UNIVERSIDADE ESTADUAL DE LONDRINA

MASTERS DISSERTATION

Cosmological constraints from galaxy clusters

Author:
Henrique CARDOSO NAVES LETTIERI

Supervisor:
Dr. Sandro DIAS PINTO VITENTI

A thesis submitted in fulfillment of the requirements for the degree of Masters

in the

Theoretical Physics Group
Physics Department

May 15, 2023

UNIVERSIDADE ESTADUAL DE LONDRINA

Abstract

Center for Exact Sciences

Physics Department

Masters

Cosmological constraints from galaxy clusters

by Henrique CARDOSO NAVES LETTIERI

The Vera Rubin Observatory is one of the new experiments which will provide galaxy cluster catalogs, these catalogs will be larger and with more precise data than those currently available. Naturally, with these catalogs arises the need for development of software able to deal with the new size and precision. Moreover, the need to test different physical hypotheses requires modularized cosmological tools. Using this software we can test each model independently and include each aspect necessary to the final analyses. In our work we tested two of these cosmological frameworks, NumCosmo and CCL, which will be used by the team of researches in the LSST-DESC collaboration. These tests were be done in two parts, the first was a cross-check of the cluster counts and halo bias predictions from the two cosmological libraries. While the second part was performing a complete MCMC analysis, applying NumCosmo tools, to constraint the cosmological parameters used to create the $DC2$ simulation.

Acknowledgements

The first people I want to thank are my parents Scheilla and Deverson, without you I wouldn't be able to get where I am today. Next I would like to thank my brothers Thais and Junior, you were my first friends and second parents. I also want to thank all my friends who gave me great moments during my master's degree, especially Caio for sharing this experience of moving to another state and for all the discussions and help in the most complicated moments. I would like to thanks UEL for the masters opportunity and CAPES for the financial support. Thanks to Nersc for the resources made available for the analysis. Thanks the LSST-DESC collaboration which allowed me to participate in a relevant project. Finally my biggest thanks go to my advisors Sandro and Mariana for all the attention and guidance in this project.

Contents

Chapter 1

Introduction

In the last decades, there have been many large-scale cosmological experiments aimed to improve our understanding of the universe. Some of them are: the Chandra observatory [1], Sloan Digital Sky Survey (SDSS) [2], Wilkinson Microwave Anisotropy Probe (WMAP) [3], South Pole Telescope (SPT) [4], PLANCK satellite [5]. The SPT, WMAP and PLANCK experiments focused on measuring the Cosmic Microwave Background (CMB) anisotropies and from it constraining cosmological parameters. Beyond those constraints, we can use the CMB information for other purposes, for example, detect and measure the mass of galaxy clusters using the Sunyaev-Zeldovich effect. The Chandra experiment main focus is to study galaxy clusters and their constraints on cosmological parameters, focused on measurements in X-ray. SDSS is used to study the large-scale structure of the universe through Baryonic Acoustic Oscillations (BAO) and Supernova measurements. These are examples of successful experiments that measured the pillars of observational cosmology in this era, that is: CMB, Large Scale Structure through BAO peak, galaxy clusters, and Supernova. All these experiments yield information that can be used to understand dark matter and dark energy. To better understand the physics behind the dark sector, we can use this data to constrain cosmological models. The following years will bring the results of new experiments with new challenges.

Two examples of major large surveys that are being carried out are, EUCLID [6] and *Legacy Survey of Space and Time* (LSST) [7]. They will map a large fraction of the sky in the optical bands and will provide galaxy cluster catalogs much larger than the previous experiments. Weak lensing surveys and larger galaxy cluster catalogs are two examples of such observables. Naturally from a large amount of data on those observables arises the need for cosmological software to analyze them. There are many steps to connect our cosmological models to the data we measure. To fill this gap, we need to improve our physical modeling and construct various phenomenological models. For example, for galaxy cluster cosmology, we need to understand the collapse process of the initial matter density field into galaxy clusters, model the effects of Milk Way dust when measuring the photometric redshift, solve the linearly perturbed Boltzmann equations, just to name a few. Since these models can change when using different observations or physical assumptions, it is important to construct our cosmological software in a modular fashion. One way to solve this is to write a library containing modules with independent functionality. In this way, we can execute and change them separately allowing consistent testing of each model individually. Once these modules are ready, we need to test them to see if they run as expected and then calibrate these modules using simulated datasets so we can apply them in real data analyses. The only way of testing such modules is by comparing two independent implementations. In this work we will compare the results of two cosmological libraries, *Numerical Cosmology* (NumCosmo) [8] and *Core Cosmology Library* (CCL) [9], to validate their routines so we can calibrate them later. The development of these two libraries is mainly focused on analyzing the data of the LSST survey which will be discussed in the next section. To develop the tools necessary to analyze LSST data the *Dark Energy Science Collaboration* (DESC) was created and this work is an extension of two DESC projects. The numerical libraries tested here and their statistical tools will be used by the collaboration when data become available.

In section 1.1, we explain the goals and structure of the LSST-DESC collaboration and where our work fit in. In section 1.2, we talk about the N-body simulation created by LSST-DESC to calibrate the tools developed by the collaboration. In section 1.3, we present NumCosmo and CCL which will be the two softwares used by the collaboration to model the background cosmology. In chapter 2, we construct the theoretical background to do galaxy cluster cosmology. In section 2.1, will be discussed how we can use galaxy clusters as cosmological probes. Section 2.2, we will construct the halo mass function from its formal definition to the expressions resulting from N-body simulation. Section 2.3 is where we construct the cluster number counts which is the observable we are mainly interested in here. In section 2.4 we define the halo bias function which will be used in future work to improve the precision of cluster counts and add cluster clustering as a new observable. Finally, in section 2.5 we will discuss how to relate the observed quantities of the cluster with its mass and redshift. In chapter 3, we will develop the statistical tools used to do the analyses, from the maximum likelihood estimators in section 3.1 to Bayes theorem and the Markov Chain Monte Carlo method (MCMC) in section 3.2. Then in section 3.3, we will present the likelihoods and proxies used in this work. Chapters 4 and 5 will contain the results obtained in this work. First, the results of CCL and NumCosmo cross-check from the true table in section 4.1 to the table with the mass proxy in section 4.2 and finally for the halo bias in 4.3. In chapter 5 we present the MCMC for a mock catalog, section 5.1, and $DC2$ in 5.2. Finally, in chapter 6 we will present our conclusions.

1.1 LSST-DESC

The Vera C. Rubin Observatory is under construction in Cerro Pachón ridge in north-central Chile. The Simonyi Survey Telescope is designed with a 8.4m main mirror and 3.2 giga-pixel camera. The observatory will observe the sky in six optical bands with wavelength ranging from $320 - 1050nm$ [7]. The telescope will cover up $18000deg^2$ 800 times over its ten-year lifetime and generate around 500 petabytes of imaging data and over 50 petabytes in catalog databases [10]. The main science achievements expected for the observatory are

- Taking an inventory of the solar system.
- Exploring the transient optical sky.

- Mapping the Milky Way.
- Probing dark energy and dark matter.

where the last is the important one for the purpose of this work.

The Legacy Survey of Space and Time (LSST) Dark Energy Science Collaboration (DESC) was established with the purpose of developing tools to analyze its data focusing on the dark sector physics such as neutrinos, dark matter and dark energy. DESC is built around five dark energy probes. They are [11]:

- Weak Gravitational Lensing (WL), that is the deflection of light rays of distant sourceswhich can be used to determine cluster mass by the change in shape of background galaxies or map the matter density distribution in the universe.
- Large Scale Structure (LSS), in particular the large scale power spectrum of the matter density field.
- Type Ia Supernova (SN), which measure the luminosity distance as a function of the redshift.
- Strong Gravitational Lensing (SL), that is the morphological distortion and angular displacement of a source objective due to the presence of a massive object in the line of sight.
- Galaxy Clusters (CL), their spatial distribution, number counts, and mass.

In order to perform the individual and joint analyses of these probes, the DESC is divided in 18 working groups with their main priorities presented in [11].

One of them is the Cluster Galaxies group which is concerned with extracting robust cosmological constraints from this probe. To do this, we need to understand the systematic uncertainties associated with their observed properties and matter distribution. In order to improve the constraints with the cluster number counts observable, we need to run through four crucial steps. The first is predicting the mass function for the halos, the second is constructing the observed catalog of clusters. The third is understanding the statistical relation between cluster observables and the halo mass and the fourth is using a self-consistent likelihood analysis to constrain the cosmological parameters. The Cluster Galaxies group is mainly responsible for performing the 1-4 steps while the 2-3 will be done using the LSST data [11]. To fulfill this purpose, the group established the following research priorities,

- Emulating the halo-mass and halo-halo functions;
- Developing high-quality photometric cluster finders;
- Developing models to determine the mass-observable relations for different observables such as richness and weak lensing profiles;
- Combining the above results to construct a single pipeline to constrain cosmological parameters both in real data and simulated datasets;

This work is an extension of the DESC projects number 139 and 203 in the Galaxy clusters group. The first is a comparison between the constraints on cosmological parameters in binned and unbinned likelihood and the self-calibration on the mass-observable relation following the fourth aspect of research priority. The latter is a cross-check between NumCosmo and CCL cluster counts and halo bias prediction to test the cosmological software to be used by the collaboration.

1.2 Second Data Challenge

To test the tools developed by DESC they created the Data Challenges which are a set of simulated data that mimics the LSST data. These data challenges are [11]:

- $DC1$: Which simulates the ten year data of LSST taking an $40deg^2$ sky area but only in the r-band. The main focus of $DC1$ was to investigate the systematic uncertainties relevant to large scale structure measurements and it's end-to-end workflow
- $DC2$: The second challenge consider five year data of LSST but with an area of $300deg^2$ and with all the six optical bands of LSST $ugrizy$. Beyond $DC1$ and $DC2$ it is included pipelines for weak lensing correlations, large scale structure statistics, cluster abundance, supernova light curve recovery and inferences on redshift distribution from photometric redshifts.

In order to simulate these data products the $DC2$ workflow is divided in four areas. First is the construction of the extragalactic catalog, second is the generation of input catalogs for the image simulations, then is done the image simulations themselves and finally the processing of these images with LSST-DESC pipelines [12].

The extragalactic catalog created is called CosmoDC2 [13] and is based on the N-body Outer Rim simulation which contains about a trillion particles with mass around $2.6 \times 10^9 .M_\odot$. The halo finder is a friends-of-friends algorithm with a linking length $b = 0.168$ with a minimum of 20 particles. Afterwards, the halo light-cone is built and the properties of galaxies are assigned leading to the finishing of the catalog [13].

Latter for the input catalog we need to provide input instances catalogs which consider the objects in the extragalactic catalog in a field of view at a specific time. In this part the $DC2$ introduces difficulties in the observation like the earth's motion influence, galactic dust and other details. Despite the instance catalog which will be used to simulate the image the second component also produces the truth catalogs. This catalog will contain all the true information of the measured quantities in order to compare with the output of the imaging processes and other pipeline analyses.

In possession of the instance catalogs the imSim software package simulates pixel data. The hardware and telescope specific information to do this will be provided using LSST Science Pipelines and LSST Simulation Framework software environment [14]. The imSim software was conceived to allow multiple models to study the same effect. This opens the possibility to vary the precision of the features in the simulation depending on the need of the group, allowing the most precise and slow simulation and the simpler ones [12]. Finally the LSST pipelines processes the images and delivery data products similar to the Rubin Observatory.

This work will deal with the first end of the $DC2$ catalog data products. We will deal with the simpler case using the extra-galactic catalog to test the developed tools with all the clusters true information available. In a future work, we will do an analyses with the data products of the processed images to consider the measurements of richness and test the mass-proxy and perform the self-calibration.

1.3 NumCosmo and CCL

Two tools developed by members of the LSST-DESC group are the cosmological libraries NumCosmo[1] and Core Cosmology Library (CCL).[2] Both libraries provide a set of tools to calculate many different cosmological quantities from the background cosmological model to the observables like cluster number counts, halo mass function, supernova data, weak lensing shear, etc. These libraries also use these data to perform statistical analyses to constraints the cosmological parameters [8, 9]. Both libraries are constructed in C and provide a python interface (more recently CCL moved much of the newer implementation directly to python). They also allow the integration with other libraries like CAMB [15] and CLASS [16] to improve the precision of the analyses as the need of the user.

In this work we will make use of both libraries tools to predicted the cluster number counts and halo mean bias and compare their precision and time efficiency. The simplest mock catalog used to make these comparison was built using NumCosmo. The constraints of the cosmological parameters and the cross-check in the mock catalog and in the $DC2$ catalog were done using only NumCosmo.

[1] https://github.com/NumCosmo/NumCosmo
[2] https://github.com/LSSTDESC/CCL

Chapter 2

Galaxy Clusters

Galaxy Clusters are the largest bounded structure in the universe, their mass are mainly composed by dark matter halos which governs their dynamics turning them into crucial cosmological probes. This fact allows a simple modeling for N-body simulations, enabling the cross-check with their theoretical properties like their internal structure and abundance as a function of the mass [17–21]. This turns clusters into useful tools to constraint cosmological models. Their mass is divided in three components. The first is the dark matter halo composing most of the cluster mass. Secondly, we have the Intra-Cluster Medium (ICM) that is a hot ionized gas which is mainly composed by hydrogen and a small fraction o helium. The third part are the galaxies which composes the smallest fraction of the cluster.

Each of these components provides important information about cosmological evolution and the properties of the galaxy clusters. In this chapter we will, first, explain how to use galaxy clusters as cosmological probes and the challenges in doing it. Second we will construct the halo mass function with the Press & Schechter ansatz and all the ingredients necessary to calculate the mass function. Then we will use the mass function to predict the number of halos in a mass interval at a survey region in the sky. However, these halos are biased tracers of the dark matter density distribution and to improve the cosmological constraints we will study its bias. Finally, in the last part of this chapter we will connect the halo formalism with clusters observables.

2.1 Clusters as cosmological probes

There are two ways of using galaxy clusters as cosmological probes. The first method is by their number counts. Here, we use the working assumption that all the matter in the universe is contained into galaxy clusters. Then, we can construct a theoretical model to predict the amount of clusters in a survey region given that its matter density is the mean density of the universe. Finally, we can compare with the observed number of clusters to constrain our model. This method is the one we will be mainly interested in this work. The rest of this chapter is dedicated to explain how to develop it. The second method is the spatial distribution of these galaxy clusters in the survey region, called cluster clustering. The latter will not be addressed in this work. However, ignoring cluster clustering may result in bias in the number of clusters observed. The bias generated by the cluster clustering on the number counts likelihood is discussed here but the correct modeling of this effect is left to future work.

To compare our predictions of cluster counts or cluster clustering we need to construct cluster catalogs using one or more of its observables. The three main electromagnetic bands we use to detect clusters are x-ray, optical and microwaves. The first is by measuring the thermal emission of the plasma in the inter-cluster medium. In the optical band the number of galaxies with a specific color-brightness relation is measured called richness. While in microwave we measure the Sunyaev-Zeldovich effect. There is still detection of clusters by gravitational lensing. There are two main challenges in constructing galaxy cluster catalogs, the completeness and purity of the catalog. Independently of the property we measure of the catalog there will always be clusters we cannot detect, this makes the catalog incomplete. In a realistic analyses we must insert a function to model this effect. The second difficulty in constructing catalogs is to falsely detect a cluster. This is called the purity of the catalog and we also need to modeled this in a real catalog. However in this work to test the tools developed we will deal only with 100% pure and complete simulated catalogs.

2.2 Halo Mass Function

The halo mass function $n(M, z)$ is the number density of halos at redshift z in a mass interval $[M, M + \mathrm{d}M]$. The mass function possesses a crucial value in cosmology. Some of its application are the constraining the amplitude of the power spectrum and its evolution can constrain dark matter and dark energy parameters [22, 23] and can even test our gravity theory [24]. In this work its main application will be calculating the cluster number counts and the halo bias.

To count the number of halos in a mass interval we consider that all matter in the universe is inside halos hence the number of halos is only the ratio of the halo mass and the total mass in that patch of the sky M_T/M. However, not all halos have mass M so we need to account only the fraction of halos that lie in this range $F(M, M + \mathrm{d}M)$, and by taking that the universe is approximately homogeneous and isotropic we can simply write the total mass as $M_T = \bar{\rho}_m V$, where $\bar{\rho}_m$ is the mean matter density of the universe, leading to the halo mass function

$$\frac{N}{V}\mathrm{d}M = n(M, z)\mathrm{d}M = \frac{\bar{\rho}_m}{M}F(M, M + \mathrm{d}M). \tag{2.1}$$

The final piece of information left is the fraction of halos collapsed in the mass range required. To model F we use the Press-Schechter ansatz which states that the probability of the halo collapse can be traced back to the amplitude of the density fluctuations in that scale [25].

The density fluctuation field $\delta(\vec{x}, z)$ is a random variable defined as

$$\delta(\vec{x}, z) \equiv \frac{\rho_m(\vec{x}, z) - \bar{\rho}_m(z)}{\bar{\rho}_m(z)}, \tag{2.2}$$

from this definition is clear that δ has zero mean. Initially these density fluctuations are very small and we can model their time evolution as a linear perturbation of an ideal fluid (pressureless in the case of baryons and dark matter) following the two hydrodynamic plus Poisson equations [22]

- Conservation equation
$$\dot{\delta} + \vec{\nabla} \cdot [(1+\delta)\vec{V}] = 0, \tag{2.3}$$

- Euler equation
$$\dot{\vec{V}} + H(t)\vec{V} = -\vec{\nabla}\phi, \tag{2.4}$$

- Poisson Equation
$$\nabla^2\phi = 4\pi G\bar{\rho}_m\delta, \tag{2.5}$$

where $\dot{} = \frac{\mathrm{d}}{\mathrm{d}t}$ the $\vec{V}$ is the fluid peculiar velocity, ϕ the gravitational potential of the fluid and H is the Hubble function given by

$$\boxed{H(z) = H_0[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda]^{1/2}}, \tag{2.6}$$

where H_0 is the Hubble constant today, $\Omega_r, \Omega_m, \Omega_k, \Omega_\Lambda$ are the density parameters for radiation, matter, curvature and dark energy respectively. Taking the time derivative of 2.3 and neglecting the terms above order one in $\vec{V}$ and δ we get

$$\ddot{\delta} + \vec{\nabla} \cdot \dot{\vec{V}} = 0, \tag{2.7}$$

and substituting equations 2.4 and 2.5 into the previous one we get,

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}_m\delta = 0. \tag{2.8}$$

Now we define the growth function as $D(t) = \delta(t)/\delta(t_0)$ which compares the density fluctuation today with that in a time t. In cosmological terms is more convenient to write the equation 2.8 in terms of the redshift using

$$\frac{\mathrm{d}}{\mathrm{d}t} = -H(z)(1+z)\frac{\mathrm{d}}{\mathrm{d}z}, \tag{2.9}$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} = -H(z)\frac{\mathrm{d}H}{\mathrm{d}z}(1+z)\frac{\mathrm{d}}{\mathrm{d}z} - H(z)\frac{\mathrm{d}}{\mathrm{d}z} - H(z)(1+z)\frac{\mathrm{d}^2}{\mathrm{d}z^2}. \tag{2.10}$$

Finally we can put the evolution of the density fluctuation as

$$\boxed{\frac{\mathrm{d}^2D}{\mathrm{d}z^2} - \left(\frac{1}{1+z} - \frac{1}{E(z)}\frac{\mathrm{d}E}{\mathrm{d}z}\right)\frac{\mathrm{d}D}{\mathrm{d}z} - \frac{3\Omega_m}{2E^2(z)}(1+z)D = 0}, \tag{2.11}$$

where $E \equiv H(z)/H_0$ and we used $\bar{\rho}_m = \Omega_m(1+z)^3\frac{3H_0^2}{8\pi G}$.

However in order to use the Press-Schechter formalism we need to filter the density fluctuation throwing away all the information in scales smaller than the typical scale R of the filter $W\left(|\vec{x}-\vec{y}|, R\right)$ leading to

$$\delta_R(\vec{x},z) = \int \mathrm{d}\vec{y} W(|\vec{x}-\vec{y}|, R)\delta(\vec{y},z). \tag{2.12}$$

In this work we will always use a Top-Hat filter, given by

$$W_{TH}(r,R) = \frac{3}{4\pi R^3}\begin{cases} 1, \ r \le R \\ 0, \ r > R. \end{cases} \tag{2.13}$$

Finally we say that if the region of size R have a density contrast δ_R greater than a critical density δ_c then this region will collapse and form a cluster. The critical density depends on the details of the collapse model, for the spherical collapse model $\delta_c(0) = 1.686$ and will have a small dependence on the matter density parameter [26]. Note that the typical density of theses regions are not in the linear regime ($\delta_R \ll 1$), however, the idea behind this extrapolation is that exists a one-on-one relation between that initial state from linear perturbation with the collapsed structure.

