UNIVERSITY NAME

DOCTORAL THESIS

Non-linear structure formation in models beyond LCDM

Author:
Santiago CASAS CASTRO

Supervisor:
Dr. Valeria PETTORINO

A thesis submitted in fulfillment of the requirements for the degree of Doctor in Physics

in the

Research Group Name
Department or School Name

March 16, 2017

Declaration of Authorship

I, Santiago CASAS CASTRO, declare that this thesis titled, "Non-linear structure formation in models beyond LCDM" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

UNIVERSITY NAME

Abstract

Faculty Name
Department or School Name

Doctor in Physics

Non-linear structure formation in models beyond LCDM

by Santiago CASAS CASTRO

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too…

Contents

List of Figures

List of Tables

List of Abbreviations

LSS	**L**arge **S**cale **S**tructure
BAO	**B**aryon **A**coustic **O**scillations
RSD	**R**edshift **S**pace **D**istortions
SPT	**S**tandard **P**erturbation **T**heory
RPT	**R**enormalized **P**erturbation **T**heory
EFT	**E**ffective **F**ield **T**heory
CDM	**C**old **D**ark **M**atter
GNQ	**G**rowing **N**eutrino **Q**uintessence
GR	**G**eneral **R**elativity
RDE	**R**adiation **D**ominated **E**ra
MDE	**M**atter **D**ominated **E**ra
EdS	**E**instein-**d**e **S**itter
MG	**M**odified **G**ravity
DE	**D**ark **E**nergy
WL	**W**eak **L**ensing
GC	**G**alaxy **C**lustering

Physical Constants

Speed of Light $c = 2.997\,924\,58 \times 10^{8}\,\mathrm{m\,s^{-1}}$ (exact)

List of Symbols

a	scale factor	
H_0	Hubble parameter	($\mathrm{km\,s^{-1}Mpc^{-1}}$)
z	redshift	
Ω_b	Fraction of energy density in baryons	
Ω_c	Fraction of energy density in CDM	
Ω_m	Fraction of energy density in non-relativistic matter	
Ω_ν	Fraction of energy density in neutrinos	
Ω_r	Fraction of energy density in radiation	

Dedicated to ...

Chapter 1

Overview of Standard Cosmology

1.1 The framework of General Relativity

1.1.1 The equivalence principle and geometry

We can probably delete this section

- The equivalence principle
- Manifolds and local flatness
- Physical relation equivalence principle -> local flatness
- The metric formalism
- The Levi-Civita connection
- Dipheomorphism invariance

1.1.2 The Einstein-Hilbert action and the field equations

$$S = \int \mathrm{d}x^4 \sqrt{-g}(R - \Lambda) \tag{1.1}$$

$$R_{\mu\nu} + \frac{1}{2} g_{\mu\nu} R + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu} \tag{1.2}$$

- Bianchi identities and the conservation of $T_{\mu\nu}$
- Lovelock's theorem where do you need this?

1.1.3 Gravitational potentials

- Gravitational potentials Φ and Ψ.
- Very simple explanation of Gauge choice probably don't need it
- Newtonian Gauge
- Smallness of gravitational potentials what is this?
- Remark for later that in GR $\Psi = \Phi$
- Weyl potential (lensing potential, null geodesics)
- The Newtonian limit I put this here, because all items above are generally defined, not just in the Newtonian limit

1.2 The standard cosmological model

- small discussion on how to define a cosmology
- Observer along timelike geodesic and foliation of spacetime

1.2.1 Friedmann equations

We can probably make this section very short: 1 page

- Friedmann-Lemaître-Robertson-Walker metric
- 00 and ii components of Einstein equation
- First and second Friedmann equations
- Several ways of writing Hubble function for different species
- Background expansion, , RDE, MDE, w(z)

1.2.2 Distances

We can probably delete this whole section

- definition of distances
- definition of Hubble time
- Horizons?

