

# Euclid sensitivity forecasts for neutrino mass and modified gravity

Sefa Pamuk

The present work was submitted to the

Institute for Theoretical Particle Physics and Cosmology  
Rheinisch-Westfälische Technische Hochschule Aachen  
Germany  
29.09.2023

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Masterarbeit selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Aachen, den 29.09.2023

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Neutrinos and the Large Scale Structure</b>	<b>4</b>
2.1	Massive Neutrinos . . . . .	4
2.2	Massless Neutrinos . . . . .	7
<b>3</b>	<b>The Nonlinear Matter Power Spectrum</b>	<b>9</b>
3.1	The Halo Model . . . . .	10
3.2	The <code>HMCODE</code> Implementation . . . . .	15
<b>4</b>	<b>The Spectroscopic Probe</b>	<b>19</b>
4.1	Extracting the Power Spectrum from Observation . . . . .	20
4.2	Extracting the Power Spectrum from Theory . . . . .	24
4.3	The Effect of Massive Neutrinos . . . . .	26
<b>5</b>	<b>The Photometric Probe</b>	<b>29</b>
5.1	The Formalism of Cosmic Shear Observations . . . . .	30
5.2	Extracting the Covariance from Theory . . . . .	34
5.3	The Effect of Massive Neutrinos . . . . .	36
<b>6</b>	<b>Methodology</b>	<b>40</b>
6.1	The Markov Chain Monte Carlo Method . . . . .	41
6.2	The Fisher Information Method . . . . .	43
<b>7</b>	<b>Validation of the Forecasting Pipeline</b>	<b>48</b>
7.1	The Einstein-Bolzmann Solvers . . . . .	50
7.2	Comparison of the different Fisher Information Methods . . . . .	53
7.3	The validity of the Fisher approximation . . . . .	56
7.4	Bias from Modeling . . . . .	58
<b>8</b>	<b>Results</b>	<b>65</b>
<b>9</b>	<b>Conclusion</b>	<b>69</b>

# Chapter 1

## Introduction

The neutrino is well known for how difficult it is to measure. It hardly partakes in the fundamental forces that govern the physics of our universe. Without a charge or a colour charge, for a long time, neutrinos were believed to only partake in the weak force. Yet one of the biggest revelations of the last 40 years was the experimental proof that neutrinos have a mass. This was done by the observation of neutrino oscillation. This is because neutrinos from reactors seem to vanish after travelling long distances. It was later understood that they did not vanish but rather converted from one flavour to another. This had to mean that the neutrinos have a mass difference and thus also a non-zero mass. However, these oscillation measurements are only able to measure the mass difference between the different mass eigenstates of the neutrino. To measure the absolute mass scale this was not enough. This gives however a minimum mass that the sum of the neutrinos would need to have. We find

$$\sum m_\nu > 0.06\text{eV}$$

To get a handle on the neutrino mass we have to look at the largest particle accelerator that we know, the big bang. The Neutrinos are believed to have decoupled from the thermal plasma a second after the Big Bang when typical energies were of the order of MeV. But even after decoupling they have a distinct signature on the observables of cosmology. From Big Bang nucleosynthesis up to structure formation in recent times the neutrinos have left their fingerprints for us to search and find.

The tightest constraints on the Neutrino mass come from cosmic microwave background experiments combining the temperature and lensing anisotropies with the baryonic acoustic oscillations. The *Planck* collaboration finds an upper 95% confidence limit for the neutrino mass of

$$\sum m_\nu < 0.12\text{eV},$$

closing in with the oscillation experiments to find the absolute mass scale. Of course, the bounds from cosmology are model dependent but since there is no real evidence against  $\Lambda$ CDM the constraints from *Planck* are to be taken very seriously.

The new upcoming data from ESA's *Euclid* mission will help us to weigh in on the challenge of measuring the neutrino mass. It is planned to be the largest galaxy catalogue to date with a sky

convergence of 15000 square degrees and an estimated one billion galaxies used to build a tomographic three-dimensional map of the sky. To compare with the sky coverage of *DES* and *KiDS* is 5000 and 1000 square degree respectively. With all it's data, *Euclid* is forecast to measure the matter power spectrum up to 1% level and give valuable insight into the many open questions of cosmology.

In this work, we will not only try to gauge how well the *Euclid* satellite will be able to measure the mass of the neutrinos but also question how well *Euclid* be able to measure new massless particles that are not included in the standard model parametrized through the famous  $N_{\text{eff}}$ . On top of that, we will try to forecast constraints on generic dark energy parametrizations to put limits on accelerated expansion models. Finally, we will check how well *Euclid* will be able to measure a certain class of models of modified gravity theories.

In this work, we will use different statistical methods, build a validation for our forecasting pipeline, and explore the deep and rich physics in modelling the large-scale structure of the universe.

## Chapter 2

# Neutrinos and the Large Scale Structure

In this chapter, we will briefly discuss the effects that neutrinos have on the matter power spectrum. For a more in-depth discussion of all the effects that neutrinos have on our universe, we refer to the main source of this chapter [Lesgourgues et al., 2013]. In the first section, we will give a phenomenological overview of the effect of massive neutrinos on the large-scale structure(LSS). Then we will remark on the effect that additional massless neutrinos have on the matter power spectrum. These discussions will also be representative of other simple dark matter models. Simply put any relativistic, collisionless, and thermalized particle acts like a neutrino on the large-scale structure of the universe.

### 2.1 Massive Neutrinos

When discussing the effect of the massive neutrinos on the power spectrum, the main effect can be very roughly described with three phenomenological statements:

- Massive neutrinos stop clustering on scales larger than their free streaming scale which is related to their velocity.
- Massive neutrinos lead to an overall steplike suppression of the power spectrum by an amount proportional to their mass.
- Massive neutrinos change the redshift of matter-to-radiation equality essentially shifting the peak of the matter power spectrum.

The first effect leads to a suppression of the power spectrum for scales smaller than some scale  $k_*$ . This can be understood from the following reasoning. The free-streaming scale represents a scale under which collisionless particles can not be confined. If we define it analogously to the Jeans wavenumber we find a free-streaming wavenumber of

$$k_{\text{fs}} = \sqrt{\frac{3}{2}} \frac{a H}{c_\nu}. \quad (2.1)$$

For relativistic neutrinos, the thermal velocity  $c_\nu = 1$  while for nonrelativistic neutrinos we can find the thermal velocity by dividing the mean momentum by the neutrino mass. This leads to

$$c_\nu = \approx \frac{8.78 \cdot 10^{-3}}{a} \frac{0.06 \text{ eV}}{m_\nu}$$

$$k_{\text{fs}} \approx 4.66 \cdot 10^{-3} \frac{m_\nu}{0.06 \text{ eV}} a^2 \frac{H}{H_0} h \text{ Mpc}^{-1}. \quad (2.2)$$

During matter domination for the relativistic case, the free streaming wavenumber falls with  $\eta^{-1}$  and for the nonrelativistic case the wavenumber grows with  $\eta$ . This means that at the transition from relativistic to nonrelativistic we find the largest scale at which neutrinos stopped clustering. We denote this scale with  $k_{\min}$  (we will call it the minimum clustering scale), and it is precisely where the steplike suppression of the neutrinos starts. Given that the transition to the nonrelativistic regime happens during matter domination,  $k_{\min}$  can be approximated by

$$k_{\min} \approx 7.96 \cdot 10^{-3} \sqrt{\frac{\Omega_m}{0.3}} \sqrt{\frac{m_\nu}{0.06 \text{ eV}}} h \text{ Mpc}^{-1}. \quad (2.3)$$

To describe the second effect we first need to separate our matter perturbation into its three constituents. These are baryons, cold dark matter (CDM) and massive neutrinos. We can write

$$\delta_m = f_c \delta_c + f_b \delta_b + f_\nu \delta_\nu. \quad (2.4)$$

The factors  $f_X$  are the fractional contribution of the species  $X$  to the total energy density of the universe. The equations of motion drive the neutrino perturbations to an equilibrium similar to the cold constituents. However, due to the high velocity of the neutrinos, there is non-negligible pressure and anisotropic stress. This leads to

- (a) a small-scale suppression of the neutrino perturbation that is approximately given by

$$\delta_\nu \propto \left( \frac{k_{\text{fs}}}{k} \right)^2 \delta_c$$

on scales smaller than the free streaming scale.

- (b) a scale-dependent growth rate that is reduced compared to the cold matter constituents until the equilibrium has been reached.

For the following discussion, we will only look at what happens to the matter perturbations for scales much larger than the minimum clustering scale. Because of the two reasons stated above, on these scales, we use that at late times typically  $|\delta_\nu| \ll \delta_c$ . We can now expand equation 2.4 to find:

$$\langle \delta_m(k) \delta_m^*(k') \rangle \approx (f_c + f_b)^2 \left\langle \frac{(f_c \delta_c(k) + f_b \delta_b(k))(f_c \delta_c^*(k') + f_b \delta_b^*(k'))}{(f_c + f_b)^2} \right\rangle$$

$$:= (f_c + f_b)^2 \langle \delta_{cb}(k) \delta_{cb}^*(k') \rangle \quad (2.5)$$

$$\iff P_{mm}(k) \approx (1 - f_\nu)^2 P_{cb}(k). \quad (2.6)$$

Here, we have defined the perturbation of the CDM+baryon field  $\delta_{cb}$  and its respective power spectrum. After the baryon drag epoch, i.e. after baryons decouple from the radiation field, the equations of motion for this new density contrast simplify to

$$\delta''_{cb} + \frac{a'}{a} \delta'_{cb} - 4a^2 \pi G \langle \rho_c + \rho_b \rangle \delta_{cb} = 0, \quad (2.7)$$

where the prime denotes derivatives with respect to conformal time.

The third term is starting to look like the Friedmann–Lemaître equation minus the effect of massive neutrinos. If we add them we can write

$$\delta''_{cb} + \frac{a'}{a} \delta'_{cb} - \frac{3}{2} (1 - f_\nu) (a')^2 \delta_{cb} = 0. \quad (2.8)$$

If we assume that we are deep into matter domination such that the bulk of neutrinos are already non-relativistic. The neutrino density thus scales with  $a^{-3}$ . At this time the background evolution matches the neutrinoless case, and we can write that  $a \propto \eta^2$ . If we insert the conformal time into the equality we can solve the differential equation to find

$$\delta''_{cb} + \frac{2}{\eta} \delta'_{cb} - \frac{6}{\eta^2} (1 - f_\nu) \delta_{cb} = 0. \quad (2.9)$$

$$\implies \delta_{cb} \propto \eta^{\alpha \pm} \quad \text{with} \quad \alpha_\pm = -\frac{1}{2} \pm \frac{1}{2} \sqrt{25 - 24f_\nu} \quad (2.10)$$

This gives us one growing mode and a decaying mode. The result we find is different from the result of the neutrinoless universe where for the growing mode we find  $\delta \propto a$ . When we also use that neutrinos contribute very little to the total energy density, i.e.  $f_\nu \ll 1$ , we can find an approximate solution

$$\delta_{cb} = A_- a^{-\frac{3}{2} + \frac{3}{5} f_\nu} + A_+ a^{1 - \frac{3}{5} f_\nu}. \quad (2.11)$$

If we neglect the decaying mode, inserting the result for  $\delta_{cb}$  into the Poisson equation we also find

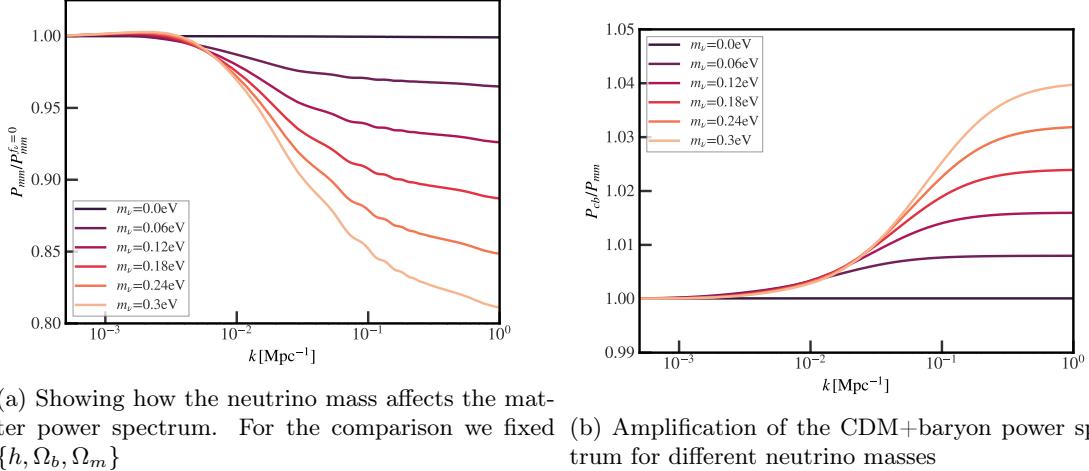
$$k^2 \Psi = -\frac{3}{2} (1 - f_\nu) (a')^2 \delta_{cb} \propto a^{-\frac{3}{5} f_\nu}. \quad (2.12)$$

Due to the massive neutrinos, not only does the density contrast of clustering matter grow slower, but also the metric perturbations decay slowly. From numerical calculations, we can find a rough estimate of the suppression on scales much smaller than the minimal clustering scale. For  $f_\nu \ll 1$  the overall suppression of the power spectrum can be estimated with

$$\frac{P_{cb}}{P_{cb}^{f_\nu=0}} \approx 1 - 6 f_\nu, \quad (2.13)$$

or for the total matter power spectrum  $P_{mm}/P_{mm}^{f_\nu=0} \approx 1 - 8 f_\nu$ . We illustrate the suppression of the power spectrum and the difference between the CDM+baryon spectrum and the total matter power spectrum in figure 2.1. We have fixed the total nonrelativistic matter density to match the low redshift expansion history between the different models, which means reducing the CDM density. These changes also lead to further suppression of the power spectrum on the smallest scales, essentially tilting the plateau further.

Figure 2.1: The effect of changing  $N_{\text{eff}}$  when fixing different quantities as explained in the Text. The Ratios were multiplied with a factor to better differentiate between them.



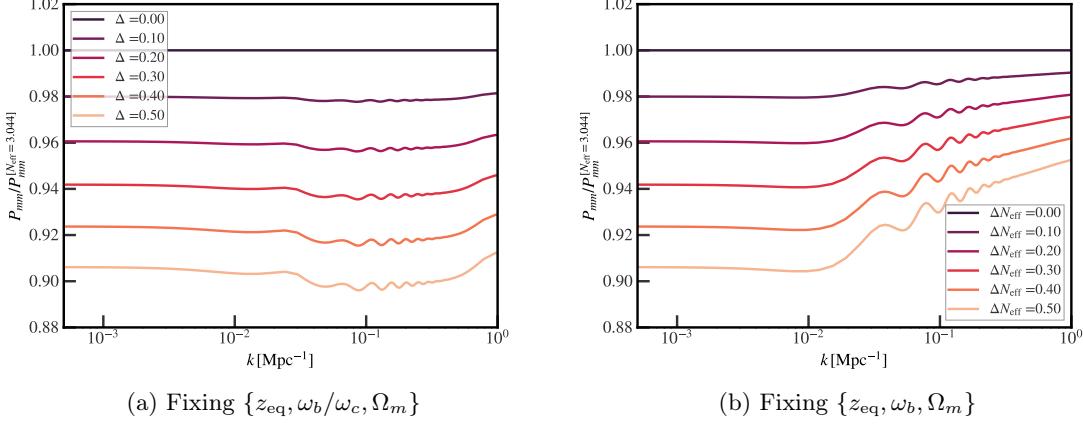
The last effect of massive neutrinos is the shifting of the redshift of equality. When fixing the density parameters of radiation, total matter, and baryons additional neutrinos shift the redshift of equality by a factor of  $(1 - f_\nu)$ . This is due to the neutrinos being ultra-relativistic at the time of equality and thus not contributing to the matter density. The shift of the redshift of equality leads to a shift in the peak of the matter power spectrum as it comes from the wavenumber that crossed the Hubble horizon precisely at equality. The slight wiggles in the ratio come from a change in the expansion history at earlier times shifting the scale of the BAO. We note that when we combine our *Euclid* probes with *Planck*, the CMB experiments strongly constrain the angular scale  $\theta_s$ , the redshift of equality  $z_{\text{eq}}$ , and the physical density parameter of baryons  $\omega_b$ . It can be shown that when fixing these quantities a change in the neutrino mass only leads to a more scale-independent suppression and a slight shift in the phase of the BAOs.

## 2.2 Massless Neutrinos

The effect of additional massless neutrinos on the matter power spectrum can be separated into background effects and perturbation effects. For the latter, we can already guess that these will be small. The main contribution of massless neutrinos to the matter power spectrum must happen during radiation domination as they become negligible during matter domination. During radiation domination on scales smaller than the Hubble horizon, it can be shown that the massless neutrinos have very little effect on the evolution of cold perturbations. This is because the massless neutrino perturbations crossing the Hubble horizon show an oscillatory behaviour. When decomposing the equations of motion for the perturbations into fast oscillating modes and slowly growing modes, we see that cold matter is primarily determined by slow modes. As a consequence, we can treat the massless neutrino perturbations as free streaming once matter domination starts.

The background effects of additional massive neutrinos have to be analysed a bit more carefully.

Figure 2.2: The effect of changing  $N_{\text{eff}}$  when fixing different quantities as explained in the Text. The Ratios were multiplied with a factor to better differentiate between them.



Depending on what quantities are fixed other quantities need to change as well. Firstly we start with the parametrization of additional massless relics, for this, we use the parameter  $N_{\text{eff}}$ . We write, that the total radiation density is

$$\rho_r = \rho_\gamma \left[ 1 + \frac{7}{8} \left( \frac{4}{11} \right)^{4/3} N_{\text{eff}} \right] \quad (2.14)$$

This means, that when fixing the density parameters of matter a change of  $N_{\text{eff}}$  coincides with a change of the redshift of equality, shifting the peak of the power spectrum. In order to fix the redshift of equality we need to scale  $\omega_m$  as  $\omega_\gamma$  is very tightly constrained by the *Firas* experiment. The overall amplitude of the matter power spectrum is related to the amplitude of scalar perturbations  $A_s$  as well as the total matter density parameter  $\Omega_m$ . The latter is proportional to the physical density of total matter by  $\omega_m = \Omega_m h^2$ . If we also want to fix the amplitude of the matter power spectrum without changing  $A_s$ , we have to scale  $h$  accordingly.

Next, if we want to subtract that effect, we need to rescale the matter density. Depending on if we switch  $\omega_b$  or  $\omega_c$  this changes the ratio of  $\omega_b/\omega_c$ . This ratio is crucial in the large-scale amplitude of the power spectrum as well as for the amplitude of the BAOs. To cancel this effect we need to equally scale the CDM and baryon density. What would be left in this set of fixed variables is a change in  $\omega_b$ , that changes the phase of the BAOs. We no longer have the freedom to subtract this effect as  $\omega_b$  had to be varied to keep the ratio fixed.

We could have also decided to fix  $\omega_b$  which would need us to change the ratio of cold dark matter to baryons and suppress the large-scale plateau. We illustrate the effects of changing  $N_{\text{eff}}$  for both cases in figure 2.2.

## Chapter 3

# The Nonlinear Matter Power Spectrum

With the Euclid mission, we will be able to deeply probe the nonlinear regime of the matter power spectrum. The nonlinear power spectrum can not be calculated using perturbation theory, so to predict the dynamics we would resort to N-body simulations. As running such simulations are computationally expensive, we have different codes to obtain the nonlinear power spectrum. In this work, we use two different codes called `HALOFIT` [Mead et al., 2021] and `HMCODE` [Takahashi et al., 2020]. Both codes agree with each other on the level of 6% but at the level of the *Euclid* mission's forecast precision, this is already too much. We will show later on how the choice of the nonlinear code will bias our parameter inference. Since `HMCODE` is better at matching N-body simulations of cosmologies with dark energy and massive Neutrinos than `HALOFIT`, we will use it for the forecast for the massive neutrinos.

We will however use `HALOFIT` for a forecast of modified gravity as the fitting formula of [?] was done with the power spectrum of `HALOFIT` as a pseudo. This means that in order to obtain the matter power spectrum in cosmologies with modified gravity, we do not need to solve the modified Einstein equations, but apply a multiplicative factor to the power spectrum to obtain the modified gravity nonlinear matter power spectrum. The difficulty remains in obtaining this so-called boost. For this, we need to employ N-body simulations where the modified gravity was already included. We will not go into further detail here, but for the thesis, we have added to our likelihood a setup to add any linear boost. This could not only be used for the modified gravity fitting formulas but also for example emulators for exotic particle dark matter models or baryonic feedback.

Both `HMCODE` and `HALOFIT` are based on the halo model to predict the nonlinear power spectrum. In the first section of this chapter we will go over the basics of the halo model.

The codes differ in that `HALOFIT` is a direct functional fit of the power spectrum from the halo model. `HMCODE` on the other hand is a semi-analytical model where the individual ingredients of the halo model are fitted. They are then combined in the context of a modified halo model to better match N-body simulations. In the second section of this chapter, we will then go over the implementation of `HMCODE` we used and explain where the critical modelling of massive neutrinos is important.

### 3.1 The Halo Model

The halo model describes how the nonlinear power spectrum is calculated using a sum of two terms

$$P_{mm}(k, z) = P^{1h}(k, z) + P^{2h}(k, z). \quad (3.1)$$

The first "one halo term",  $P^{1h}(k, z)$ , dominates at smaller scales and is calculated solely from the intrinsic properties of halos. The second term is called the "two-halo term",  $P^{2h}(k, z)$ . It dominates at larger scales and describes the spatial correlation between two halos. In the following, we will need to discuss how to extract a power spectrum from a density map.

We start by stating that all the derivation is done at one point in time, but it is the time dependence can be added afterwards. The main assumption of the halo model is that the total matter field  $\rho$  can be written as a sum over different halos, i.e. all the matter is inside halos. The second assumption is that the density profile of halos,  $\rho_H$ , is only a function of their mass,  $M_i$ , and the relative position to their centre,  $\mathbf{x}_i$ . These assumptions lead to

$$\rho(\mathbf{x}) = \sum_i \rho(\mathbf{x} - \mathbf{x}_i, M_i) = \int \sum_i \delta(M' - M_i) \delta^{(3)}(\mathbf{x}' - \mathbf{x}_i) \rho_H(\mathbf{x} - \mathbf{x}', M') d^3\mathbf{x}' dM' \quad (3.2)$$

In the last step, we have artificially added a sum over Dirac-delta distributions to pull out the dependence on the particular realization of the universe out of the halo density profile. The sum in the integral itself can be understood as a random realization of an underlying probability density,  $\langle dn/dM \rangle$ , i.e., the number density per mass interval  $[M, M + dM]$ . This can be understood in the following way. Imagine that the universe is separated into small volume bins,  $\delta V_i$ , and mass bins,  $\delta M_j$ , such that in each bin there is only the centre of one halo with its mass in that particular mass bin. This would define a random variable  $S^{ij}$  such that

$$S^{ij} = \begin{cases} 1 & \text{if a halo is in present the corresponding bins } ij \\ 0 & \text{otherwise} \end{cases}$$

The expectation value of this variable is given by the probability of finding a halo in this bin. One could either integrate the underlying probability or take the ensemble average of particular realizations. This leads to

$$\langle S^{ij} \rangle = \int_{\delta M_j} \int_{\delta V_i} \left\langle \frac{dn}{dM} \right\rangle (M) d^3\mathbf{x} dM \quad (3.3)$$

$$\stackrel{!}{=} \left\langle \int_{\delta M_j} \int_{\delta V_i} \sum_k \delta(M - M_k) \delta^{(3)}(\mathbf{x} - \mathbf{x}_k) d^3\mathbf{x} dM \right\rangle \quad (3.4)$$

$$\iff \left\langle \frac{dn}{dM} \right\rangle (M) = \left\langle \sum_i \delta(M - M_i) \delta^{(3)}(\mathbf{x} - \mathbf{x}_i) \right\rangle. \quad (3.5)$$

This can be used to find a normalization condition for the probability density by evaluating the expectation value of the density field

$$1 \stackrel{!}{=} \frac{1}{\langle \rho \rangle} \left\langle \int \sum_i \delta(M' - M_i) \delta^{(3)}(\mathbf{x}' - \mathbf{x}_i) \rho_H(\mathbf{x} - \mathbf{x}', M') d^3 \mathbf{x}' dM' \right\rangle \quad (3.6)$$

$$\begin{aligned} &= \frac{1}{\langle \rho \rangle} \int \left\langle \sum_i \delta(M' - M_i) \delta^{(3)}(\mathbf{x}' - \mathbf{x}_i) \right\rangle \rho_H(\mathbf{x} - \mathbf{x}', M') d^3 \mathbf{x}' dM' \\ &= \frac{1}{\langle \rho \rangle} \int \left\langle \frac{dn}{dM} \right\rangle (M') \rho_H(\mathbf{x} - \mathbf{x}', M') d^3 \mathbf{x}' dM' \\ &= \int \frac{M'}{\langle \rho \rangle} \left\langle \frac{dn}{dM} \right\rangle (M') dM'. \end{aligned} \quad (3.7)$$

In the first line, we have used that the halo profile of a single halo with mass  $M'$  and location  $\mathbf{x}'$  is independent of a particular realization of the universe and such can be factored out of the ensemble average. In the next step, we can use this to find an expression for the density contrast,

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \langle \rho \rangle}{\langle \rho \rangle} \quad (3.8)$$

$$= \frac{1}{\langle \rho \rangle} \int \left[ \sum_i \delta(M' - M_i) \delta^{(3)}(\mathbf{x}' - \mathbf{x}_i) - \left\langle \frac{dn}{dM} \right\rangle (M') \right] \rho_H(\mathbf{x} - \mathbf{x}', M') d^3 \mathbf{x}' dM' \quad (3.9)$$

$$:= \int \left\langle \frac{dn}{dM} \right\rangle (M') \delta_H(\mathbf{x}', M') \frac{\rho_H(\mathbf{x} - \mathbf{x}', M')}{\langle \rho \rangle} d^3 \mathbf{x}' dM'. \quad (3.10)$$

We have defined a new halo distribution contrast,  $\delta_H$ , that compares the realization of halo masses and centres to their underlying uniform distribution. The next step is to calculate the two-point correlation function,  $\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle \delta(\mathbf{x}_1) \delta(\mathbf{x}_2) \rangle$ . For this we will need to evaluate the same term, but with the halo distribution contrast. We write it as

$$\begin{aligned} \langle \delta_H(\mathbf{x}_1, M_1) \delta_H(\mathbf{x}_2, M_2) \rangle &= \left[ \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \right]^{-1} \left\langle \left( \sum_i \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_i) - \left\langle \frac{dn}{dM} \right\rangle_1 \right) \right. \\ &\quad \times \left. \left( \sum_j \delta^{(4)}(\tilde{\mathbf{x}}_2 - \tilde{\mathbf{x}}_j) - \left\langle \frac{dn}{dM} \right\rangle_2 \right) \right\rangle \end{aligned} \quad (3.11)$$

$$\begin{aligned} &= \left[ \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \right]^{-1} \left[ \left\langle \sum_i \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_i) \sum_j \delta^{(4)}(\tilde{\mathbf{x}}_2 - \tilde{\mathbf{x}}_j) \right\rangle \right. \\ &\quad \left. + \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \right]. \end{aligned} \quad (3.12)$$

We have used the shorthand notations  $\delta^{(4)}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_i) = \delta^{(3)}(\mathbf{x} - \mathbf{x}_i) \delta(M - M_i)$  for the four dimensional Dirac delta and  $\langle dn/dM \rangle_i = \langle dn/dM \rangle(M_i)$  for the halo mass function.