Since the density contrast δ is a random field it follows that the δ_R is also a random field. Hence, we say that the fraction of objects collapse with mass M or greater is given by $P(\delta_R > \delta_c)$ where the mass of the cluster is linked with R by

$$M = \frac{4\pi}{3} f_W R^3 \bar{\rho}_{ref}, \tag{2.14}$$

where f_W will depend on the choice of filter function, in the case of the Top-hat filter $f_W = 1$ and ρ_{ref} is our reference density to define the mass, it can be the mean density of the universe $\bar{\rho}_m$ or the critical density ρ_c. The probability, $P(\delta_R > \delta_c)$, is

$$P(\delta_R > \delta_c) = \int_{\delta_c}^{\infty} F(\delta_R)\mathrm{d}\delta_R, \tag{2.15}$$

where now we use the mass as the variable to define the cluster density by the relation 2.14. Now, to calculate the fraction of objects in the mass range $M, M+\mathrm{d}M$ we just need to calculate $P(\delta_{R+\mathrm{d}R} > \delta_c)$ which computes the fraction of collapsed objects with mass $M+\mathrm{d}M$ and above and take the difference. Physically this means that when we calculate $P(\delta_R > \delta_c)$ we may include regions with larger and larger scales. However, we want only to consider dense regions surrounded by underdense regions to assure they are an isolated structure. So we can write the fraction of halos formed in the mass range $M, M+\mathrm{d}M$ $F(M, M+\mathrm{d}M)$ as the probability of a region with mass between M and $M+\mathrm{d}M$ collapsing into a cluster, that is,

$$F(M, M+\mathrm{d}M) = P(M) - P(M+\mathrm{d}M) = P(M) - P(M) - \frac{\mathrm{d}P}{\mathrm{d}M}\mathrm{d}M = -\frac{\mathrm{d}P}{\mathrm{d}M}\mathrm{d}M, \tag{2.16}$$

finally we can write the mass function as

$$n(M,z) = -\frac{\bar{\rho}_m}{M}\frac{\mathrm{d}P}{\mathrm{d}M}. \tag{2.17}$$

Since the initial density perturbations follow approximately a Gaussian distribution, Press and Schechter assumed (as a first approximation) that the filtered density fluctuations also follows a Gaussian distribution with mean zero and variance $\sigma_R^2(z)$, hence using equation 2.15, we have

$$P(M,z) = \int_{\delta_c}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_R} \exp\left(-\frac{\delta_R^2}{2\sigma_R^2}\right) \mathrm{d}\delta_R = \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\delta_c(z)}{\sqrt{2}\sigma_R(z)}\right)\right). \tag{2.18}$$

Finally we can calculate the Press-Schechter(PS) mass function as

$$n_{PS}(M,z) = -2\frac{\bar{\rho}_m}{M}\frac{\delta_c(z)}{\sigma_R(z)}\frac{1}{\sqrt{2\pi}}\left(\frac{\mathrm{d}\ln\sigma_R}{\mathrm{d}M}\right)\exp\left(\frac{-\delta_c^2(z)}{2\sigma_R^2(z)}\right), \tag{2.19}$$

Where the factor 2 was latter introduced by Press and Schechter to correct the normalization of equation 2.18 when we take $\delta_c = 0$. An interesting property of the PS mass function with the spherical collapse model in an Eistein De Sitter universe is that both $\sigma_R(z)$ and $\delta_c(z)$ will evolve with the growth function $D(z)$ and consequently their ratio is constant in z, that is, $\sigma_R(z)/\delta_c(z) = \sigma_R(0)/\delta_c(0)$. For this reason, we can decouple the details of the collapse model in the multiplicity function defined as

$$f_{PS}(\sigma_R) = \frac{2}{\sqrt{2\pi}}\frac{\delta_c(0)}{\sigma_c(0)}\exp\left(\frac{-\delta_c^2(0)}{2\sigma_R^2(0)}\right), \tag{2.20}$$

from the cosmological information presented in $\sigma_R(z)$ and $\bar{\rho}_m$. This property of the multiplicity function is called universality and repeats itself in different collapsed models, one example is the one proposed by Sheth and Tormen (ST) in [27] an Ellipsoidal collapsed model and obtaining the following multiplicity function

$$f_{ST}(\sigma_R) = A\sqrt{\frac{2a}{\pi}}\left[1 + \left(\frac{\delta_c}{\sigma_R}\right)^{-p}\right]\left(\frac{\delta_c}{\sigma_R}\right) exp\left[-\frac{1}{2}\left(a\frac{\delta_c}{\sigma_R}\right)^2\right]. \tag{2.21}$$

This universality property arises from the consideration that the collapse process happens in time scale much smaller than the rate of the universe expansion, however, this is not exactly true and we expect a dependence on the redshift in multiplicity functions as shown in N-body simulations. Another important aspect which is not considered in this model is that the formation process depends on the structures already formed around the collapsing object inducing the dependency on z since in higher redshifts less objects have been formed. We can circumvent these issues by fitting N-body simulations, which provides us with a mass function with fitting formula expressed by [28]

$$\boxed{n(M,z) = -\frac{\bar{\rho}_m}{M}\left(\frac{\mathrm{d}\ ln\sigma_R}{\mathrm{d}M}\right) f(\delta_c(z),\sigma_R(z))}\,, \tag{2.22}$$

with the multiplicity function left to be constructed from a simulation or a theoretical collapse model.

2.2.1 Filtered density fluctuation variance

From equation 2.22 we can see that all the cosmological information comes from the variance of the filtered power spectrum. This happens because we assumed a Gaussian distribution with zero mean and consequently all the information lies in the variance. If we introduce non-gaussianity, we would have all the distributions moments. However, this degree of approximation has shown to be enough in cosmological simulations. We need to understand the connection between the cosmological model and σ_R, so we consider its formal definition

$$\sigma_R^2(z) \equiv \left\langle \delta_R^2 - \langle\delta_R\rangle^2 \right\rangle = \langle\delta_R^2\rangle = \int \mathrm{d}^3y \int \mathrm{d}^3y' W(\vec{y}',R)W(\vec{y},R)\left\langle \delta(\vec{x}+\vec{y}',z)\delta(\vec{x}+\vec{y},z)\right\rangle, \tag{2.23}$$

now we can expand the density fluctuations in their Fourier modes $\delta(\vec{k},z)$ leading to

$$\sigma_R^2(z) = \int \mathrm{d}^3y \int \mathrm{d}^3y' W(\vec{y}',R)W(\vec{y},R) \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, e^{-i\vec{k}\cdot(\vec{x}+\vec{y})} \int \frac{\mathrm{d}^3k'}{(2\pi)^3}\, e^{i\vec{k}'\cdot(\vec{x}+\vec{y}')} \left\langle \delta(\vec{k},z)\delta^*(\vec{k}',z)\right\rangle, \tag{2.24}$$

and the variance of the Fourier modes are connected to the matter linear power spectrum $P(k,z)$ by the relation [29, 30]

$$\left\langle \delta(\vec{k},z)\delta^*(\vec{k}',z)\right\rangle = (2\pi)^3\delta_D(\vec{k}-\vec{k}')P(k,z), \tag{2.25}$$

where $\delta_D(\vec{k}-\vec{k}')$ is the three dimensional Dirac delta function. The delta function eliminates the integral in k' leading to

$$\sigma_R^2 = \int \frac{d^3k}{(2\pi)^3} P(k,z) \left(\int d^3y W(\vec{y},R) e^{-i\vec{k}\cdot\vec{y}}\right)\left(\int d^3y' W(\vec{y}',R) e^{i\vec{k}\cdot\vec{y}'}\right), \tag{2.26}$$

The terms in the parenthesis are the inverse Fourier transform of the filter function hence the filtered density fluctuation variance is given by

$$\boxed{\sigma_R^2(z) = \int \frac{d^3k}{(2\pi)^3} |W(k,R)|^2 P(k,z)}\,. \tag{2.27}$$

For the top-hat filter, its Fourier transform is

$$W(k,R) = \frac{3}{kR} j_1(k,R), \tag{2.28}$$

where $j_1(k, R)$ is the first kind spherical Bessel function. The linear matter power spectrum can be calculated using the relation (see for example [31])

$$\boxed{P(k,z) = Ck^{n_s}T(k)^2D(z)^2}\,, \tag{2.29}$$

where $D(z)$ is the growing solution, known as growth function given by equation 2.11, Ck^{n_s} is the Harrison-Zel'dovich-Peebles primordial spectrum which comes as a prediction from inflation and the n_s parameter is determined by measurement (Plank's 2018 measurement was $n_s = 0.967 \pm 0.004$ [5]), while C is it's a normalization which is usually chosen by the parameter $\sigma_8 = \sigma_8(0)$, that is the filtered density fluctuation in a scale of $R = 8h^{-1}Mpc$ so C becomes

$$\boxed{C = \frac{\sigma_8^2}{\int_0^\infty \frac{dk}{2\pi^2} k^{n_s+2} T(k)^2 W(k, 8h^{-1})^2}}\,, \tag{2.30}$$

and σ_8 will also be treated as a cosmological parameter to be adjusted from the data. Finally, $T(k)$ is the transfer function that measures the evolution of the material components from the primordial perturbations until the decoupling. This is done by solving the coupled Boltzmann's equations for every component in the universe, such as dark matter, photons, barions, neutrinos, etc. This is a rather complicated task and for this reason there exists numerical codes specialized only in this problem. The two main codes are CAMB [15] and CLASS [16], both descendants of CMBFAST [32]. There exits however some analytical expressions for specific models in a constrained area of the parameter space. These expressions give precise enough prediction for the power-spectrum such that we can use this faster version to calibrate the software. Here we use the fitting formula proposed by Eisenstein-Hu (EH) unless stated otherwise [33].

2.2.2 Mass Function from N-body simulations

The best way of determining the mass function from clusters is from N-body cosmological simulations. First we count the number of clusters in a volume element in the simulation. To do this we need to define the cluster and its mass. There are two main methods

- Friends-of-Friends (FoF): In this method we choose a particle in the simulation and if another particle is within a certain distance, called the correlation distance given by $bn^{1/3}$, from the original particle we add it to the group of particles composing the cluster. Variable n is the particle number density and b is a parameter depending on the simulation. In the case of $DC2$ we have $b = 0.168$ [13]. The cluster mass is the product of one particle mass times the number of particles [34, 35].

- Spherical Overdensity (SO): In this method we choose a center and construct spheres with varying size until the region mean density reaches a reference density, given a value of Δ as $\Delta\rho_{ref}$ with size R_{ref} and calculates the mass by $M = \frac{4\pi}{3}\Delta\ \rho_{ref}R_\Delta^3$ [36]. The reference density ρ_{ref} is usually either the universe matter mean density $\bar{\rho}_m$ or the critical density

ρ_c and Δ is the number of times this region mean density needs to be greater than the reference density, usually is chosen 200 or 500.

One difficulty of the SO method is how to determine the center of the sphere, there are different procedures that can be used. One possible is by using the FoF to find the most dense region and from there we set the center and apply the SO method [37]. In the *DC*2 catalog we are provided with clusters mass defined by both methods and we will use the SO mass.

Once we know the clusters mass and redshift in the simulation we can count them in a small volume of the simulation determining the left hand side of equation 2.22. Hence we can determine the multiplicity function as

$$f(\delta_c, \sigma_R) = -\frac{M}{\bar{\rho}_m(z)} \left(\frac{\mathrm{d} \ln \sigma_R}{\mathrm{d}M} \right)^{-1} n(M, z). \tag{2.31}$$

Different simulations show that the universality of the PS multiplicity function is repeated in other cases for different cosmological models and different initial conditions, leading to similar shapes [38, 39] (different from the PS shape). However, in more recent simulations is shown an explicitly dependence of z when we look in higher Redshift intervals, indicating a small dependence on the cosmological model. The multiplicity function mainly used in this work is the one by Tinker [28]

$$\boxed{f(\delta_c(z), \sigma_M(z)) = A \left[\left(\frac{\sigma_R(z)}{b} \right)^{-a} + 1 \right] \exp\left(-\frac{c}{\sigma_R(z)^2} \right)}, \tag{2.32}$$

$$A(z) = A_0(1+z)^{-0.14}, \tag{2.33}$$

$$a(z) = a_0(1+z)^{-0.06}, \tag{2.34}$$

$$b(z) = b_0(1+z)^{-\alpha}, \tag{2.35}$$

$$\log(\alpha) = -\left(\frac{0.75}{\log(\Delta/75)} \right)^{1.2}. \tag{2.36}$$

The Tinker multiplicity function is constructed for the SO mass definition and will not depend strongly on the cosmological model. All z independent parameters A_0, a_0, b_0, α, c depends on the choice of reference density critical or mean and on Δ, where here we will mainly use $\Delta = 200$ and critical density, all the values are given in reference [28].

2.3 Cluster Number Counts

Once we know the mass function we can find the expected number of clusters in mass interval $[M_{min}, M_{max}]$ within a co-moving volume of the sky searched. In order to do this, we only need to calculate the integral

$$N = \int_{M_{min}}^{M_{max}} \int n(M, z) \mathrm{d}V \mathrm{d}M, \tag{2.37}$$

where $\mathrm{d}V$ is the co-moving volume element that depends on the cosmology. It is convenient to express the volume in spherical coordinates as $\mathrm{d}V = r_A^2(z)\mathrm{d}r_A \mathrm{d}\Omega$ where the angular part $\mathrm{d}\Omega$ is independent of the cosmology due to homogeneity and isotropy and r_A is the angular distance seen by a co-moving observer and related to the coordinate distance $r(z)$ by

$$r_A = \begin{cases} \frac{\sin\left(\sqrt{-\Omega_k}\, c/H_0 r(z)\right)}{\sqrt{-\Omega_k}\, c/H_0} & \Omega_k < 0 \\ r(z) & \Omega_k = 0 \\ \frac{\sinh\left(\sqrt{\Omega_k}\, c/H_0 r(z)\right)}{\sqrt{\Omega_k}\, c/H_0} & \Omega_k > 0, \end{cases} \tag{2.38}$$

where Ω_k is the curvature parameter and the co-moving radial coordinate is

$$r(z) = \int_0^z \frac{c}{H(z')} \mathrm{d}z'. \tag{2.39}$$

The next step is to take the differential form of equation 2.38 that is

$$\mathrm{d}r_A = \frac{c}{H(z)} \mathrm{d}z \begin{cases} \cos\left(\sqrt{-\Omega_k}\, c/H_0\, r(z)\right) & \Omega_k < 0 \\ 1 & \Omega_k = 0 \\ \cosh\left(\sqrt{\Omega_k}\, c/H_0 r(z)\right) & \Omega_k > 0. \end{cases} \tag{2.40}$$

In order to lead to a more compact expression for cluster number counts we rewrite the previous equation as

$$\mathrm{d}r_A = \frac{c}{H(z)} F(\Omega_k, z) dz, \qquad F(\Omega_k, z) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\sqrt{-\Omega_k}\, c/H_0 r(z)\right)^{2n}}{2n!}. \tag{2.41}$$

The angular part on the integral 2.37 is trivial and yields the survey area A_s of our instrument. Therefore, we can put the cluster number count in the mass and redshift intervals $[M_{min}, M_{max}]$, $[z_{min}, z_{max}]$

$$N = A_s \int_{M_{min}}^{M_{max}} \int_{z_{min}}^{z_{max}} n(M, z) r_A(z)^2 \frac{c}{H(z)} F(\Omega_k, z) \mathrm{d}z \mathrm{d}M. \tag{2.42}$$

Fortunately, the universe has a very small curvature allowing us to neglect Ω_k and leading to the simpler expression (nevertheless, the cosmological codes take the curvature correctly into account when necessary)

$$\boxed{N = A_s \int_{M_{min}}^{M_{max}} \int_{z_{min}}^{z_{max}} n(M, z) \frac{c}{H(z)} \left(\int_0^z \frac{c}{H(z')} \mathrm{d}z' \right)^2 \mathrm{d}z \mathrm{d}M}. \tag{2.43}$$

From equation 2.43 it is clear that the number of clusters is strongly dependent on the cosmological model, different from the mass function alone where its dependence comes only from σ_R. Here, it comes explicitly from $H(z)$ which makes the cluster number count a valuable tool for constraining the cosmological parameters.

2.4 Halo Bias function

It's noticed that places with a high primordial density fluctuations generates more massive clusters [40]. Hence, this tells us that clusters are biased cosmological probes of the underlying matter distribution. For this reason, the clustering of clusters is a complementary information to the number counts to constrain cosmological parameters. Also, the clustering information is valuable for self-calibration of mass proxies which will be crucial when we perform a realistic analyses with unknown parameters of the mass-observable relations [41]. In [42] it's shown that we can understand the clustering process by studying the clusters bias. So, now we construct an analytical model for the determination of the linear bias based on [43] and then, we present the bias given by N-body simulations.

The idea follows a principle similar to the construction of the mass function. However, now we consider the number density of clusters with mass between $M_1, M_1 + \mathrm{d}M_1$ and redshift z_1 collapsed from a initial region in redshift z_0 with size R_0 and linear density perturbation δ_0 that hasn't collapsed yet. By this construction we expect that $M_0 > M_1$ and $\delta_0 < \delta_c(z) = (1 + z_1)\delta_c(0)$. Following the Gaussian distribution from the PS formalism as a first model the fraction of objects with mass between $M_1, M_1 + \mathrm{d}M_1$ inside this region is given by

$$f(\sigma_1, \delta_c|\sigma_0, \delta_0)\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\mathrm{d}M_1 = \frac{1}{(2\pi)^{1/2}}\frac{\delta_c - \delta_0}{(\sigma_1^2 - \sigma_0^2)^{3/2}}\exp\left[-\frac{(\delta_c - \delta_0)^2}{2(\sigma_1^2 - \sigma_0^2)}\right]\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\mathrm{d}M_1, \tag{2.44}$$

where σ_0, σ_1 are the variance on the density fluctuations on the not collapsed and collapsed region respectively. By equation 2.44 is easy to see that if we take the case were $\delta_0 = \sigma_0 = 0$ we get back to part of equation 2.19. Basically we are repeating that treatment but considering that the region of the sky has a density different from the average of the universe. This will lead to number density of clusters given by

$$\mathfrak{n}\,\mathrm{d}M_1 = -\frac{M_0}{M_1}f(\delta_c, \sigma_1|\delta_0, \sigma_0)\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\mathrm{d}M_1. \tag{2.45}$$

Naturally we can take the relative difference between the average value of cluster number density (halo mass function) with respect to the number density of cluster inside region R_0 and calculate the cluster overdensity as

$$\delta_h(\sigma_1, \delta_c|\sigma_0, \delta_0) = \frac{\mathfrak{n}(\sigma_1, \delta_c|\sigma_0, \delta_0)}{n(M_1, z_1)V_0} - 1, \tag{2.46}$$

where V_0 is the volume of that spherical region. In fact, we want to compare the halo overdensity with the fluctuation density at the time we are seeing the cluster hence we need to evolve δ_0 to some δ. This evolution depends on the collapse model, in the case of the spherical model in a De-Sitter universe the mass inside the initial sphere with size R_0 is conserved hence we have $R_0 = R(1 + \delta)^{1/3}$ so the volume in the time we are looking is $V_0 = V(1 + \delta)$.