1.2.3 The ΛCDM model

- Reasons for CDM
- Reasons for Lambda
- Baryons, neutrinos and other species

1.2.4 The cosmological constant problem

- small discussion on old and new cosmological constant problem
- fine tuning, naturalness, basic definitions

1.2.5 Linear perturbations

- Differentiate between linear perturbations in different eras, just shortly
- Linear perturbations in matter dominated eras, newtonian gauge
- Fluid equations, full and linearized

1.3 Early Universe

We can probably delete this whole section

1.3.1 Cosmic Microwave Background Radiation

- Short introduction and importance of CMB
- Important constraints on parameters coming from CMB
- Constrain CDM alone
- Constrain initial power amplitude and tilt
- Constrain relativistic degrees of freedom

1.3.2 Inflation

I would delete this

- Inflation as a paradigm
- Flatness and horizon problems
- Inflation produces almost scale invariant spectrum

Chapter 2

Dark Energy and Modified Gravity

2.1 Universal coupling to matter

- define Jordan and Einstein frame for universal couplings these coincide; you should move this to non-universal coupling part
- write action in both cases
- write that in Universal coupling matter couples to the full $\sqrt{-g}$

2.1.1 Quintessence

- small introduction to quintessence, Ratra, Peebles, Wetterich, Luca
- different potentials, Brans-Dicke, Exponential potential

2.1.2 f(R) Theories

delete this paragraph

- small intro to f(R), Amendola, Felice, Tsujikawa
- write f(R) as conformal transformation of quintessence (Valeria's paper)

2.1.3 Horndeski Theory

not Horndeski, unless you have results for it; call it 'parameterizing Modified Gravity' and refer to the paper inside; don't refer to QS limit in this chapter but only in the point where you actually make this assumption

- Most general single scalar field theory, second order equations of motion, 4 dimensions, ghost-free
- 5 Lagrangians which can be parametrized with the α functions that give more physical insight.
- Quasistatic limit yields modifications of gravity that depend only on time and k^2.

In a homogeneous and isotropic universe, the distance element assumes the form:

$$ds^2 = a(\tau)[-(1+2\Psi)d\tau^2 + (1-2\Phi)\delta_{ij}dx^i dx^j]\,, \tag{2.1}$$

where the conformal time τ is related to the cosmic time t by $d\tau = dt/a(\tau)$ and Ψ and Φ are the two scalar gauge invariant gravitational potentials, related (modulo sign) to the Bardeen potentials.

Horndeski models include (almost) all theories with second order equations of motion in the fields and can be generically described with five functions of time. They include a variety of particular models studied in literature, characterised by a universal coupling to matter fields.

The Lagrangian can be written as ...

where ... is the matter Lagrangian and ... are arbitrary functions of the scalar field ϕ, whose canonical kinetic term is $X = -\frac{1}{2}\phi^{;\mu}\phi_{;\mu}$; the latter is also coupled to the Ricci scalar and the Einstein tensor.

In the following, we will adopt the conformal Newtonian gauge in the notation of (**MaBertschinger**), so that the gauge invariant potentials Ψ and Φ are also equal to the gravitational potentials in the Newtonian limit. In this formulation, the cold dark matter (CDM, c) perturbation equations for the density contrast δ and the velocity divergence θ can be written in Fourier space in the usual way:

$$\dot{\delta}(\tau, \mathbf{k}) = -(\theta(\tau, \mathbf{k}) - 3\Phi(\tau, \mathbf{k})) \ , \tag{2.2}$$

$$\dot{\theta}(\tau, \mathbf{k}) = -H\theta(\tau, \mathbf{k}) + k^2\Psi(\tau, \mathbf{k}) \ . \tag{2.3}$$

Derivating the first of the above equations with respect to τ and eliminating $\dot{\theta}$ and θ using both equations, one obtains the usual second order equation in time for the density contrast:

$$\ddot{\delta}(\tau, \mathbf{k}) = -(H\dot{\delta}(\tau, \mathbf{k}) - 3H\dot{\Phi}(\tau, \mathbf{k}) + k^2\Psi(\tau, \mathbf{k}) - 3\ddot{\Phi}(\tau, \mathbf{k})) \ . \tag{2.4}$$

In the following, we will stay within the quasi-static limit, i.e restrict to scales much smaller than the cosmological horizon ($k/aH \gg 1$) and well inside the Jeans length of the scalar field $c_s k \gg 1$, so that terms containing k dominate over terms containing time derivatives. In this case, Eq. 2.4 reduces to:

$$\ddot{\delta}(\tau, \mathbf{k}) + H\dot{\delta}(\tau, \mathbf{k}) + k^2\Psi(\tau, \mathbf{k}) = 0 \, . \tag{2.5}$$

The gravitational potential is then obtained from the Poisson equation, that relates the gravitational potential to matter and relativistic species in the universe, by perturbing the Einstein field equations to first order.