In the next step we will separate the integrals that show up in the calculation of the two-point correlation function into bins like before, this allows us to find the random variables  $S^{ij}$  again. For

notation's sake, we will slice up the integration space into four-dimensional volumes with one index. We will then find integrals like in the following

$$\mathcal{I}_g = \iint g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) \left\langle \sum_i \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_i) \sum_j \delta^{(4)}(\tilde{\mathbf{x}}_2 - \tilde{\mathbf{x}}_j) \right\rangle d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \quad (3.13)$$

$$= \sum_{\mu} \sum_{\nu} g(\tilde{\mathbf{x}}_{\mu}, \tilde{\mathbf{x}}_{\nu}) \langle S^{\mu} S^{\nu} \rangle. \quad (3.14)$$

$g$  is an arbitrary test function. In the easy case the indices  $\mu$  and  $\nu$  are the same, and we find

$$\langle S^{\mu} S^{\mu} \rangle = \langle S^{\mu} \rangle = \int_{\delta V_{\mu}^{(4)}} \left\langle \frac{dn}{dM} \right\rangle_1 d^4\tilde{\mathbf{x}}_1. \quad (3.15)$$

For the term with two different indices, we need to use the two-point correlation function of halo seed densities of equation 3.11. This leads to

$$\langle S^{\mu} S^{\nu} \rangle = \int_{\delta V_{\mu}^{(4)}} \int_{\delta V_{\nu}^{(4)}} (\delta_H(\tilde{\mathbf{x}}_1) \delta_H(\tilde{\mathbf{x}}_2) - 1) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \quad (3.16)$$

$$:= \int_{\delta V_{\mu}^{(4)}} \int_{\delta V_{\nu}^{(4)}} (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2. \quad (3.17)$$

With this, we can find the value of  $\mathcal{I}_g$  and relate it to the original integral by joining the bins.

$$\mathcal{I}_g = \sum_{\mu \nu} g(\tilde{\mathbf{x}}_{\mu}, \tilde{\mathbf{x}}_{\nu}) \langle S^{\mu} S^{\nu} \rangle \quad (3.18)$$

$$= \sum_{\mu \neq \nu} \int_{\delta_{\mu}^{(4)}} \int_{\delta_{\nu}^{(4)}} g(\tilde{\mathbf{x}}_{\mu}, \tilde{\mathbf{x}}_{\nu}) (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \\ + \sum_{\mu} \int_{\delta_{\mu}^{(4)}} g(\tilde{\mathbf{x}}_{\mu}, \tilde{\mathbf{x}}_{\mu}) \left\langle \frac{dn}{dM} \right\rangle_1 d^4\tilde{\mathbf{x}}_1 \quad (3.19)$$

$$= \sum_{\mu \neq \nu} \int_{\delta_{\mu}^{(4)}} \int_{\delta_{\nu}^{(4)}} g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \\ + \sum_{\mu} \int_{\delta_{\mu}^{(4)}} g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_1) \left\langle \frac{dn}{dM} \right\rangle_1 d^4\tilde{\mathbf{x}}_1 \quad (3.20)$$

$$= \iint g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \\ + \iint g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) \left\langle \frac{dn}{dM} \right\rangle_1 \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2) d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \quad (3.21)$$

$$= \iint g(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) \left\langle \frac{dn}{dM} \right\rangle_1 \left[ \left\langle \frac{dn}{dM} \right\rangle_2 (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) + \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2) \right] d^4\tilde{\mathbf{x}}_1 d^4\tilde{\mathbf{x}}_2 \quad (3.22)$$

In the fourth step we use that the function  $g$  is varying slowly in one bin and thus can be approximated as the value at the bin center. We then combine the microcells in equation 3.20 artificially add an integration over  $\tilde{\mathbf{x}}_2$  to find the final result. When comparing equations 3.13 and 3.22 we find an expression for the halo seed density two-point correlation function:

$$\left\langle \sum_i \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_i) \sum_j \delta^{(4)}(\tilde{\mathbf{x}}_2 - \tilde{\mathbf{x}}_j) \right\rangle = \left\langle \frac{dn}{dM} \right\rangle_1 \left[ \left\langle \frac{dn}{dM} \right\rangle_2 (\xi_H(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) - 1) + \delta^{(4)}(\tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_2) \right] \quad (3.23)$$

$$\langle \delta_H(\mathbf{x}_1, M_1) \delta_H(\mathbf{x}_2, M_2) \rangle = \xi_H(\mathbf{x}_1, \mathbf{x}_2, M_1, M_2) + \left[ \left\langle \frac{dn}{dM} \right\rangle(M_1) \right]^{-1} \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2) \delta(M_1 - M_2). \quad (3.24)$$

With this, we can find the two-point correlation function.

$$\begin{aligned} \xi(\mathbf{x}_1, \mathbf{x}_2) &= \iint \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \langle \delta_H(\tilde{\mathbf{x}}'_1) \delta_H(\tilde{\mathbf{x}}'_2) \rangle \frac{\rho_H(\mathbf{x}_1 - \mathbf{x}'_1, M'_1)}{\langle \rho \rangle} \frac{\rho_H(\mathbf{x}_2 - \mathbf{x}'_2, M'_2)}{\langle \rho \rangle} d^4 \tilde{\mathbf{x}}'_1 d^4 \tilde{\mathbf{x}}'_2 \\ &= \iint \xi_H(\mathbf{x}'_1, \mathbf{x}'_2, M'_1, M'_2) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \frac{\rho_H(\mathbf{x}_1 - \mathbf{x}'_1, M'_1)}{\langle \rho \rangle} \frac{\rho_H(\mathbf{x}_2 - \mathbf{x}'_2, M'_2)}{\langle \rho \rangle} d^4 \tilde{\mathbf{x}}'_1 d^4 \tilde{\mathbf{x}}'_2 \\ &\quad + \int \left\langle \frac{dn}{dM} \right\rangle_1 \frac{\rho_H(\mathbf{x}_1 - \mathbf{x}'_1, M'_1)}{\langle \rho \rangle} \frac{\rho_H(\mathbf{x}_2 - \mathbf{x}'_1, M'_1)}{\langle \rho \rangle} d^4 \tilde{\mathbf{x}}'_1 \end{aligned} \quad (3.25)$$

In the next step, we note that the halo density contrast,  $\delta_H$ , is a biased tracer of the matter density contrast  $\delta_m$ . We assume that the bias is a function of the mass of the halo but not its location due to the isotropy of space. We also assume linear biasing.

$$\xi_H(\mathbf{x}_1, \mathbf{x}_2, M_1, M_2) = b(M_1) b(M_2) \xi_m(\mathbf{x}_1, \mathbf{x}_2)$$

Our initial assumptions that all the matter should be in halos argues that the bias function should be close to unity and thus the integral

$$\int b(M) \left\langle \frac{dn}{dM} \right\rangle \frac{\rho_H(\mathbf{x}, M)}{\langle \rho \rangle} d^3 \mathbf{x} dM \approx 1. \quad (3.26)$$

To find the power spectrum we start with the Fourier transformation of the density contrast.

$$\delta(\mathbf{k}) = \frac{1}{(2\pi)^3} \int \left\langle \frac{dn}{dM} \right\rangle(M') \delta_H(\mathbf{x}', M') \frac{\rho_H(\mathbf{x} - \mathbf{x}', M')}{\langle \rho \rangle} \exp[-i\mathbf{k} \cdot \mathbf{x}] d^3 \mathbf{x} dM' d^3 \mathbf{x} \quad (3.27)$$

$$= \int \left\langle \frac{dn}{dM} \right\rangle(M') \delta_H(\mathbf{x}', M') \frac{\rho_H(\mathbf{k}, M')}{\langle \rho \rangle} \exp[-i\mathbf{k} \cdot \mathbf{x}'] d^3 \mathbf{x}' dM' \quad (3.28)$$

In the next step, we need to calculate the expectation value of the density contrast. We will then

insert the two-point correlation of halo density contrast from equation 3.24 to find the following.

$$\begin{aligned} \langle \delta(\mathbf{k}_1) \delta^*(\mathbf{k}_2) \rangle &= \int \xi(\mathbf{x}_1 - \mathbf{x}_2) b(M_1) b(M_2) \left\langle \frac{dn}{dM} \right\rangle_1 \left\langle \frac{dn}{dM} \right\rangle_2 \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} e^{-i\mathbf{k}_1 \cdot \mathbf{x}_1} \\ &\quad \times \frac{\rho_H(\mathbf{k}_2, M_2)}{\langle \rho \rangle} e^{i\mathbf{k}_2 \cdot \mathbf{x}_2} d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 dM_1 dM_2 \\ &\quad + \int \left\langle \frac{dn}{dM} \right\rangle_1 \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} \frac{\rho_H(\mathbf{k}_2, M_1)}{\langle \rho \rangle} e^{-i\mathbf{x}_1 \cdot [\mathbf{k}_1 - \mathbf{k}_2]} d^3 \mathbf{x}_1 dM_1 \end{aligned} \quad (3.29)$$

$$\begin{aligned} &= \int P_{mm}(\mathbf{k}') b(M_1) \left\langle \frac{dn}{dM} \right\rangle_1 \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} e^{-i\mathbf{x}_1 \cdot [\mathbf{k}_1 - \mathbf{k}']} \\ &\quad \times b(M_2) \left\langle \frac{dn}{dM} \right\rangle_2 \frac{\rho_H(\mathbf{k}_2, M_2)}{\langle \rho \rangle} e^{i\mathbf{x}_2 \cdot [\mathbf{k}_2 - \mathbf{k}']} d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 dM_1 dM_2 d^3 \mathbf{k}' \\ &\quad + (2\pi)^3 \int \left\langle \frac{dn}{dM} \right\rangle_1 \left[ \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} \right]^2 \delta(\mathbf{k}_1 - \mathbf{k}_2) dM_1 \end{aligned} \quad (3.30)$$

$$\begin{aligned} &= (2\pi)^6 \int P_{mm}(\mathbf{k}_1) b(M_1) \left\langle \frac{dn}{dM} \right\rangle_1 \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} \delta(\mathbf{k}_1 - \mathbf{k}_2) \\ &\quad \times b(M_2) \left\langle \frac{dn}{dM} \right\rangle_2 \frac{\rho_H(\mathbf{k}_2, M_2)}{\langle \rho \rangle} dM_1 dM_2 \\ &\quad + (2\pi)^3 \int \left\langle \frac{dn}{dM} \right\rangle_1 \left[ \frac{\rho_H(\mathbf{k}_1, M_1)}{\langle \rho \rangle} \right]^2 \delta(\mathbf{k}_1 - \mathbf{k}_2) dM_1 \end{aligned} \quad (3.31)$$

In the first step, we have inserted that the two-point correlation function of matter is the reverse Fourier transformed matter power density. Then we have factored out plane waves such that after integration over real space we are left with Dirac delta distributions, we then use them to simplify the integrals in the last step. After we factor out the remaining Dirac deltas we find the expression of the power spectrum is the sum of the one-halo and two-halo terms. The first term is the two-halo term and is given by

$$P^{2h}(k) := P_{mm}(k) \left[ (2\pi)^3 \int b(M) \left\langle \frac{dn}{dM} \right\rangle(M) \frac{\rho(\mathbf{k}, M)}{\langle \rho \rangle} dM \right]^2. \quad (3.32)$$

This is the correlation of two separate halos. The integral itself can be understood as a mean over the galaxy shapes with a small bias stemming from the fact that very massive halos cluster much more strongly than normal matter. It is also proportional to the matter power spectrum because these halos are formed from seeds of initial matter perturbations. We infer from this that their spatial correlation is given by the matter power spectrum.

The second term is the one-halo term.

$$P^{1h}(k) := (2\pi)^3 \int \left\langle \frac{dn}{dM} \right\rangle(M) \left[ \frac{\rho(\mathbf{k}, M)}{\langle \rho \rangle} \right]^2 dM \quad (3.33)$$

In the context of galaxy clustering the interpretation of this term would be the shot noise. In this context, it is a bit more subtle as it now correlates the matter in one halo with itself. This gives it the functional dependence of  $\rho^2$ . The integral itself can again be understood as a mean over the halo distribution such that this term becomes the mean self-correlation.

### 3.2 The HMCODE Implementation

Both terms of the halo model need firstly the Fourier transformation of the halo density profile

$$\rho_H(k, z, M) = \int e^{ik \cdot r} \rho_H(r, z, M) d^3r = \int_0^{r_{\text{lim}}} 4\pi r^2 \frac{\sin(kr)}{kr} \rho(r, z, M) dr. \quad (3.34)$$

The cutoff radius,  $r_{\text{lim}}$ , is defined via the mass of the halo. It is calculated by requiring that the mass enclosed by a sphere of that radius is given by

$$M = \frac{4\pi}{3} r_{\text{lim}}^3 \Delta_H(z) \bar{\rho}, \quad (3.35)$$

where  $\Delta_H$  is a density contrast, describing how much higher the halo density is compared to the mean matter density of the universe  $\langle \rho \rangle$ . It was calculated in Einstein-de Sitter (EdS) cosmologies that the Halo after virialization has an overdensity of  $\approx 178$ . To account for cosmologies with  $\Omega_m \neq 1$  and massive neutrinos, we use a fitting formula from Mead [Mead, 2016]. It is fitted to use only parameters that govern the formation of structure like the (integrated) growthrate and the denisity parameter of matter. To handle massive neutrinos, an additional correction factor as a function of  $f_\nu$  is added to the fitting formula. Furthermore, to keep the fitting parameters for cosmologies with massive neutrinos and dark energy the same, the growthrates and mass fraction that enter the fitting formula are calculated in an equivalent cosmology where the massive neutrinos and the dark energy are converted to CDM and  $\Lambda$  respectively.

The halo density profile  $\rho$  is taken as a modified FNW [Navarro et al., 1997] profile

$$\rho(r, z, M) \propto \frac{1}{\frac{r}{r_s \nu^\eta} \left( 1 + \frac{r}{r_s \nu^\eta} \right)}, \quad (3.36)$$

where an additional parameter  $\nu^\eta$  was added. It parametrizing any additional bloating of the halo shape. The proportionality factor in front is used to fix the mass of the halo. The parameter  $\nu$  is the peak-height variable defined as

$$\nu = \frac{\delta_c(z)}{\sigma_M^{cb}(z)}. \quad (3.37)$$

The mass  $M$  defines a radius  $R_\sigma$  over which we calculate the variance of the smoothed field. The relation is setting the scale to the radius of a sphere of background matter density with mass  $M$ . This leads to a definition similar to  $r_{\text{lim}}$  defined in 3.35.

$$M = \frac{4\pi}{3} R_\sigma^3 \bar{\rho}$$

The variable  $\delta_c$  is the critical overdensity that matter distributions must reach to collapse into a halo. It was computed in EDS to be close to 1.686. To more accurately model spherical collapse  $\delta_c$ , is taken as a fitting function from [Nakamura and Suto, 1997] and modified to account for the presence of massive neutrinos through a correction factor. Similar to  $\Delta_H$ , the fitting formula is a function only of the (integrated) growthrate and the total matter density calculated in an equivalent cosmology in the presence of massive neutrinos and dark energy.

The parameter  $\nu$  essentially quantifies how likely it is for given matter overdensities to collapse. That is why we use the variance of the CDM+baryon field to calculate it, as we know that neutrinos

are not slow enough to cluster on scales of halos. It is often encountered in the context of the peak-background split formalism and spherical collapse models like the ones from Press-Schächter [Press and Schechter, 1974]. The parameter  $\eta$  is a free parameter in the HMCODE model and is fitted to be

$$\eta(z) = \mathcal{A}_\eta \times [\sigma_8^{cb}(z)]^{\alpha_\eta} \quad (3.38)$$

The functional form is chosen like this because it should be a function of a parameter that describes the collapse of the halo and not cosmological parameters.

Finally, the halo shape radius,  $r_s$ , parametrizes the radial profile of the halo density and is related to the boundary of the halo,  $r_{\text{lim}}$ , via the halo concentration,  $c$ ,

$$r_{\text{lim}} = c r_s.$$

This parameter affects the innermost structure of the halo thus affecting the matter power spectrum on the smallest scales. We take its functional form to be

$$c(M, z) = \mathcal{B} \left[ \frac{1 + z_f(M, z)}{1 + z} \right] \frac{D(z_c)}{D^{\text{eqq}}(z_c)} \frac{D(z)}{D^{\text{eqq}}(z)}. \quad (3.39)$$

The factor  $\mathcal{B}$  is a free constant that was fitted to the N-body simulations. The term  $z_f$ , that appears in the equation 3.39, is the redshift at which the halo is formed. We consider a halo of mass  $M$  to be formed when its innermost part crosses the critical density. If we then assume normal growth we find the implicit equation

$$\frac{D(z_f)}{D(z)} \sigma_{\gamma M}^{cb}(z) = \delta_c(z), \quad (3.40)$$

with the free parameter  $\gamma$  chosen to be 1%, defining what innermost part actually means. The formula is equivalent to asking when a smaller halo of mass  $\gamma M$  would have formed. The redshift of formation  $z_f$  often ends up being very high, indicating that the properties of halos are set very early in their evolution. Still in some cases  $z_f$  could end up being earlier than the redshift  $z$ . In these cases, we set  $z_f = z$ . The next two factors account for the effect of dark energy further modifying the dynamics of collapse. The first is an empirical correction factor that was obtained from fits of the NFW profile to N-body simulations by [Dolag et al., 2004]. The redshift of collapse  $z_c$  would be calculated similarly as  $z_f$ , but we can use that the ratio becomes constant at high redshifts and just use a high redshift of  $z_c = 10$ . For high redshifts  $z$ , this factor would need to vanish, as then the effect of dark energy is negligible. The second fraction was added for this purpose in HMCODE. The next ingredient to calculate the one-halo term and the two-halo term is the halo mass function,  $\frac{dn}{dM}$ . As we know from equation 3.7 it has to be normalized. That is why it is more convenient to use the normalized halo mass function  $F(\nu, z)$ . We take it as

$$F(\nu, z) d\nu = \frac{M}{\bar{\rho}} \frac{dn}{dM} dM = A \left[ 1 + \frac{1}{(q \nu^2)^p} \right] e^{-q \nu^2 / 2} d\nu. \quad (3.41)$$

It is a modified version of the Press Schechter function with free parameters  $q$  and  $p$ . The parameter  $A$  is calculated from the normalization. We use the standard values for these.

$$p = 0.3 \quad q = 0.707 \quad A = 0.21616$$

The final ingredient that we would need to calculate the two-halo term would be the halo mass bias. To calculate it, one would need to follow the peak-background split formalism, where we assume that all matter perturbations, that cross the critical overdensity, collapse into halos. The bias itself quantifies how much stronger more massive halos cluster and can thus be calculated from the derivative of the halo mass function with respect to the peak-height variable. In reality, we do not need the bias to calculate the two-halo term. As stated, the bias is a function close to unity and thus the integral that goes into the two-halo term is approximately 1, as seen in equation 3.26. To better match simulations, the final formula for the two-halo term is modified in two regards. Firstly the power spectrum in the two-halo term is replaced by a power spectrum where the wiggles of the BAO have been smoothed out to an extent. This is a well-understood effect of gravitational evolution. The exact formulation of the "de-wiggling" is a topic of current debate, but a simple approach is followed in `HMCODE`.

We start with the linear power spectrum,  $P_{mm}^{\text{lin}}$ , obtained by our Boltzmann solver. We then divide it with the Eisenstein-Hu no-wiggle power spectrum approximation [?]. We then smooth the ratio by coevolving it with a broad Gaussian filter. After multiplying with the Eisenstein-Hu approximation we find the smoothed power spectrum,  $P_{\text{smt}}$ . The "de-wiggled" power spectrum is a weighted sum of the no-wiggle and the linear power spectrum.

$$P_{dw} = e^{-g(k,z)} P_{mm}^{\text{lin}}(k,z) + (1 - e^{-g(k,z)}) P_{\text{smt}}(k,z) \quad (3.42)$$

$$g(k,z) = k^2 \frac{1}{6\pi^2} \int_0^\infty P_{mm}^{\text{lin}}(k,z) dk := k^2 \sigma_v^2(z) \quad (3.43)$$

The second modification of the two-halo spectrum is that by default it does not contain an early dip of the power spectrum that occurs in the one-loop correction in the effective field theory approach of the LSS. Its physical interpretation is that the growth of voids, in reality, is slower than predicted in linear perturbation theory and leads to an overestimate of the power spectrum on very large scales. This effect is modelled by a multiplication of the de-wiggled power spectrum with a function fitted to dampen the power spectrum by a factor  $(1 - f_D)$  past a scale  $k_D$ .

The final two-halo power spectrum in `HMCODE` is given by

$$P^{2h}(k,z) = P_{dw}(k,z) \left[ 1 - f_D(z) \frac{k^{n_D}(z)}{k_D^{n_D}(z) + k^{n_D}} \right]. \quad (3.44)$$

The functions  $k_D(z)$  and  $f_D(z)$  have the same functional form as the shape parameter,  $\eta(z)$ , that appears in the FNW profile. We again use the variance of the CDM+baryon field in the functional form to fit as massive neutrinos should not affect the universal collapse dynamics. The parameter  $n_D$  is also a free parameter and is fitted to N-body simulations.

The one halo term also has two modifications done to it compared to the one obtained from the halo model. Firstly in equation 3.8 we assume that halos make up the matter density contrast of the universe. Since only cold matter clusters in halos, we need to correct this and multiply the halo density profile with a factor  $(1 - f_\nu)$ .

The second modification we need to do is a suppression of the one halo term on large scales. It turns out that the one halo term in the standard halo model becomes constant on the largest scales, but due to energy conservation, a virialized halo should lead to a power spectrum growing  $\propto k^4$ . To account for this, we multiply the one halo power spectrum with an empirical correction. Our final formula is given by

$$P^{1h}(k,z) = (1 - f_\nu)^2 (2\pi)^3 \int_0^\infty \left\langle \frac{dn}{dM} \right\rangle(M) \left[ \frac{\rho(k,M)}{\langle \rho \rangle} \right]^2 dM \frac{k^4}{k_*(z)^4 + k^4}. \quad (3.45)$$

The transition wave number,  $k_*$ , is fitted to a functional form like the transition wave number of the two-halo term  $k_D$ .

A final problem of the halo model is, that it is not well suited to describe the correlation in the transition region from the one-halo term to the two-halo term. This is due to neglecting that very close halos or overlapping subhalos correlate with each other. Neglecting these effects leads to an up to 20% underestimate of the power spectrum. To better model the transition we have added a smoothing parameter  $\alpha(z)$ . As this parameter affects the spectral index of the power spectrum, its functional form is not a function of the variance of the CDM+baryon field but of the effective spectral index of its power spectrum,  $n_{\text{eff}}$ . To calculate it we can use the definition of the variance in real space to find

$$3 + n_{\text{eff}}(z) = -\frac{d \log \sigma_{cb}(R, z)}{d \log R}(R_c, z). \quad (3.46)$$

Here,  $R_c$  is defined as the radius where the standard derivation of the CDM+baryon field crosses the critical overdensity  $\delta_c(z)$ . The final fitting formula for  $\alpha$  is then

$$\alpha(z) = \mathcal{A}_\alpha [\mathcal{C}_\alpha]^{n_{\text{eff}}(z)}. \quad (3.47)$$

With the transition smoothing we define our nonlinear power spectrum as

$$P_{mm}^{\text{nl}}(k, z) = \left[ (P^{1h}(k, z))^{\alpha(z)} + (P^{2h}(k, z))^{\alpha(z)} \right]^{1/\alpha(z)}. \quad (3.48)$$

## Chapter 4

# The Spectroscopic Probe

is one of the main two instruments of the *Euclid* mission. It is supposed to measure the redshift of galaxies to high precision and obtain from that a three-dimensional map of galaxies. One could use this map of galaxies to define a density field

$$\rho_g(\mathbf{r}) := \sum_i \delta^{(3)}(\mathbf{r} - \mathbf{r}_i), \quad (4.1)$$

where we use a Dirac-delta centred around the measured position of the galaxies  $\mathbf{r}_i$ . This is achieved by doing a long-time exposure of the galaxies and using some characteristic emission lines to measure the redshift of the galaxy. By assuming some arbitrary reference cosmology one would then translate the redshift and the angles into a position. The measurement of the redshift is much more precise than with the method of the photometric probe, but it is limited by a smaller sample size and a smaller survey volume. For the *Euclid* mission, it is planned to further bin the galaxies into four separate redshift bins with redshifts between 0.9 and 1.8. Each bin will have its separate survey volume  $V_i$  that can be approximated by multiplying the comoving volume by the planned sky fraction  $f_{\text{sky}} = 0.35$ .

The final likelihood is modelled as a Gaussian likelihood where the observed power spectrum  $P^{\text{obs}}$  is compared with a theoretically predicted one  $P^{\text{th}}$ :

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} \sum_i \frac{1}{2V_{\mathbf{k},i}} \int_{\Delta V_{\mathbf{k},i}} \frac{(P^{\text{th}} - P^{\text{obs}})^2}{(\sigma_{P^{\text{obs}}})^2} d^3\mathbf{k} \right], \quad (4.2)$$

where the summation is done over each redshift bin. The factor  $1/2$  in front of the integral stems from the fact, that we want to integrate over the entire three-dimensional  $\mathbf{k}$ -space, but we should only integrate over independent modes. Since the power spectrum is a real quantity it must have  $P(\mathbf{k}) = P(-\mathbf{k})$ , thus the factor is needed to account for this.  $V_{\mathbf{k},i}$  is the volume of  $\mathbf{k}$ -space that each mode has and is given through

$$V_{\mathbf{k},i} := \frac{(2\pi)^3}{V_i}.$$

This chapter will be structured in the following way. First, we will briefly explain how the observed power spectrum is extracted from the galaxy catalogue. Then we will talk about the modeling of the predicted power spectrum. We will have a separate section to explain the effects

that massive neutrinos will have on the predicted power spectrum, and explain our changes to the method of calculating the predicted power spectrum described in [Sprenger et al., 2019] and [Euclid Collaboration: Blanchard et al., 2023]. We do this to have a robust forecast of the neutrino mass. The survey specifications needed for the calculation of the likelihood are found at the end of the chapter.

## 4.1 Extracting the Power Spectrum from Observation

The derivation of how to extract the matter power spectrum from observation is very close to the derivation of the halo model discussed in the section 3.1. We will thus only give the main steps and refer the reader to that section if one wants step-by-step instruction on the results. In the derivation, one needs to assume that all galaxies have the same mass and have a delta-peak density profile. The simplification is later dropped when discussing the finite resolution of the instrument. One could then extract from this density field the underlying probability density field  $p_g(\mathbf{r})$ . This would define a local deviation  $\delta_g(\mathbf{r})$  from a homogenous distribution by factorizing out the mean number density  $\tilde{n}$  and some local detection probability  $p_{\text{det}}(\mathbf{r})$

$$p_g = \tilde{n} p_{\text{det}}(\mathbf{r}) (1 + \delta_g(\mathbf{r})). \quad (4.3)$$

This quantity is the equivalent of the distribution density contrast from section 3.1. For the *Euclid* mission, we assume that the detection probability is just a constant. For simplicity, we will assume in the following derivation a constant probability  $p_{\text{det}}(\mathbf{r}) = 1$ , but it should be noted, that as long as it is sufficiently constant, every step could be done with a non-constant probability.