Since we want to study clustering we need our region R to be much bigger than the cluster size, that is, $R \gg R_1$ and consequently the density fluctuation is much smaller $\sigma \ll \sigma_1$. Also we consider that the density in this region is close to the mean density of the universe hence $\delta \ll \delta_c$ so we can neglect terms depending on σ_0 and perform a Taylor expansion in first order on the density. In this approximation we have that $M_0 \approx \bar{\rho}_m V_0$ and get

$$\mathfrak{n}(\sigma_1, \delta_c|\sigma_0, \delta) \approx -V_0 \frac{\bar{\rho}_m}{M_1} \frac{2}{\sqrt{2\pi}} \frac{\mathrm{d}\ln \sigma_1}{\mathrm{d}M_1} \frac{(\delta_c - \delta)}{\sigma_1} \exp\left[-\frac{\delta_c^2}{2\sigma_1^2}\right] \exp\left[\frac{\delta_c \delta}{\sigma_1^2}\right] \exp\left[-\frac{\delta^2}{2\sigma_1^2}\right], \tag{2.47}$$

neglecting the last term and Taylor expanding the second exponential we get

$$\mathfrak{n}(\sigma_1, \delta_c|\delta) \approx -V_0 \frac{\bar{\rho}_m}{M_1} \frac{2}{\sqrt{2\pi}} \frac{\mathrm{d}\ln \sigma_1}{\mathrm{d}M_1} \exp\left[-\frac{\delta_c^2}{2\sigma_1^2}\right] \frac{(\delta_c - \delta)}{\sigma_1} \left(1 + \frac{\delta_c \delta}{\sigma_1^2}\right), \tag{2.48}$$

leading to

$$\mathfrak{n}(\sigma_1, \delta_c|\delta) \approx -V_0 \frac{\bar{\rho}_m}{M_1} \frac{2}{\sqrt{2\pi}} \frac{\mathrm{d}ln\sigma_1}{\mathrm{d}M_1} \exp\left[-\frac{\delta_c^2}{2\sigma_1^2}\right] \delta_c \left[1 + \delta\left(\frac{1-\nu_1^2}{\delta_c}\right)\right] = nV_0 \left[1 + \delta\left(\frac{\nu_1^2 - 1}{\delta_c}\right)\right], \tag{2.49}$$

where $\nu_1 = \delta_c/\sigma_1$.

Finally the halo overdensity in a region of the sky is given by

$$\delta_h(\sigma_1, \delta_c|\delta) = \left(\frac{\nu_1^2 - 1}{\delta_c} + 1\right)\delta = b(\nu_1)\delta. \tag{2.50}$$

where in the last equality we are defining the PS linear bias function $b(\nu_1)$. This linear bias definition depends on the choice of multiplicity function. Since we know that the PS do not give precise results in simulations, this bias has the same problem. However, we can use equation 2.50 to define the linear bias in suitable way to be determined in N-body simulations. This is done by calculating the halo variance in Fourier space

$$\langle \delta_h^2 \rangle = \int \mathrm{d}^3 k P_h(k) W_R^2(k) e^{ikR} = \int \mathrm{d}^3 k b(k)^2 P_m(k) W_R^2(k) e^{ikR} = b^2(\nu_1) \langle \delta^2 \rangle, \tag{2.51}$$

since both integrands must be equal we define the linear halo bias to be

$$\boxed{b^2(k) \equiv \frac{P_h(k)}{P_m(k)}}. \tag{2.52}$$

In this work we will mainly use the Tinker linear bias function [44]

$$\boxed{b(\nu) = 1 - A\frac{\nu^a}{\nu^2 + \delta_c^a} + B\nu^b + C\nu^c}, \tag{2.53}$$

where the parameters A, a, B, b, C, c are defined by the fitting and will depend on the cluster mass definition both by type (mean or critical) and by choice of Δ and the values used will be addressed in the analyses. Beyond this, we also impose a condition coming from our ansatz that all matter lies in halo and consequently

$$\int b(\nu)n(\nu)d\nu = 1. \tag{2.54}$$

A final consequence of the halo bias is that the actual number of clusters won't be given by equation 2.43 because will be slightly biased by a term of

$$\delta N = A_s \int_{M_{min}}^{M_{max}} \int_{z_{min}}^{z_{max}} n(M,z)\frac{c}{H(z)}\left(\int_0^z \frac{c}{H(z')}\mathrm{d}z'\right)^2 b(M,z)\delta \, \mathrm{d}z\mathrm{d}M. \tag{2.55}$$

This effect is known as super sample covariance and affects mainly survey regions with small angular size since the background density fluctuation is bigger. Super sample covariance will be crucial for future constraints on cosmological parameters given the data precision, however this will only be addressed in future work.

2.5 Proxies

In order to use expression 2.43 to constraint cosmological parameters we need to know the clusters mass and Redshift. For the redshift, we can use different techniques to measure it directly, however, the mass is not directly obtained. The solution is to associate the mass with one or more cluster observables. In both cases relations between the observed quantities and true mass and redshift are not deterministic, that is there exists only a probability $P(obs|true)$ that a cluster with mass M and redshift z yields a measured value M_{obs} and z_{obs}. Such a relation is called a proxy and plays a central role in connecting the theoretical developments in cluster cosmology and its observational counterparts.

There are two main methods for measuring cluster redshift. The first is by spectroscopy, that measures a cluster spectrum and by measuring the shift in its elements spectral lines determines the cluster redshift. This method yields a precise measurement of the cluster true redshift, however, it has a potential drawback. This measurement is expensive and requires more telescope time, consequently, it is very challenging to use it in large surveys. Therefore, we usually have spectroscopic catalogs with fewer objects and therefore may result in increased statistical error on the cosmological parameters. The second method is the photometric one in which the redshift is estimated by looking the light emitted by the clusters in different bands of the electromagnetic spectrum. This method allows us to measure more objects allowing the construction of larger catalogs. However, this introduces uncertainties in the determination of z making necessary the construction of a probabilistic relation between the true redshift z of the cluster with the photometric one z_p with a relation $P(z_p|z)$.

For the cluster mass the problem is more complicated since we cannot access directly, so we construct a probabilistic relation between one cluster observable and the cluster mass called mass-proxy $P(M_{obs}|M)$. This function gives the probability of a cluster with mass M yields an observable M_{obs}. There are different observables that can be used to infer mass, for example, the X-ray emissions from the plasma present in the the inter-cluster medium, its temperature is related to the cluster mass [45, 46]. Another example is the Sunyaev-Zel'dovich effect, that is the inverse Compton scattering of CMB photons within the plasma in the inter-cluster medium [47]. Finally, the weak lensing shear of background galaxies when their light passes through the cluster [48, 49]. This last method is the only one in which the dark matter interacts with the observable directly. There also the cluster richness, the number of bright galaxies in the cluster which is the mass proxy used is this work. $DC2$ uses the red-sequence Matched-filter Probabilistic Percolation (redMaPPer) algorithm to detect and determine the photometric redshift and richness of the clusters on the final catalogs [50].

Chapter 3

Statistical Developments for Cluster Counts analyses

Like every area o physics, astrophysical measurements are subject to uncertainties from the apparatus or systematic effects from the spatial environment before light arrives at us. In addition to these errors, there are some quantities in cosmology which comes from quantum fluctuations of the primordial universe like the fluctuations on the matter density field and that are fundamentally random. Naturally, this leads to a need of statistical tools for studying this uncertainties and relate the observations with our theoretical models.

In this chapter, we will present the statistical definitions and methods necessary to develop the results of this work. First, we will define what are the likelihood functions of a random variable and state the maximum likelihood method to find the best fit, that is the most probable value for a parameter given our models. However this information is not enough and in the second section we will state Bayes theorem and the Markov Chain Monte Carlo (MCMC) method to determine the parameters confidence region. Next, we will present the binned and unbinned likelihoods and cluster mean bias for the cluster number counts, the mass-observable relation adopted.

3.1 Maximum Likelihood estimators

In statistical theory there are two points of view to interpret the results, first is the frequentist and second the Bayesian. The first understand the results as the frequency a given dataset can be found given a probability distribution. While the Bayesian thinks the probability as a degree of confidence one has in a hypothesis. The formalism is very similar for both strands. In this work, we will adopt the Bayesian point of view.

Given a random event we call the space of all possible outcomes the sample space, for example a dice has six possible results, one through six, so the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. A subset of the sample space is called an event. In the dice roll case, an example of event is the result of falling at one or two, that is, $E = \{1, 2\}$. So, we define the probability of an event happening using a measure function that associates a real number for each event. The sum of such measurements for all disjoints events must be one. In our dice roll example, the size of the sample space is 6 while the event in our example has size 2, thus, if we assume all events are equiprobable the probability of the dice lands with size 1 or 2 up is $P(1\ or\ 2) = 2/6 = 1/3$. For a formal definition and construction for the size of general sample spaces the reader should see [51].

What is usually done is assign each element of the sample space to a real number such a map is called a random variable. For the dice roll example we can define a random variable x as a function given by $x : \Omega \to R| \{x(i) = i\}$, where i is the number written on the side up. Thus, instead of writing the probability in terms of the event with one element, we will write in terms of the random variable associated with that event $P(x_i)$. If we want to calculate the events with more than one element we just add the probabilities $P(x_i\ or\ x_j) = P(x_i) + P(x_j)$. Also we can impose that the probability of any event on the sample space happening to be one $P(x(\Omega)) = 1$.

Of course our sample space can be uncountable like the height distribution of a population. In this case, it's meaningless to talk about the probability of measuring a single height. Instead, we calculate the probability of a random variable to be in a interval $x, x + \mathrm{d}x$ by

$$P(x, x + \mathrm{d}x) = p(x)\mathrm{d}x, \tag{3.1}$$

and $p(x)$ is the probability density function (pdf). Generalizing for finite intervals the probability of a random variable be found in a interval $[x_0,\ x_f]$ is

$$P(x_0 < x < x_f) = \frac{\int_{x_0}^{x_f} p(x)dx}{\int_{-\infty}^{\infty} p(x)dx}, \tag{3.2}$$

and we set the denominator integral to one. In the Bayesian point of view, we understand the expression 3.2 as the confidence we have that the random variable will lie in that interval. Two important pdf's in statistics are the Gaussian and Poisson given respectively by

$$p_g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad p_p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \tag{3.3}$$

note that in the Poisson distribution x is a discrete variable. μ, σ, λ are fixed parameters.

Another key notion in statistics is the distribution moments which are defined as

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n p(x)\mathrm{d}x, \tag{3.4}$$

where n is the number of the moment, they yield information of p and we can reconstruct the full pdf from all of it's moments [52]. In practice, we will be interested in the first moment which is only the mean of the distribution and the variance given by $\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$. For the Gaussian and Poisson distribution the mean is respectively $\langle x_g \rangle = \mu$, $\langle x_p \rangle = \lambda$ and variance is $\sigma_g^2 = \sigma^2, \sigma_p^2 = \lambda$.

In more realistic analyses we usually do not have much information about the pdf that describes a phenomenon we have only a sample $\{x_1, x_2, ..., x_N\}$ of N measurements. For example, we do not know the distribution of a populations height, we have only a sample of the height of people we measured. Naturally the question arises how we can extract pdf information from these samples? One method is by using estimators of the distribution parameters. These estimators are function of the data, the two most common estimators are the arithmetic mean

$$\hat{x} = \frac{\sum_{i=1}^{N} x_i}{N}, \tag{3.5}$$

which estimates the mean distribution μ and sample variance

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (\hat{x} - x_i)^2}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i^2 - \hat{x}^2, \tag{3.6}$$

which estimates the distribution variance σ^2. We can construct estimators for any parameter of our distribution, however, in order for this construction to be useful the estimator must have three properties, they need to be unbiased, consistent and efficient. To understand these properties we need to consider an estimator $\hat{\theta}$ of a parameter θ as random variable itself with its pdf $p(\hat{\theta})$. So the estimator is said to be unbiased if it's expectation value is the true value of the parameter that is

$$\langle \hat{\theta} \rangle = \int_{-\infty}^{\infty} \hat{\theta} p(\hat{\theta}) \mathrm{d}\hat{\theta} = \theta. \tag{3.7}$$

The arithmetic mean in equation 3.5 is an example of unbiased estimator since

$$\langle \hat{x} \rangle = \frac{\sum_{i=1}^{N} \langle x_i \rangle}{N} = \frac{\sum_{i=1}^{N} \int p(x_i) x_i dx_i}{N} = \frac{\sum_{i=1}^{N} \mu}{N} = \mu \tag{3.8}$$

where μ is the true mean of the distribution. While the sample variance estimator is biased since

$$\sigma_{\hat{x}}^2 = \frac{1}{N} \left(\sum_{i=1}^{N} \langle \hat{x}_i^2 \rangle - \mu^2 \right) = \frac{1}{N} \left(\sum_{i=1}^{N} \int p(\hat{x}_i) \hat{x}_i^2 d\hat{x}_i - \mu \right) = \frac{1}{N} \left(\sum_{i=1}^{N} \sigma^2 + \mu^2 - \mu^2, \right) = \frac{\sigma^2}{N} \tag{3.9}$$

$$\langle \hat{\sigma}^2 \rangle = \frac{\sum_{i=1}^{N} \langle x_i^2 \rangle - \langle \hat{x}^2 \rangle}{N} = (\sigma^2 + \mu^2) - (\sigma_{\hat{x}}^2 + \mu^2) = \left(\frac{N-1}{N} \right) \sigma^2 \tag{3.10}$$

where σ^2 is the true variance of the distribution and $\sigma_{\hat{x}}$ is the variance of the mean estimator, the first equality comes from the definition of the distribution variance.

The second property is consistency this means that if the number of measurements N in the sample increases the probability of the estimator yields the true value of the parameter also increases that is if

$$N \to \infty \Rightarrow P(\hat{\theta} = \theta) \to 1. \tag{3.11}$$

One way to practically check the consistency of the estimator is by looking at their variance, for instance by equation 3.9 we can see that the if N growths the mean estimator variance goes to zero and hence it's consistent. Finally, the last property is efficiency, this is measured in terms of the estimator variance. Unlike the other two properties one cannot usually find the most efficient estimator only compare the efficiency of two estimators.

The choice of estimators for a given distribution parameters are rather arbitrary. A natural question arises, how can we construct a good estimator? One way to do this requires that we first define the likelihood of the dataset. A likelihood $L(\{x_i\}|\vec{\theta})$ is the conditional probability of obtain a sample of measurements $\{x_i\}$ given that the value of your parameters model are $\vec{\theta}$. Note that the likelihood is a pdf in the measurements $\{x_i\}$ not in $\vec{\theta}$. Since L yields the probability of obtaining that data ,we naturally want to know what values of $\vec{\theta}$ makes our data most likely to happen. Hence, we will choose our estimators to be the ones that maximizes the likelihood, that is

$$\hat{\vec{\theta}} = max\ L(\{x_i\}|\vec{\theta}). \tag{3.12}$$

This method is called maximum likelihood estimator (MLE). Let us consider for example a sample $\{x_i\}$ from a Gaussian distribution with variance one and want to find the MLE for the mean μ. In this case we need only to calculate the derivative of equation 3.3 and make it zero resulting in

$$\frac{\mathrm{d}p(\{x_i\}|\mu)}{\mathrm{d}\mu} = \frac{1}{\sqrt{2\pi}} exp\left[-\frac{\sum_{i=1}^{N}(x_i-\hat{\mu})^2}{2}\right]\sum_{i=1}^{N}(x_i-\hat{\mu}) = 0 \Rightarrow \hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}x_i, \tag{3.13}$$

which is just the arithmetic mean. In general, we cannot assure that the estimators generated from this method have all the three properties we want (however the theory of the MLE estimators prove that all three are asymptotically true given mild assumptions about the likelihood smoothness). For the likelihoods used in this work the estimators are consistent. However, the estimators from these likelihoods can have significant bias depending of the sample size and mass observable relation as shown in [53].

The shape of the cluster counts likelihood will be given in the followings sections and we will use this method to state the most likely value for the cosmological parameters. However, we will not test if the shape of the likelihood function is correct or not. Such an analyses require further tools which goes beyond the scope of this work.

3.2 Bayes Theorem and Markov Chain Monte Carlo method

In the last section, we show the MLE method to find the most likely value for a parameter given a likelihood. Every measurement has uncertainties associated with it, hence, we need a full region of possible values to estimate the error bars in our cosmological parameters. In the Bayesian viewpoint, we want to update our confidence region on the parameters value after collecting the data, therefore, we want the probability of the parameters have value $\vec{\theta}$ given our dataset $P(\vec{\theta}|\{x_i\})$, also called the posterior distribution.

We calculate the posterior probability using Bayes theorem given by

$$P(\vec{\theta}|\{x_i\}) = \frac{L(\{x_i\}|\vec{\theta})P(\vec{\theta})}{P(\{x_i\})}, \tag{3.14}$$

where L is our likelihood function, $P(\vec{\theta})$ is our prior distribution and $P(\{x_i\})$ is the probability of obtaining the data independent of our model. The prior distribution represents our degree of confidence in a hypothesis before we collect any data stating our prejudices with respect to the model. There are two problems when choosing the prior that one must be aware. First and most obvious is choosing a prior to be too restrictive. For example choosing a Gaussian prior with a small variance will make the posterior dominated by the prior information, yielding the wrong confidence region for the parameter. The second problem is the opposite choosing a prior that has too little information, in that case we may be unable to reach any conclusion. For example, if we choose a prior to be close to zero in equation 3.14 the value of the posterior will be close to zero, that is we have not gained much new information. In the analyses performed here we will use an unnormalized flat prior, that is all the values in a fixed interval will be equally likely

$$P(\theta_i) = 1, \text{ if } \theta_{i,min} \leq \theta \leq \theta_{i,max}. \tag{3.15}$$

At a first look appears this prior has the least information possible, however when choosing a flat prior we know the scale of the quantity. For instance we know the matter density parameter Ω_m is close to one hence we can use a flat prior. If we deal with quantities that may vary widely in scale like the mass of clusters we could use a flat log, that is the $\ln(\theta_i)$ is equally likely. But since the cosmological parameters we are going to estimate have a well limited scale we will stick to the flat prior.

The denominator of Bayes theorem is marginalized likelihood, mathematically is only the posterior normalization and given by

$$P(\{x_i\}) = \int L(\{x_i\}|\theta)P(\vec{\theta})\mathrm{d}\theta. \tag{3.16}$$

Usually we cannot calculate the normalization explicitly since it's computationally expensive, however. when running the MCMC we won't need the normalization constant. For this reason, we do not need to worry about it. In order to define the confidence interval for a parameter θ_i of our model marginalize the posterior in all the other parameters that is

$$P(\theta_1|\{x_i\}) = \int P(\vec{\theta}|\{x_i\})d\theta_2...d\theta_N. \tag{3.17}$$

And from the last equation we can determine the error bar of the parameter as the region $[\theta_{1,min}, \theta_{1,max}]$ encapsulating $1-\alpha$ of the probability of the parameter be in that interval by solving

$$1-\alpha = \int_{\theta_{1,min}}^{\theta_{1,max}} P(\theta_1|\{x_i\})d\theta_1. \tag{3.18}$$

We say that θ is in $[\theta_{1,min}, \theta_{1,max}]$ with $1-\alpha$ confidence level. Usually we will use the 68.2%, 95.4% and 99.7% confidence levels. We can extend this notion to higher dimensional confidence regions. We integrate over all other parameters obtaining an integral as equation 3.18, but now in the dimension of our confidence region. Usually, we are interested in the two dimensional regions. This integrals however are costly to be performed hence we turn to numerical methods to estimate the posterior and one method is the Markov Chain Monte Carlo (MCMC).

The idea of the MCMC method is to construct a sequence of samples of a target distribution which is difficult to sample from. This sequence must converge to the target distribution, in our case the posterior distribution of our cluster counts models. First, we need to construct a Markov Chain which is a sequence of realizations of our random variables $\{\vec{\theta}^1, \vec{\theta}^2, ...\vec{\theta}^N\}$ where the current element depends only on the last that is $\vec{\theta}^i$ depends only on $\vec{\theta}^{i-1}$. The idea is that we start sampling with a random choice based on some criteria and as the chain evolves we converge to the posterior distribution. From there, we use the sample the posterior distribution to construct the confidence intervals. The first elements in the MCMC samples are usually dominated by the movement towards the target distribution center and is called the burn in. In order to the chain converge to the desired distribution we need to satisfy

$$\int P(\vec{\theta}^i|\{x_i\})K(\vec{\theta}^i \to \vec{\theta}^{i+1})d^N\theta^i = P(\vec{\theta}^{i+1}|\{x_i\}), \tag{3.19}$$

where $K(\vec{\theta}^i \to \vec{\theta}^{i+1})$ is the transition kernel. This condition is only the requirement that the point $\vec{\theta}^{i+1}$ is sampled from the same distribution as $\vec{\theta}^i$.