In Horndeski theories, in the quasi-static limit, the deviation of the gravitational potentials from General Relativity has a known scale dependence in the Poisson equation, given by (**...**):

$$\mu(k, a) \equiv -\frac{2k^2\Psi}{3\Omega_m\delta_m} = h_1 \left(\frac{1 + (k/k_*)^2 h_5}{1 + (k/k_*)^2 h_3}\right) \tag{2.6}$$

where h_1, h_3, h_5 are functions of time only. A similar expression, with different time dependent coefficients h_2, h_4, holds for the gravitational slip η:

$$\eta(k, a) \equiv -\frac{\Phi}{\Psi} = h_2 \left(\frac{1 + (k/k_*)^2 h_4}{1 + (k/k_*)^2 h_5}\right) \ . \tag{2.7}$$

In this work we will consider all h_i functions as constants, for simplicity, although in general they are all functions of time. k_* is an arbitrary pivot scale, chosen at large scales close to the horizon, in order to fulfil the quasi-static limit.

As it can be seen from eqns. 2.6 and 2.7, the functions h_5 and h_3 are degenerate as well as h_4 and h_5. ΛCDM case is recovered when $\mu = \eta = 1$, which implies $h_5 = h_4 = h_3$ and an amplitude $h_1 = h_2 = 1$.

In the following, we will focus on CDM perturbations and deal only with the function $\mu(k, a) = \mu(k)$ (although perturbations require two functions of time and scale to fully specify the model). In particular, we will test the following cases:

1. Scale independent μ with modified amplitude: $h_5 = h_3$ and $h_1 > 1$ or $h_1 < 1$.

2. Scale dependent μ with unity amplitude: $h_1 = 1$ and $h_5 > h_3$ or $h_5 < h_3$.

3. Scale dependent μ with modified amplitude: $h_1 \neq 1$ and $h_5 > h_3$ or $h_5 < h_3$.

In logarithmic space $\mu(k)$ is very close to a step function, as shown in figure **??**, so that we will only use values for the h_i in which the difference of the minimum and the maximum is of the order of 15%, since otherwise we would impose a rather unrealistic structure formation.

2.1.4 Effective Field Theory of Dark Energy

delete unless there is any result you have on EFT

- Another approach to construct a general model for dark energy based on the allowed symmetries in the action.

- Can be parametrized with the α functions.

- Allows in a simple way to create theories beyond Horndeski and to accomodate non-universal couplings.

2.2 Non-universal coupling

- Motivate dark energy, the coincidence problem and a dark sector interaction

- Explain why baryons should be uncoupled (Solar system constraints)

2.2.1 Coupled Dark Energy

- Introduce coupling at the energy momentum tensor

- Linear perturbations, gravitational bias DM-baryons, signatures in LSS, non-linear perturbations

- Talk about present constraints

2.2.2 Growing Neutrino Quintessence

- Motivate coupling to neutrinos (energy scale of DE and coincidence problem), particle physics seesaw mechanism perspective, non-relativistic neutrinos

- Background equations, linear perturbations and instabilities, failure of linear theory to take into account structures, lumps and backreaction

- Non-linear perturbations, simulations, neutrino lump dynamics, heating of neutrinos

2.2.3 Non-universal couplings in EFT

delete this paragraph Paper by Gleyzes

Chapter 3

Statistics in Cosmology

3.1 The relation between Gaussianity, linearity and homogeneity

3.1.1 Gaussian random fields

In cosmology, we are interested in studying observables which are the product of a large number of random processes. The best examples are the fluctuations of temperature in the CMB $\delta T/T$ or the fluctuations of the matter density in the Universe, usually called δ, which originate from complicated interactions between different matter species, gravity and other fundamental forces. Using the Central Limit Theorem **(cite...)** we can show that the sum of a large number of random variables, that originate from independent random processes, will yield a variable that it is almost exactly Gaussian distributed. The Gaussian probability distribution $p(x)$ for a random variable x, with mean μ and variance σ is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3.1}$$

As we sill see, Gaussian distributed variables will be very important in observations and analysis of cosmological observables. In the case of structure formation, δ is a random field that describes inhomogeneities, however, due to the assumption that it originates from independent random processes in a homogeneous and isotropic universe, its statistical properties must be homogeneous and isotropic too **(cite...Peacock; Bjoern; ?)** That means that the probability distribution $p(\delta(\vec{x}))$ must be invariant under translations and rotations of space. The expectation value of $\langle\delta^n\rangle$ is formally speaking defined by an ensemble average, meaning that we should actually observe this process in many random realizations of the Universe. Since this is not possible, in cosmology we postulate the ergodicity principle **(cite Adler...)** which states that for sufficient large volumes the ensemble average is equal to the volume average and therefore, we can write:

$$\langle\delta^n\rangle = \frac{1}{V}\int_V \mathrm{d}^3x\, \delta^n(\vec{x}) p(\delta(\vec{x})) \tag{3.2}$$