The two-point correlation function  $\xi_g(\Delta\mathbf{r})$  of this quantity of the deviation is the Fourier transform of the power spectrum  $P^{\text{obs}}(\mathbf{k})$ :

$$\xi_g(\Delta\mathbf{r}) = \langle \delta_g(\mathbf{x}) \delta_g(\mathbf{x} + \Delta\mathbf{r}) \rangle_{\mathbf{x}} \quad (4.4)$$

$$P^{\text{obs}}(\mathbf{k}) = \int_V \xi_g(\Delta\mathbf{r}) e^{-i\mathbf{k}\cdot\Delta\mathbf{r}} d^3\Delta\mathbf{r}. \quad (4.5)$$

The quantity that was called the one-halo term of section 3.1 will be added later. In this case, it will be just a constant Poisson noise. The measurement of the galaxies does not give us their real comoving coordinates  $\mathbf{r}$  but the observational coordinates  $\mathbf{x} = (z, \phi, \theta)$ , where  $z$  is their redshift and  $\phi$  and  $\theta$  their sky angles. In observation space, we can define the power spectrum as

$$P(\mathbf{u}) := \int \xi(\Delta\mathbf{x}) e^{i\mathbf{u}\cdot\Delta\mathbf{x}} d^3\Delta\mathbf{x}, \quad (4.6)$$

where the quantity  $\xi(\Delta\mathbf{x})$  is extracted similarly to  $\xi_g(\Delta\mathbf{r})$ . Using this observed power spectrum we can then derive the likelihood

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} \sum_i \frac{1}{2V_{u,i}} \int \frac{(P^{\text{th}}(\mathbf{u}) - P(\mathbf{u}))^2}{(\sigma_P)^2} d^3\mathbf{u} \right]. \quad (4.7)$$

$V_{x,i}$  is the unitless volume of a redshift bin in observation space. Since the calculation of the observed power spectrum is not done in observation space but in real space, we need to translate the quantities in the likelihood. For this, we need to assume a reference cosmology, the effects of

choosing this are discussed below. To translate from observation space to comoving space we can use the transformation of measures

$$\begin{aligned} dr_{\parallel} &= \frac{c}{H(z)} dz \\ dr_{\perp} &= (1+z) D_A(z) d\theta, \end{aligned} \quad (4.8)$$

where the subscripts  $\parallel$  and  $\perp$  stand for the components parallel and orthogonal to the line of sight respectively. This directly tells us how to translate  $V_x$  into  $V$  by doing the variable transformation

$$V_x \approx \frac{H(\bar{z})V}{c(1+\bar{z})^2 D_A(\bar{z})^2},$$

where  $\bar{z}$  denotes the center of a redshifts bin.

Since we work with binned redshifts  $z$  in a flat sky approximation, the typical separations  $\Delta x$  are small, such that we can approximate the transformation also for the coordinate separations to find

$$\Delta r_{\parallel} = \frac{c}{H(\bar{z})} \Delta x_{\parallel} \quad (4.9)$$

$$\Delta r_{\perp} = (1+\bar{z}) D_A(\bar{z}) \Delta x_{\perp} \quad (4.10)$$

The conjugated variables for  $\Delta r$  and  $\Delta x$  are  $\mathbf{k}$  and  $\mathbf{u}$  respectively. They scale inversely

$$d^3\mathbf{k} = \frac{1}{c(1+\bar{z})^2} \frac{H(\bar{z})}{(D_A(\bar{z}))^2} d^3\mathbf{u}. \quad (4.11)$$

We can use the transformation from observation space to real space we find how the power spectra are related:

$$P^{\text{obs}}(\mathbf{k}) = c(1+\bar{z})^2 \frac{(D_A(\bar{z}))^2}{H(\bar{z})} P(\mathbf{u}). \quad (4.12)$$

When doing the standard FKP[Feldman et al., 1994] method of extracting a minimum uncertainty estimator of the two-point correlation and power spectrum one finds that the error  $\sigma_P$  is given by

$$\sigma_P^2(\mathbf{u}) = \frac{(2\pi)^3}{V_x V_u} P(u)^2. \quad (4.13)$$

The resolution of the instrument is finite. The angles and redshifts have some error associated with them  $\sigma_z$  and  $\sigma_\theta$  respectively. Due to this, the galaxy density in real space is no longer a summation over Dirac-deltas but over Gaussians. In Fourier space, the power spectrum is then suppressed by a Gaussian such that

$$\begin{aligned} P(k, z) &\rightarrow P(k, z) \exp \left[ -\sigma_{\parallel}^2 k_{\parallel}^2 - \sigma_{\perp}^2 k_{\perp}^2 \right] := q_\sigma(k, \mu, z) P(k, z) \\ \sigma_{\parallel} &= \frac{c}{H} \sigma_z, \quad \sigma_{\perp} = (1+z) D_A \sigma_\theta. \end{aligned} \quad (4.14)$$

The final formula is this simple as the Fourier transform of a Gaussian is again a Gaussian. This leaves us with the equivalent quantity to the two-halo term from chapter 3. Its relation to the bias is discussed in the next section. Finally, we can add the one-halo term to this. But its interpretation is quite different, since the galaxy count is a counting experiment, we can identify it as shot noise. The contribution is exactly like for Poisson noise such that for each bin we add

$$P(k, z) \rightarrow P(k, z) + 1/n_i^{\text{ref}}, \quad (4.15)$$

where  $n_i^{\text{ref}}$  is the mean number density of galaxies in the redshift bin  $i$  calculated at the reference cosmology.

If we insert everything now into the likelihood we find that

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} \sum_i \frac{V_i}{2(2\pi)^3} \int_{\Delta V_{k,i}} \frac{(P^{\text{th}}(\mathbf{k}) - P^{\text{obs}}(\mathbf{k}))^2}{(P^{\text{th}}(\mathbf{k}))^2} d^3k \right]. \quad (4.16)$$

In the next step, we use that the power spectrum can only be a function of the modulus of the wave vector  $k := \|\mathbf{k}\|$  as well as the angle between the line of sight  $\hat{\mathbf{n}}$  and the wave vector  $k \mu := \mathbf{k} \cdot \hat{\mathbf{n}}$ . We can use this to integrate over the azimuth angle of the wave vector and find that the likelihood is given by

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} \sum_i \frac{V_i}{8\pi^2} \int_{k_{\min}}^{k_{\max}} \int_{-1}^1 k^2 \frac{(P^{\text{th}}(k, \mu) - P^{\text{obs}}(k, \mu))^2}{(P^{\text{th}}(k, \mu))^2} d\mu dk \right]. \quad (4.17)$$

It should be noted, that the quantities  $k$  and  $\mu$  are very much still dependent on the reference cosmology. When we do an Markov Chain Montecarlo parameter inference, we calculate the theoretical power spectrum  $P^{\text{th}}(k, \mu)$  that is predicted by some other cosmology that we will call sample cosmology. We then use the likelihood to test how likely it is to measure the observed data given the sample cosmology. But since all the integration is done in the reference cosmology we need to account for that during the transformation from observation space to real space, described by equations 4.8 as well as when we translate the power spectrum from  $\mathbf{u}$ -space to  $\mathbf{k}$ -space 4.12. This leads to the equations describing the *Alcock-Paczynski* effect. The components of the wave vector parallel and orthogonal to the line of sight are transformed into

$$\begin{aligned} k_{\parallel}^{\text{smp}} &:= k^{\text{smp}} \mu^{\text{smp}} = \frac{H^{\text{smp}}(z)}{H^{\text{ref}}(z)} k^{\text{ref}} \mu^{\text{ref}}, \\ k_{\perp}^{\text{smp}} &:= k^{\text{smp}} \sqrt{1 - (\mu^{\text{smp}})^2} = \frac{D_A^{\text{ref}}(z)}{D_A^{\text{smp}}(z)} k^{\text{ref}} \sqrt{1 - (\mu^{\text{ref}})^2}, \\ P^{\text{smp}}(k^{\text{ref}}, \mu^{\text{ref}}) &= \left( \frac{D_A^{\text{ref}}(z)}{D_A^{\text{smp}}(z)} \right)^2 \frac{H^{\text{smp}}(z)}{H^{\text{ref}}(z)} P^{\text{smp}}(k^{\text{smp}}, \mu^{\text{smp}}), \end{aligned} \quad (4.18)$$

where we have used the index 'smp' to note quantities calculated in the sample cosmology and ref for quantities calculated in the reference cosmology. This can be used to find equations describing the transformation of  $k$  and  $\mu$  to find the final equations.

$$\begin{aligned} k^{\text{smp}} &= k^{\text{ref}} \sqrt{\left[ \frac{H^{\text{smp}}(z)}{H^{\text{ref}}(z)} \mu^{\text{ref}} \right]^2 + \left[ \frac{D_A^{\text{ref}}(z)}{D_A^{\text{smp}}(z)} \sqrt{1 - (\mu^{\text{ref}})^2} \right]^2} \\ \mu^{\text{smp}} &= \mu^{\text{ref}} \frac{H^{\text{smp}}(z)}{H^{\text{ref}}(z)} \sqrt{\left[ \frac{H^{\text{smp}}(z)}{H^{\text{ref}}(z)} \mu^{\text{ref}} \right]^2 + \left[ \frac{D_A^{\text{ref}}(z)}{D_A^{\text{smp}}(z)} \sqrt{1 - (\mu^{\text{ref}})^2} \right]^2}^{-1} \end{aligned} \quad (4.19)$$

Using these transformations, we can write now the effective  $\chi^2$  using only functions of the reference cosmology

$$\chi^2 = \sum_i \int_{k_{\min}}^{k_{\max}} \int_{-1}^1 \frac{V_i k_{\text{ref}}^2}{8\pi^2} \frac{(q_{\parallel} q_{\perp}^2 P_{\text{smp}}^{\text{th}}(k^{\text{smp}}(k^{\text{ref}}, \mu^{\text{ref}}, \bar{z}_i), \mu^{\text{smp}}(\mu^{\text{ref}}, \bar{z}_i), \bar{z}_i) - P_{\text{ref}}^{\text{obs}}(k^{\text{ref}}, \mu^{\text{ref}}, \bar{z}_i))^2}{(q_{\parallel} q_{\perp}^2 P_{\text{smp}}^{\text{th}}(k^{\text{smp}}(k^{\text{ref}}, \mu^{\text{ref}}, \bar{z}_i), \mu^{\text{smp}}(\mu^{\text{ref}}, \bar{z}_i), \bar{z}_i))^2} d\mu^{\text{ref}} dk^{\text{ref}},$$

where we have defined the redshift dependent parameters  $q_{\parallel} := H^{\text{smp}}/H^{\text{ref}}$  and  $q_{\perp} := D_A^{\text{ref}}/D_A^{\text{smp}}$ .

## A Footnote about Units

When working in cosmology there are two different units we use for the wave number  $k$ , you could either pass it with units  $\text{Mpc}^{-1}$  or  $h \text{ Mpc}^{-1}$ . There are arguments to use each unit but one consequence of the second way of doing it is that now the value of the wave number is dependent on the value of the cosmological parameter  $h$ . This has the following effect: If we use the units with  $h$ , then the integration in the likelihood is done over  $k^{\text{ref}}$  in  $h^{\text{ref}} \text{ Mpc}^{-1}$ . The wave number of the sample cosmology  $k^{\text{smp}}$  would still be in units  $h^{\text{smp}} \text{ Mpc}^{-1}$ . To fix this we need to do a substitution of  $k^{\text{smp}} \rightarrow k' = \frac{h^{\text{ref}}}{h^{\text{smp}}} k^{\text{smp}}$ , fixing the unit of the integration variable to the unit in the reference cosmology and multiplying a factor  $h_{\text{smp}} h_{\text{ref}}^{-1}$  in front of  $P_{\text{smp}}^{\text{th}}$ . By absorbing the substitution in equations 4.18 we find the modified equation

$$\begin{aligned} k'_{\text{smp}} &= k_{\text{ref}} \frac{h_{\text{smp}}}{h_{\text{ref}}} \sqrt{q_{\perp}^2 + (q_{\parallel}^2 - q_{\perp}^2) \mu_{\text{ref}}^2} \\ \mu'_{\text{smp}} &= k_{\text{ref}} q_{\parallel} \sqrt{q_{\perp}^2 + (q_{\parallel}^2 - q_{\perp}^2) \mu_{\text{ref}}^2}^{-1}. \end{aligned}$$

Also, the unit of the power spectrum could be either  $\text{Mpc}^3$  or  $\text{Mpc}^3 h^{-3}$ . Again the second choice has the effect that when we compare the power spectra the units do not match. One would need to absorb this unit change again by multiplying  $P_{\text{smp}}^{\text{th}}$  with a factor  $h_{\text{ref}}^3 h_{\text{smp}}^{-3}$ . In the final likelihood we choose the units of  $k^{\text{ref}}$  to be  $h \text{ Mpc}^{-1}$  while the unit of the power spectrum is  $\text{Mpc}^3$  such that the formula for the effective  $\chi^2$  is given as

$$\chi^2 = \sum_i \int_{k_{\min}}^{k_{\max}} \int_{-1}^1 \frac{V_i k_{\text{ref}}^2}{8\pi^2} \frac{(q_{\parallel} q_{\perp}^2 \frac{h^{\text{smp}}}{h^{\text{fid}}} P_{\text{smp}}^{\text{th}}(k'_{\text{smp}}, \mu'_{\text{smp}}, \bar{z}_i) - P_{\text{ref}}^{\text{obs}}(k^{\text{ref}}, \mu^{\text{ref}}, \bar{z}_i))^2}{(q_{\parallel} q_{\perp}^2 \frac{h^{\text{smp}}}{h^{\text{fid}}} P_{\text{smp}}^{\text{th}}(k'_{\text{smp}}, \mu'_{\text{smp}}, \bar{z}_i))^2} d\mu^{\text{ref}} dk^{\text{ref}}. \quad (4.20)$$

## 4.2 Extracting the Power Spectrum from Theory

For the calculation of the likelihood, we will compare the observed power spectrum and a power spectrum that we calculate from theoretical modelling. Since we have no real data yet, for the forecast we will use the same recipe that we use to calculate the theoretical power spectrum for the observed power spectrum. For this, we will choose some fiducial cosmology and let it coincide with the reference cosmology for simplicity.

The power spectrum of galaxies is assumed to be a tracer of the matter power spectrum. In the most naive way, we say that the density contrast of galaxies is related to the density contrast of matter via a linear bias  $b$ , like we did in the halo model. We can thus write

$$\delta_g(k, z) = b(z, k) \delta(k, z).$$

Calculating the variance of this on both sides extracts the power spectrum. We find that the bias just factors out.

$$P_g(k, z) = b^2(z, k) P_{mm}(k, z) \quad (4.21)$$

If we assume the bias to be scale independent (which is known to be valid in cosmologies without scale-dependent growth and on small intermediate scales) we could treat the bias as one free parameter for each redshift bin.

The matter power spectrum begins to deviate from the linear power spectrum on scales that are of interest to us. This is also discussed in chapter 3. Most interesting for the spectroscopic probe is that the baryonic acoustic oscillations are damped by the individual  $k$ -modes interacting with each other. To model this we use the prescription of [Euclid Collaboration: Blanchard et al., 2020], which "de-wiggles" the power spectrum by taking a weighted sum of the linear power spectrum  $P_{mm}^{\text{lin}}$  and a "no-wiggle" power spectrum  $P_{nw}$ .

$$P_{mm}(k, \mu, z) = P_{mm}^{\text{lin}}(k, z) e^{-g} + P_{nw}(k, z) (1 - e^{-g}) \quad (4.22)$$

$$g(k, \mu, z) = \sigma_v(z)^2 \left( k_\perp^2 + k_\parallel^2 (1 + f^{\text{fid}}(z))^2 \right) \quad (4.23)$$

The parameter  $\sigma_v(z)$  this time is a nuisance parameter that describes the velocities of the galaxies on large scales. Like in the case of HMCODE, it can be modelled as the variance of the velocity divergence field, which on linear scales is given as an integral over the matter power spectrum. As this is very much open for debate, in the final forecast we will vary  $\sigma_v$  and use the value inferred from the integral as our fiducial value

$$[\sigma_v^{\text{fid}}]^2 = \frac{1}{6\pi^2} \int_{k_{\min}}^{k_{\max}} P_{mm}^{\text{lin}}(k, z) dk. \quad (4.24)$$

In the formula of the de-wiggling weight  $g$ , we have also fixed the linear growth rate  $f(z)$  to its fiducial value. This is done such that the modelling of the de-wiggling weight is not cosmology-dependent and thus does not lead to further constraining power. To obtain the "no-wiggle" spectrum we apply a Savitzky- Golay filter to the linear spectrum.

Notice the similarity between the calculation of the de-wiggled power spectrum in HMCODE to the calculation presented in this section. At the fiducial, the only difference is the addition of the factor  $f$  in the exponential and the numeric calculation of the "no-wiggle" spectrum.

The "de-wiggled" power spectrum is further modified by other observational effects that we have

discussed like the *Alcock-Paczynski*-effect and the finite resolution of the instrument. One effect we have not discussed yet is redshift space distortions (RSD). These stem from the fact, that the redshift inferred from galaxies is not only of cosmological origin but also part of it is from their particular velocities. The effect can be described by the following example: Imagine two mass points in space moving towards each other due to their attractive force. The point that is further away from the observer will have its projected velocity towards the observer and thus appear to have a smaller redshift. Since the opposite is true for the point closer to the observer the points will move towards each other in redshift space. If the collapse is slow enough, this would lead to an overall amplification of the power spectrum, as the points look like they are closer to each other. The effect is strongest when the two points align with the line of sight and vanishes when they are perpendicular to the line of sight, thus the function describing this effect has to be a function of  $\mu^2$ . The effect would also be proportional to the clustering parameter <sup>1</sup>  $\sigma_8 f$  to have these spherically collapsing matter distributions. As the observation is done in redshift space we multiply the redshift space power spectrum with the Kaiser correction factor

$$q_{\text{RSD}} = (b(z) + f(z) \mu^2)^2. \quad (4.25)$$

To find the clustering parameter, which is better constrained by clustering experiments, we expand with  $\sigma_8$  and find

$$q_{\text{RSD}} = (b(z)\sigma_8(z) + f(z)\sigma_8(z) \mu^2)^2 \frac{1}{\sigma_8^2(z)}.$$

The combination  $b\sigma_8$ , or to be more precise the log of this parameter, is varied as a nuisance parameter in the forecast for each bin.

If the matter points happen to approach each other very quickly, then at some point in observation space the point further away in real space will have a smaller redshift than the closer one. This effect is called the "fingers of God" (FOG) effect, as a spherically collapsing matter distribution can end up looking like very elongated ellipses. As the two points look like they would be further away, the overall power spectrum will be reduced. The effect should have the same angle dependence, should be stronger on smaller, more nonlinear scales and again be stronger in stronger clustering cosmologies. One possible parametrization of this effect is given as

$$q_{\text{FOG}} = \frac{1}{1 + [f^{\text{fid}}(z) \sigma_p(z) k \mu]^2}. \quad (4.26)$$

The parameter  $\sigma_p$  describes the variance of the peculiar velocities of the galaxies and is again modelled as an integral over the linear velocity divergence power spectrum. As this is again a source of theoretical modelling uncertainty, we will fix  $f$  to its fiducial value and vary  $\sigma_p$ . The fiducial value of  $\sigma_p$  is chosen to coincide with  $\sigma_v$ .

$$\sigma_p^{\text{fid}} = \sigma_v^{\text{fid}} \quad (4.27)$$

---

<sup>1</sup>More often the term clustering parameter refers to the parameter  $S_8 = \sqrt{\Omega_m/0.3} \sigma_8$  and not the combination  $f\sigma_8$ . These two quantities are very similar to each other as in  $\Lambda$ CDM the growth rate can be approximated as  $f \approx \Omega_m^\gamma$  with  $\gamma = 0.55$ . When looking at the 2-D contour of similar galaxy clustering experiments we see that the definition with  $\gamma = 0.55$  describes the degeneracy in the  $\sigma_8$ - $\Omega_m$  plane better than  $\gamma = 0.5$ , so we will continue calling  $f\sigma_8$  the clustering parameter.

Finally, we add some additional shot noise  $P_i^{\text{shot}}$  for each bin as an additional nuisance parameter. Our final model for the power spectrum of some cosmology evaluated at the fiducial cosmology is thus

$$\begin{aligned} P^{\text{th}}(k^{\text{fid}}, \mu^{\text{fid}}, z_i) &= q_{\parallel} q_{\perp}^2 q_{\text{RSD}} q_{\text{FOG}} q_{\sigma} P_{mm} + \frac{1}{n_i^{\text{fid}}} + P_i^{\text{shot}} \\ &= \left[ \frac{H(z_i)}{H^{\text{fid}}(z_i)} \left[ \frac{D_A^{\text{fid}}(z_i)}{D_A(z_i)} \right]^2 \frac{(b_i \sigma_8(z_i) + f(z_i) \sigma_8(z_i) \mu^2)^2}{1 + [f^{\text{fid}}(z_i) \sigma_p(z_i) k \mu]^2} \right. \\ &\quad \times \exp \left[ -\sigma_{\parallel}^2 k^2 \mu^2 - \sigma_{\perp}^2 k^2 (1 - \mu^2) \right] \frac{P_{mm}(k, z)}{\sigma_8^2(z)} \\ &\quad \left. + \frac{1}{n_i^{\text{fid}}} + P_i^{\text{shot}} \right] \frac{H^{\text{ref}}}{c [(1 + z_i) D_A^{\text{ref}}]^2}, \end{aligned} \quad (4.28)$$

where we have left out the arguments  $k(k^{\text{fid}}, \mu^{\text{fid}}, z_i)$  and  $\mu = \mu(\mu^{\text{fid}}, z_i)$ . We also note that the factor  $H^{\text{ref}}/c [(1 + z_i) D_A^{\text{ref}}]^2$ , that converts the power spectrum from real space to redshift space, cancels out in the final likelihood but is listed here for completeness's sake.

### 4.3 The Effect of Massive Neutrinos

The main sources of this section are the papers [Raccanelli et al., 2018] and [Vagnozzi et al., 2018] which I will be referencing throughout. One assumption we made in the last section was that the galaxy bias  $b$  would be scale-independent. The reason this cannot be true is that since the neutrinos become free streaming at some scale they contribute to the power spectrum like CDM at large scales and do not contribute to the matter power spectrum on small scales. This adds an overall scale dependence to the bias as now there is a scale at which the content of the clustered matter changes. One method of getting rid of the scale dependence in the bias is to understand what this bias is describing. It is a light-to-mass bias meaning that it associates some matter overdensity to the galaxy or halo observed. Since these structures are much smaller than the free streaming scale, they are not tracers of the total matter spectrum  $P_{mm}$ , but the power spectrum of CDM+baryons  $P_{cb}$ . It has been shown in [Villaescusa-Navarro et al., 2014] and the follow-up papers thereof, that by using the CDM+baryon power spectrum the bias is again scale-independent large scales. In real space, we can write

$$\begin{aligned} P_g(k) &= b^2(k, z) P_{mm}(k, z) = \hat{b}^2(z) P_{cb}(k, z), \\ \hat{b}(z) &:= b(k, z) \frac{T_{mm}(k, z)}{T_{cb}(k, z)}, \end{aligned} \quad (4.29)$$

where we used the transfer function  $T_{mm}$  and  $T_{cb}$ . We define the bias with the transfer functions and not with the power spectra as the bias originally connects the density contrasts and not the power spectra. The difference is a bit subtle but will become apparent in section 5.3 when we discuss the neutrino effect in the photometric probe.

In the presence of RSD, we have to modify the Kaiser correction factor to account for the change

from the matter power spectrum to the CDM+baryon power spectrum. The factor becomes

$$q_{\text{RSD}} = \left( \hat{b}(z) + f_{cb}(z, k) \mu^2 \right)^2. \quad (4.30)$$

We can notice two changes here, first the aforementioned change from  $b$  to  $\hat{b}$  and the change from the growth rate  $f$  to the related quantity  $f_{cb}$ . It is calculated exactly like  $f$  but using the growth of the CDM+baryon perturbations. To again find the clustering parameter of the CDM+baryon distribution we also have to multiply with the quantity  $\sigma_8^{cb}$  that is calculated again using the CDM+baryon power spectrum. The reason is simple, if we remind ourselves of the definition of  $\sigma_R$  to be the variance of perturbations that are smoothed in real space over balls of radius  $R$ . We argued that the Kaiser correction should be a function of  $f \sigma_8$  as this combination describes the clustering of the matter distribution. Now we want to see how much the CDM+baryon clusters, so we want to smooth the CDM+baryon field and find that this has to be proportional to  $\sigma_8^{cb}$ . This finally leads to

$$q_{\text{RSD}} = \left( \hat{b}(z) \sigma_8^{cb}(z) + f_{cb}(z, k) \sigma_8^{cb}(z) \mu^2 \right)^2 \frac{1}{(\sigma_8^{cb}(z))^2}.$$

The FOG effect remains unchanged. It could be argued that in order to describe the effect of the redshift space distortions we should change the fiducial value of  $f \sigma_p$  to quantities computed using the CDM+baryon power spectrum. However, we decided to keep them, since the effect of the FOG is coming from random peculiar velocities of the individual halos towards each other. The velocities are related to the gravitational potential inside these structures by the Virial theorem, equating their mean kinetic energy with their mean potential energy. This means that the effect of the neutrinos should still apply. Thus, we stick to the prescription with the quantities calculated using the total matter distribution.

An open question would be how to correctly calculate the nonlinear power spectrum. If we decide to stick to the method of de-wiggling, the question becomes if we want to change our model for  $\sigma_p$ . We decided against this. The physical effect of de-wiggling comes from large-scale flows of the matter structures and is thus related to the underlying gravitational potential. As the neutrinos still contribute to the potential we stick to the quantities calculated with the total matter power spectrum. We stress here that both of these choices do not matter. In reality, when we do an MCMC we do not know about any fiducial values of these nonlinear modeling parameters. This means that there should not be any dependence of our final forecast on the fiducial value of these nuisance parameters.

The other factors do not depend on the power spectrum and are as before. The final power spectrum is used to more robustly describe the effect of massive neutrinos.

Nuisance Param	Fiducial Value	Nuisance Param	Fiducial Value
$\log(\hat{b}\sigma_8^{cb})_1$	-0.3270	$\sigma_p(z_1)$	5.2554
$\log(\hat{b}\sigma_8^{cb})_2$	-0.3128	$\sigma_p(z_2)$	4.8287
$\log(\hat{b}\sigma_8^{cb})_3$	-0.3087	$\sigma_p(z_3)$	4.4606
$\log(\hat{b}\sigma_8^{cb})_4$	-0.3186	$\sigma_p(z_4)$	4.0677
$P_1^{\text{shot}}$	0	$\sigma_v(z_1)$	5.2554
$P_2^{\text{shot}}$	0	$\sigma_v(z_2)$	4.8287
$P_3^{\text{shot}}$	0	$\sigma_v(z_3)$	4.4606
$P_4^{\text{shot}}$	0	$\sigma_v(z_4)$	4.0677

Table 4.1: Fiducial values of nuisance parameters needed to calculate the spectroscopic likelihood.