To construct the Markov Chain we need to provide the information of walking to the current step to the next, this relation is probabilistic and given by the transition probability $T(\vec{\theta}^{i+1}|\vec{\theta}^i)$, once we establish this function we can sample from this distribution the next sample of the chain. Given a next step in the chain we need to accept or reject it, there are many ways of doing this. The algorithm we will use is based on the Metropolis-Hastings criteria, that is the probability of accepting a new point given the old one is

$$A\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right) = \min\left\{1, \frac{P(\vec{\theta}^{i+1}|\{x_i\})T(\vec{\theta}^i|\vec{\theta}^{i+1})}{P(\vec{\theta}^i|\{x_i\})T(\vec{\theta}^{i+1}|\vec{\theta}^i)}\right\}. \tag{3.20}$$

Since we only look at the ratio of the likelihoods to accept or reject the point we don't need to calculate the normalization of the posterior function. We will want to construct reversible Markov Chains, that is the probability of taking the step $\vec{\theta}^i \to \vec{\theta}^{i+1}$ must be the same of going back $\vec{\theta}^{i+1} \to \vec{\theta}^i$, this is called the detailed balance condition

$$P(\vec{\theta}^i|\{x_i\})T\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right)A\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right) = P(\vec{\theta}^{i+1}|\{x_i\})T\left(\vec{\theta}^i|\vec{\theta}^{i+1}\right)A\left(\vec{\theta}^i|\vec{\theta}^{i+1}\right). \tag{3.21}$$

With the detailed balance condition we can define the kernel transition as

$$K(\vec{\theta}^i \to \vec{\theta}^{i+1}) = A\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right)T\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right) + \left[1 - \int A\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right)T\left(\vec{\theta}^{i+1}|\vec{\theta}^i\right)d^N\theta^{i+1}\right]\delta^C_{\theta^{i+1},\theta^i}, \tag{3.22}$$

where the second term is to account for the case where we reject the next point. By this definition for the kernel together with the detailed balance condition we can assure condition 3.19 to be satisfied. Each algorithm has it's own way of prescribing the transitions probabilities. We will use APES algorithm proposed in [54].

Once we obtained the samples from the posterior distribution using the MCMC we can calculate the expectation value of any function of a given cosmological parameter by

$$\langle F(\theta^i) \rangle \approx \frac{1}{N} \sum_j F(\theta^i_j), \tag{3.23}$$

we usually use this to calculate the mean and the variance of the parameter distribution. Another simplification coming from the MCMC is the calculation of the marginalized distributions. Instead of performing multidimensional integrals now we only ignore the columns of the resulting matrix. And to calculate the $1D$ and $2D$ confidence regions with the MCMC we use the points to calculate the interpolation of marginalized posterior and from this interpolation we calculate level surfaces and count the fraction of points lying inside this region. When the fraction of points inside this region is 68.2%, 95.4% and 99.7% we found the 1σ, 2σ, 3σ regions.

3.3 Binned and Unbinned cases

Once we predicted the number counts of clusters we need a likelihood function to compare this prediction with the observer number of clusters, but before we need to establish the observable relation between mass and richness and for the redshift. The latter we will consider only the spectroscopic measurement hence the relation between the true redshift and observed one is given by a Dirac delta function

$$P\left(z_{obs}|z\right) = \delta^D(z_{obs} - z), \tag{3.24}$$

that is the measured value is the true redshift. For the mass we will consider two cases, the first will be the simplest one which we know the true mass and also will be given by a delta function

$$P\left(M_{obs}|M\right) = \delta^D(M_{obs} - M). \tag{3.25}$$

We will consider this simplest case for two reasons, first is to test the tools in a controlled scenario and second to compare the loss of precision when including the proxy. Then, we will include the the richness proxy following a normal distribution with mean and variance depending on the mass and redshift of the cluster given by

$$P(\ln\lambda|M,z) = \frac{1}{\sqrt{2\pi}\sigma(M,z)}\exp\left[-\frac{(\ln\lambda-\mu(M,z))^2}{2\sigma(M,z)^2}\right], \tag{3.26}$$

$$\mu(M,z) = \mu_0 + \mu_M \ln\frac{M}{M_0} + \mu_z \ln\frac{1+z}{1+z_0}, \tag{3.27}$$

$$\sigma(M,z) = \sigma_0 + \sigma_M \ln\frac{M}{M_0} + \sigma_z \ln\frac{1+z}{1+z_0}, \tag{3.28}$$

where M_0 and z_0 are respectively a fiducial mass and redshift to be chosen. In a realistic analyses, we will need to fit all the parameters in the mass-observable relation which worsen the constraints on the cosmological parameters due to the high dimensional space on the parameters. We can reduce this issue by considering the cluster clustering analyses and the bias or including different kinds of data, like supernova or BAO, but in this work we do not carry this analysis will only fit a small subset of these parameters in a controlled environment.

Once we have the proxies we marginalize the cluster number counts in the true mass and Redshift and find the distribution of clusters with observed mass $M_{obs,i}$ and Redshift $z_{obs,i}$ given by

$$\boxed{x_i = A_s \int_{z_{min}}^{z_{max}}\int_{M_{min}}^{M_{max}} \frac{\mathrm{d}^2V(z)}{\mathrm{d}z\mathrm{d}\Omega}\frac{\mathrm{d}n(M,z)}{\mathrm{d}M} P(M_{obs,i}|M,z)P(z_{obs,i}|M,z)\mathrm{d}M\mathrm{d}z}. \tag{3.29}$$

To get the full clusters in the range of the observed quantities we only integrate x_i is the remaining variables, that is.

$$N^T = \int_{z_{obs,min}}^{z_{obs,max}}\int_{M_{obs,min}}^{M_{obs,max}} x_i \,\mathrm{d}M_{obs}\mathrm{d}z_{obs}. \tag{3.30}$$

We define the linear bias of a given cluster in similar fashion

$$\boxed{b_i = \frac{A_s}{x_i}\int_{z_{min}}^{z_{max}}\int_{M_{min}}^{M_{max}} \frac{\mathrm{d}^2V(z)}{\mathrm{d}z\mathrm{d}\Omega}\frac{\mathrm{d}n(M,z)}{\mathrm{d}M} b(M,z)P(M_{obs,i}|M,z)P(z_{obs,i}|M,z)\mathrm{d}M\mathrm{d}z}\,. \tag{3.31}$$

In order to construct the unbinned likelihood we need to construct the probability of finding a cluster with observed mass $M_{obs,i}$ and redshift $z_{obs,i}$, but since we know the cluster distribution, we can find this probability normalizing equation 3.29 with equation 3.30. Then the probability of the whole sample is just the product of the individual probabilities. In addition to this we need to account for the probability of measuring a different number of clusters than the predicted one, to model this we use a Poisson distribution. Thus, the likelihood function for the cluster counts individually is

$$\boxed{51L = \frac{\left(N^T\right)^N e^{-N^T}}{N!}\prod_{i=1}^{N}\frac{x_i}{N^T}}\,. \tag{3.32}$$

We called this the cluster counts unbinned likelihood [53].

When using the unbinned likelihood, we use the full information presented by the catalog and naturally this yields the best constraint possible for the cosmological parameters. However, the MCMC method for this likelihood is very costly and could take a long time even using a supercomputer. So, it is useful derive a new likelihood which gives up some information of the catalog if it allows this constraints to be computed faster. We do this by binning the observed quantities in a small number of bins and instead of having the distribution of clusters we calculate the number of clusters in a bin α of redshift and a bin β of observed mass by

$$\boxed{N^T_{\alpha\beta} = \int\limits_{z_{obs,\alpha}}^{z_{obs,\alpha+1}} \int\limits_{M_{obs,\beta}}^{M_{obs,\beta+1}} x_i \, dM_{obs} dz_{obs}} \, . \tag{3.33}$$

As for the linear bias we now calculate the bias in the bin and is

$$\boxed{b_{\alpha\beta} = \int\limits_{z_{obs,\alpha}}^{z_{obs,\alpha+1}} \int\limits_{M_{obs,\beta}}^{M_{obs,\beta+1}} b_i \, dM_{obs} dz_{obs}} \, . \tag{3.34}$$

Obviously when binning the data we need to provide a new likelihood for the number of clusters inside the bin. Following the idea of the probability of measuring N having predicted N^T in equation 3.32 the probability of measuring $N_{\alpha,\beta}$ in a bin given the theoretical prediction is given by a Poisson distribution. Hence, the likelihood is the product over all bin pairs, that is

$$\boxed{L = \prod_{\alpha}^{n_\alpha} \prod_{\beta}^{n_\beta} \frac{\left(N^T_{\alpha,\beta}\right)^{N_{\alpha,\beta}} e^{-N^T_{\alpha,\beta}}}{N_{\alpha,\beta}!}} \, , \tag{3.35}$$

where n_α and n_β are the number of redshift and observed mass bins respectively. We should stress that we can have more bins if we consider other set of mass-observable relations, for instance if we use combined with richness lensing mass for the cluster we will have another product over the bins on the lensing masses as exemplified in [55].

Naturally when we bin the data we will be losing information from the catalog and this will depend on the size and number of bins we choose, for instance if choose only one bin we will be giving up almost all the information about mass and redshift. Another problem with the binning process is that we cannot provide uncertainties for each measurement, instead we talk only on the average value of the proxy inside the bin. Also there is another issue in consider the binning likelihood, as shown in [53] the biases on the parameters best fit is bigger than the unbinned case. However, it is faster to run the MCMC with likelihood 3.35, hence it must be consider case by case if is worth it giving up the precision.

As a final remark is worth pointing out that besides the information about the proxy a more realistic analyses would require the knowledge of the catalog completeness and purity. For this we would need to provide two more functions to correct the cluster counts distribution in equation 3.29.

Chapter 4

NumCosmo and CCL Cross-Check

The results presented in this chapter comes from the project 203 of LSST-DESC group. Here, we compared the prediction of cluster counts and mean halo bias from the libraries NumCosmo and CCL. These libraries will be used by the collaboration, therefore, it is crucial they are providing the correct results in their predictions. The only way of truly testing them is to perform a direct comparison since they were built independently and should not have the same bugs and errors.

Beyond their internal consistency, we can compare their efficiency on predicting cluster count and halo bias. It is important that these libraries perform calculations as fast as possible since they will be needed to calculate the likelihood later and since the MCMC is computationally expensive. We need these numerical integrals to be performed in the fastest way while keeping a reasonable precision. When the Vera Rubin is finished these two libraries will provide likelihoods to be used by the collaboration and to ensure their functionality, we must check it in all levels of the libraries. From their precision in calculating the background cosmology to the construction of the halo mass function and proxies until the numerical integration of the cluster counts prediction.

This chapter will first present the cross-check of the true table of cluster counts prediction, that is, when we know the true mass and redshift of the cluster in the unbinned case where we only need to calculate the mass function and the binned case where we will need to perform the numerical integration in each bin. In the second section we repeat the cross-check adding richness-mass relation to see how it changes the agreement and efficiency of each library. Finally, we will compare the mean halo bias predicted and show the NumCosmo implementation of the mean bias when we have a proxy. This case is still not implemented in CCL.

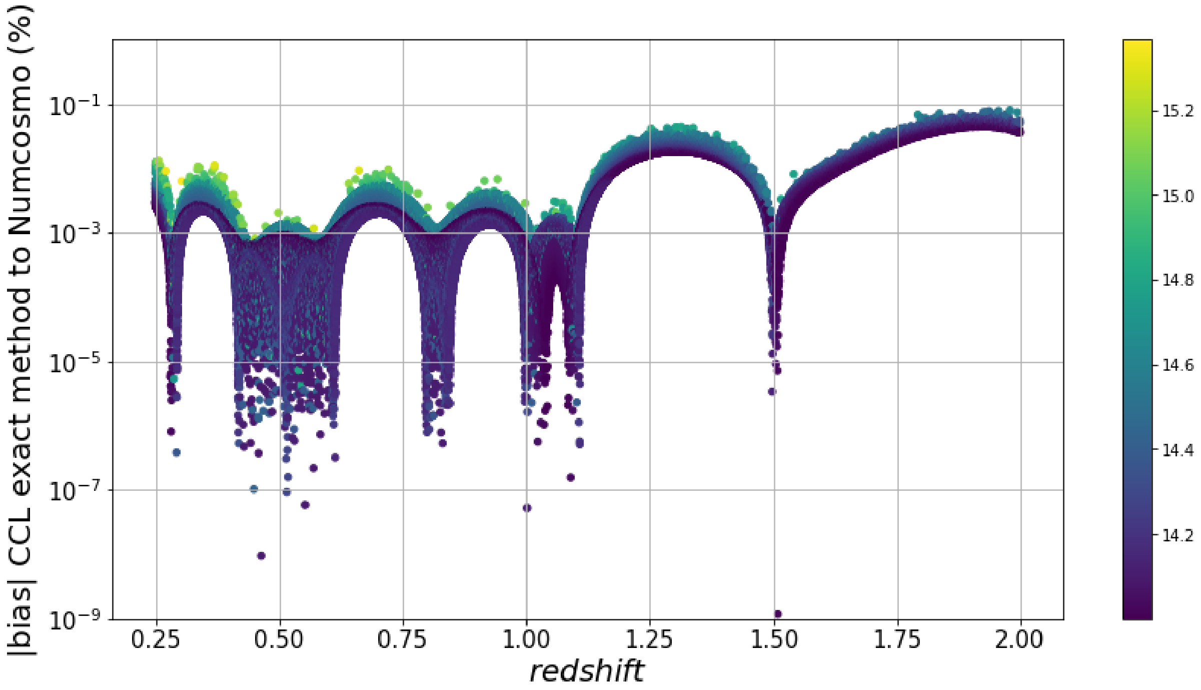

FIGURE 4.1: Relative difference between NumCosmo and CCL cluster count distribution, the color map is for the different cluster mass.

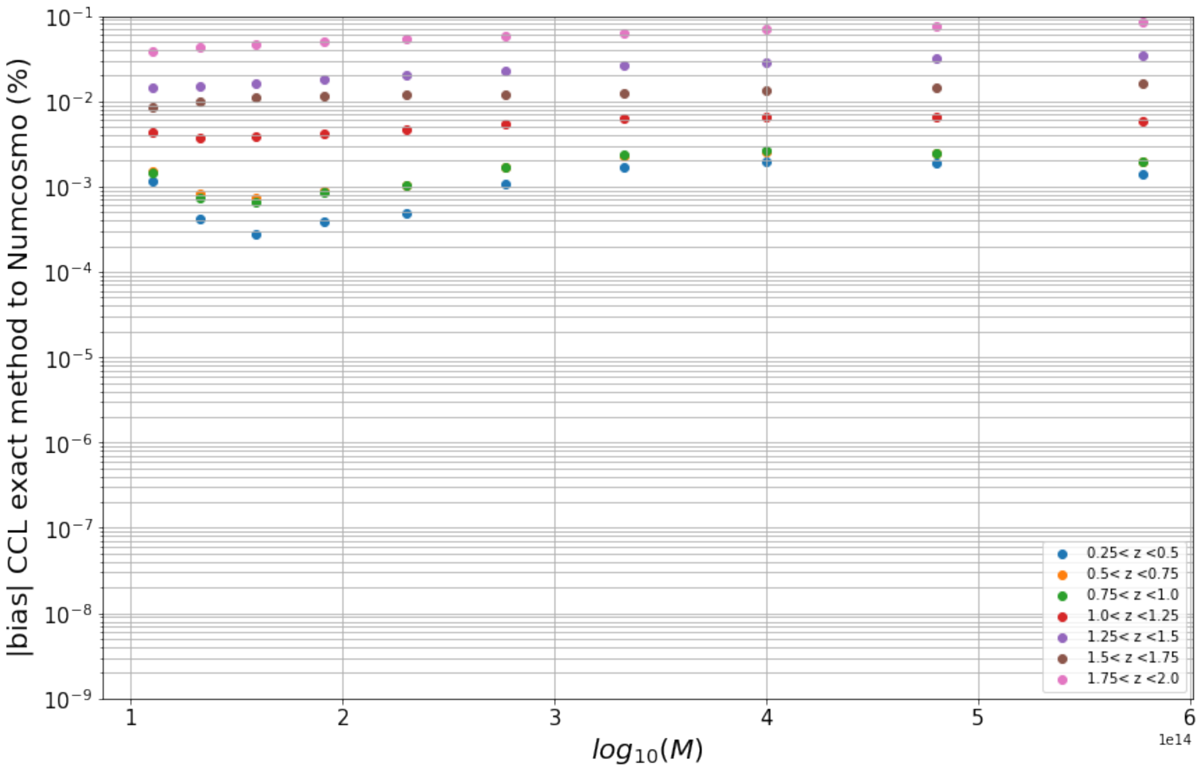

FIGURE 4.2: Relative difference between NumCosmo and CCL cluster count prediction inside the bins, each color stands for a different redshift bin.

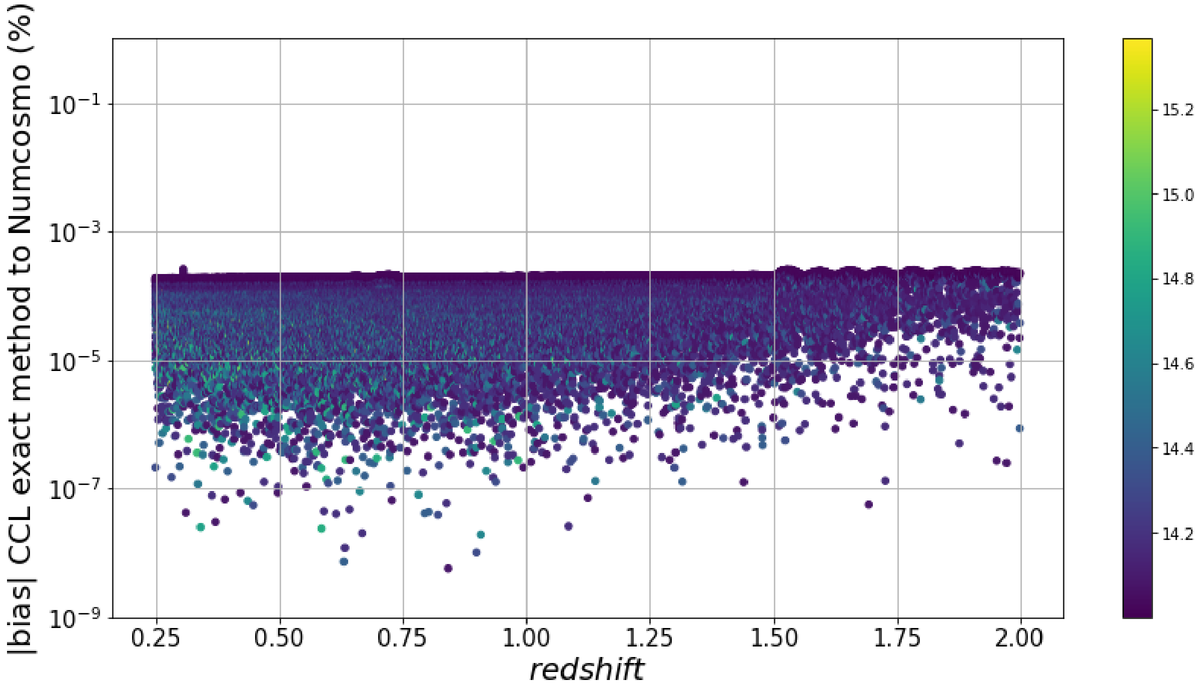

FIGURE 4.3: Relative difference between NumCosmo and CCL cluster count distribution, the color map is for the different cluster mass with set-high-precision option on.

4.1 True table

In order to perform the integrals comparison, we constructed a mock catalog with the fiducial cosmological parameters used in the $DC2$ simulation these values are given in table 4.1. We constructed the mock catalog with a survey area of 5264.05 deg^2, which is the survey area of the $DC2$ *skysim*5000_*v*1.1.1 halo catalog. Used a depth of $z \approx 2$ and mass range of $[10^{14}, 10^{16}]M_{\odot}$. We used these values to have a suitable range to compare the predictions in a wide range for possible intervals of a future analyses. For the mass definition of the cluster we choose as a reference the critical density with a $\Delta = 200$. But since the Tinker multiplicity function is fitted in terms of the mean matter density of the universe we need to convert Δ by

$$\Delta' = \frac{\Delta}{\Omega_m(z)}, \tag{4.1}$$

in order to predict the correct results for the cluster abundance. The parameters used in the multiplicity function are a function on Δ' hence we need to interpolate the parameters A_0, a_0 and b_0 to be used, the values set by our initial definition of Δ are given in [28].