The importance of this postulate in practical terms is that we can use our formal statistical tools to compute theoretically the expectation values of observables the sentence from here on is not clear, what matches what? why do you need this paragraph? and on the observational side, we just need to match this by observing over large volumes in the sky.

3.1.2 The two-point correlation function

The advantage of working with fluctuation variables in cosmology is that their mean is zero, i.e. $\langle\delta\rangle = 0$ mmm I am not sure what do you mean here; in general please only say things that you really need and that they remain small ($\delta << 1$) for most of their evolution time. Therefore what makes sense to study too colloquial, please use formal language is not its mean, but its variance, $\langle\delta^2\rangle$, or if we are measuring the fluctuations at two separated points in the universe ($\vec{x}$ and $\vec{y}$), then it makes sense to study their two-point correlation function, defined as:

$$\xi(\vec{x}, \vec{y}) = \langle\delta(\vec{x})\delta(\vec{y})\rangle \tag{3.3}$$

As we have seen before, δ is statistically homogeneous and isotropic. Therefore, due to homogeneity, the correlation function can only depend on separation $\vec{r} = \vec{x}_2 - \vec{x}_1$ x1, x2 not defined, we had x and y and r and due to invariance under rotations it can only depend on the magnitude of the separation, denoted as r. In simple words—for the case of large scale structure—, $\xi(r)$ is telling us how probable it is to find a galaxy a distance r away from a given galaxy at position $\vec{x}$. We are talking here about too informal ideal measurements for a linear and Gaussian observable, but in the later chapters, when we discuss about non-linearities and observational effects (for example redshift space distortions), we will see that the situation is in practice not so simple and the observable δ, will not be specified only by its two-point correlation function and will not be statistically isotropic, leading to different correlation functions parallel and perpendicular to the line of sight **(cite some BAO stuff)**

3.1.3 The data covariance matrix

Let us introduce at this point the *data* covariance matrix, not to be confused with the *parameter* covariance matrix, which will be introduced later. never use sentences like: since this is a concept that causes a lot of misunderstanding in the literature; just explain what you think it's good, there is no need to criticize others If we are measuring the fluctuation δ at three points in space, $\vec{x}, \vec{y}$ and $\vec{z}$, then their joint probability density will still be Gaussian:

$$p(\delta(\vec{x}), \delta(\vec{y}), \delta(\vec{z})) = \frac{1}{\sqrt{(2\pi)^2 \det \mathbf{C}}} \exp\left(-\frac{1}{2}\begin{pmatrix}\delta(\vec{x})\\ \delta(\vec{y})\\ \delta(\vec{z})\end{pmatrix}^T \mathbf{C}^{-1}\begin{pmatrix}\delta(\vec{x})\\ \delta(\vec{y})\\ \delta(\vec{z})\end{pmatrix}\right) \tag{3.4}$$

where now instead of a Gaussian dispersion σ, we have a Gaussian *data* covariance matrix $\mathbf{C}$, given by:

$$C_ij = \begin{pmatrix} \langle\delta^2(\vec{x})\rangle & \langle\delta(\vec{x})\delta(\vec{y})\rangle & \langle\delta(\vec{x})\delta(\vec{z})\rangle \\ \langle\delta(\vec{x})\delta(\vec{y})\rangle & \langle\delta^2(\vec{y})\rangle & \langle\delta(\vec{y})\delta(\vec{z})\rangle \\ \langle\delta(\vec{x})\delta(\vec{z})\rangle & \langle\delta(\vec{y})\delta(\vec{z})\rangle & \langle\delta^2(\vec{z})\rangle \end{pmatrix} \tag{3.5}$$

Notice that the covariance matrix is symmetric and positive definite. This is the data covariance matrix for three measurements in real space. If you would measure the density fluctuations at 1000 different points in space, the covariance matrix would have 1 million elements. Therefore, it is very important not only for theoretical, but for practical reasons of data analysis, that the covariance matrix is as simple and symmetric as possible.