Survey Spec	Value	Survey Spec	Value
$\bar{n}_1$	$6.68 \cdot 10^{-4} h^3 \text{Mpc}^{-3}$	$[z_{\min}, z_{\max}]_1$	[0.9, 1.1]
$\bar{n}_2$	$5.58 \cdot 10^{-4} h^3 \text{Mpc}^{-3}$	$[z_{\min}, z_{\max}]_2$	[1.1, 1.3]
$\bar{n}_3$	$4.21 \cdot 10^{-4} h^3 \text{Mpc}^{-3}$	$[z_{\min}, z_{\max}]_3$	[1.3, 1.5]
$\bar{n}_4$	$2.61 \cdot 10^{-4} h^3 \text{Mpc}^{-3}$	$[z_{\min}, z_{\max}]_4$	[1.5, 1.8]
$\sigma_\theta$	0	$z_1$	1.0
$\sigma_z$	0.001	$z_2$	1.2
$k_{\min}$	$0.001 h\text{Mpc}^{-1}$	$z_3$	1.4
$k_{\max}$	$(25/30) h\text{Mpc}^{-1}$	$z_4$	1.65

Table 4.2: Survey specifications needed to calculate the spectroscopic likelihood. For  $k_{\max}$  we noted two values that represent pessimistic and optimistic settings respectively.

$$\begin{aligned}
P^{\text{th}}(k^{\text{fid}}, \mu^{\text{fid}}, z_i) = & \left[ \frac{H(z_i)}{H^{\text{fid}}(z_i)} \left[ \frac{D_A^{\text{fid}}(z_i)}{D_A(z_i)} \right]^2 \frac{\left( \hat{b}_i \sigma_8^{cb}(z_i) + f^{cb}(z_i) \sigma_8^{cb}(z_i) \mu^2 \right)^2}{1 + [f^{\text{fid}}(z_i) \sigma_p(z_i) k \mu]^2} \right. \\
& \times \exp \left[ -\sigma_{\parallel}^2 k^2 \mu^2 - \sigma_{\perp}^2 k^2 (1 - \mu^2) \right] \frac{P_{mm}(k, z)}{(\sigma_8^{cb}(z))^2} \\
& \left. + \frac{1}{n_i^{\text{fid}}} + P_i^{\text{shot}} \right] \frac{H^{\text{ref}}}{c \left[ (1 + z_i) D_A^{\text{ref}} \right]^2} \tag{4.31}
\end{aligned}$$

In tables 5.2 and 4.2 are listed the specifications and fiducial values needed to compute the likelihood.

## Chapter 5

# The Photometric Probe

The second instrument of the *Euclid* mission that we will cover is the photometric instrument. It is supposed to measure the galaxy shape of one billion galaxies and use redshift information from ground-based observatories to conduct a weak lensing and a galaxy clustering survey. In order to quickly measure the redshift of one billion galaxies the instrument uses a photometric method. A photometric redshift is obtained by measuring the apparent brightness magnitude in different colour bands. These color brightnesses are compared to each other leading to a redshift estimate. The observable in this case is not the power spectrum as for the spectroscopic probe, but as two-dimensional overdensity field  $\Theta$  that can either correspond to the lensing map or the galaxy map. From these maps we obtain the moments  $a_{lm}$  via decomposition into spherical harmonics. Using their orthogonality relation, we find

$$a_{\ell,m} = \int_{S_2} \Theta(\Omega) Y_{\ell m}^*(\Omega) d^2\Omega. \quad (5.1)$$

We can then build a vector by calculating these moments for each  $z$  bin. The final likelihood is modelled as a Gaussian where the observed spherical moments are modelled to have a zero mean and covariance  $C_\ell^{\text{th}}$ , that can be extracted from theory. Because of the isotropy of space, the covariance matrix can not depend on the azimuth quantum number  $m$ . The covariance matrix correlates the multipoles for the same  $\ell$  and  $m$  to the ones from different  $z$  bins. The formula for the likelihood is thus given by

$$\mathcal{L} \propto \prod_{\ell m} \left[ \frac{1}{\sqrt{\det C_\ell}} \exp \left( -f_{\text{sky}} \frac{1}{2} \vec{a}_{\ell m} \cdot (C_\ell^{\text{th}})^{-1} \vec{a}_{\ell m}^* \right) \right]. \quad (5.2)$$

This expression can be further simplified by calculating an estimator of the covariance of the moments. To find the real covariance one would need to average over different realizations of the universe but by using that the covariance is independent of  $m$  we can find

$$(C_\ell^{\text{obs}})_{ij} \approx (\hat{C}_\ell^{\text{obs}})_{ij} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} (a_{\ell m})_i (a_{\ell m})_j^*. \quad (5.3)$$

After factoring out  $\sqrt{\det \hat{C}_\ell^{\text{obs}}}$  from the likelihood normalization we can write it using only products of  $\hat{C}_\ell^{\text{obs}}$  and  $C_\ell^{\text{th}}$ . For that, we also need to pull the product into the exponential and use the

summation over  $m$  to find the estimator again. After pulling the product over  $\ell$  into the exponential we are left with

$$\chi^2 = f_{\text{sky}} \sum_{\ell=\ell_{\min}}^{\ell_{\max}} (2\ell+1) \left[ \log \left( \frac{\det \hat{C}_{\ell}^{\text{obs}}}{\det C_{\ell}^{\text{th}}} \right) + \text{Tr} \left[ (C_{\ell}^{\text{th}})^{-1} \hat{C}_{\ell}^{\text{obs}} \right] - N_{\text{bin}} \right], \quad (5.4)$$

where  $N_{\text{bin}}$  is the number of redshift bins.

The structure of this chapter is the following: We will first briefly explain the basics of shearing and the shearing formalism. We will then go over how we obtain the covariance matrix from theory. Finally, we will more carefully adjust the recipe to correctly model the effect of massive neutrinos.

## 5.1 The Formalism of Cosmic Shear Observations

This section follows the work of [Kilbinger, 2015] and [Sprenger et al., 2019]. We will be citing them throughout this chapter.

Cosmic shearing is a very rich and deep topic where the slight deformations of galaxy shapes are traced back to the gravitational potential on the line of sight. These potentials are related to the matter perturbations  $\bar{\rho}\delta$  via the Poisson equation.

$$\begin{aligned} \Delta(\Psi + \Phi) &= 8\pi G a^2 \bar{\rho} \delta \\ &= 3 H_0^2 \frac{\Omega_m}{a} \delta \end{aligned}$$

This can be used to probe the underlying LSS. The deformations are a transformation that maps angular separations in observed space  $\boldsymbol{\Omega}$  to angular separations in unlensed space  $\boldsymbol{\beta}$ . We can then define the linear transformation  $A$  as

$$A = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\Omega}} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (5.5)$$

The parameter  $\kappa$  is called the convergence and describes the overall change in the size of the observed shape while the parameters  $\gamma_1$  and  $\gamma_2$  describe the shear or the overall rotation of the shape. Since in the context of cosmic shearing, the true sizes of galaxies are inaccessible by shape measurement alone, the real observable is the complex, reduced shear

$$g := \frac{\gamma}{1 - \kappa}, \quad \text{with } \gamma = \gamma_1 + i\gamma_2. \quad (5.6)$$

The reduced shear can be related to the observed ellipticity  $\epsilon$  of galaxies. We define the complex ellipticity of a galaxy we can define elliptic regions of constant intensity with minor to major axis ratio  $b/a$ , and rotation angle  $\phi$   $\epsilon = (a-b)/(a+b) \times \exp(2i\phi)$ . This leads to the observed ellipticity.

$$\epsilon = \frac{\epsilon^2 + g}{1 + g^* \epsilon^s} \quad (5.7)$$

Here  $\epsilon^s$  denotes the intrinsic ellipticity due to the random rotation of the disc-like galaxies. Typical values of the intrinsic ellipticity are in the order of  $\mathcal{O}(0.1)$  with zero mean, while the reduced shear

is typically of order  $\mathcal{O}(0.01)$ . By averaging over many galaxies and using that  $g$  is much smaller than  $\epsilon^s$  we can find the estimator of  $g$

$$\langle \epsilon \rangle = \left\langle \frac{\epsilon^2 + g}{1 + g^* \epsilon^s} \right\rangle \approx \langle \epsilon^s + g \rangle = g. \quad (5.8)$$

The next step is to relate the lensing parameters to each other and the metric perturbations to extract the overdensity field  $\Theta$ . We can use that photons travel on null geodesics to find that the travel time  $t$ .

$$t = \int (1 - (\Phi + \Psi)) dr$$

The integration is carried out on the light path of the photon. By using that photons always take a path that minimises the travel time under slight variation we find that the deflection angle  $d\alpha$  is given by

$$d\alpha = -\nabla_{\perp}(\Phi + \Psi) dr, \quad (5.9)$$

where we have denoted the derivative of the potentials perpendicular to the path with  $\nabla_{\perp}$  in real coordinates. This translates into a change of comoving separation  $dx$  and thus a change of observation angle  $d\alpha$ .

$$dx = (\eta - \eta') d\alpha = -(\eta - \eta') \nabla_{\perp}(\Phi + \Psi) d\eta' \quad (5.10)$$

$$x = \eta \Omega - \int_0^\eta (\eta - \eta') [\nabla_{\perp}(\Phi + \Psi)(x(\eta'), \eta') - \nabla_{\perp}(\Phi + \Psi)(0, \eta')] d\eta' \quad (5.11)$$

We can think of this formula geometrically. We have defined the origin of the observed photons at the comoving distance  $\eta$ . The gravitational potential is placed at the comoving distances  $\eta'$  on the path of the photon. The difficulty of this integration is, that the comoving separation of the two photon paths  $x$  appears again in the integral. Since the effect of the lens is assumed to be very weak, we can do a Born approximation where the zeroth order solution  $x^0$  is inserted into the integral. This solution is the unlensed path  $x^0 = \eta \Omega$ . With this the matrix elements of the transformation  $A$  can be calculated to

$$A_{ij} = \delta_{ij} - \int_0^\eta \frac{(\eta - \eta') \eta'}{\eta} \frac{\partial^2}{\partial x_i \partial x_j} (\Phi + \Psi)(\eta' \Omega, \eta') d\eta'. \quad (5.12)$$

The first derivative for  $x_i$  is the component of  $\nabla_{\perp}$  while the derivative for  $x_j$  appears with the factor  $\eta'$  due to the chain rule. We can reformulate this equation as well using the inverse chain rule. The derivatives can then be pulled out of the integral to define lensing potential  $\psi$ .

$$A_{ij} := \delta_{ij} - \frac{\partial^2}{\partial \Omega_i \partial \Omega_j} \psi, \quad (5.13)$$

$$\psi := \int_0^\eta \frac{(\eta - \eta')}{\eta \eta'} (\Phi + \Psi)(\eta' \Omega, \eta') d\eta' \quad (5.14)$$

The lensing parameters  $\kappa$  and  $\gamma$  can be calculated from this.

$$\begin{aligned}\kappa &= \frac{1}{2} \left[ \frac{\partial^2}{\partial \Omega_1^2} + \frac{\partial^2}{\partial \Omega_2^2} \right] \psi \\ \gamma_1 &= \frac{1}{2} \left[ \frac{\partial^2}{\partial \Omega_1^2} - \frac{\partial^2}{\partial \Omega_2^2} \right] \psi \\ \gamma_2 &= \frac{\partial^2}{\partial \Omega_1 \partial \Omega_2} \psi\end{aligned}\tag{5.15}$$

The equation for the convergence parameter  $\kappa$  can be transformed into the Poisson equation by adding a term to  $\partial^2/\partial\eta^2$ . The integral over the second derivative has negligible impact as the photons tend to travel on paths where the positive and negative contributions cancel out. We find

$$\begin{aligned}\kappa(\eta \boldsymbol{\Omega}, \eta) &= \frac{1}{2} \int_0^\eta \frac{(\eta - \eta')\eta'}{\eta} \underbrace{\left[ \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial \eta^2} \right]}_{=\Delta} (\Phi + \Psi)(\eta' \boldsymbol{\Omega}, \eta') d\eta' \\ &= \frac{3 H_0^2 \Omega_m}{2} \int_0^\eta \frac{(\eta - \eta')\eta'}{\eta} \frac{1}{a(\eta')} \delta(\eta' \boldsymbol{\Omega}, \eta') d\eta'.\end{aligned}\tag{5.16}$$

This equation describes how the convergence can be used as a density contrast field to calculate the spherical multipole moments. Since we still need to average over multiple galaxies in a redshift bin  $i$ , we need to convolve the convergence with the galaxy distribution  $n_i$  and find

$$\Theta_i^L(\boldsymbol{\Omega}) = \int_{\eta_{\min}}^{\eta_{\max}} n_i(\eta) \kappa(\eta \boldsymbol{\Omega}, \eta) d\eta\tag{5.17}$$

$$\begin{aligned}&= \frac{3 H_0^2 \Omega_m}{2} \int_{\eta_{\min}}^{\eta_{\max}} \int_0^\eta n_i(\eta) \frac{(\eta - \eta')\eta'}{\eta} \frac{1}{a(\eta')} \delta(\eta' \boldsymbol{\Omega}, \eta') d\eta' d\eta \\ &= \frac{3 H_0^2 \Omega_m}{2} \int_{\eta_{\min}}^{\eta_{\max}} \int_\eta^{\eta_{\max}} n_i(\eta') \frac{(\eta' - \eta)\eta}{\eta'} \frac{1}{a(\eta)} \delta(\eta \boldsymbol{\Omega}, \eta) d\eta' d\eta \\ &= \int_{z_{\min}}^{z_{\max}} \frac{1}{H(z)} \frac{3 H_0^2 \Omega_m}{2} (1+z) \eta(z) \int_z^{z_{\max}} \hat{n}_i(z') \frac{\eta(z') - \eta(z)}{\eta(z')} dz' \delta(\eta(z) \boldsymbol{\Omega}, \eta(z)) dz \\ &:= \int_{z_{\min}}^{z_{\max}} \frac{1}{H(z)} W_i^L(z) \delta(\eta(z) \boldsymbol{\Omega}, \eta(z)) dz.\end{aligned}\tag{5.18}$$

In the last step, we have defined the lensing window function  $W^L$ . Since galaxies have a gravitational effect on each other, the galaxy shapes get correlated additionally to the effect of lensing. This effect of intrinsic alignment (IA) is difficult to model such that we will treat it as a nuisance effect. In our forecast, we add to our lensing window function an additional window function  $W^{IA}$ . The exact modelling of the lensing window function is discussed after an intermezzo about Galaxy clustering as the derivation is done similarly. We will assume that the IA density contrast is a biased tracer of the matter density contrast like for galaxy clustering. This is following the description by [Hirata and Seljak, 2010].

We will now briefly go over the phenomenology of IA. Firstly, due to IA close galaxies will have no

random orientation anymore. Thus, in equation 5.8 the mean of  $\epsilon^s$  does not equal zero. In reality, when we do our averaging we would need to give close galaxies a smaller weight to lessen this effect. Another speciality about the IA is that its signals sign is the opposite of the lensing. It becomes apparent if one imagines two spherical galaxy configurations. One that is in the foreground where the galaxies are aligned due to their tidal forces. The foreground galaxy has a perpendicular alignment as the galaxies have their shapes pointing towards each other. The background galaxies will get lensed such that their shapes are aligned tangentially around the foreground galaxies. This means that the deformation directions are perpendicular to one another and the lensing signal is therefore reduced.

Last but not least, IA is special in that its distortion of galaxy shapes does not come with a convergence  $\kappa$  like for shearing. One could use that to lessen the effect of the IA by checking if the convergence is enough to explain the ellipticity of the galaxies or directly using the convergence as the observable. The difficulty would be, that the convergence is inaccessible by galaxy shape measurements alone. There are some proposed methods to measure the convergence by measuring the magnification, as lensed galaxies appear brighter due to the Liouville theorem. This is not further discussed here, but it is the subject of current research.

### Intermezzo: Galaxy Clustering in Angular Space

Since we can also do galaxy clustering with the photometric probe, we will do a brief revision of it's observable but this time in angular space. The idea is similar to the spectroscopic probe. We assume that the galaxy field is a biased tracer of the underlying matter field. Assuming a linear bias, we can relate the density contrast of the galaxies to the density contrast of matter

$$\delta_g(\eta \boldsymbol{\Omega}, \eta) = b(\eta) \delta(\eta \boldsymbol{\Omega}, \eta).$$

Like in the case of weak lensing, we need to average over a redshift bin  $i$  with a galaxy distribution  $n_i(\chi)$ . We can then expand with  $H(z)$  to find a similar expression of the observable using a window function

$$\begin{aligned} \Theta_i^G(\boldsymbol{\Omega}) &= \int_{\eta_{\min}}^{\eta_{\max}} n_i(\eta) \delta_g(\eta \boldsymbol{\Omega}, \eta) d\eta \\ &= \int_{\eta_{\min}}^{\eta_{\max}} n_i(\eta) b(\eta) \delta(\eta \boldsymbol{\Omega}, \eta) d\eta \\ &= \int_{z_{\min}}^{z_{\max}} \frac{1}{H(z)} H(z) \hat{n}_i(z) b(z) \delta(\eta \boldsymbol{\Omega}, z) dz \\ &:= \int_{z_{\min}}^{z_{\max}} \frac{1}{H(z)} W_i^G(z) \delta(\eta(z) \boldsymbol{\Omega}, \eta(z)) dz. \end{aligned} \quad (5.19)$$

This does not account for the observational effects like RSD that were discussed in the chapter about the spectroscopic probe. The modelling and implementation of these are left for future work. Nevertheless, a brief overview of this can be found in [Collaboration et al., 2023].

Continuing our modelling of IA, we use an observationally motivated model that the IA density contrast is given by

$$\delta_{\text{IA}}(\eta \boldsymbol{\Omega}, \eta) = \mathcal{A}_{\text{IA}} \mathcal{C}_{\text{IA}} \Omega_m \frac{\mathcal{F}_{\text{IA}}(\eta)}{D(\eta)} \delta(\eta \boldsymbol{\Omega}, \eta). \quad (5.21)$$

The factor  $\mathcal{A}_{\text{IA}}$  is the bias amplitude that we are going to vary in our analysis, while  $\mathcal{C}_{\text{IA}}$  is a fixed normalization. The factor  $D(\eta)$  is the linear growth factor. The factor  $\mathcal{F}_{\text{IA}}$  is an extension to the standard IA model and is given by the function

$$\mathcal{F}_{\text{IA}}(z) = a(\eta)^{-\eta_{\text{IA}}} \left[ \frac{\langle L \rangle(\eta)}{L_*(\eta)} \right]^{\beta_{\text{IA}}}. \quad (5.22)$$

The functions  $\langle L \rangle$  and  $L_*$  denote the time-dependent mean luminosity and characteristic luminosity of galaxies respectively. The parameter  $\beta_{\text{IA}}$  is fixed while we vary  $\eta_{\text{IA}}$  around its fiducial value. We can now follow the derivation from galaxy clustering to find our window function for IA

$$W_i^{\text{IA}} = \mathcal{A}_{\text{IA}} \mathcal{C}_{\text{IA}} \Omega_m \frac{\mathcal{F}_{\text{IA}}(z)}{D(z)} H(z) \hat{n}_i(z). \quad (5.23)$$

To deal with IA we will now use this window function to modify our old lensing window function

$$W^{\text{L}}(z) \rightarrow W^{\text{L}}(z) - W^{\text{IA}}(z), \quad (5.24)$$

where we use the minus sign as the effect of IA reduces the normal shearing signal.

## 5.2 Extracting the Covariance from Theory

We can now do the derivation of the covariance either using the observable of lensing or clustering. In reality, there is even more information than just the two observable on their own. Since the galaxies trace the matter distribution on which the lensing occurs, there is additional information in the cross-correlation of lensing and galaxy clustering. We can calculate the correlation by taking the averaging over the spherical moments for the two different observable. Since we deliberately chose our definitions with the window functions to be the same between both probes, all elements of the covariance matrix can be calculated in the same way. We define  $(C_\ell^{\text{XY}})_{ij} = \langle (a_{\ell m}^{\text{X}})_i (a_{\ell m}^{\text{Y}})_j \rangle$ , where X and Y could either be G or L. Inserting our definitions of the observables in comoving coordinates, we find

$$(C_\ell^{\text{XY}})_{ij} = \int_{\eta_{\min}}^{\eta_{\max}} \int_{\eta_{\min}}^{\eta_{\max}} \int_{S_2} \int_{S_2} W_i^{\text{X}}(\eta) W_j^{\text{Y}}(\eta') Y_{\ell m}(\Omega) Y_{\ell m}^*(\Omega') \langle \delta(\eta \Omega) \delta(\eta' \Omega') \rangle d^2\Omega d^2\Omega' d\eta d\eta'. \quad (5.25)$$

The mean over the density contrast  $\delta$  will lead to an unequal time power spectrum  $P$ . If we use the orthogonality of the spherical harmonics, we find

$$(C_\ell^{\text{XY}})_{ij} = \int_{\eta_{\min}}^{\eta_{\max}} \int_{\eta_{\min}}^{\eta_{\max}} \int_0^{k_{\max}} \frac{2k^2}{\pi} W_i^{\text{X}}(\eta) W_j^{\text{Y}}(\eta') j_\ell(\eta k) j_\ell(\eta' k) P(k, \eta, \eta') dk d\eta d\eta',$$

where the  $j_\ell(x)$  are the spherical Bessel functions. We can now use that these functions are strongly centred around  $x = \ell + 1/2$  and replace them with Dirac distributions. This is the so-called Limber approximation and it gives

$$j_\ell(k\eta) \approx \sqrt{\frac{\pi}{2\ell+1}} \delta \left( k\eta - (\ell + \frac{1}{2}) \right). \quad (5.26)$$

Inserting this into the definition from  $C_\ell$  gives the full formula for the covariance

$$\begin{aligned} (C_\ell^{\text{XY}})_{ij} &= \int_{\eta^{\min}}^{\eta^{\max}} \frac{1}{\eta^2} W_i^X(\eta) W_j^Y(\eta) P_{mm} \left[ k = \frac{\ell + \frac{1}{2}}{\eta}, \eta \right] d\eta \\ &= \int_{z^{\min}}^{z^{\max}} \frac{W_i^X(z) W_j^Y(z)}{\eta^2(z) H(z)} P_{mm} \left[ k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z \right] dz. \end{aligned} \quad (5.27)$$

To obtain the galaxy distribution per redshift bin in redshift space  $\hat{n}_i(z)$  we start with the overall distribution of galaxies  $\hat{n}(z)$ . We model it as a polynomial with an exponential cut-off

$$\hat{n}(z) = \left( \frac{z}{z_0} \right)^2 \exp \left[ - \left( \frac{z}{z_0} \right)^{1.5} \right]. \quad (5.28)$$

When there is no photometric redshift error, we could just multiply the distribution with some uniform top hat function for each bin. The redshift error can lead to galaxies being counted in the wrong bin. This leads to a bleeding of the bins into each other. To describe this we convolve the galaxy distribution with a sum of two Gaussians

$$\begin{aligned} p(z, z') &= \frac{1 - f_{\text{out}}}{\sqrt{2\pi} \sigma_b (1+z)} \exp \left[ \frac{1}{2} \left( \frac{z - c_b z' - z_b}{\sigma_b (1+z)} \right)^2 \right] \\ &\quad + \frac{f_{\text{out}}}{\sqrt{2\pi} \sigma_0 (1+z)} \exp \left[ \frac{1}{2} \left( \frac{z - c_0 z' - z_0}{\sigma_0 (1+z)} \right)^2 \right]. \end{aligned} \quad (5.29)$$

The second Gaussian is there to describe the effect of strong outliers where the redshift is shifted by a large margin  $z_0$ . This is acting similar to an overall reduction of the observable. The parameters that go into this are all fixed to experimental specifications, and briefly described in table 5.1. In the next step we can calculate the binned galaxy distribution

$$\hat{n}_i(z) = \mathcal{N} \int_{z_i^{\min}}^{z_i^{\max}} \hat{n}(z') p(z, z') dz' \approx \int_{z_i^{\min}}^{z_i^{\max}} \hat{n}(z) p(z, z') dz', \quad (5.30)$$

where we have used, that the function  $p(z, z')$  is strongly centered around  $z = z'$ . The integration limits are the edges of the redshift bins and the factor  $\mathcal{N}$  in front is a normalization that is calculated to be

$$\mathcal{N}^{-1} = \int_{z^{\min}}^{z^{\max}} \int_{z_i^{\min}}^{z_i^{\max}} \hat{n}(z) p(z, z') dz' dz \quad (5.31)$$

Finally, since this is a counting experiment we need to add shot noise  $N_{ij}^{\text{XY}}$  to the angular power spectrum. Since the noise of different bins and probes should be uncorrelated, we have

$$N_{ij}^{\text{GG}} = \delta_{ij} \frac{1}{\bar{n}_i}, \quad N_{ij}^{\text{LL}} = \delta_{ij} \frac{\sigma_\epsilon^2}{\bar{n}_i} \text{ and } N_{ij}^{\text{GL}} = N_{ij}^{\text{LG}} = 0, \quad (5.32)$$

where  $\sigma_\epsilon^2$  is the variance of measured ellipticities. The covariance is thus given by  $C_{ij}^{\text{XY}} \rightarrow C_{ij}^{\text{XY}} + N_{ij}^{\text{XY}}$ . To calculate the likelihood we can now construct a combined observable and covariance

$$\mathbf{a}(\ell, m) = \begin{pmatrix} a_1^L(\ell, m) \\ \vdots \\ a_{N_{\text{bin}}}^L(\ell, m) \\ a_1^G(\ell, m) \\ \vdots \\ a_{N_{\text{bin}}}^G(\ell, m) \end{pmatrix} \quad \text{and} \quad C_\ell = \begin{pmatrix} C_\ell^{\text{LL}} & C_\ell^{\text{LG}} \\ C_\ell^{\text{GL}} & C_\ell^{\text{GG}} \end{pmatrix} \quad (5.33)$$

In the likelihood formula this combined probe can be understood as one observable with twice the bins, so for multipole moments  $\ell$  where we have both observables, we can replace  $N_{\text{bin}}$  with  $2 \times N_{\text{bin}}$ . When we start to cut scales due to resolution or the lack of an adequate nonlinear model, we can find that both observables have a different  $\ell_{\text{max}}$ . For every summand of the likelihood with only one probe, we stick to the equation given above with the covariance of the single probe. To be able to probe higher wave numbers  $k$  we need to be able to predict the nonlinear power spectrum. To obtain our nonlinear power spectrum we use the **HMCODE** halo model that is described in chapter 3. Unlike the nonlinear prescription in the case of the spectroscopic probe, we do not need to add additional nuisance parameters to our modelling. All open parameters have already been fixed by fits to N-body simulations covering a wide range of  $w_0 w_a$ CDM cosmologies with massive neutrinos.

### 5.3 The Effect of Massive Neutrinos

In our definition of the galaxy window function  $W^G$ , due to massive neutrinos the galaxy bias becomes scale-dependent. If we follow our prescription from the spectroscopic probe, we can use a scale-dependent bias by replacing the matter power spectrum with the CDM+baryon spectrum. To calculate the angular power spectrum of the galaxy clustering we then find

$$\begin{aligned} (C_\ell^{\text{GG}})_{ij} &= \int_{z_{\text{min}}}^{z_{\text{max}}} \frac{W_i^G(k, z) W_j^G(k, z)}{\eta^2(z) H(z)} P_{mm} \left[ k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z \right] dz. \\ &= \int_{z_{\text{min}}}^{z_{\text{max}}} \frac{H(z) \hat{n}_i(z) \hat{n}_j(z)}{\eta^2(z)} b^2(k, z) P_{mm} \left[ k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z \right] dz. \\ &= \int_{z_{\text{min}}}^{z_{\text{max}}} \frac{H(z) \hat{n}_i(z) \hat{n}_j(z)}{\eta^2(z)} \hat{b}^2(z) P_{cb} \left[ k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z \right] dz. \end{aligned} \quad (5.34)$$

After the replacement of power spectra the new scale independent bias  $\hat{b}$  is approximated as a step-like function that is constant in every bin. We can then treat the value of the bias in each bin as a free nuisance parameter and marginalize over them. We can use the new bias to define a modified scale-independent window function for galaxy clustering

$$\hat{W}_i^G(z) = \hat{b}(z) \hat{n}_i(z) H(z)$$

For the shearing part of the weak lensing window function, we note that we do not need to do any

modifications. The neutrinos still contribute to the lensing power spectrum  $\Psi + \Phi$  even on scales where they do not cluster, thus the lensing power spectrum is not a probe of the CDM+baryon power spectrum but of the total matter power spectrum.