We binned the data into ten equally log spaced mass from 10^{14} to $10^{14.8}$ and seven evenly spaced redshift bins ranging from 0 to 2. Then we set NumCosmo and CCL to calculate the cluster counts prediction in both unbinned and binned cases and the results are displayed in figures 4.1, 4.2. From this two figures we can see that there are cases in which their results agree only up to 0.1% especially in the binned case. The binned case has a higher difference because is performed a numerical integration in each bin. This level of agreement is insufficient

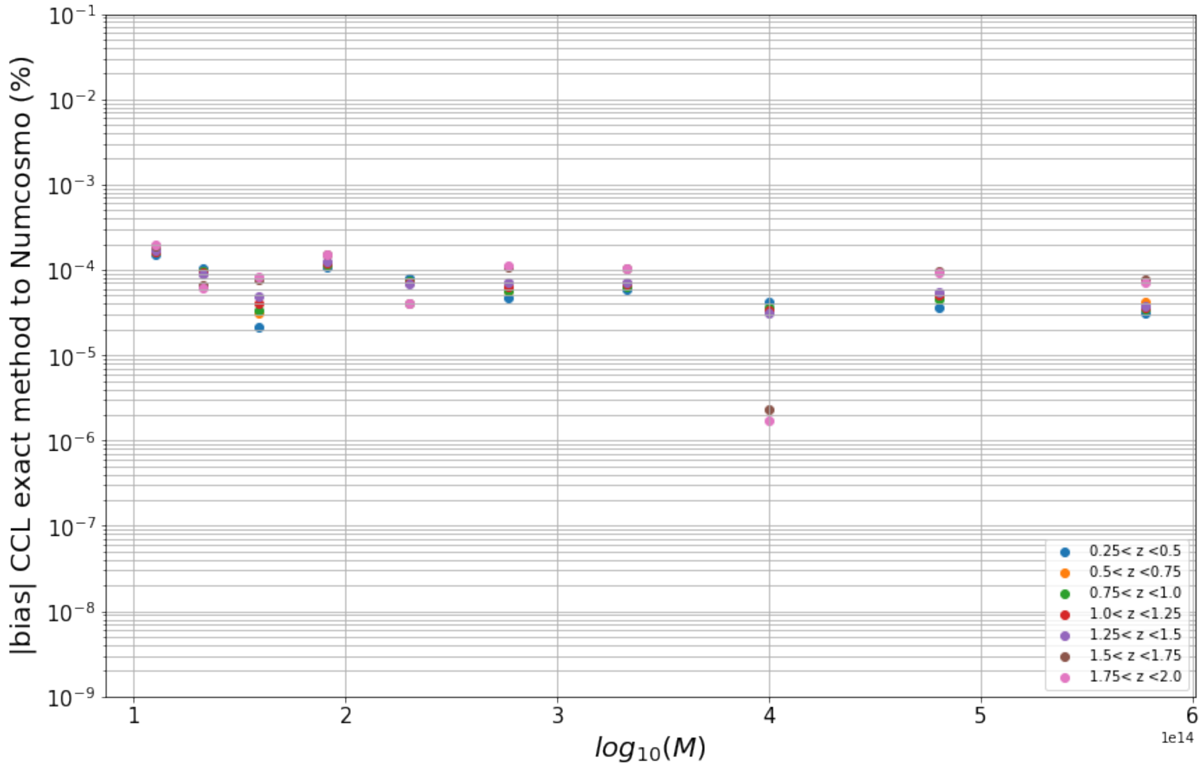

FIGURE 4.4: Relative difference between NumCosmo and CCL cluster count prediction inside the bins, each color stands for a different redshift bin with set-high-precision option on.

within the precision required to compare future analyses on constraints of the cosmological parameters. This disagreement comes from the lack of precision in CCL's default configuration when performing the splines to interpolate the mass function, Hubble function $H(z)$, cosmological distances and other cosmological functions. With that in mind we developed a python script in NumCosmo to change these configurations in CCL and match with the ones used in NumCosmo named *ccl_cosmo_set_high_prec*.

With this option on we repeated the comparison and the results are given by figures 4.3, 4.4 for the unbinned and binned case respectively. Now the two libraries agree up to 10^{-4}% even for the binned case, leading to a satisfactory result. Another important improvement when using this option is that this difference has a small variation, that is, while in the first case there were points 10^{-1} and 10^{-6}% with the high precision configuration now they tend change less allowing a greater control on this difference.

With the two libraries agreeing in their predictions, we can compare their efficiency in calculating those quantities. First we see that when we set the high precision option the cluster distribution and number counts in the bin are calculated faster in a time 15.44s and 21.21s respectively, without this option it increases to 19.23s and 26.64s. This happens because the functions are subject to a smaller error and the integrals can be calculated faster, however the time increases when CCL is preparing these interpolations from 0.3s to 23.30s making the overall time increase for the true table case. Now NumCosmo perform the same calculations in

Ω_m	0.265
H_0	71km/sMpc
Ω_k	0
σ_8	0.8
Ω_b	0.0448

TABLE 4.1: Fiducial Cosmological parameters used in *DC*2.

0.71s and 0.42s which is significantly lower when compared to CCL.

4.2 Mass Proxy table

With the agreement between the cluster counts for a true table we can introduce the richness-mass relation given by equation 3.26. We now constructed a catalog using NumCosmo as the latter but now we assign a richness for each cluster ranging from $[0, 10^3]$ and removed the clusters with richness larger than 100. The bins distribution for the redshift is the same as the last and we binned the base ten logarithm of the richness in seven equally spaced bins from 0 to 2. The proxy parameters used in constructing the mock catalog are shown in table 4.2. Here we only display the results for the set high precision option. It's important to mention however that again when setting this option the numerical calculations time are reduced from 21min45s and 8min43s to 19min10s and 5min31s in the unbinned and binned case respectively. Now even when considering the time to prepare the interpolations the code runs faster with the high precision option on.

Now when comparing the CCL and NumCosmo the agreement was close showing the precision option still solves the problem for this case. Also in this case we can see a tendency of a slightly worsen when we perform the calculation in higher redshift and lower richness value. However when it comes to the efficiency the difference was even bigger since CCL takes the times mentioned above and NumCosmo took 0.6s and 0.04s in the unbinned and binned case respectively. Here an additional difference between the two libraries when calculating the integral in the observable mass. While CCL perform a full numerical integration using *scipy*, in NumCosmo we perform the integration on the observed mass first and doing this can calculate analytically making the code more efficient.

In order to solve this efficiency problem CCL developed two approximate methods. The first is to construct a tabulated grid with the integrand of cluster counts and from this grid use the Simpson integration method calculate the cluster distribution. The second method is to use the grid to construct a interpolation from the integrand and perform the numerical integration using *scipy*. The results of these comparisons are given in figures 4.7 and 4.8. In the unbinned case we only check the grid method since we do not need to perform the numerical integration and the interpolation is not needed. We can see a good agreement between the exact method in most clusters and the execution time was reduced to 8.3s which is significant. Now in the binned case we lose precision in the grid method due to the numerical integration and the execution time was 0.6s, while the interpolation method showed to be in better agreement but the

μ_0	3.190
μ_M	0.868
μ_z	-0.304
σ_0	0.330
σ_M	-0.034
σ_z	0.0

TABLE 4.2: Richness-mass relation parameters.

execution time of $34.62s$. Both these approximate methods attenuates the efficiency problem, however they still do not perform as fast as NumCosmo code within the same precision. The notebooks done to compare CCL different methods were created by the collaboration and are presented in the GitHub. [1]

[1] https://github.com/LSSTDESC/CLCosmo_Sim/tree/issue/15/binned_cosmological_analysis/notebooks

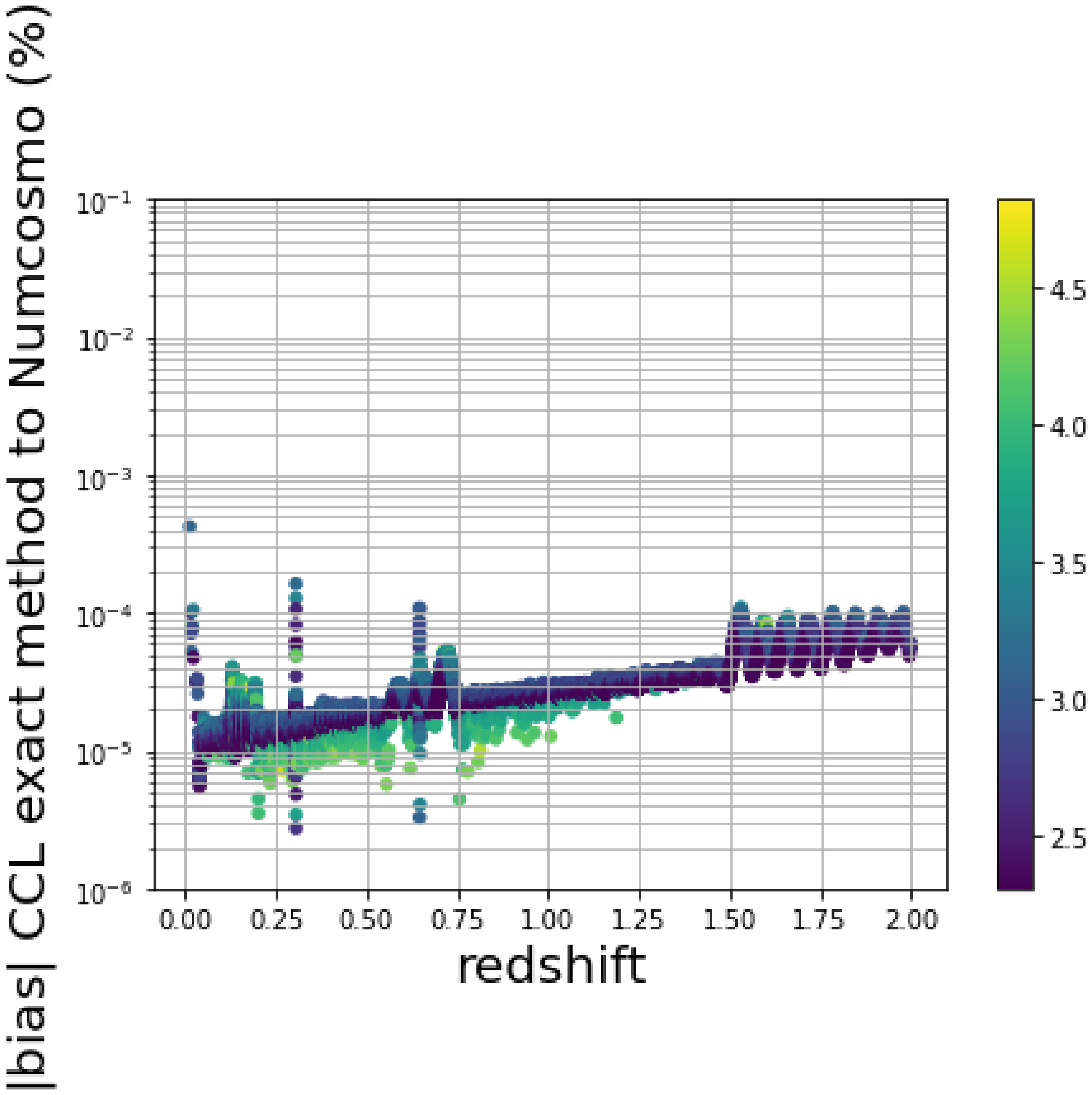

FIGURE 4.5: Relative difference between NumCosmo and CCL cluster count distribution, the color map is for the different cluster richness.

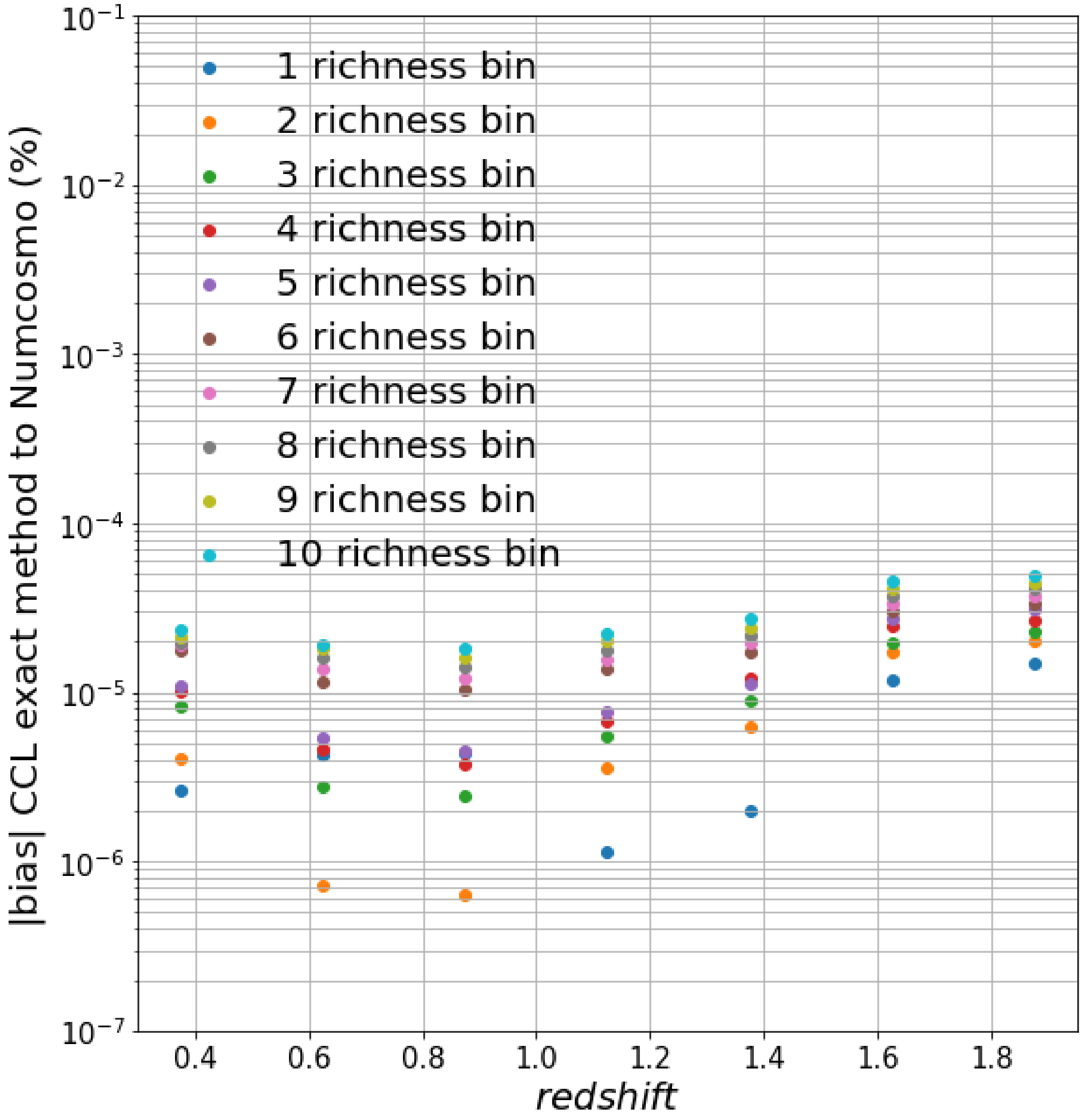

FIGURE 4.6: Relative difference between NumCosmo and CCL cluster count prediction inside the bins, each color stands for a different richness bin.

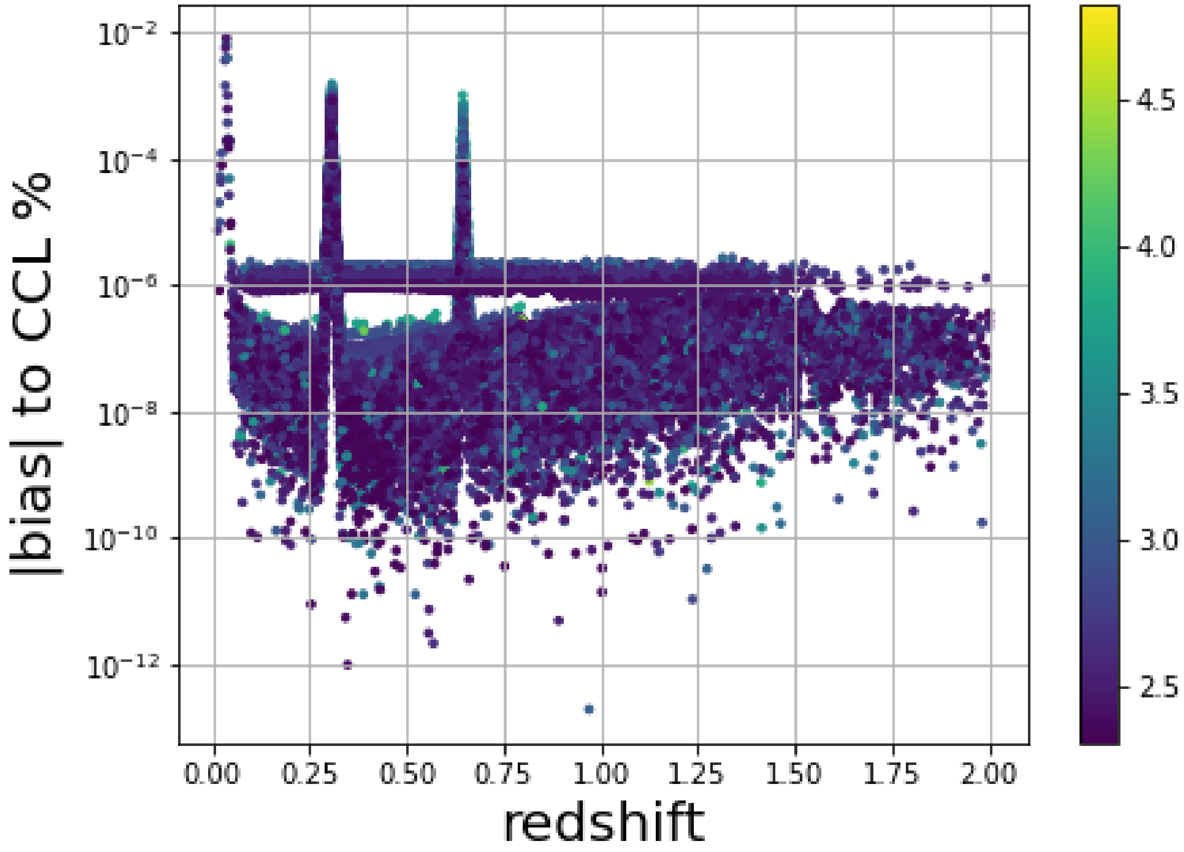

FIGURE 4.7: Relative difference between CCL grid method and CCL exact method, the color map is for the different cluster richness.

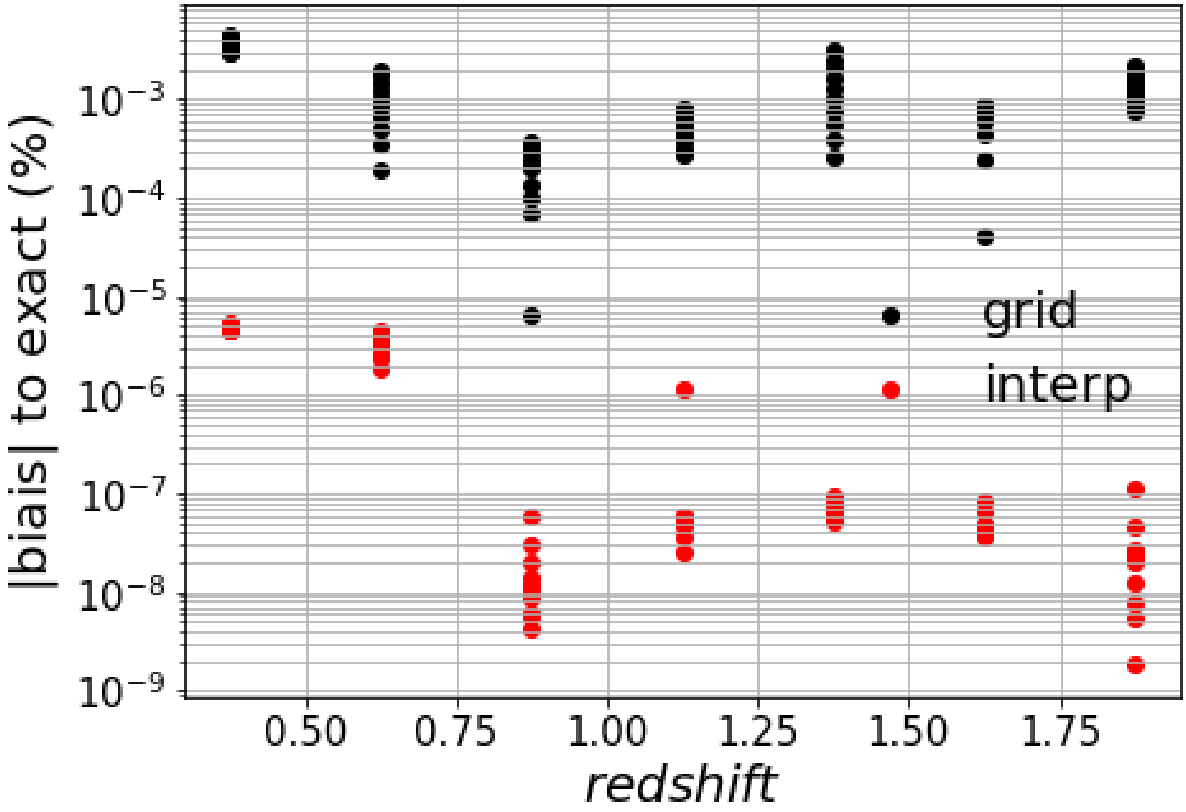

FIGURE 4.8: Relative difference between CCL grid method, CCL interpolation method and CCL exact method.