Luckily, for Gaussian random fields their statistical properties can be defined entirely by the their two-point correlation function, where the following recursion relations (obtained by Isserli's or Wick's theorem **(cite Wick or something)**) are valid:

$$\langle \delta^{2n} \rangle = (2n-1)!! \langle \delta \rangle \,, \quad \langle \delta^{2n+1} \rangle = 0 \tag{3.6}$$

Is this at a fixed position in space? Then isn't it just one of the 1 million elements in the matrix? then why is this more convenient? This is of course expected, since we have defined in Eqn. 3.4 the covariance matrix **C** as the only statistical quantity needed to specify fully the distribution of the multivariate Gaussian random field.

3.1.4 The power spectrum

Since we want observations of δ to be independent of each other, we need that the equations governing the spatial distribution of the density fluctuations are linear, otherwise non-linear terms can induce correlations between values at different positions Reformulate: we consider linear perturbations which have the advantage that they can be treated separately for each k; there is no fundamental reason why we 'want' perturbations to be independent, it's just more complicated. We can see this more clearly if we work in Fourier space. For linear equations, transforming into Fourier space (going from position $\vec{x}$ to wavevectors $\vec{k}$) is quite useful, since space derivatives become simply remove adjectives like 'simply' 'just' 'obviously' 'easy' 'straightforward' and similar factors of $\vec{k}$ and the Fourier space function $\delta(\vec{k})$ evolves in time, independently for each k-mode, but only if the field is statistically homogeneous.

The Fourier transform of the density field, denoted $\delta(\vec{k})$ is defined as:

$$\delta(\vec{k}) = \int \mathrm{d}^3 x \delta(\vec{x}) \exp(-i\vec{k} \cdot \vec{x}) \tag{3.7}$$

and its inverse is

$$\delta(\vec{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi^3)} \delta(\vec{k}) \exp(i\vec{k} \cdot \vec{x}) \quad . \tag{3.8}$$

Since now $\delta(\vec{k})$ can be complex, we can define the variance between two k-modes as:

$$\langle \delta(\vec{k}_1) \delta^*(\vec{k}_2) \rangle = \int \mathrm{d}^3 x \, \mathrm{d}^3 y \langle \delta(\vec{x}) \delta(\vec{y}) \rangle \exp(-i\vec{k} \cdot (\vec{x} - \vec{y}))(2\pi)^3 \delta_D(\vec{k}_1 - \vec{k}_2) \quad . \tag{3.9}$$

Using the definition of the correlation function 3.3 and the fact that the Gaussian random field δ is homogeneous we can define the power spectrum $P(\vec{k})$ as:

$$P(\vec{k}) = \int \mathrm{d}^3 r \, \xi(\vec{r}) \exp(-i\vec{k} \cdot \vec{r}) \tag{3.10}$$

The relation between the expected value of $\langle \delta(\vec{k}_1) \delta^*(\vec{k}_2) \rangle$ and the power spectrum $P(\vec{k})$ can be obtained by a similar straightforward calculation, **(cite Luca**; **Dodelson**; **??)** where taking into account that the δ field is real-valued, one can obtain:

$$\langle \delta(\vec{k}_1) \delta(\vec{k}_2) \rangle = (2\pi)^3 P(k) \delta_D(\vec{k}_1 - \vec{k}_2) \quad . \tag{3.11}$$

The Dirac delta δ_D in the previous equation, shows us directly what we had expected from a random Gaussian and homogeneous field; the Fourier modes $\vec{k}_1$ and $\vec{k}_2$ for the density fluctuations are mutually uncorrelated. In this case you mean: at the linear level, the data covariance matrix $\mathbf{C}$, defined in Eqn. 3.5, would be would be -> is perfectly diagonal in Fourier space, making an analysis of the observations much more tractable.

However, this is only true if the equations governing δ are linear; otherwise, non-linear terms in real space, would introduce convolutions in Fourier space, giving rise to non-vanishing correlations $\langle\delta(\vec{k}_1)\delta(\vec{k}_2)\rangle$. This will be the subject of study of the non-linear power spectrum, which will be treated with more depth in chapter 6.