For the IA, the question becomes less straightforward. Firstly the window function of IA has a growth factor in the denominator. It is scale-independent in  $\Lambda$ CDM but now due to massive neutrinos, it becomes scale-dependent, making the whole window function scale-dependent. The scale dependence propagates until equation 5.27, where the Limber approximation replaces  $k$  by the fraction  $(\ell + 1/2)/r(z)$ .

The next difficulty arises when we remind ourselves of the reason, for which we could write the contribution of IA as an additional term in a window function. We assumed that IA could be modelled as a biased tracer of the underlying density contrast. It is not so clear if the underlying density contrast is the total matter field or the CDM+baryon field. An argument for the latter would be that IA arises when close galaxies interact with each other. The galaxies themselves are biased tracers of the CDM+baryon field such that over-densities of that field form clusters that get populated with galaxies later on.

We still believe that the IA should be a probe of the total matter field as it is not the interaction of the halos with each other that aligns them but rather their interaction with their surrounding gravitational potential. As neutrinos do contribute to this field, as seen in the Poisson equation, we stick to the total matter field as the underlying overdensity. This would of course need to be checked with simulations.

The last term that is not clear is the crosscorrelation angular power spectrum of lensing and galaxy clustering. To write the expression for its observable we need to go back to equation 5.25. When we did our derivation, in the next step we used the expectation value of the product of the density contrasts in the power spectrum. But now there is a subtlety that we alluded to earlier in the chapter for the spectroscopic probe. For the cross-correlation we find

$$\begin{aligned} (C_\ell^{\text{GL}})_{ij} &= \int W_i^G(\eta) W_j^L(\eta') Y_{\ell m}(\Omega) Y_{\ell m}^*(\Omega') \langle \delta(\eta \Omega) \delta(\eta' \Omega') \rangle d^2\Omega d^2\Omega' d\eta d\eta' \\ &= \int \hat{W}_i^G(\eta) W_j^L(\eta') Y_{\ell m}(\Omega) Y_{\ell m}^*(\Omega') \langle \delta_{cb}(\eta \Omega) \delta(\eta' \Omega') \rangle d^2\Omega d^2\Omega' d\eta d\eta'. \end{aligned} \quad (5.35)$$

This means that the power spectrum that enters the cross-correlation is not the matter power spectrum nor the power spectrum of CDM+baryons. To resolve this problem we use the approximation

$$\langle \delta_{cb}(\mathbf{k}) \delta(\mathbf{k}') \rangle \approx \sqrt{P_{cb}(\mathbf{k}) P_{mm}(\mathbf{k})} \delta(\mathbf{k} - \mathbf{k}'). \quad (5.36)$$

This approximation would be correct if the power spectra were linear and had scale-independent growth. The latter is approximately true as the scale-dependent growth induced from neutrinos is very small at the level of 0.4% [Euclid Collaboration: Blanchard et al., 2020]. For the requirement of linearity, we argue that the neutrinos never become nonlinear as they are still too hot to cluster. On scales where the matter power spectra become nonlinear the neutrino perturbations have already decayed so much that the difference between the CDM+baryon power spectrum and the total matter power spectrum is just a constant factor. This means that this approximation only starts to break down at intermediate scales around the BAO. We can use the approximation to find

$$(C_\ell^{\text{GL}})_{ij} = \int_{z^{\min}}^{z^{\max}} \frac{\hat{W}_i^G(z) W_j^L\left(k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z\right)}{\eta^2(z) H(z)} \sqrt{P_{cb} P_{mm}}\left(k = \frac{\ell + \frac{1}{2}}{\eta(z)}, z\right) dz \quad (5.37)$$

Survey Spec		Value
Redshift bins	$N_{\text{bin}}$	10
Minimum redshift	$z_{\min}$	0.001
Maximum redshift	$z_{\max}$	2.5
Redshift bin edges	$z_1^{\max}, \dots, z_5^{\max}$	0.418, 0.560, 0.678, 0.789, 0.9
Redshift bin edges	$z_6^{\max}, \dots, z_9^{\max}$	1.019, 1.155, 1.324, 1.576
Galaxy distribution redshift	$z_0$	0.6363
Redshift scaling	$c_b$	1
Redshift offset	$z_b$	0
photometric error	$\sigma_0$	0.05
Outlier fraction	$f_{\text{out}}$	0.1
Outlier redshift scaling	$c_0$	1
Outlier redshift offset	$z_0$	0.1
Outlier photometric error	$\sigma_0$	0.05
Intrinsic alignment normalization	$C_{\text{IA}}$	0.0134
IA nonlinear slope	$\beta_{\text{IA}}$	2.17
Elliptisity error	$\sigma_{\epsilon}$	0.3
Integrated mean galaxy distribution	$\bar{n}_i$	$3 \text{ arcmin}^{-2}$
Minimum multipole	$\ell_{\min}$	10
Maximum multipole Lensing	$\ell_{\max}^{\text{WL}}$	3000/5000
Maximum multipole Clustering	$\ell_{\max}^{\text{GCph}}$	750/1500
Sky coverage	$f_{\text{sky}}$	0.3636

Table 5.1: Survey specifications needed to calculate the Photometric likelihood. For the Maximum multipole  $\ell_{\max}$  we noted two values that represent pessimistic and optimistic settings respectively.

The next difficulty arises when we need to predict the nonlinear CDM+baryon power spectrum. The HMCODE fit was done to predict the total matter power spectrum of cosmologies with massive neutrinos. To resolve this we remind ourselves that the power spectra are related to each other, i.e.

$$P_{mm} = f_{cb}^2 P_{cb} + 2 f_{cb} f_{\nu} P_{cb \times \nu} + f_{\nu}^2 P_{\nu}, \quad (5.38)$$

where we encounter the total matter fractions  $f_{cb}$  and  $f_{\nu}$ , which are not to be confused with the growth rates of these perturbations. The power spectrum  $P_{cb \times \nu}$  is the cross-correlation power spectrum of neutrinos and CDM+baryons while the other two are the auto-correlation spectra that we know. Next, we note that the power spectrum of neutrinos is always linear thus we can replace it in the equation with the linear version. For the cross-correlation power spectrum, we again use our approximation that the neutrino perturbations have already decayed so much on the nonlinear scales that we can approximate the cross-correlation power spectrum with its linear counterpart. Since we can predict the nonlinear total matter power spectrum with our halo model we can solve the equation for the nonlinear CDM+baryon power spectrum and find

$$P_{cb}(k, z) \approx \frac{1}{f_{cb}^2} [P_{mm}(k, z) - 2 f_{cb} f_{\nu} P_{cb \times \nu}^{\text{lin}} - f_{\nu}^2 P_{\nu}^{\text{lin}}]. \quad (5.39)$$

In tables 5.1 are the specifications and fiducial values needed to compute the likelihood.

Nuisance Param	Fiducial Value	Nuisance Param	Fiducial Value
$b_1$	1.10	$b_7$	1.44
$b_2$	1.22	$b_8$	1.50
$b_3$	1.27	$b_9$	1.57
$b_4$	1.32	$b_{10}$	1.74
$b_5$	1.36	$\mathcal{A}_{\text{IA}}$	1.72
$b_6$	1.40	$\eta_{\text{IA}}$	-0.41

Table 5.2: Fiducal values of nuisance parameters needed to calculate the Spectroscopic likelihood.

# Chapter 6

## Methodology

The main goal of a forecast is to figure out how well some experiment will be able to measure some set of parameters. In our case, the experiment will be the different probes of the *Euclid* mission while the parameters that we want to measure are cosmological parameters. For this, we generate some mock data that the experiment could have measured given an underlying fiducial model of the universe, free parameters that show up in our modelling of the data, and different survey specifications.

The two main methods of forecasting we will deploy in this work are the Markov chain Monte Carlo (MCMC) method and the Fisher Information (FI) method. They are both based in the framework of Bayesian statistics as they are methods of obtaining the posterior distributions, i.e. the probability,  $p(\boldsymbol{\theta}|\mathbf{D})$ , of observing the model parameters,  $\boldsymbol{\theta}$ , given the data,  $\mathbf{D}$ . The posterior is really what we are after, from it we can extract the mean and the measurement uncertainties of the model parameters. We can group all of our model parameters into 3 groups, cosmological parameters, nuisance parameters, and fixed parameters. While cosmological parameters are what we are after the nuisance parameters show up during modelling and represent uncertainties in the modelling of the observed data or the observation process itself. In the end, they are not of any interest to us, so we often only look at the marginal posterior. That is the posterior where the different values of nuisance parameters have been integrated over, i.e.

$$p(\boldsymbol{\theta}_c|\mathbf{D}) = \int p(\boldsymbol{\theta}|\mathbf{D}) d\boldsymbol{\theta}_n \quad (6.1)$$

In this work, we employ Bayesian statistics, where the posterior is related to the likelihood,  $\mathcal{L}$ , i.e. the probability of measuring the data given the parameters. For this we employ the central theorem of Bayesian statistics called the Bayes' theorem.

$$p(\boldsymbol{\theta}|\mathbf{D}) = \frac{p_{\text{prior}}(\boldsymbol{\theta})}{E_{\text{vidence}}(\mathbf{D})} \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) \quad (6.2)$$

The prior encompasses all of our previous knowledge. It could be either an experimental prior if there exists already a measurement of the model parameter, or a theoretical prior. A theoretical prior tells us that the theory requires parameters to be within some bounds, this could be for example the requirement that the neutrino mass is higher than zero or that there are more than zero neutrino species.

The evidence can be understood as the normalization of the likelihood function. It is not accessible with MCMC methods or FI methods, so we will not go into further details of its interpretation. We just want to note that it is used when comparing different models to explain the same data. If we obtain the evidence its value represents the goodness of the fit while also punishing adding additional parameters.

In the next sections, we will very briefly go over the two different forecasting methods.

## 6.1 The Markov Chain Monte Carlo Method

If we wanted to get the posterior from the likelihood we could just use equation 6.1. In reality, this is not feasible as most cosmological problems do not have analytical solutions. Even the most simple problems like comoving distances have no analytical expression once you add dark energy into the mix. This is why we resort to numerical methods like the MCMC method. In this method the goal is to obtain a sequence of random variables, we call them  $\{\boldsymbol{\theta}_i | i = 1, \dots, N\}$ . In our case, these variables will be a vector of our model parameters, but the idea is much more general. When this sequence or chain is marcovian it can be shown that the distributions of the random variables converge to the distributions of the parameters, i.e. the posterior. Marcovian means that each element of the chain has a probability that is only a function of the last element of the chain.

There are different algorithms to generate such a chain but the simplest one and the one that we use in this work is the Metropolis-Hastings algorithm. The Metropolis-Hastings algorithm starts with a random point in parameter space that is our first chain element  $\boldsymbol{\theta}_1$ . It then proposes a new point  $\boldsymbol{\theta}'$ . For the proposal of the new point, we typically use some multivariate Gaussian that was our initial guess of the distributions of the parameters  $\boldsymbol{\theta}$ . We denote the distribution that we sample the point from as  $q(\boldsymbol{\theta}'|\boldsymbol{\theta}_1)$ .

This new point is then accepted with an acceptance probability

$$A(\boldsymbol{\theta}_1 \rightarrow \boldsymbol{\theta}') = \min \left( 1, \frac{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}')}{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}_1)} \frac{q(\boldsymbol{\theta}_1|\boldsymbol{\theta}')}{q(\boldsymbol{\theta}'|\boldsymbol{\theta}_1)} \right). \quad (6.3)$$

The next element of the chain becomes then either  $\boldsymbol{\theta}'$  or stays  $\boldsymbol{\theta}_1$  with a probability given by the acceptance probability. This procedure is then iterated until the element of the chain have converged to follow a static distribution.

When the distribution of points,  $\pi(\boldsymbol{\theta})$ , has become static, the probabilities of going from one point to another need to fulfil the condition of reversibility. This means that the points fulfil

$$\pi(\boldsymbol{\theta}') P(\boldsymbol{\theta}' \rightarrow \boldsymbol{\theta}) = \pi(\boldsymbol{\theta}) P(\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}'), \quad (6.4)$$

where  $P$  denotes the transition probability of going from one point to the other. This condition is called detailed balance and is an expression of the condition that equilibrium states should not change under time reversal.

If we use our acceptance probability and the suggestion probability to calculate the transition probability we find

$$\pi(\boldsymbol{\theta}') q(\boldsymbol{\theta}|\boldsymbol{\theta}') A(\boldsymbol{\theta}' \rightarrow \boldsymbol{\theta}) = \pi(\boldsymbol{\theta}) q(\boldsymbol{\theta}'|\boldsymbol{\theta}) A(\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}'). \quad (6.5)$$

Next we realize that either  $A(\boldsymbol{\theta}' \rightarrow \boldsymbol{\theta}) = 1$  or  $A(\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}') = 1$  as the fraction on the right side of equation 6.3 is reciprocal under the exchange of the states. Without loss of generality we chose that

$A(\boldsymbol{\theta}' \rightarrow \boldsymbol{\theta}) = 1$  this leads to

$$\begin{aligned} \pi(\boldsymbol{\theta}') q(\boldsymbol{\theta}|\boldsymbol{\theta}') &= \pi(\boldsymbol{\theta}) q(\boldsymbol{\theta}'|\boldsymbol{\theta}) \frac{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}')}{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta})} \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{q(\boldsymbol{\theta}'|\boldsymbol{\theta})} \\ \frac{\pi(\boldsymbol{\theta}')}{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}')} &= \frac{\pi(\boldsymbol{\theta})}{\mathcal{L}(\mathbf{D}|\boldsymbol{\theta})}. \end{aligned} \quad (6.6)$$

This means that the static distribution, that the points tend to converge to, is the likelihood up to some constant factor. Inserting this into equation 6.1 gives us our posterior up to some factor. This is precisely why, when doing the Metropolis-Hastings algorithm, there is no way of obtaining the evidence.

Some detail needs to be discussed about the good convergence of Markovian chains. Firstly, to define what we mean by a converged chain. Different metrics are being used but in our case, we use the Gelman-Rubin test. In this test, we launch multiple chains from different starting points, and they all converge to the final distribution of the likelihood. Then we calculate two quantities. The first quantity is the variance of the mean of the chains, i.e. how far apart the different chain centres are from one another. The second quantity is the mean of the variance of each chain, i.e. how much the chains are varying from their centres. For a converged chain, we would expect that the ratio of these two quantities is close to one, so we use this ratio to measure chain convergence. Often denoted as the R-1 test we require a chain to fulfill

$$R - 1 := \sqrt{\frac{\text{Var}(\text{chain means})}{\text{mean}(\text{chain variances})}} - 1 \stackrel{!}{<} 0.1 \quad (6.7)$$

to be considered converged. This requires our distribution  $q$  from which we propose points should be already quite close to target distribution  $\mathcal{L}$ . The reasoning is the following: if the suggested points would be quite far from the bulk of the likelihood then they would never be accepted and the chain could not move very much. In the end, the individual chains would have very low variances. The opposite effect happens when the proposal distribution would only propose points in the very centre of the likelihood. Then the chains would accept every suggested point. This would lead to the variance of the chain average to be high. The random walk would not converge to the centre of the likelihood. Secondly, this would also have the effect that there could be regions of the likelihood that would never be explored by the chains.

To help with the convergence during each run when the chains have had enough time to find the maximum likelihood and have reached some milestones of convergence, we update the proposal distribution. For this, we calculate the covariance of the chains and use a multivariate Gaussian with the sample covariance to propose points out of. This has the effect that the main requirement of a Markovian chain stops being valid. Now the probability of going from one point to the next depends not only on the last point but also on previous points. This means that if we want to analyze the chains we need to throw away the points of the chains before the last update of the proposal distribution.

The code that we use to run our MCMC is called MontePython (MP) and has two modes that are of interest to us. The first mode is the Metropolis-Hastings mode, it uses the same strategy that was discussed earlier to generate the chains where we also update the proposal distribution to have fast convergence. The second mode is the Fisher mode which is discussed in the next section. The reason why we use both modes of MP will become clearer in the chapter 7 about Validation.

## 6.2 The Fisher Information Method

We observe that for well-constrained parameters the posterior looks often like a multivariate Gaussian. The FI method employs this in that it is an approximation where the Gaussian is tailored to the likelihood and then the error is extracted from the Gaussian. The idea of the formalism is the following: To parametrize a multivariate Gaussian you only need the covariance matrix,  $C$ , and the mean of the variables,  $\mu$ . In our case, the variables would be the cosmological parameters and the covariance then would contain the measurement uncertainties.

Assuming that the posterior can be approximated as a Gaussian we can write it as

$$p(\boldsymbol{\theta}|\mathbf{D}) \propto p_{\text{prior}}(\boldsymbol{\theta}) \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}') \approx \mathcal{N} \exp \left[ -\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\mu}) \cdot C^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right]. \quad (6.8)$$

For a Gaussian posterior, the mean is also at the best fit, this is why when we do a forecast we directly know the value of  $\mu$ . It is the value of the model parameters at the fiducial point where we generated our mock data. We can now use that the covariance can be calculated from the logarithm of the posterior by calculating the second derivative.

$$-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(p(\boldsymbol{\theta}|\mathbf{D})) \Big|_{\boldsymbol{\theta}=\boldsymbol{\mu}} \approx (C^{-1})_{ij} \quad (6.9)$$

If we assume flat priors in the region of interest we can directly calculate the derivative of the likelihood. This leads to the definition of the Fisher Information matrix as

$$F_{ij} := -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}). \quad (6.10)$$

For Gaussian priors, we are able to add the term later. We will discuss further bellow how to do this when we discuss adding different experiments. Non-Gaussian or flat priors are harder to approximate. The easiest method of dealing with exponential priors is to make it flat prior by transforming the relevant parameters into log space. This might make the contours less Gaussian depending on how well-constrained the parameters are.

The difficulty of the FI method is that often it is not so clear how to obtain the second-order derivatives of the likelihood. Some codes were purposefully built to be differentiable i.e. that you can obtain the derivates for free. The typical Einstein-Boltzmann solvers (EBS) `CLASS`[Blas et al., 2011] and `CAMB`[Lewis et al., 2000] do not have that functionality, so we need to obtain the derivatives numerically. This is what `MP` does in Fisher mode. We calculate the second-order derivates using a double-sided finite differences scheme. We note numerical derivatives with the symbol D, thus we can write

$$\begin{aligned} \frac{D^2 \log \mathcal{L}}{D \theta_i D \theta_j} &= \frac{1}{4 h_i h_j} [\log \mathcal{L}(\theta_i + h_i, \theta_j + h_j) - \log \mathcal{L}(\theta_i + h_i, \theta_j - h_j) \\ &\quad - \log \mathcal{L}(\theta_i - h_i, \theta_j + h_j) + \log \mathcal{L}(\theta_i - h_i, \theta_j - h_j)] \quad \text{for } i \neq j \end{aligned} \quad (6.11)$$

$$\frac{D^2 \log \mathcal{L}}{D^2 \theta_i} = \frac{1}{h_i^2} [\log \mathcal{L}(\theta_i + h_i) + \log \mathcal{L}(\theta_i - h_i) - 2 \log \mathcal{L}(\theta_i)], \quad (6.12)$$

where we have left out the arguments of the log-likelihood that do not vary in each line. Here, the advantage of using the FI method becomes clear. While the MCMC often needs multiple thousand

calls of the EBS the FI method needs much fewer. In the implementation we use for  $N$  parameters we have to call the likelihood  $2N$  times for the diagonal elements and  $2N(N - 1)$  times for the off-diagonal elements, making our total calls  $2N^2$ . Even then we can shave off some more time if we use the fact that when varying nuisance parameters there is no need for calling the EBS anymore. The calculation of the second derivative is numerically quite sensitive to the choice of stepsize. It was shown before in [Euclid Collaboration: Blanchard et al., 2023] that the optimal stepsize to calculate the derivatives is of the order of a few per cent of the marginalized error. And even then one needs to use very high precision settings of the EBS to be able to have stable results. This does defeat the purpose of the FI method as to obtain the marginalized error we would need to run an MCMC.

To make the computation of the FI matrix numerically stable we can use the explicit form of the photometric and spectroscopic likelihoods to write it using only the first derivates of the observables. For the photometric likelihood we find

$$F_{ij} = \frac{1}{2} \sum_{\ell=\ell_{\min}}^{\ell_{\max}} (2\ell + 1) f_{\text{sky}} \text{Tr} \left[ (C_{\ell}^{\text{fid}})^{-1} (\partial_i C_{\ell}^{\text{th}}|_{\text{fid}}) (C_{\ell}^{\text{fid}})^{-1} (\partial_j C_{\ell}^{\text{th}}|_{\text{fid}}) \right]. \quad (6.13)$$

We have replaced the observed angular power spectrum with the theoretical angular power spectrum calculated at the fiducial like in the spectroscopic probe. There are two advantages of calculating the FI matrix like this, firstly the derivatives that show up are first order and thus much easier to calculate. To calculate them we use either of two finite differences stencils

$$\frac{DC_{\ell}^{\text{th}}}{D\theta_i} = \frac{1}{2h_i} [C_{\ell}^{\text{th}}(\theta_i + h_i) - C_{\ell}^{\text{th}}(\theta_i - h_i)] \quad (6.14)$$

$$\text{or } = \frac{1}{6h_i} [-11C_{\ell}^{\text{th}}(\theta_i) + 18C_{\ell}^{\text{th}}(\theta_i + h_i) - 9C_{\ell}^{\text{th}}(\theta_i + 2h_i) + 2C_{\ell}^{\text{th}}(\theta_i + 3h_i)]. \quad (6.15)$$

By default, we will stick to the double-sided finite difference and not to the one-sided one but to check the numerical stability of our derivatives we have tested both methods. First-order derivatives have the property to be more numerically stable. That is why we do not need to tune our stepsize as much as in the case of the second-order derivatives. We always use a relative stepsize of 1% for the cosmological parameters except for the neutrino mass, where we vary it by a relative 10%. The choice for the neutrinos stems from the fact that the neutrino mass only has a small effect on our observables. For a 1% change of the neutrino mass, the change of the observables would be dominated by numerical noise, so we had to choose a higher stepsize. We also test that the choice of a 10% stepsize for massive neutrinos produces stable results.

The second advantage of this method is that to compute all derivatives one only needs to call the code  $2N$  times since the first-order derivates of the observables can then be multiplied to find the second-order derivates of the likelihood.

For the spectroscopic likelihood, we find a similar expression for the likelihood.

$$F_{ij} = \frac{1}{8\pi^2} \sum_i \int k_{\text{fid}}^2 \partial_i [\log P_{\text{th}}(k_{\text{fid}}, \mu_{\text{fid}}, z_i)|_{\text{fid}}] \partial_i [\log P_{\text{th}}(k_{\text{fid}}, \mu_{\text{fid}}, z_i)|_{\text{fid}}] V_i^{\text{fid}} dk_{\text{fid}} d\mu_{\text{fid}} \quad (6.16)$$

Again we find that the FI matrix can be written with products of first-order derivatives. To calculate the FI matrix for our forecast we will use the code CosmicFish (CF)[Raveri et al., 2016] we have modified such that our changes to the recipes of the observables are consistent with our

MP implementation. Since CF never calculates the actual likelihood this gives us a great test for our likelihood implementation if the FI matrix obtained by MP in Fisher mode agrees with the FI matrix from CF. This is further discussed in the chapter 7 about the Validation of our pipeline. Finally, the FI formalism gives us a guiding principle to make educated guesses for the proposal distributions. Firstly when we already have a covariance matrix of a previous MCMC and we want to launch another MCMC. This one should only be a submodel of the old MCMC. The naive approach of deleting the covariance matrix's corresponding rows is not optimal but was the implementation in MP. We have added a new method of fixing that leads to a better initial guess. We can calculate the covariance of the MCMC by averaging to the chains. This covariance matrix approximates the inverse of the FI matrix. The matrix now has all the parameters that the actual Fisher would have. If we remind ourselves of the definition of the FI matrix as the second derivate of the log-likelihood then fixing a parameter is the same as never adding the corresponding lines to the matrix. Inversion then leads us to the desired covariance matrix. We illustrate this procedure below.

$$\begin{array}{cc} \boldsymbol{\theta}_o & \boldsymbol{\theta}_f \\ \boldsymbol{\theta}_o \begin{pmatrix} C_{oo} & C_{of} \\ C_{fo} & C_{ff} \end{pmatrix} & \xrightarrow{\text{invert}} \begin{array}{cc} \boldsymbol{\theta}_o & \boldsymbol{\theta}_f \\ \boldsymbol{\theta}_f & \end{array} \begin{pmatrix} F_{oo} & F_{of} \\ F_{fo} & F_{ff} \end{pmatrix} \xrightarrow{\text{extract } F_{oo}} \xrightarrow{\text{invert}} \tilde{C}_{oo} \end{array}$$

We have denoted the parameters we want to keep free with  $\boldsymbol{\theta}_o$  and the parameters we want to fix with  $\boldsymbol{\theta}_f$ . In the Fisher approximation the new covariance  $\tilde{C}_{oo}$  is the covariance of the MCMC if the parameters had been fixed from the beginning. This also gives us an interpretation of what the diagonal elements of the Fisher mean. If we were to fix all parameters except one then the matrix  $F_{oo}$  is just a scalar. The inversion is then just its reciprocal value but the interpretation of this number would be the variance if every other parameter was fixed. This is precisely what the unmarginalized error is thus the diagonal of a Fisher contains the reciprocal unmarginalized errors. Next, we can ask ourselves the question of how to add multiple experiments together. On the level of the likelihood when neglecting cross-correlation the likelihood of a combined probe,  $\mathcal{L}$ , factorizes into the likelihoods of the individual probes,  $\mathcal{L}^1$  and  $\mathcal{L}^2$ . For the FI, we can write

$$F_{ij} = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log \mathcal{L} = \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\log \mathcal{L}^1 + \log \mathcal{L}^2] = F_{ij}^1 + F_{ij}^2. \quad (6.17)$$

This means that combining two uncorrelated probes is the same as adding their Fisher matrix. We can use this to make an educated guess on the proposal distribution of an MCMC with combined probes, like the spectroscopic and the photometric probes when there are previous runs. Firstly we have to separate all of our parameters into parameters that are common to both likelihoods  $\boldsymbol{\theta}_c$ , i.e. the cosmological parameters, and then the nuisance parameters of both likelihoods  $\boldsymbol{\theta}_1$  and  $\boldsymbol{\theta}_2$ . Calculating the derivate of one likelihood with respect to the nuisance parameters of the other likelihood makes little sense but could be done mathematically. The derivatives would just vanish so the Fisher Matrix would have rows of zeros in these and would be not invertible. But this helps us in the formulation of our recipe.

