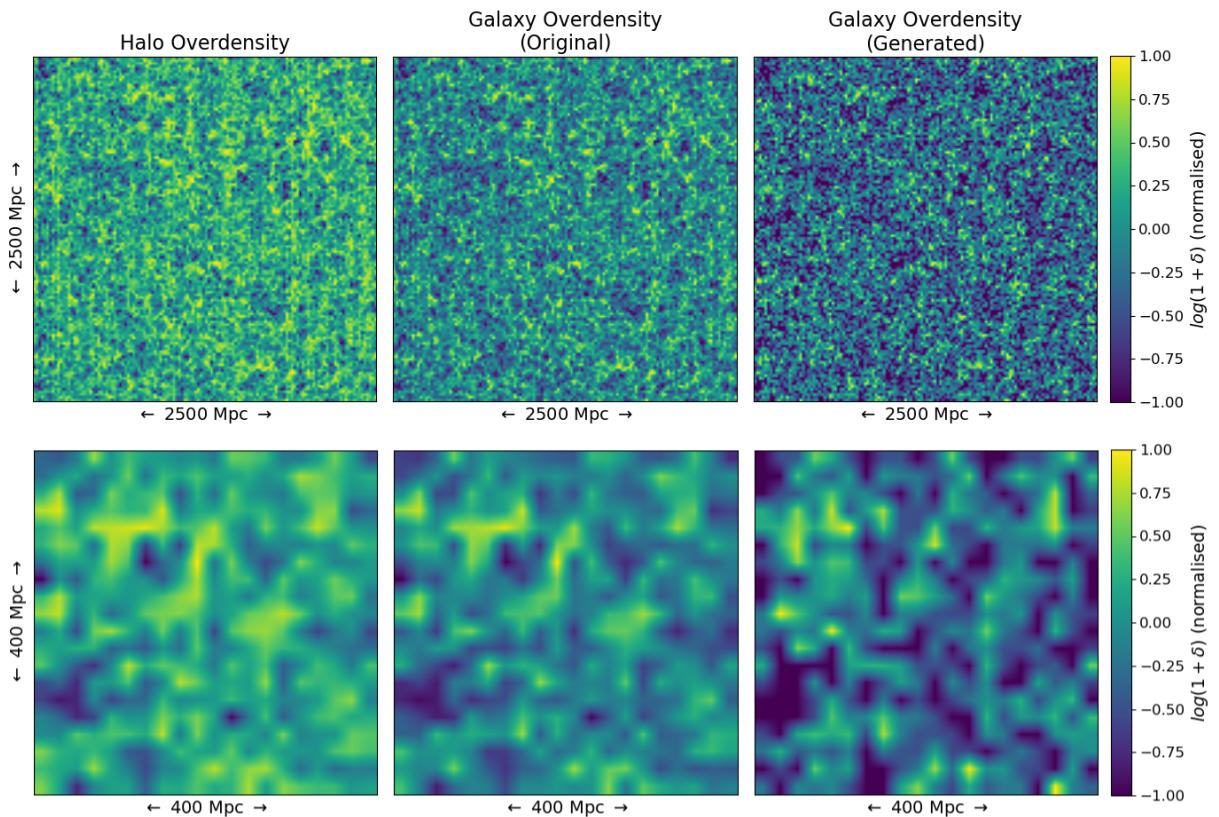


Figure 3.4: The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 1000 to $1200 \text{ Mpc } h^{-1}$ along the z-direction. The x and y axis spans the entire width of the box i.e. $2500 \text{ Mpc } h^{-1}$. The second plot is just a small patch from the above plot extending from $0 \text{ Mpc } h^{-1}$ to $400 \text{ Mpc } h^{-1}$ along both x and y directions in the box