3.1.5 Final remarks on linearity, Gaussianity and homogeneity

We have seen in the previous sections, that the concepts of linearity, Gaussianity and homogeneity are strongly related in Cosmology. The fact that observables like the matter density fluctuations originate from a large number of independent random processes, leads—thanks to the Central Limit Theorem—to a Gaussian random field. Then homogeneity and isotropy of the Universe and linear equations (in space) governing those processes, ensure that the statistical properties of the Gaussian random field are also homogeneous and isotropic and that we can use the ergodic theorem to compute ensemble averages. The Gaussianity of the field (and its vanishing average value) allows us to describe the field, solely with its two point correlation function, which directly connects us to the power spectrum and then to the fact that in Fourier space, pairs of wavemodes are mutually uncorrelated.

$$\text{Gaussianity} \Leftrightarrow \text{Linearity} \Leftrightarrow \text{Homogeneity} \tag{3.12}$$

In contrast, non-linear evolution of the density field, would yield correlations among different k-modes, leading to a non-Gaussian probability distribution and the loss of homogeneity (or the appearance of k-dependent growth). This is what happens at later stages of structure formation, where fluctuations are limited to $-1 < \delta$ in voids, but can grow to very high values $\delta \gg 1$ inside the cosmic web structure. This in turn will make decide the tense; preferably se present everywhere the data covariance matrix $\mathbf{C}$ in Fourier space not diagonal anymore and the analysis of cosmological observations more difficult.

3.2 Likelihood and the Bayesian approach

The subject of Bayesian statistics is a very complex, but very important topic in modern statistics, data science and the physical sciences **cite Gregory and others** and it is outside of the scope of this work to try to even review its more fundamental properties. We will just comment on the simple but very powerful concept of the likelihood function and how the Bayesian approach to statistical analysis is best suited in a field like cosmology.

I would start the paragraph from here; avoid this very general comment on what is important, they are too vague and subjective If an experiment yields a vector of random variables x^i, we have to try to build a theoretical model, which depends on some free parameters θ^i that can give us the probability distribution $p(x^i|\theta^i)$ to find the variable x^i inside an interval Δx for a specific measurement.

Then what physicists are interested in, is in constraining the model parameters θ and try to find the set of parameters (or the model) that best explains the data.

There are two possibilities of analyzing the results of that experiment, the *frequentist* and the *Bayesian* approach.

In the *frequentist* approach, the scientist would repeat many times the experiment and even change the settings and the theoretical parameters θ in the lab, in order to find the distribution of x^i's ($p(x^i|\theta^i)$) and from there, infer the distribution of the parameters θ. So this scientist would be able to say (as it is usually done in particle physics) that a there is a 97% probability that her data is distributed according to the model she has found as best fit (where one parameter is for example the Higgs mass).

In the *Bayesian* approach, the reasoning is reversed. The scientist would look for the probability of having a certain model (with model parameters) given the data that has already been taken ($p(\theta^i|x^i)$). This is more suitable for cosmology, since we cannot repeat many times the same experiment and we cannot change the parameters of the Universe.

How do we then "invert" $P(D|T)$ —the probability of the data given a model— to obtain $P(T|D)$? The solution is given by Bayes' theorem **Bayes**; **Dodelson**; **Luca**; **many**

$$P(T|D) = \frac{P(D|T)P(T)}{P(D)} \tag{3.13}$$

$P(T|D)$ is usually called the posterior probability (of having a theory T given the data D), $P(T)$ is the prior probability on the theory (which quantifies our previous knowledge and prejudices) and $P(D)$ is the evidence (the probability of the data), which for all purposes we can ignore, since it represents just a normalization factor. Also here, a discussion on the significance and the effects of the prior is out of our scope. **cite Amendola**; **Dodelson**; **Gregory**; **reviews** However, in our case it will be an efficient way of combining results from previous experiments into our predictions for future experiments. The prior can be as simple as the theoretical expectation that a parameter cannot be zero (for example the matter density of the Universe) or it can be the entire probability distribution of a previous experiment. After this point we will roughly call the posterior $P(T|D)$ by its more common name: the likelihood function $L(\theta|x)$.

From the likelihood we can obtain some useful quantities:

- The maximum likelihood estimators $\hat{\theta}_i$, which are found by solving $\partial L(\theta_i)/\partial\theta_i = 0$. This would amount to the 'best fit' parameters, but they are in general different tp the frequentist ones, if we have used a prior.

- The confidence regions for the parameters, denoted $R(\alpha)$, for which the normalized likelihood integrated in that region has a specific value:

$$\int_{R(\alpha)} L(\theta_i)\mathrm{d}^n\theta = \alpha \quad . \tag{3.14}$$

  If $\alpha = 0.683, 0.954, 0.997$, this regions are called 1, 2 and 3σ regions respectively.