We start by inverting our MCMC covariance matrices to have the approximate FI matrices. We then add columns and rows of zeros that correspond to the nuisance parameters of the other likelihood. Finally, we can add the matrices and invert the result to get our new guess for the proposal covariance. Typically, this works very well for complimentary probes since breaking degeneracies makes the final posterior Gaussian enough for the approximation. We illustrate the final combination

recipe below

$$\begin{array}{c}
 \theta_c \begin{pmatrix} \theta_c & \theta_1 \\ C_{cc}^1 & C_{c1}^1 \\ \theta_1 & C_{1d}^1 \\ C_{1d}^1 & C_{11}^1 \end{pmatrix}, \quad \theta_c \begin{pmatrix} \theta_c & \theta_2 \\ C_{cc}^2 & C_{c2}^2 \\ \theta_2 & C_{2c}^2 \\ C_{2c}^2 & C_{22}^2 \end{pmatrix} \xrightarrow{\text{invert}} \theta_c \begin{pmatrix} \theta_c & \theta_1 \\ F_{cc}^1 & F_{c1}^1 \\ \theta_1 & F_{1d}^1 \\ F_{1d}^1 & F_{11}^1 \end{pmatrix}, \quad \theta_c \begin{pmatrix} \theta_c & \theta_2 \\ F_{cc}^2 & F_{c2}^2 \\ \theta_2 & F_{2c}^2 \\ F_{2c}^2 & F_{22}^2 \end{pmatrix} \\
 \\
 \xrightarrow{\substack{\text{add zeros} \\ \hookrightarrow}} \theta_1 \begin{pmatrix} \theta_c & \theta_1 & \theta_2 \\ F_{cc}^1 & F_{c1}^1 & 0 \\ F_{1d}^1 & F_{11}^1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \theta_1 \begin{pmatrix} \theta_c & \theta_1 & \theta_2 \\ F_{cc}^2 & 0 & F_{c2}^2 \\ 0 & 0 & 0 \\ F_{2d}^2 & 0 & F_{22}^2 \end{pmatrix} \xrightarrow{\text{add}} \theta_1 \begin{pmatrix} \theta_c & \theta_1 & \theta_2 \\ F_{cc}^1 + F_{cc}^2 & F_{c1}^1 & F_{c2}^2 \\ F_{1c}^1 & F_{11}^1 & 0 \\ F_{2d}^2 & 0 & F_{22}^2 \end{pmatrix} \\
 \\
 \xrightarrow{\text{invert}} \tilde{C}
 \end{array}$$

This also explains how to add a Gaussian prior to a FI matrix afterwards when looking at the Bayes theorem 6.1. The prior can be pulled into the likelihood and act like an effective second experiment. The covariance matrix of the multivariate Gaussian prior directly defines an FI matrix or rather it's inverse. We can add them like we did for experiments afterwards. This of course needs the prior to be centred on the fiducial cosmology of the forecast.

A script to merge covariance matrices for MP was added by us. To show how well this works we have combined two test MCMCs of the two main *Euclid* probes and present the results in figure 6.1. We generated this plot with the use of GetDist <sup>1</sup> [Lewis, 2019].

---

<sup>1</sup><https://getdist.readthedocs.io>

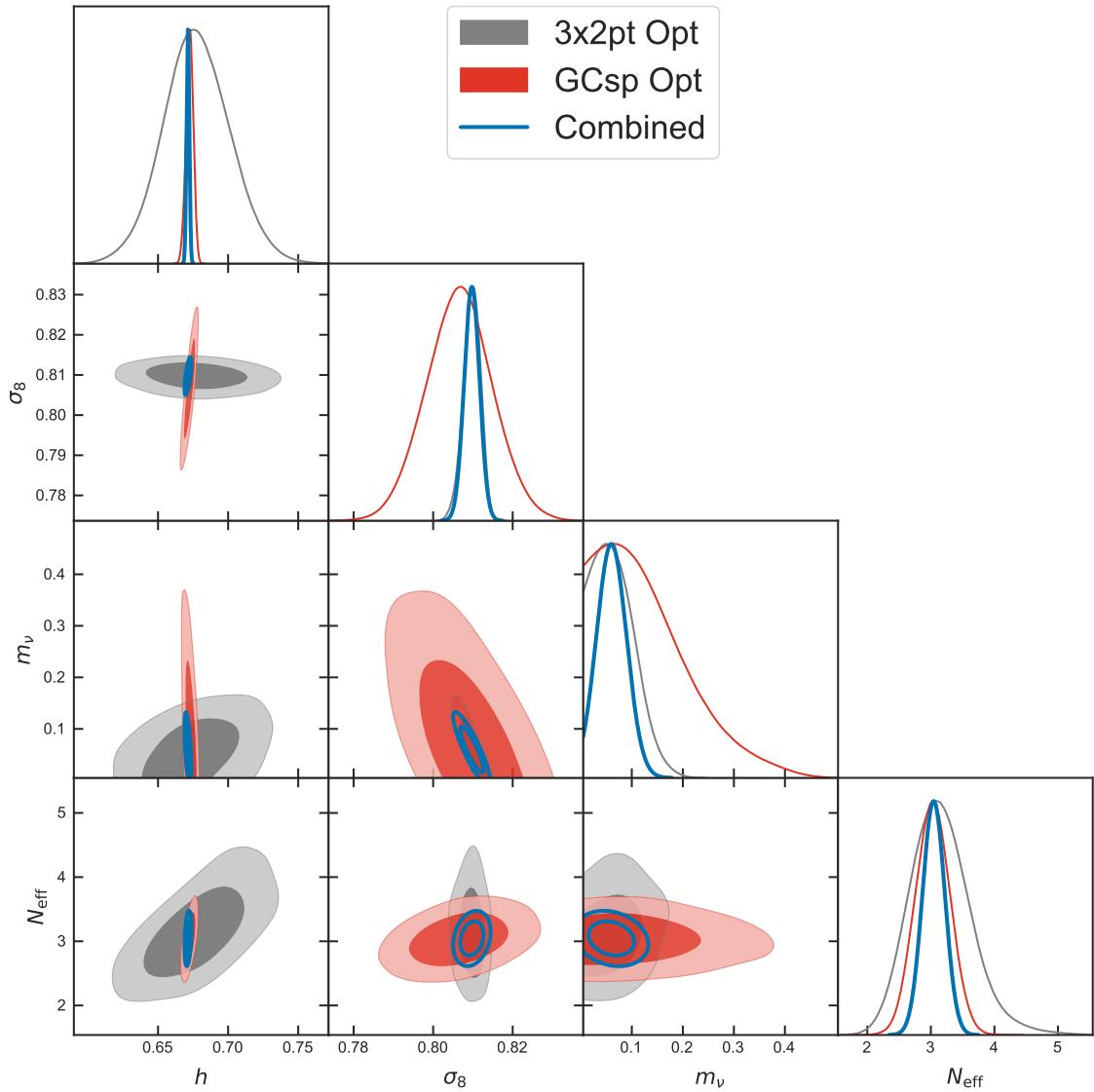


Figure 6.1: Demonstration of the covariance combination recipe. The MCMCs are done with the settings of the Validation section. The combined probe is not an MCMC, but the ellipses from the Gaussian approximation obtained with the recipe. This also demonstrates very well the complementarity of the *Euclid* probes. The terms 3x2pt and GCsp are from [Euclid Collaboration: Blanchard et al., 2020] and describe the photometric and spectroscopic probes respectively.

## Chapter 7

# Validation of the Forecasting Pipeline

Our main goal is a validated and robust sensitivity forecast for the neutrino parameters of *Euclid*. For this, we have adopted a validation pipeline from [Euclid Collaboration: Blanchard et al., 2023]. It starts with the code CF that was validated in the *Euclid* project [Euclid Collaboration: Blanchard et al., 2020]. In this work, we have modified the recopies of the nonlinear modelling and modelling of the neutrino-induced scale-dependent growth to better describe the effect of massive neutrinos. Thus, we decided to do a re-validation of the MP pipeline. We will also check how well the Fisher formalism works for the neutrino mass that is close to a theoretical prior at zero mass.

In this section, we choose a different model that we will use in the final forecasting results to have posteriors that are Gaussian enough to be able to be approximated using the FI methods. The parameters are given in table 7.1.

The extensions to the  $\Lambda$ CDM models in this section are varied in pairs of two. We first look at a model where we vary  $w_0 + w_a$ . These parameters are an approximation to a dark energy equation of state that is slowly varying. It is parametrized as having  $w(a) = w_0 + w_a(1 - a)$ . We note that when fixing the parameters to their fiducial values, we recover  $\Lambda$ CDM. This  $w_0 w_a$ CDM model was the original model that was validated in the work of We kept it in our validation pipeline to be sure that we did not break the previous validation with our changes to the pipeline. The second model that we look at is a model where we vary the constant pressure-to-density ratio of dark energy  $w_0$  (fixing  $w_a$ ) and the approximated neutrino mass  $m_\nu$ . We will use the symbol  $m_\nu$  in this section rather than  $\sum m_\nu$  to denote this quantity as in our parametrization we approximate that all the neutrino mass is concentrated into one massive neutrino with a temperature of

$$\frac{T_\nu}{T_\gamma} = \left(\frac{4}{11}\right)^{1/3} \left(\frac{3.044}{3}\right)^{1/4} = 0.716369. \quad (7.1)$$

This is different from the parametrization of [Euclid Collaboration: Blanchard et al., 2023]. In their parametrisation, the neutrino temperature had a factor of  $N_{\text{eff}}$  in their second factor instead of its fiducial value. In that sense their parameter  $N_{\text{eff}}$  is different from ours as it was partly accounting for additional  $N_{\text{eff}}$  to a higher neutrino temperature. For us,  $N_{\text{eff}}$  is accounting for additional massless degrees of freedom and thus a higher value of  $N_{\text{eff}}$  does not also mean a higher energy density of massive neutrinos.

We also note that this approximated mass of massive neutrinos is also not exactly the mass of a massive neutrino because it is the mass of a neutrino in an instantaneous decoupling approximation.

Table 7.1: Fiducial values of the parameters varied in the validation section. The  $\Lambda$ CDM parameters were always varied during the Validation runs while the extended Models were only varied two at a time. The meaning parameters are described in the text.

Fiducial								
$\Lambda$ CDM					Extensions			
$\Omega_{m,0}$	$100 \times \Omega_{b,0}$	$h$	$n_s$	$\sigma_8$	$m_\nu$ (meV)	$N_{\text{eff}}$	$w_0$	$w_a$
0.314571	4.92	0.6737	0.9661	0.81	60	3.044	0	-1

It corresponds to a neutrino density fraction,  $\Omega_\nu$ , of

$$\Omega_\nu h^2 = \frac{m_\nu}{94.07 \text{ eV}} \left( \frac{3.044}{3} \right)^{3/4} = \frac{m_\nu}{0.06 \text{ eV}} 6.44826 \cdot 10^{-4} \quad (7.2)$$

In reality the factor of 94.07 eV in the denominator of the first factor would be given through an integral over the neutrino phase space distribution. To better match the parametrization between our two EBS `CLASS` and `CAMB`, we stick to the instantaneous decoupling approximation done in `CAMB`. We do this by internally converting the input neutrino mass in `CLASS` to  $\Omega_\nu$  with the approximate equation. The change to the neutrino temperature propagates to this equation, as the second factor is proportional to the neutrino temperature.

In the third model, we fix the dark energy equation of state to its  $\Lambda$ CDM value and vary the approximated neutrino mass and the effective relativistic degrees of freedom. As we stated before our changes to the neutrino parametrization make the value of  $N_{\text{eff}}$  correspond to the actual numbers of massless species  $N_{\text{ur}}$ . We can translate these quantities in this case to another by

$$N_{\text{ur}} = N_{\text{eff}} - \frac{3.044}{3}, \quad (7.3)$$

where the second term stands for the one massive neutrino. In the actual forecasting results, we will change our neutrino parametrization to have three massive neutrinos with degenerate masses. In this case, we will subtract three times that from the  $N_{\text{eff}}$ , such that the parameter will correspond to any additional species of massless relics particles. We will denote this quantity as  $\Delta N_{\text{eff}}$  with a theoretical prior bound at zero.

In all the validation models we have fixed the modelling parameters of nonlinear effects in the spectroscopic probe, namely  $\sigma_p$  and  $\sigma_v$ , to their fiducial values. This is done because especially  $\sigma_p$  is very degenerate with the cosmological parameter, breaking the Gaussian approximation needed for the FI method.

We will validate our pipeline separately for both main probes of the *Euclid* mission, for both the optimistic and pessimistic settings and all three models totalling twelve total tests. The tests itself are done in three steps

- We will first validate our EBS implementation by validating the FI results of `CF` for the two codes `CAMB` and `CLASS`.
- We will then validate our likelihood implementation by validating those two results with the FI obtained by `MP` in Fisher mode.

- We will test for the validity of the Gaussian approximation by comparing the obtained FI matrices with the posteriors obtained by an MCMC run in `MP` with the Metropolis-Hastings algorithm.

In the last step, by construction, there is no possibility of a theoretical error showing up as the same likelihood is used to generate the MCMC to calculate the FI matrix. We decide that our implementation is validated when the marginalized and unmarginalized one-dimensional errors are within 10% of the median. This validation criterion is taken from the previous *Euclid* validation standard from [Euclid Collaboration: Blanchard et al., 2023].

This chapter will be structured in the following way. Firstly, we discuss the precision settings for the two Einstein-Boltzmann solvers(EBS) `CAMB` and `CLASS` that we need to set to get their fiducial power spectra to match to sub-percent level. We will then show the agreement of the three different FI matrices and discuss their stability under change of stepsizes and derivation methods. In the next step, we will then compare the FI matrices with the MCMC results and try to explain the derivations. In the last section, we discuss the effect of nonlinear modelling and analyze possible biasing.

## 7.1 The Einstein-Boltzmann Solvers

To match the results from the two EBS we needed to first understand the differences between codes. The first difference was already alluded to in the introduction to this section. `CAMB` uses the phase space integral for the neutrinos in the limit of instantaneous decoupling while `CLASS` uses a slightly different value taking into account relic scattering and neutrino oscillations. To circumvent this problem we added a new parameter `m_nu_camb` that is converted to the neutrino density fraction  $\Omega_\nu$  via equation 7.2, before each call of `CLASS`.

The total matter density parameter  $\Omega_m$  is also defined slightly differently in `CAMB` and `CLASS`. While `CAMB` defines the total matter density as CDM+baryons+massive neutrinos, `CLASS` does not include the massive neutrinos. To match the codes we defined a new parameter `Omega_m_camb` that is converted to the density parameter of cold dark matter,  $\Omega_c$ , before the run by subtracting the other two ingredients.

Another detail in our implementation was that `CAMB` is not able to use  $\sigma_8$  as an input parameter. This is because it is computed from the linear matter power spectrum. In `CLASS`, we have a shooting functionality where we use a reference value of  $A_s$  to calculate  $\sigma_8$ . As  $\sigma_8$  is directly proportional to  $A_s$ , by rescaling  $A_s$  we can match the desired value of  $\sigma_8$ . Since this functionality does not exist in `CAMB`, we mimic the shooting by doing a first reference run of `CAMB` to obtain the matter power spectrum and then rescale to match  $\sigma_8$  inside `CF`. With these three modifications, we were able to match the input parameters between `CLASS` and `CAMB`.

The next step in the validation of the EBS is a careful choice of the accuracy settings. This is done such that

- A: The power spectra are numerically stable enough such that their derivatives are not dominated by noise,
- B: Small differences in the approximation schemes of the EBS do not propagate to the final forecasting results.

The choice of the accuracy settings inside `CAMB` is simple. There are generic parameters that boost

the accuracy of the **CAMB** code in multiple different places. We decided to use the following accuracy settings in **CAMB**.

Listing 7.1: **CAMB**:HP precision settings

```
do_late_rad_truncation = T
high_accuracy_default = T
transfer_interp_matterpower = T
accurate_reionization = F
accuracy_boost = 3
l_accuracy_boost = 3
accurate_massive_neutrino_transfers = T
```

In the rest of the discussion, these settings will be called **CAMB**:HP. The intrinsic accuracy settings in the code, like integration cutoffs, truncation of the Boltzmann hierarchy, and stepsizes, are all multiplied by the corresponding accuracy boost. The parameter `do_late_rad_truncation` truncates the Boltzmann hierarchy for radiation at high  $\ell$  after matter domination. The option for `accurate_massive_neutrino_transfers` allows us to obtain the neutrino transfer functions that are needed in the modelling of our photometric probe.

The high-precision settings that we chose in **CLASS** are summarized in the following.

Listing 7.2: **CLASS**:HP precision settings

```
k_per_decade_for_bao = 50
k_per_decade_for_pk = 50
l_max_g = 20
l_max_pol_g = 15
radiation_streaming_approximation = 2
radiation_streaming_trigger_tau_over_tau_k = 240.
radiation_streaming_trigger_tau_c_over_tau = 100.
tol_ncdm_synchronous = 1.e-5
background_Nloga = 6000
thermo_Nz_log = 20000
thermo_Nz_lin = 40000
tol_perturbations_integration = 1.e-6
halofit_tol_sigma = 1.e-8
l_max_ncdm = 25
ncdm_fluid_trigger_tau_over_tau_k = 100.
```

The approximation schemes of **CLASS** are described in [Blas et al., 2011]. In **CLASS**, we truncate the Boltzmann hierarchy for our ultra-relativistic species, i.e. photons and neutrinos, like in **CAMB**. This cutoff is governed by the precision parameters `l_max_xxx`, where `xxx` stands for the species.

Once deep inside the Hubble horizon we further simplify the equations of radiation and light relics by doing a fluid approximation. A fluid approximation means that we approximate the species as an imperfect fluid with anisotropic stress. This essentially truncates the Boltzmann hierarchy at 2 by exponentially suppressing higher moments. For our massive neutrino, the time of this approximation is governed by the parameter `ncdm_fluid_trigger_tau_over_tau_k`. The fluid approximation is not present in **CAMB**.

To match the output of the two codes we have to deactivate the fluid approximation. This is done

by passing the parameter `ncdm_fluid_approximation=3`. We further found that the effect of the fluid approximation is negligible in the highly nonlinear regime. This is why we decided to use a different set of precision parameters in the validation of the different probes. For the photometric probe, we stick to the `CLASS:HP` settings while for the spectroscopic probe, we have adjusted our setting to the ultra-high precision settings `CLASS:UHP`.

Listing 7.3: `CLASS:UHP` precision settings

```
k_per_decade_for_bao = 50
k_per_decade_for_pk = 50
l_max_g = 20
l_max_pol_g = 15
radiation_streaming_approximation = 2
radiation_streaming_trigger_tau_over_tau_k = 240.
radiation_streaming_trigger_tau_c_over_tau = 100.
tol_ncdm_synchronous = 1.e-5
background_Nloga = 6000
thermo_Nz_log = 20000
thermo_Nz_lin = 40000
tol_perturbations_integration = 1.e-6
halofit_tol_sigma = 1.e-8
l_max_ncdm = 40
ncdm_fluid_appoxmation = 3.
evolver = 0
```

The two precision settings in listings 7.2 and 7.3 differ by three parameters. Firstly, the truncation of the Boltzmann hierarchy for massive neutrinos happens at a higher  $\ell = 40$ . This is done to better match the cutoff of `CAMB`. With our precision settings, the truncation happens at a high  $\ell$  of 75. The second change is the deactivation of the neutrino fluid approximation and the last one is the change of the evolver. The change from our high precision setting to the ultra-high precision settings makes `CLASS` very slow. The standard evolver for a typical call of the code takes the order of tens of minutes. By switching the evolver from the stiff `ndf15` evolver to a standard Runge-Kutta integrator we could cut down the computation time to approximately 20%.

Finally, when running the MCMCs we can set the precision parameters back to the `CLASS` defaults. The changes that the precision parameters make only affect the final power spectrum at the per mille level. As this only propagates to small changes in the final likelihood, the MCMC is unaffected by our precision parameters. The default precision parameters are denoted by us with `CLASS:DP`. We present our Comparison between the different power spectra obtained by our EBS in figure 7.1. For the linear power spectrum, we can see that the fluid approximation leads to an underprediction of 0.05% on the matter power spectrum at intermediate  $k$ . At the smallest scales, it has a much smaller effect such that the `CLASS:HP` and `CLASS:UHP` settings agree very well with each other. What can also be seen, is that in the sensitivity region the power spectra of `CLASS:HP` and `CLASS:DP` only differ by a small bump at large scales, but then start diverging at small scales.

For the nonlinear spectrum we can see a similar discrepancy between the different power spectra on large and intermediate scales, but then start converging again on the smallest scales, where the power spectra are dominated by the one halo term. We believe that the reason they start converging comes from the fact that the effective scalar index enters the smoothing region between one and two halo terms is slightly different. This can be explained by the fact that we observe a stronger

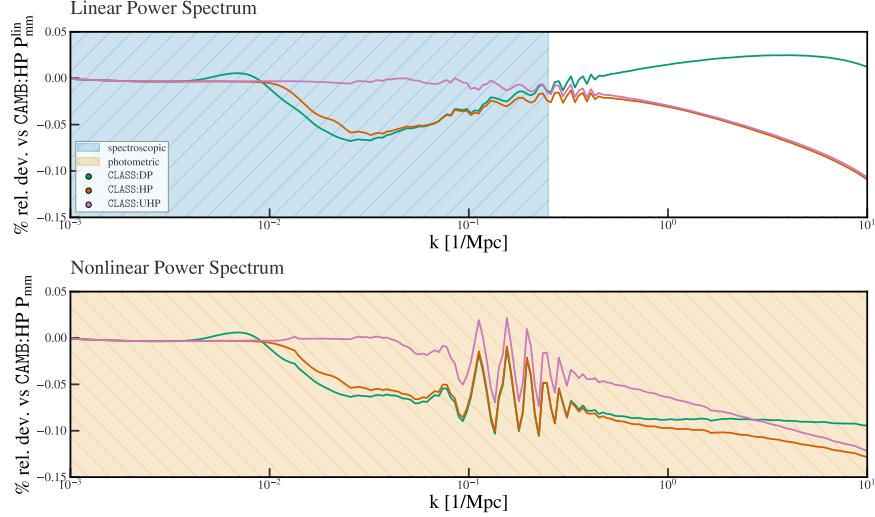


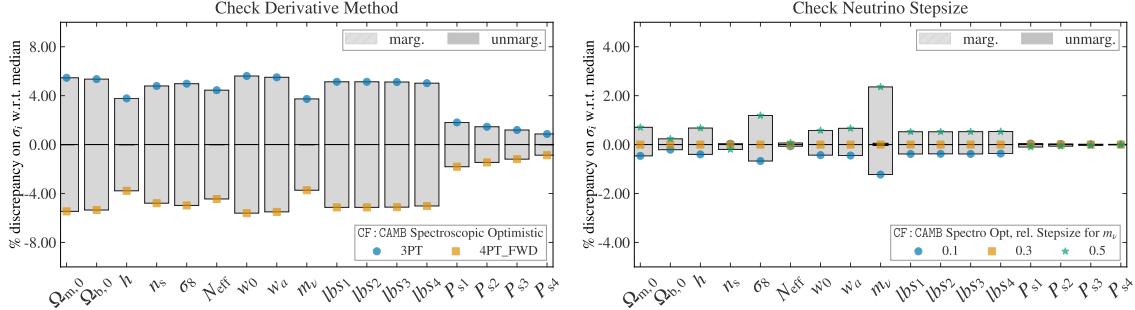
Figure 7.1: Comparison of the different Class precision settings to our 'default' high precision setting for **CAMB**. On the  $y$  axis we plot the difference to the power spectra obtained from **CAMB**, normalized to the mean value. We have marked the regions where the individual probes are sensitive to their corresponding colours.

neutrino-induced suppression on intermediate scales when doing the fluid approximation.

## 7.2 Comparison of the different Fisher Information Methods

After careful setting of the input parameters and precision settings of our two EBS, the next step is to validate our forecasting results. For this, we will first discuss the choices we made in our forecasting pipeline. The first choice we have is for the derivate method. As we have discussed, in CF the first-order derivate of the observable can either be calculated using an equal-sided three-point derivative (3PT), or a four-point forward derivative (4PT\_FWD). They both represent the two numerical derivatives that were introduced last chapter. For this test, we have compared the one-dimensional marginalized and unmarginalized errors obtained via CF. We use **CAMB** as the underlying EBS. We denote the CF pipelines as either CF:**CAMB** or CF:**CLASS** depending on the underlying EBS. The comparison is found in figure 7.2a. The derivative methods validate each other as the total deviation is within 5% of each other for all parameters. We can see a systematic underprediction of the error when using the 4PT\_FWD derivatives instead of the 3PT derivatives. The agreement of the errors for the forecast is within 10% of the mean for the case of varying all 9 cosmological parameters, thus the derivative is stable for all cosmological parameters. Due to the high precision of the 3 and 4-point derivates the difference in the unmarginalized errors is tiny hard to see on the plot. The error of those derivatives ar of the order of  $h^2$  with  $h$  already at 1%. We can use two derivative methods for the rest of our forecast. We will use the 3PT derivatives in this work.

Figure 7.2: Comparisons of the one-dimensional marginalized (in light grey) and unmarginalized errors (in dark grey). The *Euclid* probe is the spectroscopic probe with the optimistic settings. We use the abbreviation  $lbs_i$  stands for the nuisance parameter  $\ln(\hat{b}\sigma_8)_i$ , where  $i$  denotes the redshift bin. We plot the percentage deviation from the median error obtained by CF:CAMB.



(a) Comparison of the errors when switching the derivative method (b) Comparison of the errors when changing the relative stepsize for the neutrino mass.

Figure 7.3: In this figure we compare the results from CF using either of the two different Einstein Boltzmann solvers and MontePython in fisher mode, denoted with MP:Fisher.