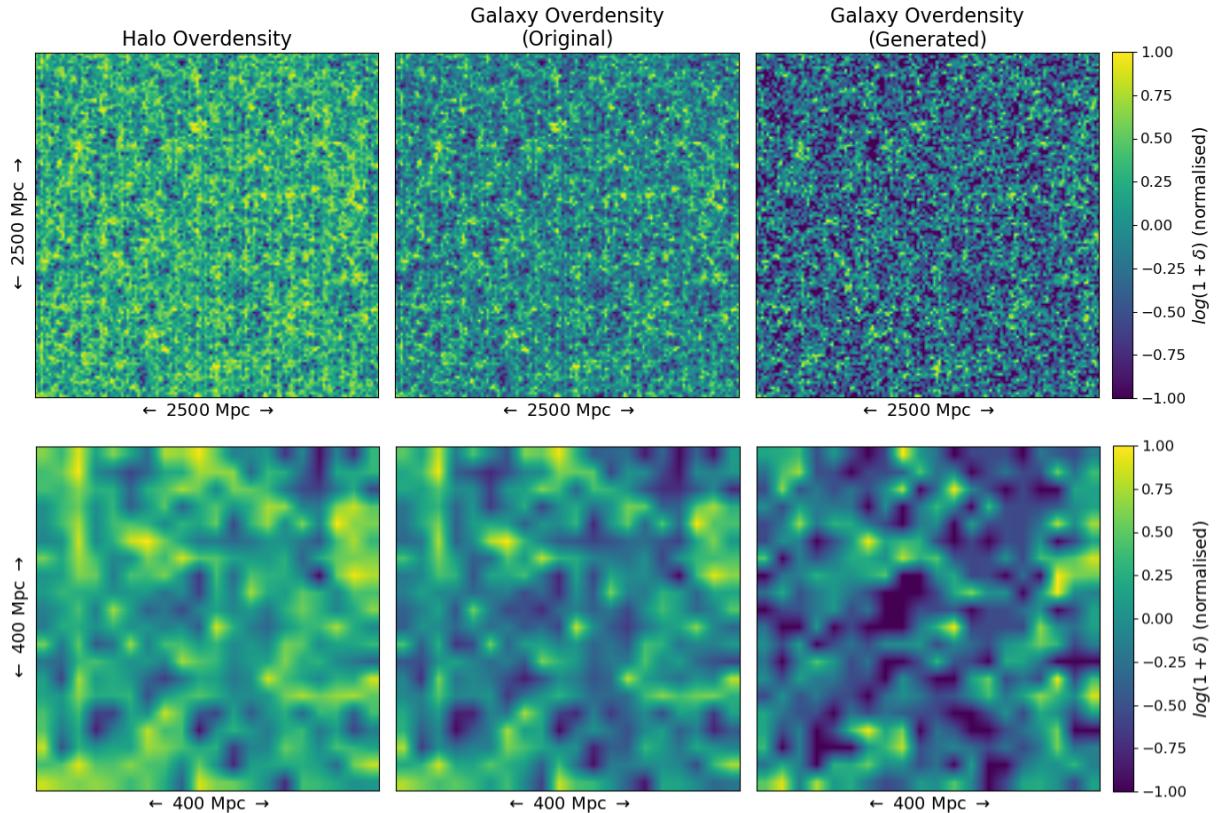


Figure 3.5: The first plot above represents the average of $\log(1 + \delta)$ for all the galaxies lying from 2000 to $2200 \text{ Mpc } h^{-1}$ along the z-direction. The x and y axis spans the entire width of the box i.e. $2500 \text{ Mpc } h^{-1}$. The second plot is just a small patch from the above plot extending from $0 \text{ Mpc } h^{-1}$ to $400 \text{ Mpc } h^{-1}$ along both x and y directions in the box.

All the plots in previous pages suggest that the generated overdensity follows the same pattern as the original one. The overdensity regions match the original's corresponding regions; the same goes for the underdense regions. Since the bias is about 0.5, we have such a difference between the halo overdensity and the galaxy overdensity. The plots clearly show that galaxies are tracers of the bulk matter of the universe.

3.8 Populating the Galaxies

Once we have the density of galaxies in each cube of our simulation box, we will populate galaxies randomly within each box. For this, we first iterate through each of the boxes and check its extent in all three directions (x , y , and z). Let one of the boxes extend in x -direction from x to $x + \Delta x$, in y -direction from y to $y + \Delta y$ and in z -direction from z to $z + \Delta z$, where all the deltas are the individual box length i.e $(2500/128) Mpc h^{-1}$. Now, within this range, we randomly select points N points and store them as coordinates of our new galaxies. N is the number of galaxies that should be present in that cube and is given as:

$$N = n \times V \quad (3.10)$$

where $V = (2500/128)^3 Mpc^3$.