4.3 Halo Bias

The last part of the cross-check is the the halo bias comparison. We kept the limits from the mass and redshift from the first analyses as well as the bins construction. In this case we perform only the comparison using the true table for mass and Redshift since CCL does not provide the option to implement proxies for calculating the halo bias. The bias comparison results are presented in figures 4.9 and 4.10 for the unbinned and binned cases respectively.

Again NumCosmo were more efficient to calculate the bias taking 1.3*s* and 0.05*s* while CCL takes 9.6*s* and 21.96*s* for the unbinned and binned case respectively. This time however their agreement was about 10^{-1}% in both cases even with the high precision option. The difference is still small for LSST-DESC standards but this increase will be investigated in a future work to return to the 10^{-4}% case.

Then we implemented in NumCosmo the possibility of calculating the bias of a cluster given some mass or redshift proxy. Here we kept the same as the one used in the proxy case comparison but now we reduced the number of bins to four for the redshift and three for the richness and the results are presented in figures 4.11 and 4.12 for the unbinned and binned case respectively. In the unbinned case NumCosmo took 6.30*s* while the binned took 0.90*s* to be calculated. This implementation will be crucial when will introduce the super sample covariance in the cluster counts likelihood in a future work.

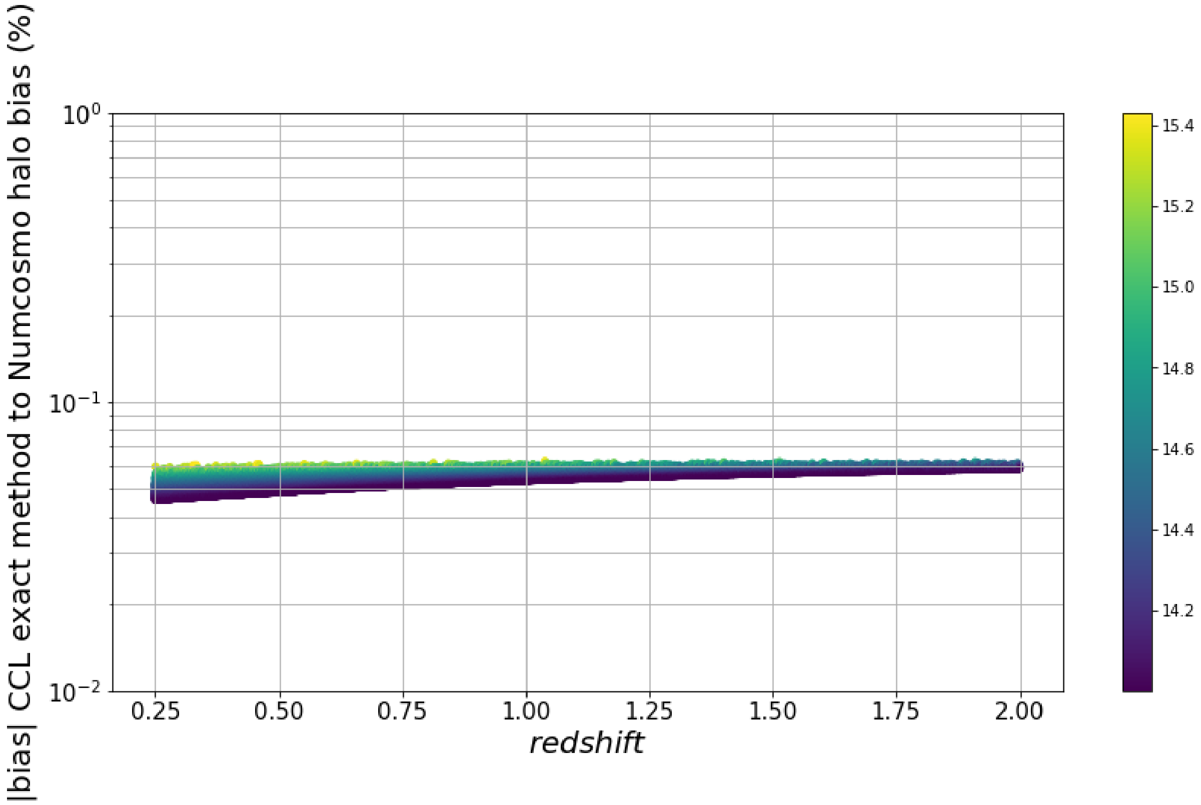

FIGURE 4.9: Relative difference between NumCosmo and CCL cluster bias distribution, the color map is for the different cluster mass.

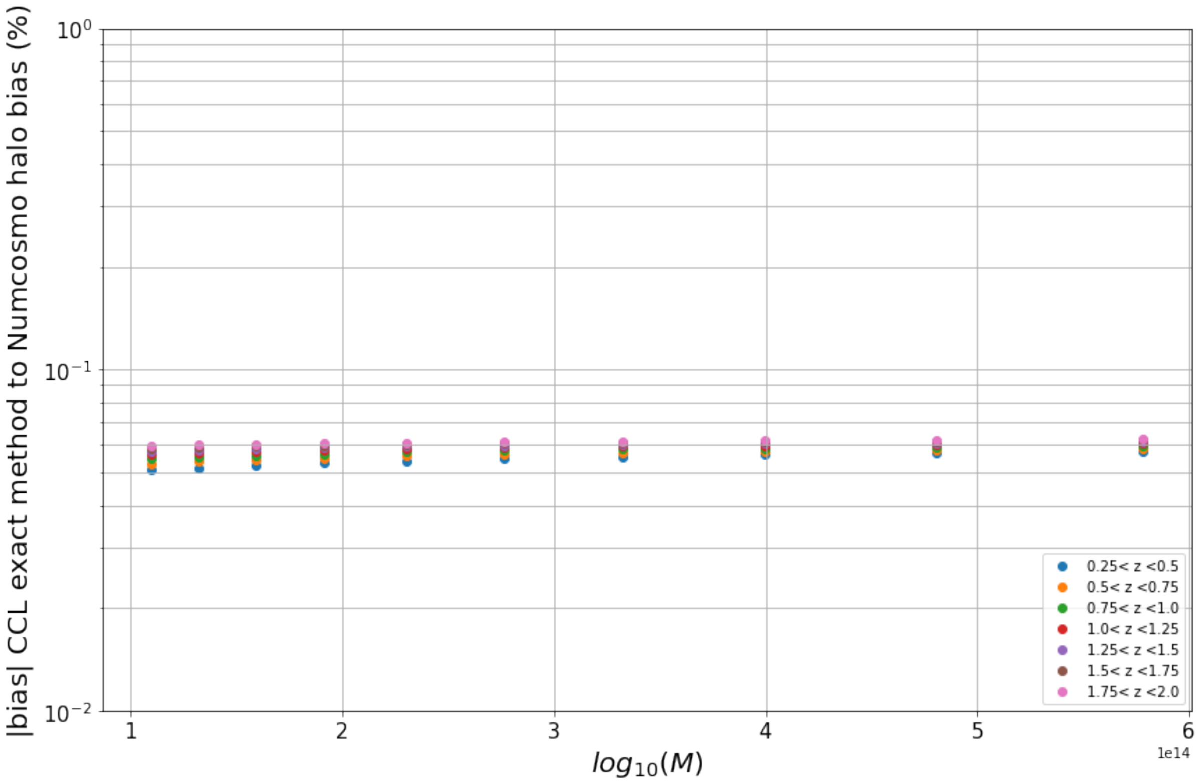

FIGURE 4.10: Relative difference between NumCosmo and CCL cluster bias prediction inside the bins, each color stands for a different redshift bin.

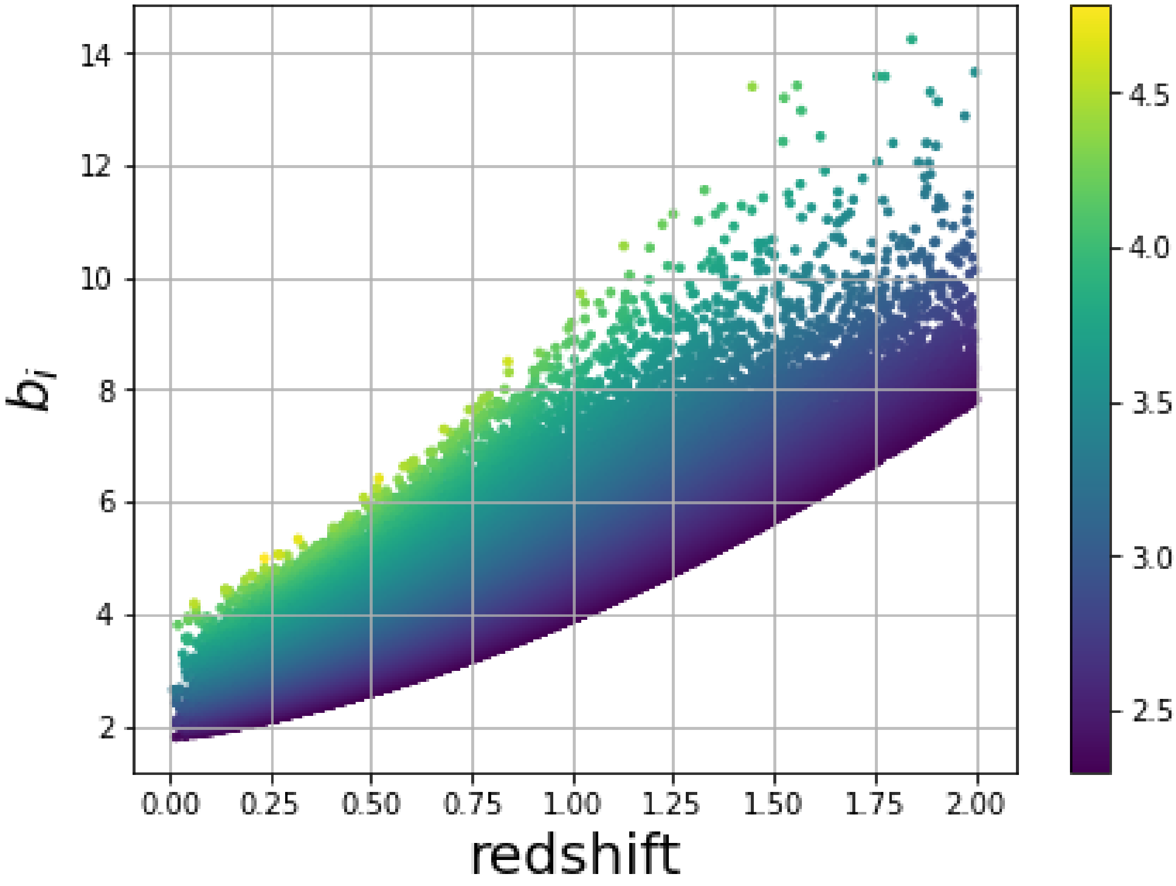

FIGURE 4.11: Bias distribution when we have richness-mass proxy, the color map is for the different cluster richness.

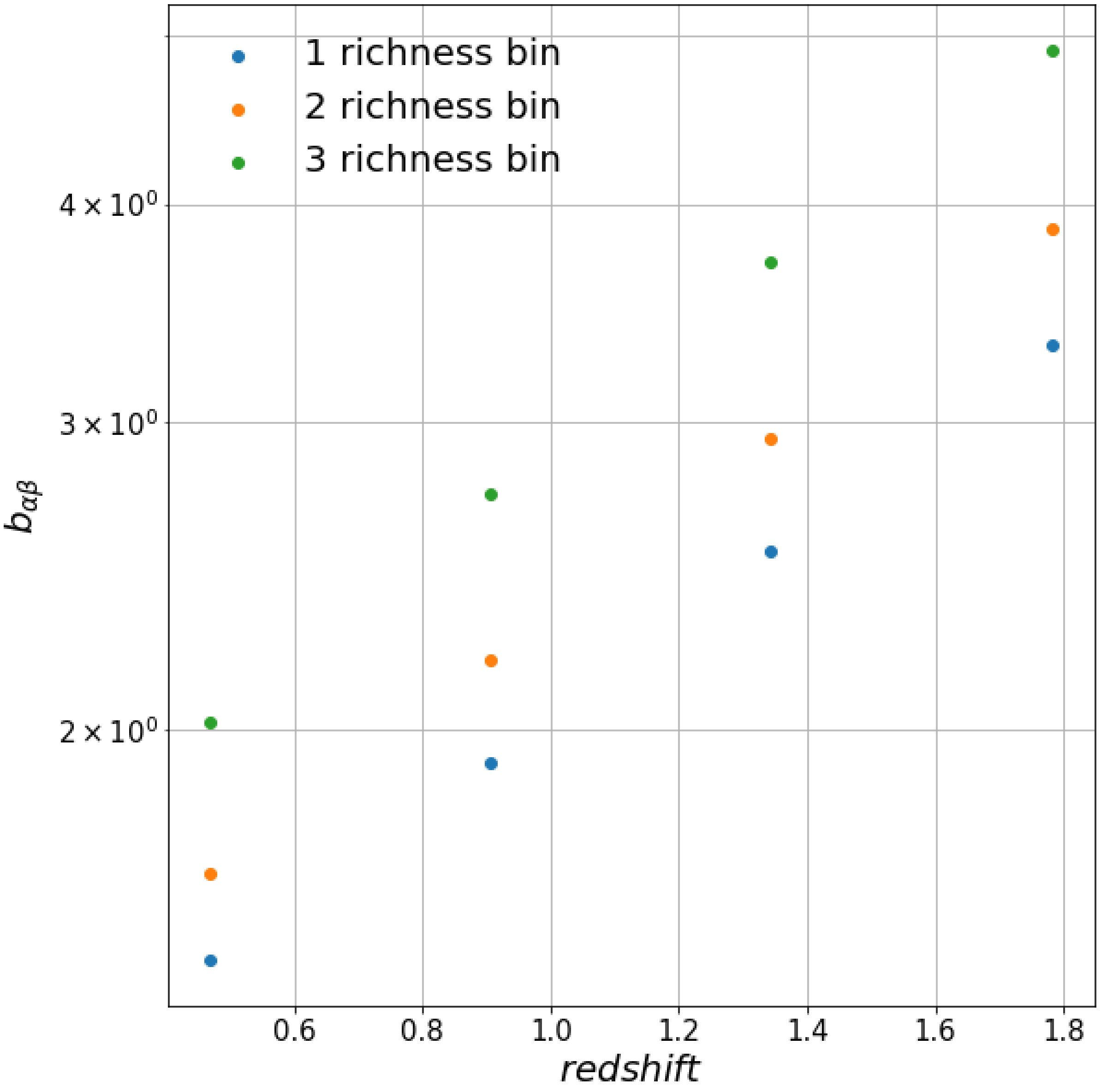

FIGURE 4.12: Bias values inside the bins of redshift and richness, each color stands for a different redshift bin.

Chapter 5

Cosmological Constraints

The results presented in this chapter come from the project 139 of LSST-DESC group. From the results of the last chapter we use the clusters distributions predicted by NumCosmo, which have the desired precision and efficiency, to test the likelihoods 3.32 and 3.35. These two functions will be used by the collaboration in the end of our project to constrain the data from the Vera Rubin Observatory. Therefore, we need to study all of its limitations and advantages to understand the best option in each situation.

The first part of this chapter consists of testing the likelihoods in the mock catalog. First we will show the complications in the analyses when we include uncertainties in the cluster mass with and without the self calibration. The next step is to compare the results from the binned and unbinned likelihoods and test the precision they retrieve and time it takes to constrain the catalog parameters. After performing these comparisons, we will apply our unbinned likelihood to the $DC2$ simulation and retrieve its cosmological parameters. All analyses in this chapter were done in the *Nersc* supercomputer using 150 threads. So keep in mind that these analyses could take much more time in a regular computer.

5.1 Binned and Unbinned comparison in the Mock catalog

To construct the mock catalogs we used the same values as the last chapter's parameters (see tables 4.1 and 4.2). This time we set the survey area to $439.79deg^2$ to match the $cosmoDC2\_v1.1.4\_redmapper\_v0.8.1$ catalog in $DC2$. In all catalogs we set the redshift z interval to $[0, 1]$, the mass M interval was $[10^{13}, 10^{15}]M_{\odot}$ and for the richness λ interval we used $[0, 100]$. One remark is that when we consider galaxy clusters with low mass clusters a realistic catalog will be incomplete and not pure and to avoid those issues we may model these two phenomena or choose a higher lower mass limit. Our mass definition still follows $200\rho_c$. In the binned analyses we binned our data into four equally spaced bins in z, while the richness and mass we binned into three equally log spaced bins. We choose to fit the σ_8 and Ω_m cosmological parameters which have a high correlation when determined by galaxy cluster number counts data. For the self-calibration we fit μ_0.

The best fit and error bars for our parameters in all cases are given in table 5.1, where the error bars are the 1σ confidence region, that is, we are 68% confident our parameters lies in

	Ω_m	σ_8	μ_0
True unbinned	$0.264^{+0.001}_{-0.001}$	$0.801^{+0.002}_{-0.002}$	No fit
True binned	$0.264^{+0.001}_{-0.001}$	$0.800^{+0.002}_{-0.003}$	No fit
Proxy unbinned	$0.266^{+0.001}_{-0.002}$	$0.797^{+0.004}_{-0.003}$	$3.191^{+0.005}_{-0.004}$
Proxy binned	$0.265^{+0.002}_{-0.002}$	$0.798^{+0.003}_{-0.004}$	$3.192^{+0.004}_{-0.005}$
Proxy unbinned noself	$0.265^{+0.002}_{-0.001}$	$0.798^{+0.003}_{-0.003}$	No fit

TABLE 5.1: Best fit and MCMC results.

that interval. First we want to compare the change in precision due to the presence of the mass proxy alone and with the self-calibration and the results are presented in figure 5.1. For the case with the richness proxy the volume of the confidence region increased 8.6% when compared to the true data case. This effect is only due to the presence of the proxy we also can see comparing the case with and without the self-calibration it's effect. By the corner plot we can see the larger size of the blue curve in comparison with the green one. We can't compare the volumes of the confidence regions between these two cases because the parameter space in the blue curve has a higher dimension, this leads to a smaller volume and NumCosmo still doesn't provide a method to calculate the sub volumes of the corner plot. This method will be implemented in the future.

The second aspect we want to compare on the MCMC results is the difference between the unbinned and binned likelihoods. In the case we have the true table, figure 5.2, for the mass and redshift the MCMC took $14min44s$ for the unbinned analyses while for the binned the time was $13min56s$. However the volume of the 1σ increased 15.64% when we binned the data meaning we lost this amount of precision. In this case the binning significantly worsen our result and doesn't provide much improvement in execution time. The last comparison for the likelihoods is when we need to perform the self-calibration, figure 5.3. This time the unbinned analyses took $1h43min43s$ to fit one additional parameter while the binned analyses took $42min58s$, the time difference already is significantly smaller. The 1σ confidence region this time increased 7.69%. We can see that the variance of the 1σ confidence region has decreased from the true case to the proxy case while the time difference of the executions has increased by 50 times. Hence when performing the full calibration it may be unavoidable to use the binned analyses since we will need to fit five more parameters and the time execution increase is much bigger than the variance increase.