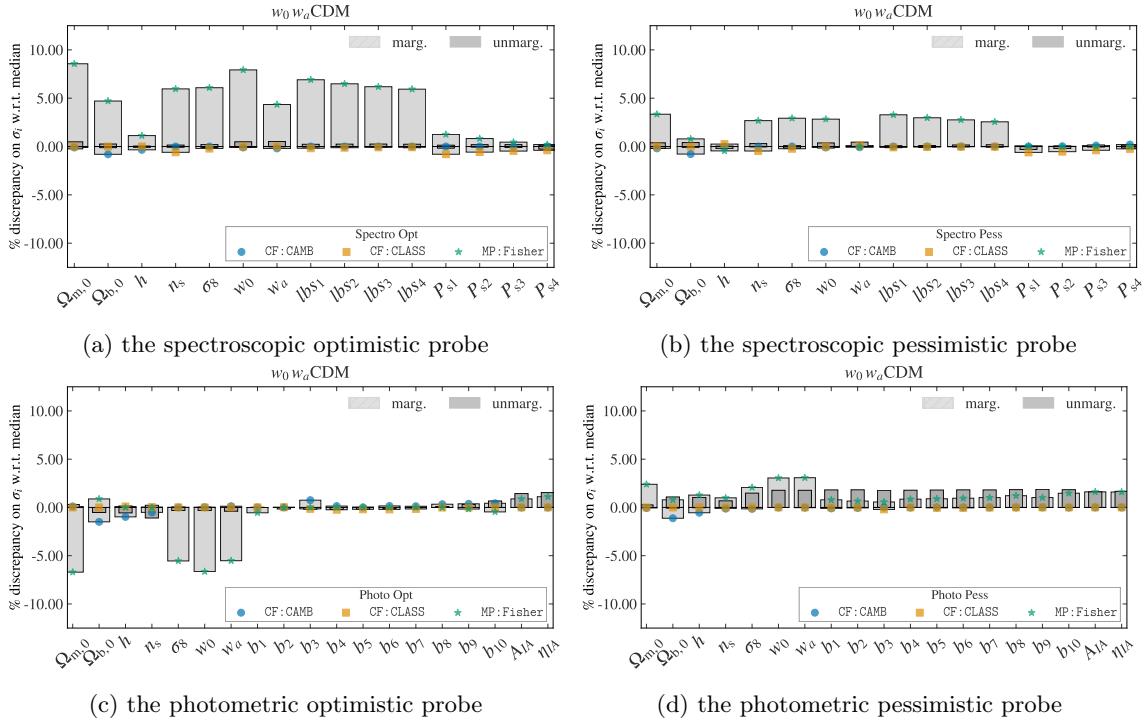
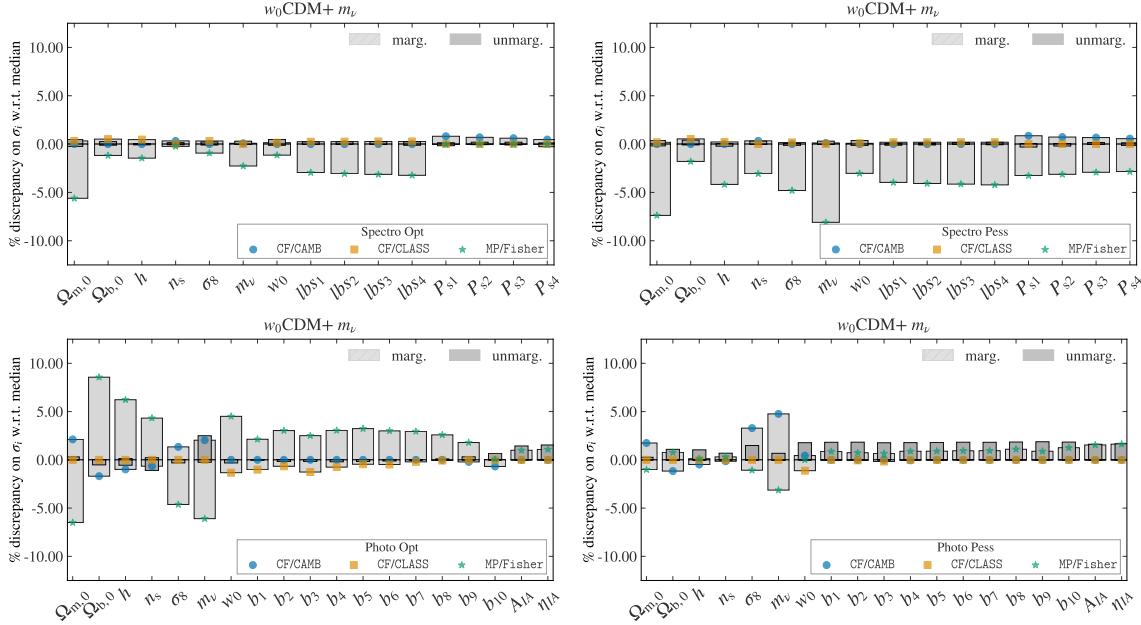
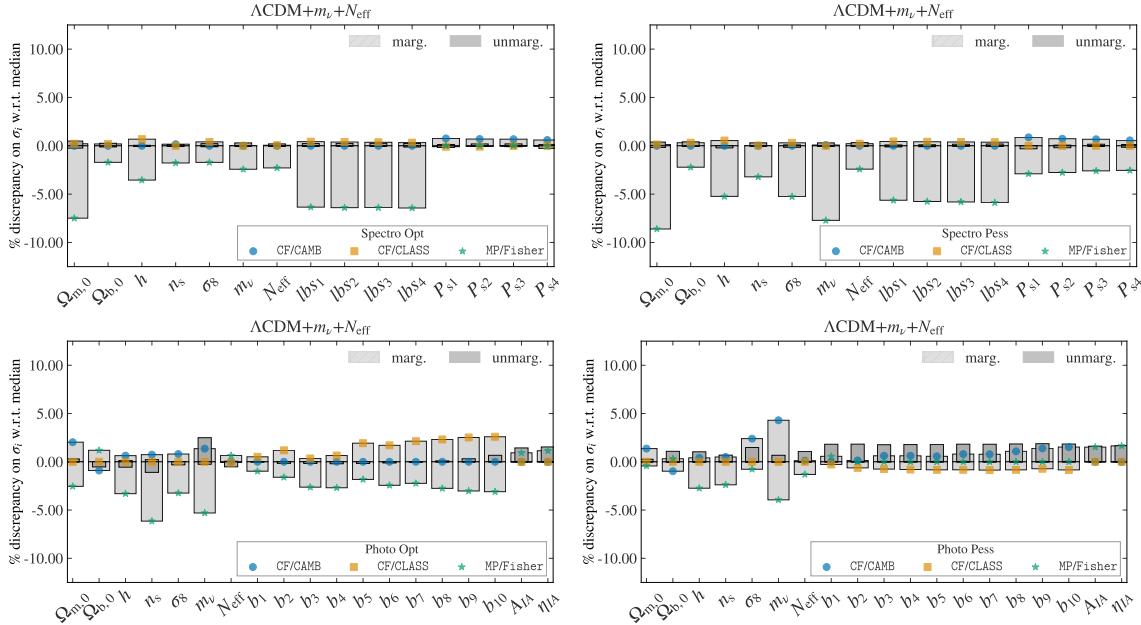


Figure 7.4: Same as figure 7.3 but for the  $w_0\text{CDM}+m_\nu$  modelFigure 7.5: Same as figure 7.3 but for the  $\Lambda\text{CDM}+m_\nu + N_{\text{eff}}$  model

The next free choice we had in our forecasting was the stepsize for the approximate neutrino mass. As we have said before the relative stepsizes for the cosmological parameters were chosen to be 1% of its fiducial value (or simply 1% for  $w_a$ ). We chose a higher stepsize of 10% for the massive neutrino as the neutrino mass only affects the matter power spectrum slightly. Too small steps always come with the risk that the change of the matter power spectrum is dominated by numerical noise. To check if the derivative with a relative stepsize of 10% is numerically stable we did the forecast with higher stepsizes of 30% and 50%. The result of the comparison can be found in figure 7.2b. We can see that the derivative with a stepsize of 10% is very stable as the errors of the higher stepsize choices are within 5% of each other. Since we have shown that the validation holds for the case with all 9 cosmological parameters, it is also valid for our validation cases with fewer parameters. As the error of a three-point numerical derivative is proportional to the stepsize squared, we will use a relative stepsize of 10% for the approximate neutrino mass.

For each of the validation cases, we present the comparison for both survey settings and both probes. The comparisons are found in figures 7.3-7.5. We show the errors of all cosmological parameters and nuisance parameters. We can see that all forecast errors are within 10% of the median value and thus the validating is successful. With our careful setting of the precision parameters, we obtain very good agreement between the two Boltzmann codes. The biggest discrepancy between the different Cosmicfish FI methods is a 5% difference in the neutrino mass. We believe this is due to residual small differences in the truncation of the Boltzmann hierarchy.

The comparison between MP and CF is a bit worse with typical discrepancies of 5-10%. The main contribution to this is that second-order derivatives are very sensitive to the choice of stepsize and there is no reasonable way to find the optimal one. The prescription for the stepsize was taken from [Euclid Collaboration: Blanchard et al., 2023] where they optimized the FI element for  $w_0 - w_a$ . This was chosen as that particular element was very sensitive to stepsizes due to the strong correlations. With this, we show that the prescription still works well enough for the validation but might be suboptimal when also varying the neutrino mass and the number of massless relics. Another reason for the worse agreement will be seen in the next section. It is the deviation from Gaussianity. This might be a subleading effect, as the deviation of the unmarginalized errors tends to be much smaller than marginalized ones. Typically, when there is non-gaussianity, the double-sided second-order derivatives get an additional contribution which becomes visible in the unmarginalized errors.

### 7.3 The validity of the Fisher approximation

As we have stated before, the Validation of the FI results from MP:Fisher validates the same likelihood that we use for the MCMC. The next step is to check the validity of the FI method by comparison with the MCMC. We do the comparison only for the optimistic cases as the pessimistic cases have the same general tendencies. To compare the results we compare the one-dimensional marginalized posteriors and the two-dimensional contours. We obtain the MCMC results from MP running in Metropolis-Hastings mode and analyze them using GetDist. The comparison can be seen in figures 7.6-7.8.

We can see in the contours of the  $w_0 w_a$ CDM model that the MCMCs and the FI match very well. With this, we recover the validation from [Euclid Collaboration: Blanchard et al., 2023] with our changes to the modelling and nonlinear corrections. This was expected as our modelling should only affect the contours of dark energy slightly due to the modelling of the nonlinear corrections. For the other parameters, the switch from HALOFIT to HMCODE came with the widening of the contours

Figure 7.6: Comparison of the one and two-dimensional marginalized contours obtained by MontePython in MCMC mode (MP:MCMC) with the contours of CF:CAMB and MP:Fisher. The contours depict the 68% and 95% confidence intervals for the  $w_0 w_a$ CDM model respectively. We plot only the cosmological parameters for the different probes. On the left, we show the spectroscopic probe and on the right, we show the photometric probe.

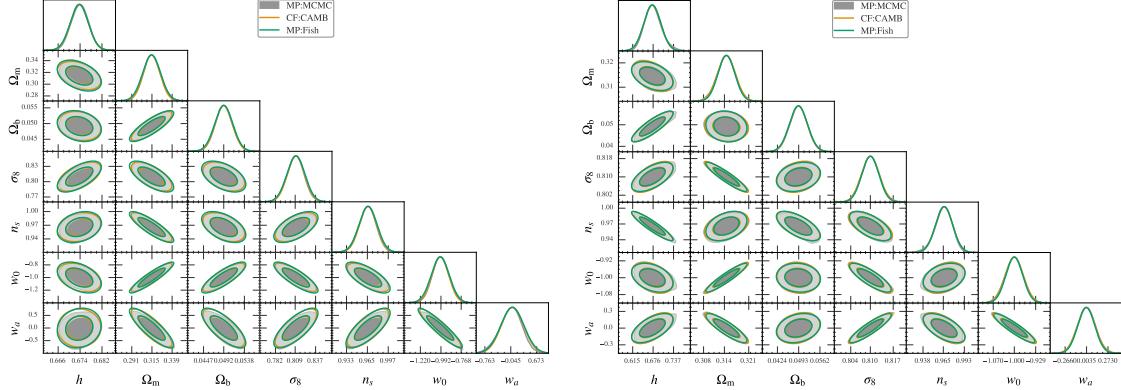


Figure 7.7: Same as figure 7.6 but for the  $w_0$ CDM+ $m_\nu$  model.

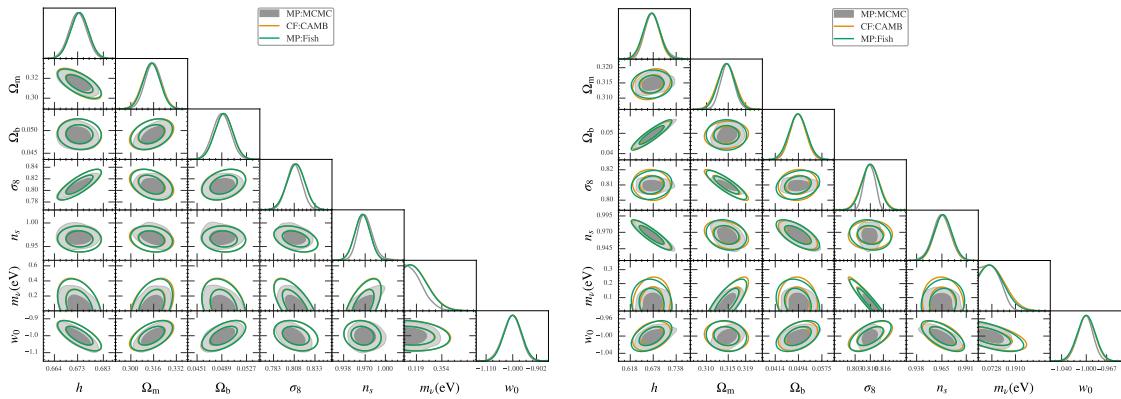
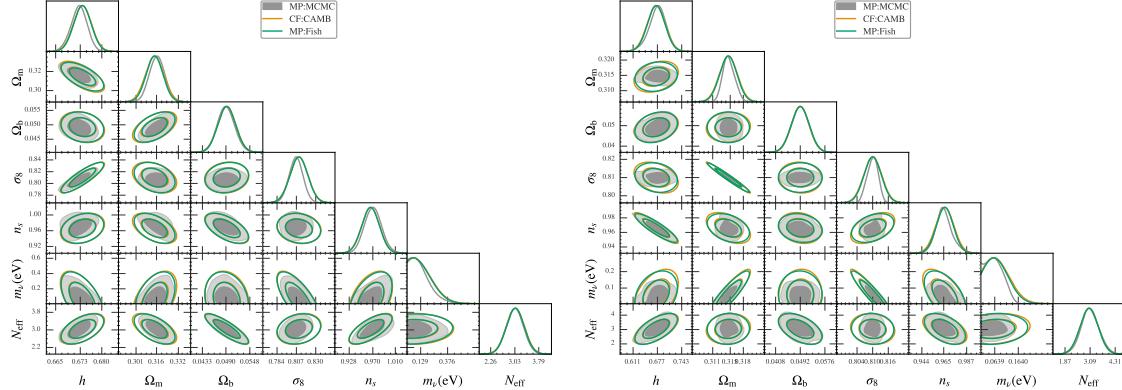


Figure 7.8: Same as figure 7.6 but for the  $\Lambda$ CDM +  $m_\nu$  +  $N_{\text{eff}}$  model.

which is discussed further in the next section.

In the other two cases, we can see strong deviations in the MCMC from the FI approximation. Most striking in this are the posteriors of  $m_\nu$ . The one-dimensional posteriors of  $m_\nu$  show that the Fisher approximation holds well at the fiducial value, but starts to deviate in both directions. For lower values of  $m_\nu$ , the posterior hits the theoretical prior, while for higher values it falls off more quickly than a Gaussian. We believe that the deviations of the other parameters can be explained solely by this.

In the spectroscopic probe of figure 7.7, we can see clearly that the parameters that are uncorrelated with  $m_\nu$  have next to no deviations from the Gaussian approximation while the stronger correlated parameters are affected more strongly. We see that due to the theoretical prior at zero mass important degeneracy directions get cut off and thus the posteriors deviate from the Gaussian approximation.

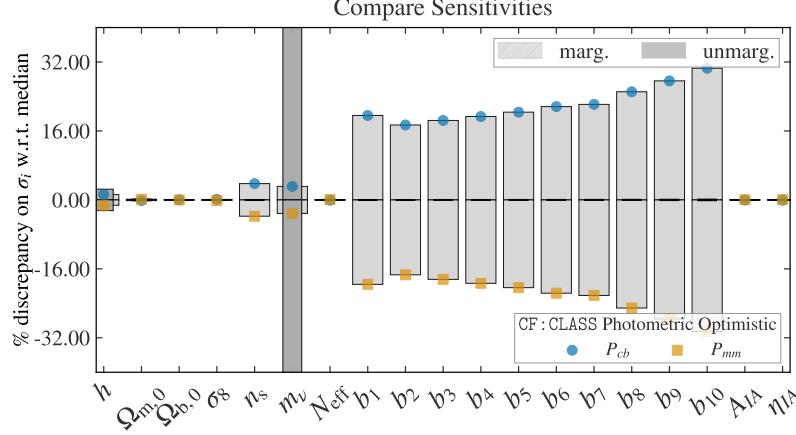
For higher values of the neutrino mass, we can also see a faster decay of the neutrino mass posterior. This is further amplified for the photometric probe. For that probe, the likelihood falls off so dramatically that the inside the 68% confidence bound of the fisher coincides with the 95% confidence bound of the MCMC. This has the effect that the strongly correlated parameters namely  $\sigma_8$  and  $\Omega_m$  not only get cut off at higher or lower values respectively, but also fall off quicker on the other side. We can also see a slight rotation of the FI contours for these parameters as the Gaussian approximation starts to fail even at the maximum likelihood.

We can see the same tendencies in the comparisons of figure 7.8. This further confirms our hypothesis that all deviations from Gaussianity are due to the prior and the deviation from a Gaussian normal distribution for the parameter  $m_\nu$ . With these results, we think that our implementations are fully validated, and also prove that to have a good forecast for neutrino parameters we have to move on from the standard FI formalism.

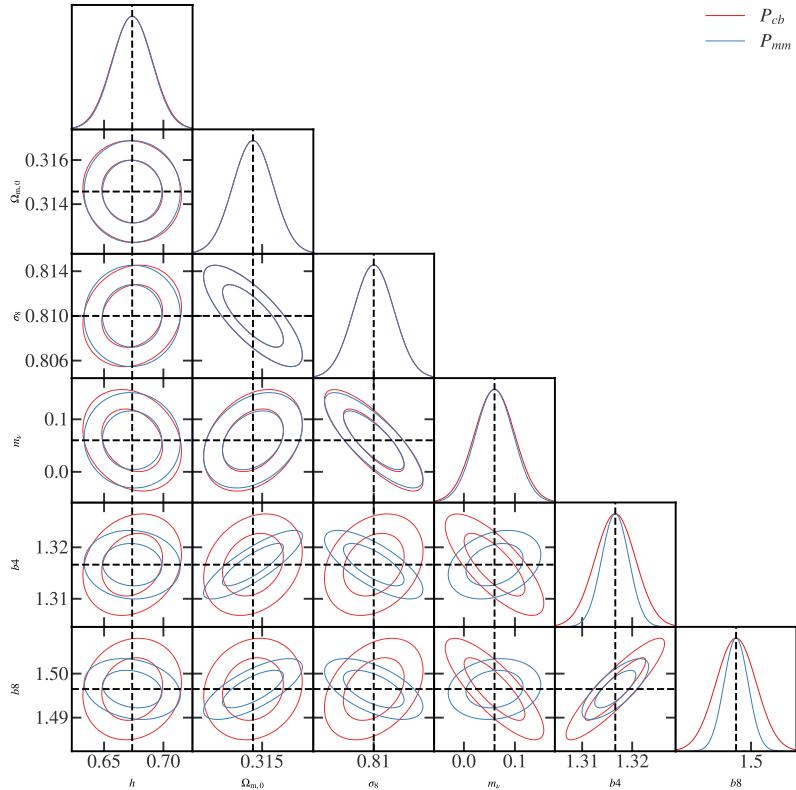
## 7.4 Bias from Modeling

Before we come to the forecasting results, we wanted to have a brief discussion about the effects of our modelling. We first discuss our switch from the total matter power spectrum to

Figure 7.9: We show the effect of switching the formulation of our observables to use  $P_{mm}$  instead of  $P_{cb}$ . We compare the results of the fisher forecast for the photometric optimistic using CF:CLASS as our code.



(a) Comparision of the one-dimensional marginalized and unmarginalized errors obtained by either using  $P_{mm}$  to formulate our observables or  $P_{cb}$ . For the parameter  $m_\nu$  the difference in the unmarginalized error goes outside of the frame and is of the order of 43% away from the mean.



(b) Comparision of the one and two-dimensional marginalized FI contours for the two formulations. We marginalized over other cosmological and nuisance parameters as their trends are the same.

the CDM+baryon power spectrum. As we have discussed before due to massive neutrinos the CDM+baryon matter power spectrum is suppressed by a factor  $\sim (1 - 6f_\nu)$  on scales smaller than the minimum clustering scale  $k_{\min}$ . We have also discussed how this translates into a shift of the order  $\sim (1 - 8f_\nu)$  for the total matter power spectrum. This means that overall the formulation of the likelihood is less sensitive to the neutrino mass when changing from  $P_{mm}$  to  $P_{cb}$ . When doing a parameter inference it would have two effects. Firstly, due to power spectra less depending to the neutrino mass, their constraints get wider. Secondly, if we assume that our description of the galaxy power spectrum would be the truth of the underlying data, then trying to fit the data with the 'wrong' description would bias our parameters.

To exemplify this claim we can think about the following gedankenexperiment. Let's say we have a galaxy clustering probe that is only sensitive to the smallest scales. Then the data power spectrum with an underlying neutrino mass  $m_\nu$  would be suppressed by a factor  $(1 - 6f_\nu)$ . When we would fit the data our model power spectrum would try to different values of  $m'_\nu$  until the power spectrum matches the data. The problem is that the suppression of the model power spectrum is given by  $(1 - 8f'_\nu)$ . Since the probe only measures the suppressed power spectrum the inference code will be able to find a matching  $f'_\nu$  to perfectly fit the data. We will find

$$6f_\nu \stackrel{!}{=} 8f'_\nu \implies m'_\nu = \frac{3}{4}m_\nu, \quad (7.4)$$

as the neutrino fraction is directly proportional to the mass. This effect is lessened by the fact that the minimum clustering scale is also proportional to the root of the neutrino mass. This can be used to break the biasing but since the dependence is less strong this biasing effect is still there. To test both of these effects we did two separate analyses. We first wanted to check how much the inferred errors change when changing the prescription to use  $P_{mm}$  instead of  $P_{cb}$ . As a simplification, we presume that the fiducial value of the galaxy bias does not change between the two runs, i.e.  $\hat{b} = b$ . This can be done as the forecast error should not depend on the actual value of the biases. To do our comparison, we compare the results for the photometric probe in optimistic settings, and we chose the  $\Lambda$ CDM+ $m_\nu + N_{\text{eff}}$  model.

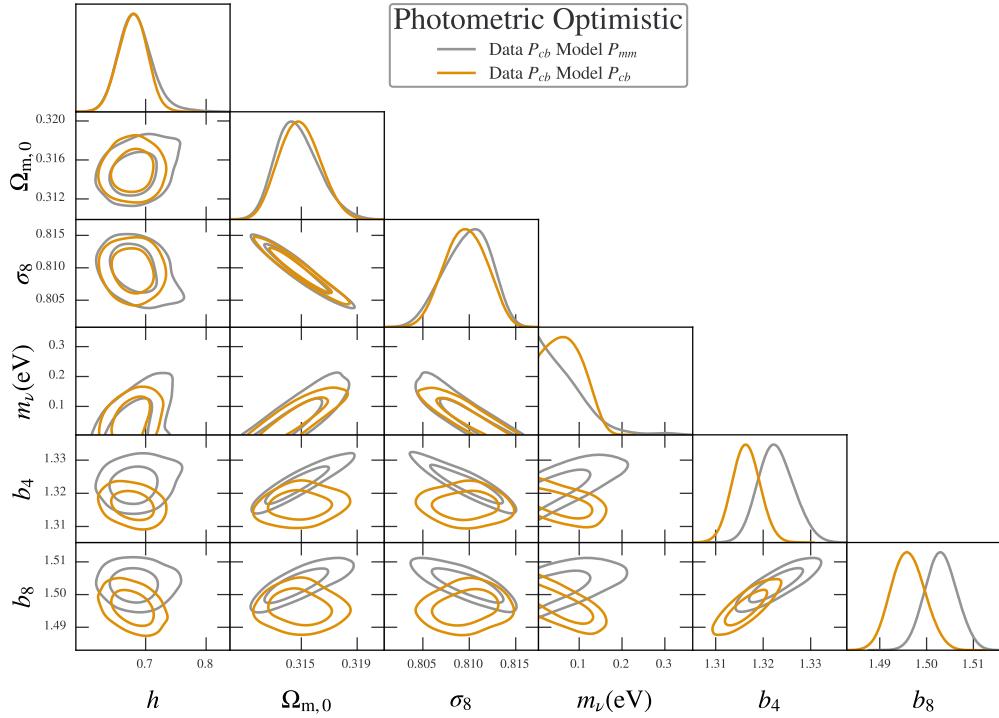
The results of this comparison are found in figure 7.9. In the upper figure 7.9a, we can see that the change in the inferred error for the cosmological parameters apart from the neutrino mass is negligible. This was expected as the effect of those parameters on the matter power spectrum and the CDM+baryon power spectrum is very similar. For the neutrino mass, we find a strong discrepancy of 80% in the unmarginalized error while the marginalized error is more comparable. We also see a strong 40%-80% shift in the errors of the galaxy biases. We believe that these two effects are related to one another. Viz., the difference between the matter power spectrum and the CDM+baryon power spectrum can be absorbed by the biases. This is due to the definition of the galaxy bias. If we remember how it was defined we can see that on scales much smaller than  $k_{\min}$

$$\begin{aligned} P_{gg}(k, z) &= b^2(k, z) P_{mm}(k, z) = \hat{b}^2(z) P_{cb}(k, z) \\ \xrightleftharpoons{\approx} & b^2(k, z) (1 - f_\nu)^2 P_{cb}(k, z) = \hat{b}^2(z) P_{cb}(k, z). \end{aligned}$$

This means that the strength of the suppression of the matter power spectrum can be absorbed by accordingly shifting all galaxy biases enough to mimic the missing suppression. Furthermore, this leads to a change in the degeneracy directions of the biases concerning the other cosmological parameters. This effect of compensation can be seen in figure 7.9b

For our next check, we wanted to perform a naive biasing test, where we generate our data vector

Figure 7.10: The biasing in the parameter inference when switching the prescription of the observables but fixing the data to  $P_{cb}$ . This is using the photometric probe in optimistic settings in a  $\Lambda\text{CDM} + m_\nu + N_{\text{eff}}$  model. In yellow, we show the posteriors if we try to fit the data using our  $P_{cb}$  formulation. In grey, we show how we fit the data using our  $P_{mm}$  formulation.



using our more correct description using  $P_{cb}$  and then fit it using the  $P_{mm}$  description. The results of this test are found in figure 7.10 where we have picked out some indicative parameters. We can see how the posteriors of  $m_\nu$  shift to lower values and hit the theoretical prior. This is due to the discrepancy between the suppression of the two power spectra. This effect is partially broken by adding the shearing probe. The angular power spectrum of weak lensing is calculated by the total matter power spectrum alone and thus independent of our modelling of the galaxy bias. The remaining shift of the neutrino mass induces a slight shift in  $\Omega_m$  and  $\sigma_8$  in their respective correlation directions. As we discussed, we can see how the bias parameters try to compensate for the reduction of  $P_{mm}$  on the smallest scales by shifting to higher values. Their correlation with the cosmological parameters  $\Omega_m$ ,  $\sigma_8$  and  $m_\nu$  shift again since they will need to compensate for additional shifts of the power spectra. This degeneracy is broken by the fact that the probes are also sensitive to intermediate ranges where the neutrino-induced suppression has not flattened out yet. To take care of this effect, we would have needed to switch to a scale-dependent galaxy bias in the  $P_{mm}$  case.

Our next discussion will be about the switch from `HALOFIT` to `HMCODE`. As we had stated earlier, the difference between them is that `HMCODE` is a semi-analytical model and `HALOFIT` is a direct fit.

This means they react slightly differently to a change in cosmological parameters. We will analyze this change in our first test as before.

All codes trying to predict the nonlinear power spectrum give slightly different results. It is unclear which of them describes the reality better. We know that `HMCODE` matches better simulations with massive neutrinos like the *Euclid* flagship simulations so we use it for our forecast. To check how much this can bias, our results for our second test we fit different data with the same model. This essentially answers the question of how inferred parameters would shift if our nonlinear prescription did not match reality.

The results of the first test are found in the figure 7.11. We can see large discrepancies between the forecast errors for all parameters. The largest discrepancy is in the unmarginalized errors for the neutrino mass. There we have a 15% discrepancy between the two codes. We believe this is due to the many different places where the `HMCODE` algorithm treats massive neutrinos, that do not get caught in the `HALOFIT` fitting formula. In the `HALOFIT` model the effect of massive neutrinos enters the total matter power spectrum while many effects of nonlinear collapse really should be dictated by the dynamics of the CDM+baryon power spectrum. Additionally `HMCODE` also uses correction factors that are directly proportional to the neutrino mass to better match simulations with massive neutrinos.