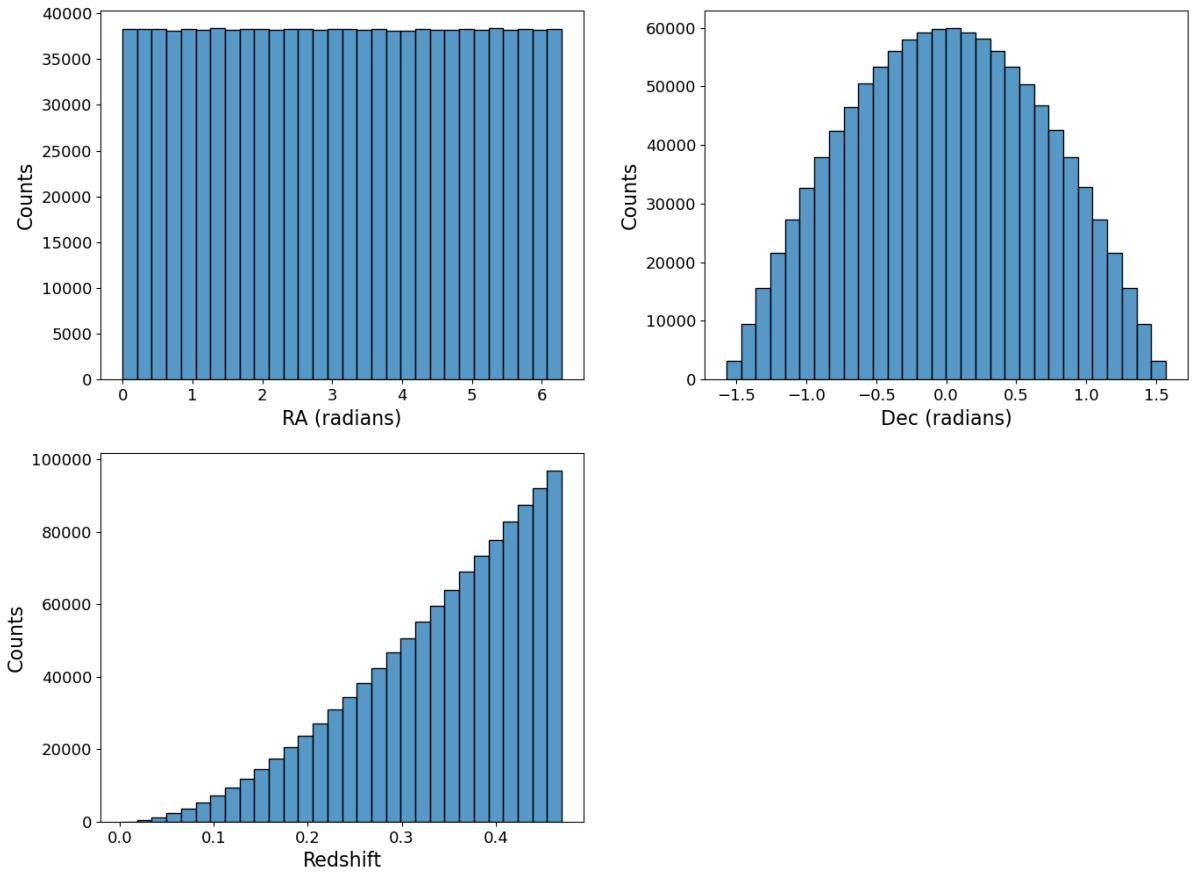


Figure 3.6: Distribution of various parameters for the populated galaxies.

Chapter 4

Future Work

The first part of the project was to populate the galaxies below a certain limit, which yielded satisfactory results. Now, the next steps are to use these galaxies, associate some GW events with them, and perform cosmological inference.

ICAROGW and **GWCOSMO** are two Python pipelines to estimate cosmological parameters using gravitational wave sirens. Using our generated catalog of newly populated galaxies, we will use both of these to infer the value of Hubble's constant.

1. GWCOSMO

GWCOSMO contains two main methods for cosmological inference: the EM counterpart method ("bright siren" method) and the galaxy catalog method (also termed the "dark siren" method). We will use the galaxy catalog method for our purpose. We will generate posterior samples and skymaps for each event and generate line-of-sight redshift prior. For this, we need to calculate the Schechter function for our catalog.

After all the requirements are fulfilled, we will run the GWCOSMO script using our catalog in these three ways:

1. The entire catalog (the entire simulated box i.e 100% complete catalog)
2. The flux-limited catalog (with only resolved galaxies)
3. The reconstructed catalog (catalog with the galaxies above the limit plus those generated below)

The same will be used with the **ICAROGW** package. Once the estimation of cosmological parameters works with the above-mentioned methods using simulated galaxy catalogs, we will then use real galaxy catalogs to test the efficiency of this method.

Bibliography

- [1] Michael R Blanton, Julianne Dalcanton, Daniel Eisenstein, Jon Loveday, Michael A Strauss, Mark SubbaRao, David H Weinberg, John E Anderson Jr, James Annis, Neta A Bahcall, et al. The luminosity function of galaxies in sdss commissioning data. *The Astronomical Journal*, 121(5):2358, 2001.
- [2] Rachel Gray. *Gravitational wave cosmology: measuring the Hubble constant with dark standard sirens*. PhD thesis, University of Glasgow, 2021.
- [3] P. Schechter. An analytic expression for the luminosity function for galaxies. , 203:297–306, January 1976.
- [4] Eric Thrane and Colm Talbot. An introduction to bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36:e010, 2019.