The results of this section also highlight the importance of the cross-check done in the last chapter. In the first part we see the increase of our uncertainties when we implement the proxy hence we need to reduce all the numerical error to the smallest possible in order to retrieve the cosmological parameters in the required precision. In the second part we show the increase in the execution time to perform the self-calibration, hence is crucial our likelihood calculations to be done as fast as possible to avoid becoming a bottleneck for the MCMC analyses.

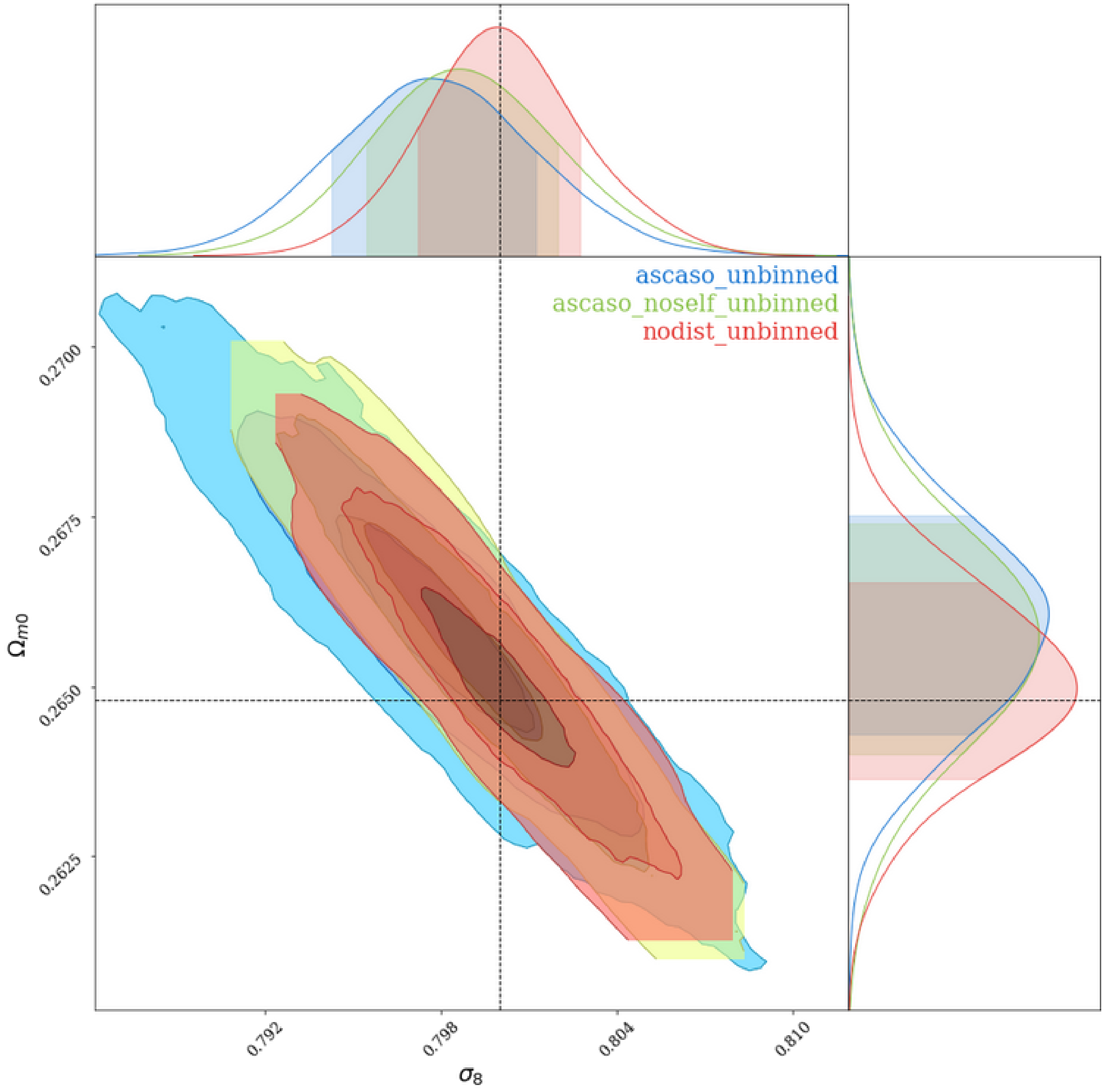

FIGURE 5.1: Corner plot from the NumCosmo mock catalog with an unbinned likelihood. The red curve is the case where we know the clusters true table, while the blue and green are the corner with the richness proxy(3.26), with and without self-calibration respectively.

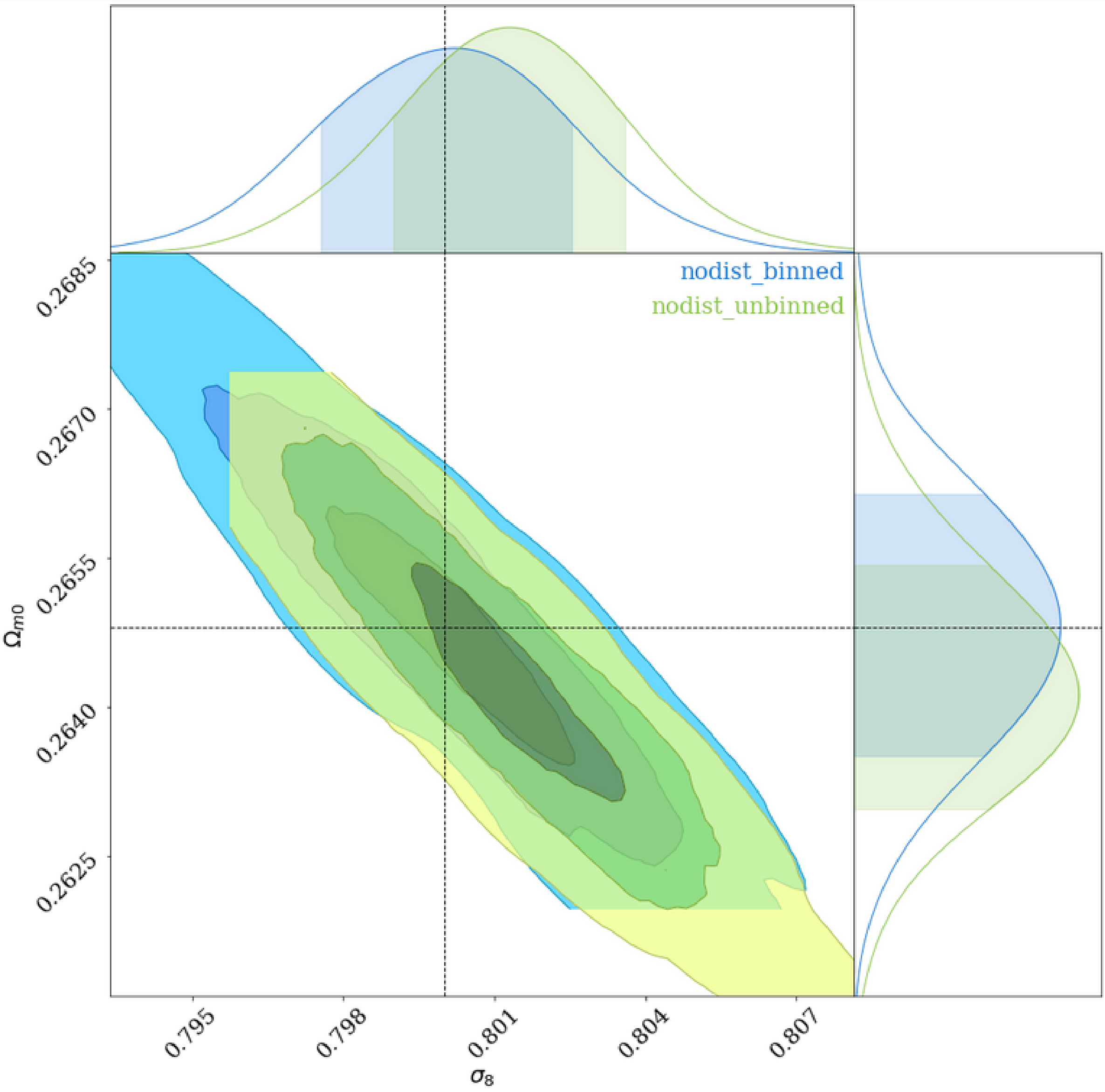

FIGURE 5.2: Corner plot from the NumCosmo mock catalog with the true table. The green curve is for the unbinned likelihood and the blue curve for the binned likelihood.

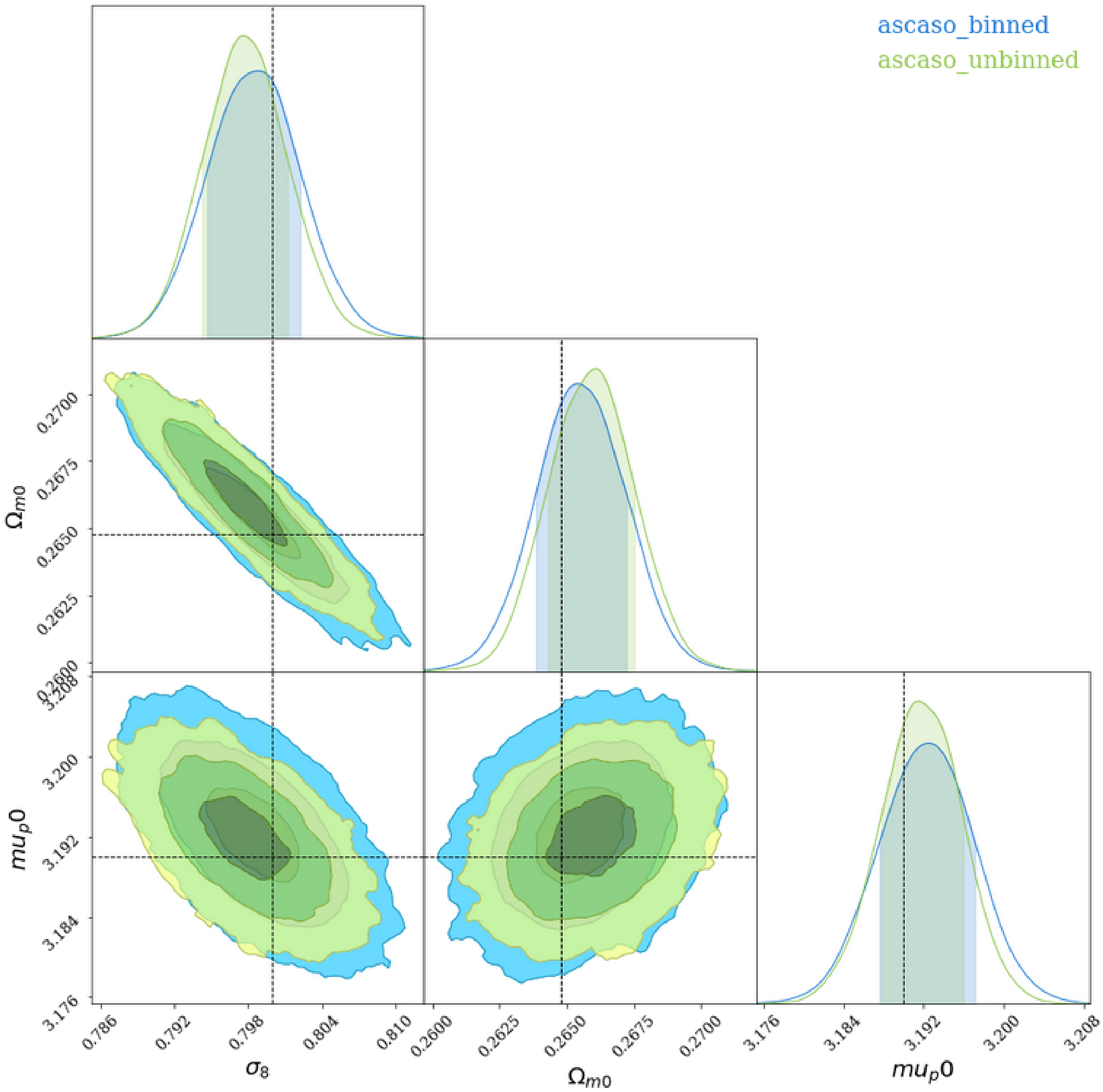

FIGURE 5.3: Corner plot from the NumCosmo mock catalog with the richness table. The green curve is for the unbinned likelihood and the blue curve for the binned likelihood.

5.2 True Table constraint from $DC2$

After testing the likelihoods we would like to retrieve the cosmological parameters used in the $DC2$ simulation, as a first test we used the $skysim5000\_v1.1.1$, which contains the true table for the simulated halos in a area of $5264.05 deg^2$. To assure the catalog completeness we set our analyses in the mass interval $M \in [10^{14}, 10^{15}]M_\odot$ and redshift range of $z \in [0, 1]$. As mentioned before the $DC2$ halo masses are defined using a FOF algorithm then they provide a catalog with the mass converted to a SOD definition using $200\rho_c$. We used the SOD mass definition so we could keep using the Tinker multiplicity function. We fit the σ_8 and Ω_m parameters.

The results of the first analysis is in figure 5.4. We couldn't recover the correct values for the parameters even in a 3σ region. To understand the reason this happened we first reconstruct the halo mass function for the catalog. In figure 5.5 we marginalized the mass function in redshift and fit it as function of the halo mass. The red curve is the case where we used the fitted cosmological parameters to construct the mass function while the black curve we used the catalog fiducial parameters. We can see that the mass part of the mass function is in agreement with the catalog given by the blue area, which is just the number of clusters inside the bins of mass. Then we repeated the analyses but now we reconstructed the mass function as a function of the redshift. From the results of figure 5.6 we spotted that for higher redshifts there is a disagreement between the fitting of the mass function and the data.

With that discrepancy in mind, we repeated the MCMC analyses in two different redshift sub intervals of the original analyses $z \in [0, 0.5]$ and $z \in [0.5, 1]$. The results are showed in figure 5.7. Indeed for the first sub interval, blue curve, we could get the correct cosmological parameters within the 2σ confidence region, which has improved the result but we still don't find it satisfactory. For the second sub interval, green curve, the correct parameters also fell out of the 3σ confidence region. This discrepancy can be due to the super sample variance, since the bias of the clusters grows with the redshift as can be seen in figure 4.11. To investigate this possibility we will implement a new likelihood in NumCosmo in a future work. Also it could be the choice of multiplicity function and we will implement in NumCosmo and test the analyses with the Despali multiplicity function [56] in the future.

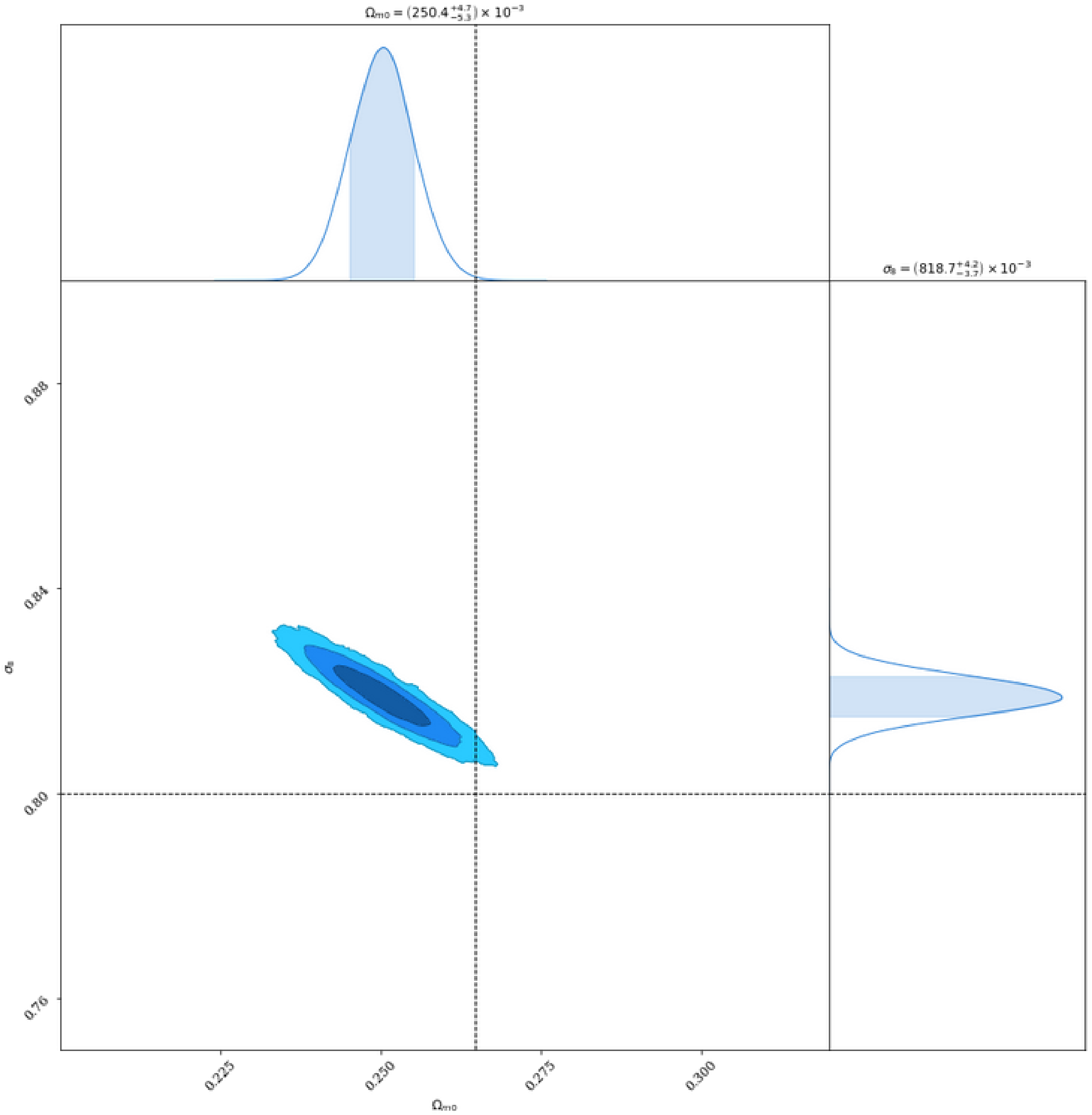

FIGURE 5.4: Corner plot from the *DC*2 catalog with an unbinned likelihood.

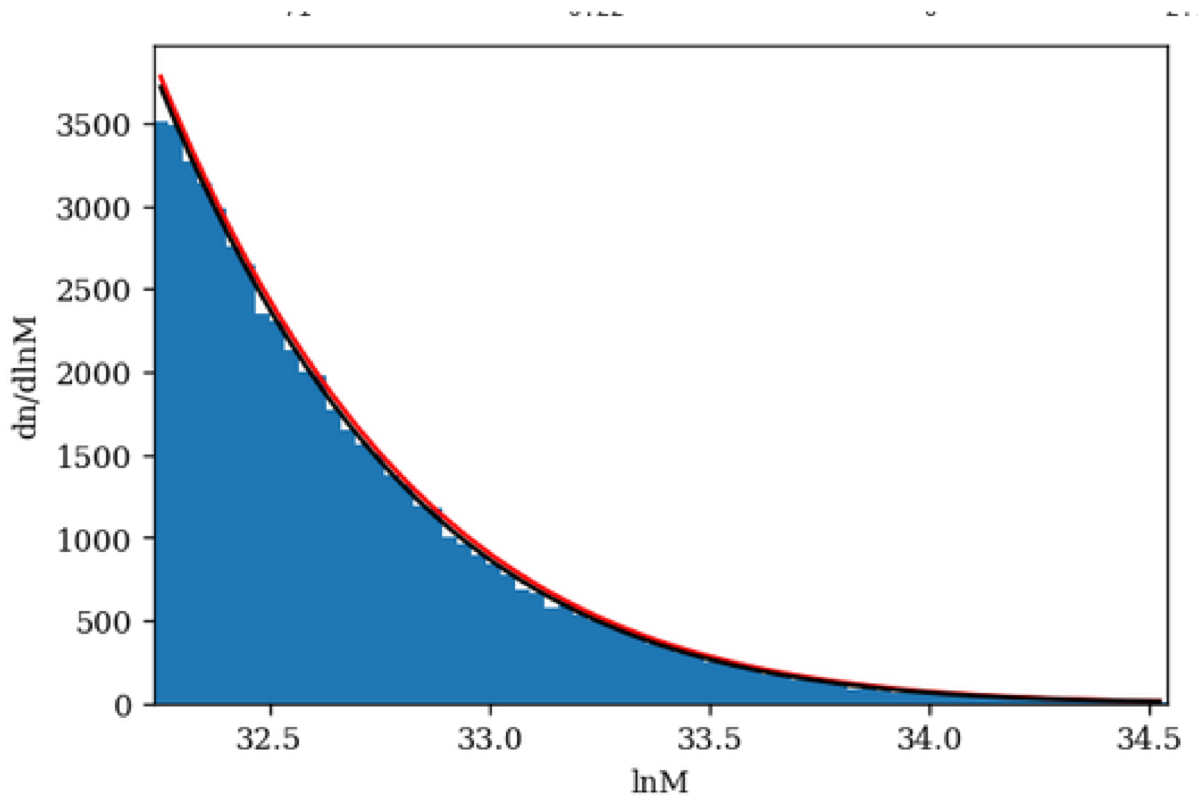

FIGURE 5.5: Mass function as a function of mass marginalized over the redshift.