The largest discrepancy in the marginalized errors is in the parameters  $\Omega_m$ ,  $\sigma_8$  and  $n_s$  where the discrepancy is of the order of 60-70%. In the two-dimensional marginalized contours, we can also see a rotation of the degeneracy directions of these parameters concerning  $h$ . The forecast errors of `HALOFIT` are systematically below the errors from `HMCODE` for the cosmological parameters implying that the fitting formula description of `HALOFIT` depends more strongly on the values of cosmological parameters than the collapse physics behind it.

Our next test is regarding the bias of `HMCODE`. For this, we generate the data using `HALOFIT` and then try fitting it with `HMCODE`. The result of this test can be seen in figure 7.11. Compared to the first biasing test the difference is that we fixed the model and changed the data and not the other way around. We believe that the bias shift for both methods should be of the same order of magnitude. Testing it this way keeps the posteriors of similar size such that they can be compared more easily.

We can see in figure 7.11 that the shifts in bias are much more prominent than in the first biasing test. To explain these shifts we have to look into [Mead et al., 2021]. There we can see that `HALOFIT` overpredicts the matter power spectrum by up to 15% on the smallest scales. To compensate for this we go to smaller values of  $h$  (essentially reducing  $\omega_b$ ) and higher values of  $n_s$  and  $m_\nu$  to fix the intermediate ranges. Fixing the amplitude needs to vary  $\Omega_m$  and  $\sigma_8$ . The degeneracy of this is broken by the shearing and cross-correlation probe essentially fixing  $\Omega_m$  and  $\sigma_8$ . Due to the higher neutrino mass, the galaxy biases need to shift up matching the different suppressions of matter and CDM+baryon spectrum.

Both tests indicate that our nonlinear modelling strongly influences our parameter inference. This is why it is important to have a good handle on N-body simulations to fit our halo models. Through extensive comparisons of different nonlinear models, it was found that `HMCODE` matches the *Euclid* flagship simulations better than `HALOFIT` [Euclid Collaboration: Adamek et al., 2023]. On scales  $k < 1 \text{ hMpc}^{-1}$  `HMCODE` matches the simulations on a 2% level while `HALOFIT` for late times starts diverging at a 5-6% level. Our second test shows how this discrepancy is already leading to multiple standard deviation shifts in the inferred parameters.

We also believe that our modelling of the neutrino-induced scale-dependent bias is important to not underestimate our errors. With these conclusions, we can consider our forecast pipeline as

thoroughly validated and checked against biasing.

Figure 7.11: The biasing in the parameter inference when switching the data vector from `HMCODE` to being generated using `HALOFIT`. We fit the data using `HMCODE` in both cases. In this figure, we depict the photometric probe with optimistic settings in a  $\Lambda\text{CDM}+m_\nu+N_{\text{eff}}$  model. In yellow, we show the posteriors if our model `HMCODE` would describe the underlying universe. In grey, we show how our model performs when the truth would be `HALOFIT`.

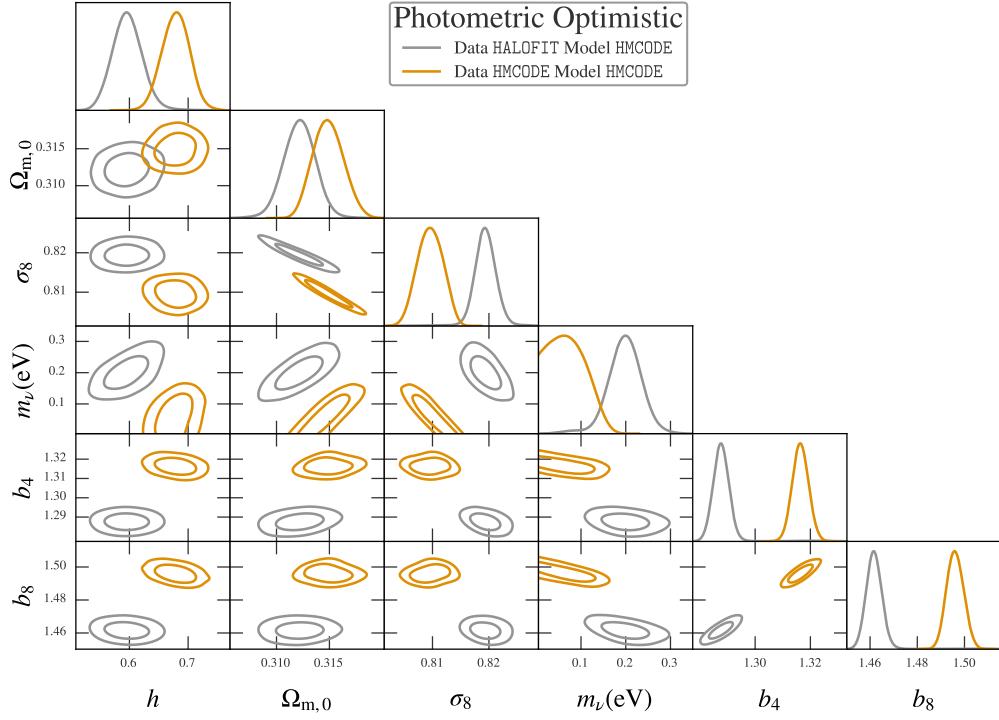
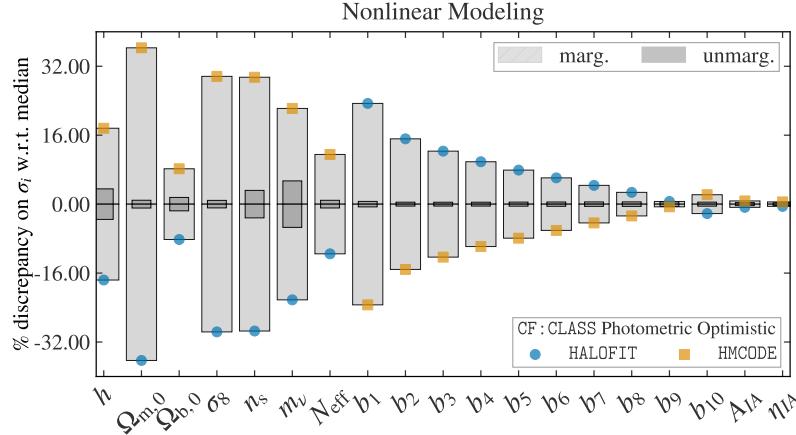
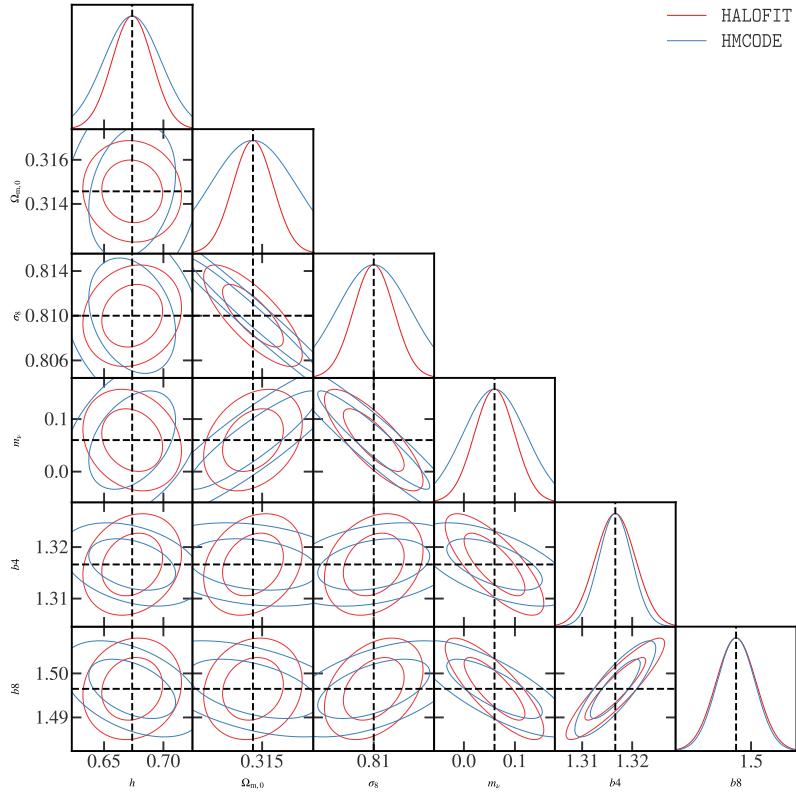


Figure 7.12: Same as figure 7.9 but switching the nonlinear model from HMCODE to HALOFIT.



(a) Comparision of the one-dimensional marginalized and unmarginalized errors obtained by using either of the two nonlinear correction codes.



(b) Comparision of the one and two-dimensional marginalized contours for the two different nonlinear correction codes. We marginalized other cosmological and nuisance parameters as their trends are the same.

# Chapter 8

## Results

In this chapter, we present our forecast results for neutrino parameters and dark energy. We remind ourselves that the validation runs were not actual forecasts. For them, we used extra optimistic settings to be able to find good agreement for the FI matrixes. For the forecast, we will switch to more realistic and conservative settings. Compared to our validation runs there are multiple differences in our parameters. Firstly the neutrino mass  $m_\nu$  is no longer condensed into one massive neutrino but evenly split between three neutrino species. To differentiate this quantity from the neutrino mass from the last chapter we will call it  $\sum m_\nu$ . This is done by setting the number of degenerate non-cold dark matter particles ' $\text{deg}_n \text{cdm} = 3$ '.

Because of this we also change our definition of  $N_{\text{eff}}$ . Before it could have been understood as a change in the temperature of the massless neutrinos and could vary freely to higher and lower values. Now all neutrinos have the same temperature of

$$T_\nu = T_{\text{CMB}} \left( \frac{3.044}{3} \right)^{1/4} \left( \frac{11}{4} \right)^{1/3}. \quad (8.1)$$

This parameter now just parametrises any additional massless relic particle. Typical models only predict additional species that contribute positively to  $N_{\text{eff}}$ . From this, we define a new parameter  $\Delta N_{\text{eff}}$ . This parameter is essentially equivalent to the parameter  $N_{\text{ur}}$  inside of **CLASS**.

Our next change is, that we vary the parameters  $\sigma_v$  and  $\sigma_p$  that govern the nonlinear corrections of the spectroscopic probe. In order to be more conservative we also stick to the pessimistic settings of the probes. We will also vary all 9 cosmological parameters for our MCMC.

For  $\sum m_\nu$  and  $\Delta N_{\text{eff}}$  we chose prior edges with a theoretical prior at 0 as a lower bound and an upper bound high enough to not change our results. The dark energy parameters have a bit tighter prior edges to not probe unphysical regions of the parameter space. These parameters are an approximation to a wider set of theories where dark energy has an equation of state that is slowly varying. When the posterior hits the prior edges for the dark energy parameters we will consider them as unconstrained. For the other cosmological and nuisance parameters we have chosen arbitrary priors that will not be reached but speed up convergence.

A summary of our cosmological parameters, their fiducial and their prior edges can be seen in the table 8.1. The results can be seen in figure 8.1. We see that the photometric and spectroscopic probes are sensitive to different cosmological parameters. The spectroscopic probe dominates the sensitivity for  $\Delta N_{\text{eff}}$  and  $h$ . These are the parameters that mainly affect the BAO. Since the

Figure 8.1: One and two-dimensional marginalized posteriors for the different *Euclid* probes. We depict the 68% and 95% confidence intervals for the cosmological parameters of the full nine-parameter model. We depict the photometric probe in purple, the spectroscopic probe in cream and the combined probe in red.

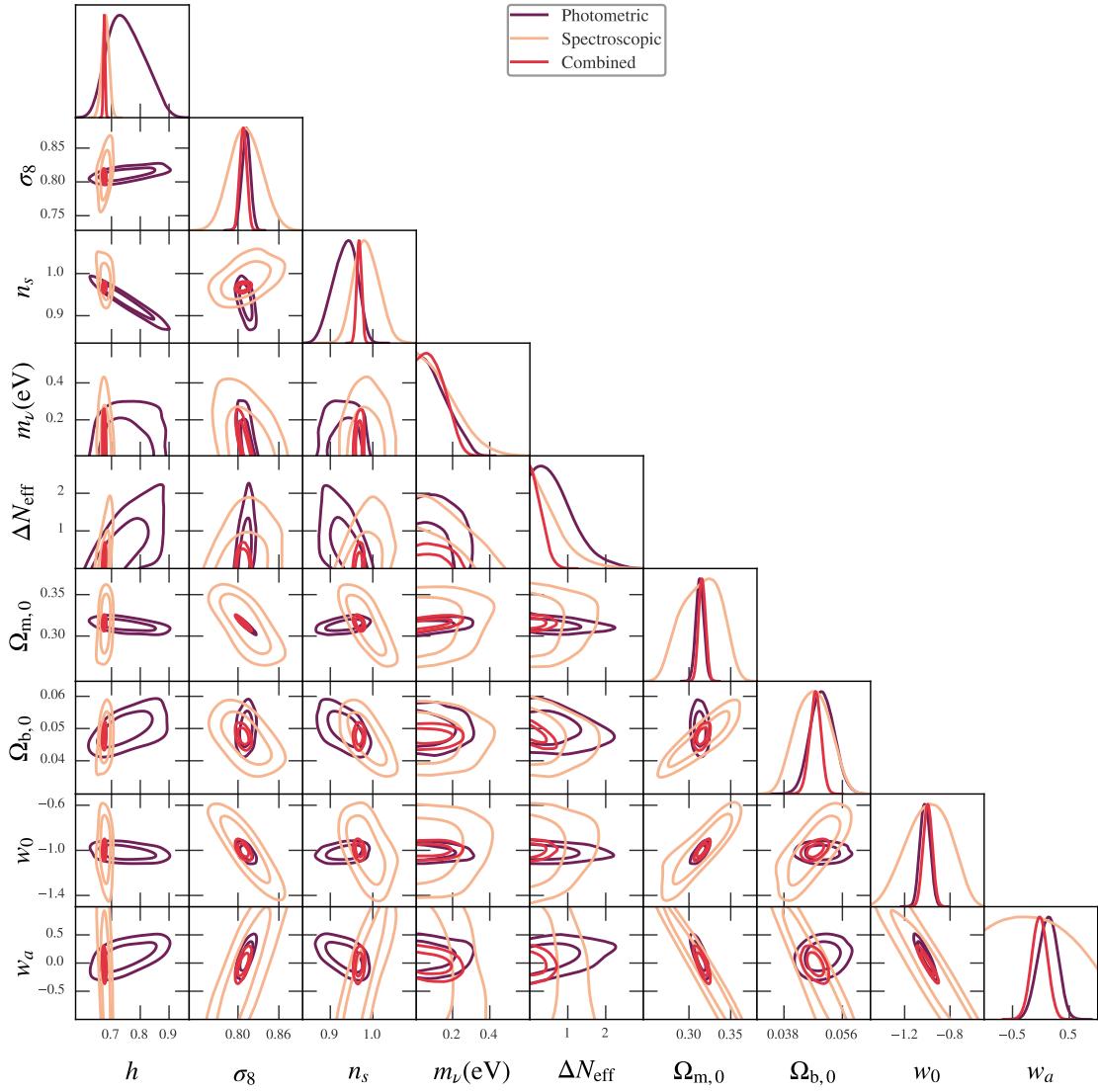


Table 8.1: Settings for the final forecast MCMCs. The fiducial values of the Nuisance parameters have been listed in the section for the respective probes.

Varied parameters								
$\Omega_m$	$100 \times \Omega_b$	$h$	$n_s$	$\sigma_8$	$\sum m_\nu$ (meV)	$\Delta N_{\text{eff}}$	$w_0$	$w_a$
Fiducial Value								
0.314571	4.92	0.6737	0.9661	0.81	60	0	0	-1
Prior Edges								
[0.005, 1]	[0.5, 100]	[0.1, 1.5]	[0.8, 1.2]	[0.7, 0.9]	[0, 1000]	[0, 5]	[-1.5, -0.5]	[-1, 1]

spectroscopic probe is sensitive to that region, it measures these well. The spectroscopic probe is nearly insensitive to the amplitude of the power spectrum as it is multiplied by the galaxy biases. To break this, the spectroscopic probe needs the RSD where the clustering parameter  $f\sigma_8$  enters. We also see that it alone is not able to constrain the time-dependent dark energy equation of state parameter  $w_a$ . This parameter controls only slightly the time evolution of the amplitude of the matter power spectrum. Its sensitivity to the baryon density parameter  $\Omega_b$  is greatly reduced by varying  $\sigma_v$ . This is due to the effect of  $\Omega_b$  in shifting the phase of the BAO can also be mimicked by varying  $\sigma_v$ .

The photometric probe is very sensitive to the amplitude of the power spectrum and dominates the dark energy parameters, the neutrino mass, the density parameter of total matter, and the variance of the matter perturbations. Its sensitivity to  $\Omega_b$  comes also from the BAO that leaves a small impact on the actual  $C_l$ . Since the scale of the BAO is fixed at around  $5 \cdot 10^{-2} \text{Mpc}^{-1}$  the imprint shows up at different multipoles  $\ell$  for different redshift bins. This leaves a very clear signature of  $\Omega_b$ . It loses its sensitivity to  $h$  and  $\Delta N_{\text{eff}}$  for the same reason. The actual scale of the BAO is washed out by the integration over  $z$  for each bin. Essentially this makes the photometric probe lose its sensitivity to scales.

In that sense, both probes are very complimentary to each other. In the figure, one can see how the correlation directions of the different probes are often perpendicular to one another breaking correlation directions and drastically improving the constraints on  $h$ ,  $n_s$ , and  $\Delta N_{\text{eff}}$ .

Nevertheless, judging from this forecast, Euclid alone will not be able to detect the neutrino mass. This is why in our final forecasting results in tabel we can only give a 95% confidence interval. It should be noted that by adding CMB data to our forecast, we can achieve a measurement of the neutrino mass on the 68% confidence level for this nine-parameter model. If we go to a smaller model with only  $\Lambda\text{CDM}+m_\nu$  we can even achieve a 99% confidence detection. We will not discuss these results further in this work as we have not discussed the CMB as an additional probe.

The marginalized errors are found in the table 8.2. The constraints of *Euclid* for the cosmological parameters are tighter than for *Planck*. *Planck* does not have the constraining power to constrain dark energy and neutrinos at the same time. Nevertheless, we can compare the constraints on the cited sensitivities of the submodels. This is because the parameters  $w_0 + w_a$  are not strongly correlated with the other two parameters in our case. This makes it so that the constraints stay similar, because we are not opening up correlation directions. This means that we expect similar errors for the parameters if we have run the smaller models. In an analysis of the  $w_0w_a\text{CDM}$  *Planck*

+ BAO + Supernovae gives constraints on the dark energy parameters

$$w_0 = -0.957 \pm 0.08 \quad \text{and} \quad w_a = -0.29 \pm 0.3$$

This means that we can constrain these parameters better with the bigger model. The main part of the constraints of *Planck* to these parameters are from contributions in additional late integrated Sachs-Wolf effect. This essentially comes from the fact that during dark energy domination, the metric perturbations start to decay adding some additional terms to the low multipole  $C_\ell$  of CMB. With *Euclid* we can see the same redshift-dependent reduction of the amplitude on all scales due to its tomographic nature. This is strongly constraining for these dark energy parameters.

The parameters  $\sum m_\nu$  and  $\Delta N_{\text{eff}}$  are constrained by *Planck* +BAO+lensing to

$$\sum m_\nu < 0.12 \quad \text{and} \quad \Delta N_{\text{eff}} < 0.34. \quad (8.2)$$

These are both tighter than in our *Euclid* forecast. The main sensitivity to these parameters comes from their background effects shifting the angular scale of recombination and the redshift of equality. Both of these parameters are very tightly constrained by the CMB measurements.

As both experiments measure probes that are very complimentary to one another as well as having additional information in their cross-correlation, we believe that a combined analysis will bring us a new milestone in precision cosmology.

Table 8.2: Forecast 68% confidence levels for the different *Euclid* probes and the combined probe. For the parameters  $\sum m_\nu$  and  $\Delta N_{\text{eff}}$  we only state the 95% upper limit as they are bound from below by their theoretical prior. The spectroscopic probe alone was not able to constrain  $w_a$  within our prior edges.

Forecast Results									
	$\Omega_m$	$100 \times \Omega_b$	$h$	$n_s$	$\sigma_8$	$w_0$	$w_a$	$\sum m_\nu (\text{meV})$	$\Delta N_{\text{eff}}$
Probe	68% Sensitivity							95% confidence limit	
Photometric	0.0049	0.38	0.065	0.029	0.0065	0.05	0.18	< 260.	< 1.70
Spectroscopic	0.0258	0.56	0.013	0.031	0.024	0.20	—	< 350.	< 1.50
Combined	0.0043	0.18	0.0030	0.0060	0.0054	0.04	0.14	< 220.	< 0.57

# Chapter 9

## Conclusion

In summary, this thesis has sought to demonstrate the importance of modelling physical effects in order to obtain realistic and robust sensitivity forecasts for neutrino mass. In the era of precision Cosmology, it is critical to address the different details of all modelling steps in order to not bias the final inference's results.

In brief, the central themes of this paper include the modelling of the neutrino-induced scale-dependent growth of structure, the handling of nonlinear corrections for the observables, and the critical role of the neutrino in the galaxy bias. We have used different statistical methods to compare our forecasts with and validate our findings. For the validation we had to look into the numerical difficulties of handling neutrinos, understand the different approximation schemes to solve non-analytical problems and compute stable numerical derivatives.

This study underscores the importance of going past the common Fisher approximation to have a correct forecast in models with massive neutrinos. Their strong correlation with the different cosmological parameters induces cuts and projection effects and intrinsic deviations from the Gaussian posterior. These can not be described by Fisher Information alone and can change our constraints. Our findings show the strengths of the *Euclid* mission and its role in unpacking the rich treasure trove of Information in the large-scale structure.

Returning to the initial question of how to do a well-modelled, unbiased, and realistic sensitivity forecast, we have also found strong biases in our parameter inference that show up if the modelling is done in an inconsistent or ignorant way.

While our research has yielded valuable insights, it is not without its constraints. In our modelling, we have yet to discuss the observational effects like redshift space distortions on the angular power spectrum, corrections from baryonic feedback and mode coupling due to masking out the galactic plane.

Moving forward, further investigation could delve into these open questions and the generalization of our neutrino modelling for other non-cold dark matter problems.

Ultimately this study shows that the large-scale structure is the next milestone in precision Cosmology.

# Bibliography

- [Blas et al., 2011] Blas, D., Lesgourgues, J., and Tram, T. (2011). The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes. *JCAP*, 07:034.
- [Collaboration et al., 2023] Collaboration, E., Tanidis, K., et al. (2023). Euclid preparation. tbd. the effect of linear redshift-space distortions in photometric galaxy clustering and its cross-correlation with cosmic shear.
- [Dolag et al., 2004] Dolag, K., Bartelmann, M., Perrotta, F., Baccigalupi, C., Moscardini, L., Meneghetti, M., and Tormen, G. (2004). Numerical study of halo concentrations in dark-energy cosmologies. *Astronomy & Astrophysics*, 416:853–864.
- [Euclid Collaboration: Adamek et al., 2023] Euclid Collaboration: Adamek, J. et al. (2023). Euclid: Modelling massive neutrinos in cosmology – a code comparison. *JCAP*, 06:035.
- [Euclid Collaboration: Blanchard et al., 2020] Euclid Collaboration: Blanchard, A. et al. (2020). Euclid preparation: VII. Forecast validation for Euclid cosmological probes. *Astron. Astrophys.*, 642:A191.
- [Euclid Collaboration: Blanchard et al., 2023] Euclid Collaboration: Blanchard, A. et al. (2023). Euclid: Validation of the montepython forecasting tools.
- [Feldman et al., 1994] Feldman, H. A., Kaiser, N., and Peacock, J. A. (1994). Power-spectrum analysis of three-dimensional redshift surveys. *The Astrophysical Journal*, 426:23.
- [Hirata and Seljak, 2010] Hirata, C. M. and Seljak, U. c. v. (2010). Erratum: Intrinsic alignment-lensing interference as a contaminant of cosmic shear [phys. rev. d 70, 063526 (2004)]. *Phys. Rev. D*, 82:049901.
- [Kilbinger, 2015] Kilbinger, M. (2015). Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901.
- [Lesgourgues et al., 2013] Lesgourgues, J., Mangano, G., Miele, G., and Pastor, S. (2013). *Neutrino Cosmology*. Cambridge University Press.
- [Lewis, 2019] Lewis, A. (2019). GetDist: a Python package for analysing Monte Carlo samples.
- [Lewis et al., 2000] Lewis, A., Challinor, A., and Lasenby, A. (2000). Efficient computation of CMB anisotropies in closed FRW models. *ApJ*, 538:473–476.

- [Mead, 2016] Mead, A. J. (2016). Spherical collapse, formation hysteresis and the deeply non-linear cosmological power spectrum. *Monthly Notices of the Royal Astronomical Society*, 464(2):1282–1293.
- [Mead et al., 2021] Mead, A. J., Brieden, S., Tröster, T., and Heymans, C. (2021). hmcode-2020: improved modelling of non-linear cosmological power spectra with baryonic feedback. *MNRAS*, 502(1):1401–1422.
- [Nakamura and Suto, 1997] Nakamura, T. T. and Suto, Y. (1997). Strong gravitational lensing and velocity function as tools to probe cosmological parameters: Current constraints and future predictions. *Progress of Theoretical Physics*, 97(1):49–81.
- [Navarro et al., 1997] Navarro, J. F., Frenk, C. S., and White, S. D. M. (1997). A universal density profile from hierarchical clustering. *The Astrophysical Journal*, 490(2):493.
- [Press and Schechter, 1974] Press, W. H. and Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *The Astrophysical Journal*, 187:425–438.
- [Raccanelli et al., 2018] Raccanelli, A., Verde, L., and Villaescusa-Navarro, F. (2018). Biases from neutrino bias: to worry or not to worry? *Monthly Notices of the Royal Astronomical Society*, 483(1):734–743.
- [Raveri et al., 2016] Raveri, M., Martinelli, M., Zhao, G., and Wang, Y. (2016). CosmicFish Implementation Notes V1.0.
- [Sprenger et al., 2019] Sprenger, T., Archidiacono, M., Brinckmann, T., Clesse, S., and Lesgourges, J. (2019). Cosmology in the era of Euclid and the Square Kilometre Array. *JCAP*, 02:047.
- [Takahashi et al., 2020] Takahashi, R., Nishimichi, T., Namikawa, T., Taruya, A., Kayo, I., Osato, K., Kobayashi, Y., and Shirasaki, M. (2020). Fitting the nonlinear matter bispectrum by the Halofit approach. *ApJ*, 895(2):113.
- [Vagnozzi et al., 2018] Vagnozzi, S., Brinckmann, T., Archidiacono, M., Freese, K., Gerbino, M., Lesgourges, J., and Sprenger, T. (2018). Bias due to neutrinos must not uncorrect'd go. *JCAP*, 09:001.
- [Villaescusa-Navarro et al., 2014] Villaescusa-Navarro, F., Marulli, F., Viel, M., Branchini, E., Castorina, E., Sefusatti, E., and Saito, S. (2014). Cosmology with massive neutrinos i: towards a realistic modeling of the relation between matter, haloes and galaxies. *Journal of Cosmology and Astroparticle Physics*, 2014(03):011–011.