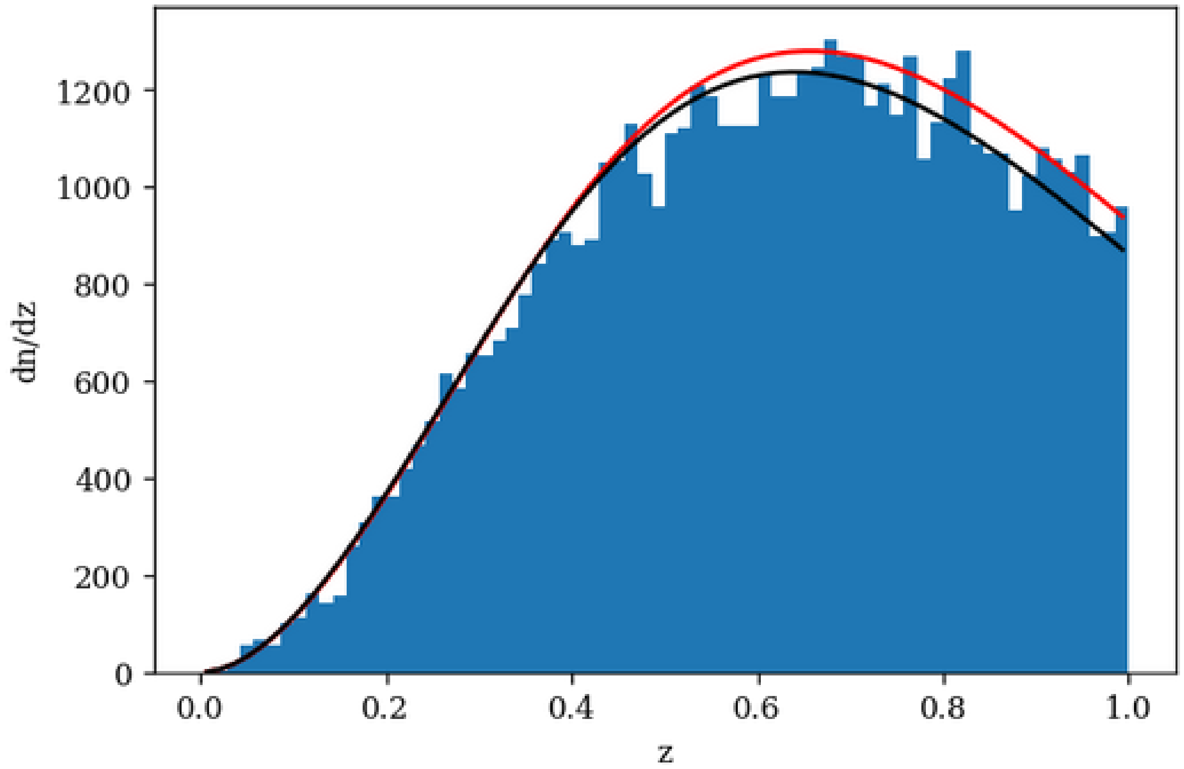

FIGURE 5.6: Mass function as a function of redshift marginalized over the mass.

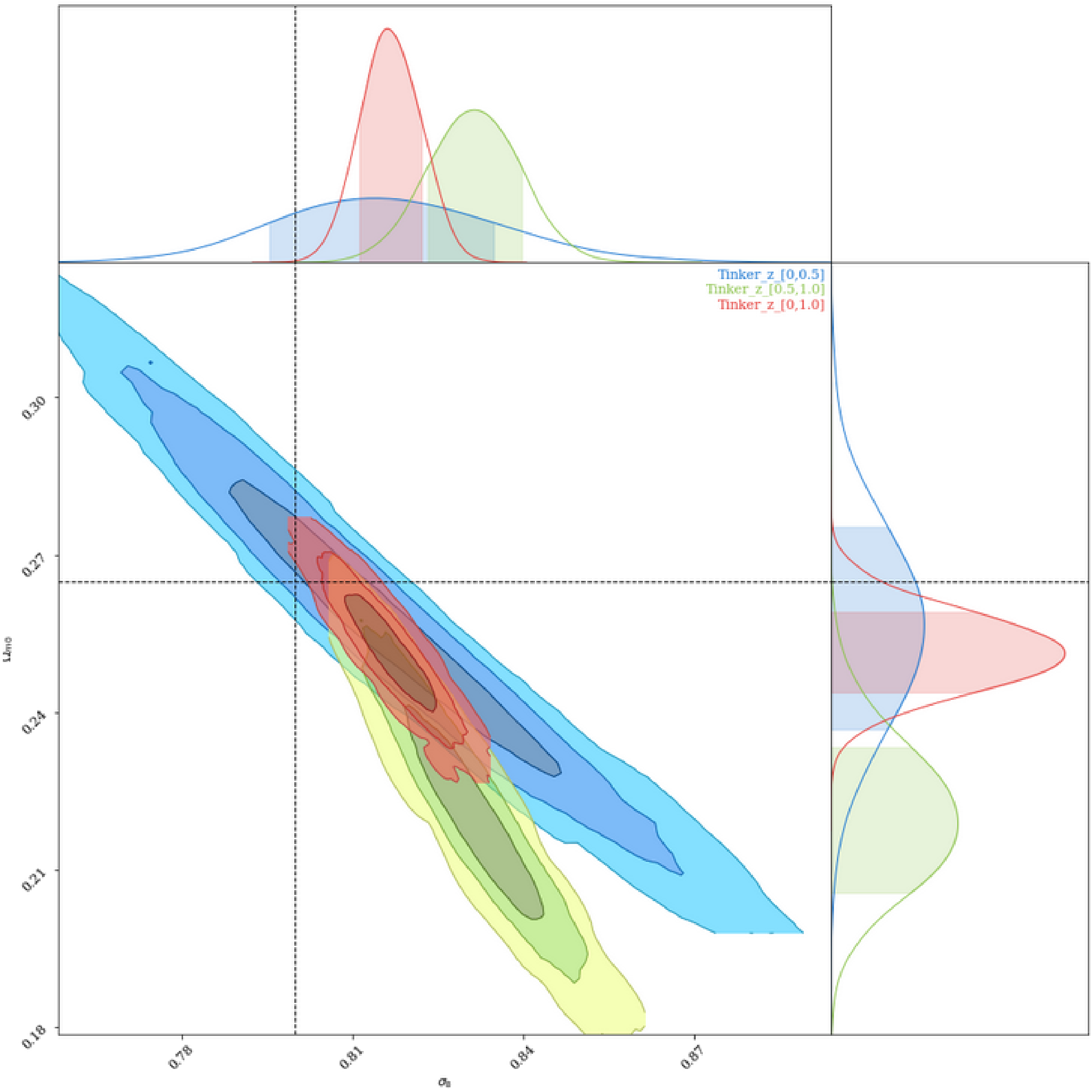

FIGURE 5.7: Corner plot from the NumCosmo mock catalog with an unbinned likelihood for different redshift intervals. The red curve is for $z \in [0, 1]$, the blue is for $z \in [0, 0.5]$ and green is for $z \in [0.5, 1]$.

Chapter 6

Final Remarks

In this work we first performed the test of two modularized cosmological software, NumCosmo and CCL. They were developed so we can test every model present in the Vera Rubin Observatory. From the background cosmological model, to the statistical relations between observable and the clusters properties or systematics in the measurement process. We could only test these libraries by comparing their results with each other since they were built independently. Our test were focused on their cluster counts prediction and the halo bias in unbinned and binned cases given by equations 3.29, 3.33, 3.31, 3.34. We initially found that the CCL cluster counts prediction did not agree within our required precision. This happened due to lack of precision in CCL splines to prepare the mass function. To solve this we implemented the *set_high_precision* function to match the precision used in NumCosmo. With that option both libraries were consistent. However we measured the efficiency for the calculations of these quantities and CCL, using the maximum precision option, were significantly slower than NumCosmo. However CCL has approximate calculations which reduced the time needed to calculate the cluster counts but it loses precision. When we compared the results of the mean halo bias the agreement has decreased to the 10^{-1}%. The reason to this is still under investigation but it is still a good agreement but significantly lower than what was obtained for the mass function. Finally, we presented NumCosmo new implementation to calculate the halo bias with proxies. In particular, we use the richness 3.26.

After the consistency test performed in these modules of NumCosmo and CCL, we turned to the test in the statistical tools developed in NumCosmo to analyze clusters catalogs. The first test was done in a Mock catalog to highlight the increase in the variance of the parameters fitted in different scenarios. First we compared the results of the MCMC using the unbinned likelihood 3.32 for the cases where we have the true table and the richness proxy with and without the self-calibration. Then with that difference known, we compared the constraints from the unbinned and binned likelihoods for the true table case and proxy case with self-calibration and quantified the increase of the posterior volume. Finally, we used these tools to run the MCMC on the $skysim5000\_v1.1.1$ catalog in the $DC2$ simulation. In the first analyses we could not find the correct parameters used in the simulation. But we found a difference between the fitted mass function and the data for higher redshifts. Repeating the analyses in a smaller redshift we retrieved the parameters in the 2σ region and the one made in higher redshift was worse than the original as expected. The possible causes to this disagreement in high redshift is still under investigation. We believe that some important part comes from the sample variance effect which is not taken into account, or the choice of the multiplicity function. Also could be the case of an error in the simulation itself.

Naturally there still much to be done in a future work to improve the tools to be used by LSST-DESC collaboration. The main topic to study in a future project is the effects of sample variance in the cluster counts likelihoods. With that new implementation, it will also be crucial to repeat the cross-check between CCL and NumCosmo to test their consistency at all levels to make them best tools possible to be used by the collaboration.

Bibliography

[1] Martin C Weisskopf et al. “Chandra X-ray Observatory (CXO): overview”. In: *X-Ray Optics, Instruments, and Missions III* 4012 (2000), pp. 2–16.

[2] Michael R. Blanton et al. “Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe”. In: *Astron. J.* 154.1, 28 (July 2017), p. 28. DOI: 10.3847/1538-3881/aa7567. arXiv: 1703.00052 [astro-ph.GA].

[3] Charles L Bennett et al. “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: final maps and results”. In: *The Astrophysical Journal Supplement Series* 208.2 (2013), p. 20.

[4] WB Everett et al. “Millimeter-wave Point Sources from the 2500 Square Degree SPT-SZ Survey: Catalog and Population Statistics”. In: *The Astrophysical Journal* 900.1 (2020), p. 55.

[5] Nabila Aghanim et al. “Planck 2018 results-I. Overview and the cosmological legacy of Planck”. In: *Astronomy & Astrophysics* 641 (2020), A1.

[6] Rene Laureijs et al. “Euclid definition study report”. In: *arXiv preprint arXiv:1110.3193* (2011).

[7] LSST Dark Energy Science Collaboration et al. “Large synoptic survey telescope: dark energy science collaboration”. In: *arXiv preprint arXiv:1211.0310* (2012).

[8] Sandro Dias Pinto Vitenti and Mariana Penna-Lima. “NumCosmo: Numerical Cosmology”. In: *Astrophysics Source Code Library* (2014), ascl–1408.

[9] Nora Elisa Chisari et al. “CCL: Core Cosmology Library”. In: *Astrophysics Source Code Library* (2019), ascl–1901.

[10] Željko Ivezić et al. “LSST: from science drivers to reference design and anticipated data products”. In: *The Astrophysical Journal* 873.2 (2019), p. 111.

[11] The LSST Dark Energy Science Collaboration. *LSST DESC Science Roadmap*. Version 2.6. Immediately after the executive summary, there is a brief "How to use this document" guide that concisely describes how to navigate it. Please start there to find your way through the document! Sept. 2021. DOI: 10.5281/zenodo.5527255. URL: https://doi.org/10.5281/zenodo.5527255.

[12] Bela Abolfathi et al. “DESC DC2 Data Release Note”. In: *arXiv preprint arXiv:2101.04855* (2021).

[13] Danila Korytov et al. “CosmoDC2: A synthetic sky catalog for dark energy science with LSST”. In: *The Astrophysical Journal Supplement Series* 245.2 (2019), p. 26.

[14] Andrew J Connolly et al. “An end-to-end simulation framework for the Large Synoptic Survey Telescope”. In: *Modeling, Systems Engineering, and Project Management for Astronomy VI*. Vol. 9150. SPIE. 2014, pp. 414–421.

[15] Antony Lewis and Anthony Challinor. "Camb: Code for anisotropies in the microwave background". In: *Astrophysics source code library* (2011), ascl–1102.

[16] Thomas Essinger-Hileman et al. "CLASS: the cosmology large angular scale surveyor". In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*. Vol. 9153. SPIE. 2014, pp. 491–513.

[17] CS Frenk et al. "The Santa Barbara cluster comparison project: a comparison of cosmological hydrodynamics solutions". In: *The Astrophysical Journal* 525.2 (1999), p. 554.

[18] August E Evrard et al. "Galaxy clusters in hubble volume simulations: cosmological constraints from sky survey populations". In: *The Astrophysical Journal* 573.1 (2002), p. 7.

[19] Volker Springel et al. "Simulations of the formation, evolution and clustering of galaxies and quasars". In: *nature* 435.7042 (2005), pp. 629–636.

[20] Katrin Heitmann et al. "Robustness of cosmological simulations. I. Large-scale structure". In: *The Astrophysical Journal Supplement Series* 160.1 (2005), p. 28.

[21] Stefano Borgani and Andrey Kravtsov. "Cosmological simulations of galaxy clusters". In: *Advanced Science Letters* 4.2 (2011), pp. 204–227.

[22] Stefano Borgani. "Cosmology with clusters of galaxies". In: *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*. Springer, 2008, pp. 287–334.

[23] Piero Rosati, Stefano Borgani, and Colin Norman. "The evolution of X-ray clusters of galaxies". In: *arXiv preprint astro-ph/0209035* (2002).

[24] Enrique Gaztanaga and J Alberto Lobo. "Nonlinear gravitational growth of large-scale structures inside and outside standard cosmology". In: *The Astrophysical Journal* 548.1 (2001), p. 47.

[25] William H Press and Paul Schechter. "Formation of galaxies and clusters of galaxies by self-similar gravitational condensation". In: *The Astrophysical Journal* 187 (1974), pp. 425–438.

[26] Ding-fang Zeng and Yi-hong Gao. *Spherical Collapse Model*. 2004. DOI: 10.48550/ARXIV.ASTRO-PH/0412628. URL: https://arxiv.org/abs/astro-ph/0412628.

[27] Ravi K Sheth, HJ Mo, and Giuseppe Tormen. "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes". In: *Monthly Notices of the Royal Astronomical Society* 323.1 (2001), pp. 1–12.

[28] Jeremy Tinker et al. "Toward a halo mass function for precision cosmology: the limits of universality". In: *The Astrophysical Journal* 688.2 (2008), p. 709.

[29] Scott Dodelson and Fabian Schmidt. *Modern cosmology*. Academic Press, 2020.

[30] Viatcheslav Mukhanov. *Physical foundations of cosmology*. Cambridge university press, 2005.

[31] Patrick Peter and Jean-Philippe Uzan. *Primordial cosmology*. Oxford University Press, 2009.

[32] U Seljak and M Zaldarriaga. "CMBFAST: A microwave anisotropy code". In: *Astrophysics Source Code Library* (1999), ascl–9909.

[33] Daniel J Eisenstein and Wayne Hu. "Baryonic features in the matter transfer function". In: *The Astrophysical Journal* 496.2 (1998), p. 605.

[34] Marc Davis et al. "The evolution of large-scale structure in a universe dominated by cold dark matter". In: *The Astrophysical Journal* 292 (1985), pp. 371–394.

[35] Jaan Einasto, Mihkel Jôeveer, and Enn Saar. “Structure of superclusters and supercluster formation”. In: *Monthly Notices of the Royal Astronomical Society* 193.2 (1980), pp. 353–375.

[36] Cedric Lacey and Shanu Cole. “Merger rates in hierarchical models of galaxy formation–II. Comparison with N-body simulations”. In: *Monthly Notices of the Royal Astronomical Society* 271.3 (1994), pp. 676–692.

[37] Darren Reed et al. “Evolution of the mass function of dark matter haloes”. In: *Monthly Notices of the Royal Astronomical Society* 346.2 (2003), pp. 565–572.

[38] Michael S Warren et al. “Precision determination of the mass function of dark matter halos”. In: *The Astrophysical Journal* 646.2 (2006), p. 881.

[39] Adrian Jenkins et al. “The mass function of dark matter haloes”. In: *Monthly Notices of the Royal Astronomical Society* 321.2 (2001), pp. 372–384.

[40] Nick Kaiser. *On the spatial correlations of Abell clusters*. Vol. 284. 1984.

[41] Marcos Lima and Wayne Hu. “Self-calibration of cluster dark energy studies: Observable-mass distribution”. In: *Physical review D* 72.4 (2005), p. 043006.

[42] Idit Zehavi et al. “The luminosity and color dependence of the galaxy correlation function”. In: *The Astrophysical Journal* 630.1 (2005), p. 1.

[43] HJ Mo and Simon DM White. “An analytic model for the spatial clustering of dark matter haloes”. In: *Monthly Notices of the Royal Astronomical Society* 282.2 (1996), pp. 347–361.

[44] Jeremy L Tinker et al. “The large-scale bias of dark matter halos: numerical calibration and model tests”. In: *The Astrophysical Journal* 724.2 (2010), p. 878.

[45] Alexey Vikhlinin et al. “Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation”. In: *The Astrophysical Journal* 640.2 (2006), p. 691.

[46] Andrey V Kravtsov, Alexey Vikhlinin, and Daisuke Nagai. “A new robust low-scatter X-ray mass indicator for clusters of galaxies”. In: *The Astrophysical Journal* 650.1 (2006), p. 128.

[47] John E Carlstrom, Gilbert P Holder, and Erik D Reese. “Cosmology with the Sunyaev-Zel’dovich effect”. In: *arXiv preprint astro-ph/0208192* (2002).

[48] Erin Scott Sheldon et al. “Weak-lensing measurements of 42 SDSS/RASS galaxy clusters”. In: *The Astrophysical Journal* 554.2 (2001), p. 881.

[49] Erin S Sheldon et al. “Cross-correlation weak lensing of SDSS galaxy clusters. I. Measurements”. In: *The Astrophysical Journal* 703.2 (2009), p. 2217.

[50] ES Rykoff et al. “redmapper. i. algorithm and sdss dr8 catalog”. In: *The Astrophysical Journal* 785.2 (2014), p. 104.

[51] Krishna B Athreya and Soumendra N Lahiri. *Measure theory and probability theory*. Vol. 19. Springer, 2006.

[52] Robert Lupton. *Statistics in theory and practice*. Princeton University Press, 1993.

[53] Mariana Penna-Lima, Martın Makler, and Carlos Alexandre Wuensche. “Biases on cosmological parameter estimators from galaxy cluster number counts”. In: *Journal of Cosmology and Astroparticle Physics* 2014.05 (2014), p. 039.

[54] Sandro D. P. Vitenti and Eduardo J. Barroso. “APES: Approximate Posterior Ensemble Sampler”. In: *in preparation* (2015).

[55] Mariana Penna-Lima et al. “Calibrating the Planck cluster mass scale with CLASH”. In: *Astronomy & Astrophysics* 604 (2017), A89.

[56] Giulia Despali et al. “The universality of the virial halo mass function and models for non-universality of other halo definitions”. In: *Monthly Notices of the Royal Astronomical Society* 456.3 (2016), pp. 2486–2504.