- Marginalization over a parameter. If the likelihood depends on three parameters $\theta^1, \theta^2, \theta^3$, but we are not interested in θ^3, because it is a nuisance parameter or we have no clear idea how to relate it to the theory, we can marginalize over it by integrating it out: $L(\theta_1, \theta_2) = \int L(\theta^1, \theta^2, \theta^3)\,\mathrm{d}\theta^3$

Despite the great usefulness of the likelihood, its evaluation represents a challenge both numerically and computationally, since it is a complicated multi- dimensional function. Evaluating the likelihood for a relatively coarse grid of points in parameter space, becomes unfeasible for more than 7 or 8 parameters. So, techniques like Markov Chain Monte Carlo (MCMC) **cite some basic MCMC literature** have become more and more sophisticated. In this techniques, the basic idea is to explore the parameter space using a random walk, which is itself guided by the steepness or flatness of the likelihood at that point. In this way there are much less evaluations in places where the likelihood is smooth and many evaluations where the likelihood changes very rapidly. Two very well known codes in the cosmological community are MONTEPYTHON and COSMOMC **cite Cosmomc and Montepython** which are used to evaluate the likelihood of many large and modern experiments like *Planck* **cite Planck**

Therefore many approximations can be made in order to estimate the likelihood or at least to find its maximum. For a Gaussian distributed data vector $\boldsymbol{x}$ with n components, with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{C}$ of dimension $n \times n$ we can write the likelihood as $L = \exp(-\chi^2/2)$, which results in:

$$L = \frac{1}{(2\pi)^{n/2} \det(\boldsymbol{C})} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \tag{3.15}$$

and we can define the matrix of data points as $\boldsymbol{D} = (\boldsymbol{x}-\boldsymbol{\mu})^T(\boldsymbol{x}-\boldsymbol{\mu})$ and in full generality the covariance matrix $\boldsymbol{C}$ is just given by the expected value of $\boldsymbol{D}$, namely $\boldsymbol{C} = \langle \boldsymbol{D} \rangle$.

We will see in the next section that a very useful approximation is to say that the likelihood is Gaussian not only in the data, but also in the parameters. In this case, a useful quantity will be the log-likelihood $\mathcal{L} \equiv \ln L$, since we can get rid of the exponential function. For (approximately) Gaussian distributed parameters, the parameter covariance matrix will be of extremely importance to analyze and forecast the results of large future experiments.

3.3 Fisher Matrix formalism

We have seen in the previous section, how the log-likelihood function $\mathcal{L}$ is a powerful tool to estimate the probability distribution of model parameters, given some previously measured data. But what if we don't have data avilable yet? For example when we want to know with which accuracy a certain future experiment is going to be able to constrain certain parameters of a model. This is where the Fisher formalism **(cite Fisher**; **Tegmark**; **etc.)** enters. The Fisher matrix is defined as the expectation value of the *curvature* of the log-likelihood:

$$F_{ij} = -\left\langle \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \theta_i \theta_j} \right\rangle \tag{3.16}$$

If we assume that the errors on the data measurements are Gaussianly distributed (where then we can write the likelihood as $L = e^{-\chi^2/2}$) then we can also think of the Fisher matrix as a Taylor expansion of the log-likelihood around its maximum $\hat{\theta}$, or equivalently where the χ^2 is a minimum:

$$\mathcal{L}(\theta) = \mathcal{L}(\hat{\theta}) + (\theta - \hat{\theta})\frac{\partial \mathcal{L}}{\partial \theta} + \frac{1}{2}(\theta - \hat{\theta})^2 \frac{\partial^2 \mathcal{L}}{\partial \theta^2} \tag{3.17}$$

The linear derivative term in the above equation vanishes, since we are expanding around the maximum and we are considering here for simplicity just one parameter θ; but the generalization to more parameters is straightforward. The fact that the Fisher matrix is related to the curvature of the log-likelihood around its maximum, tells us that the Fisher matrix is an indication of how fast the likelihood changes around the peak **cite Dodelson...** If the curvature is high, then the likelihood changes fast and the experiment is very constraining, allowing for just small changes of the parameters, before the likelihood becomes too small. If the curvature is low, then the likelihood is very flat and the experiment is not very constraining in parameter space [1].

3.3.1 The Fisher matrix for a galaxy power spectrum

In order to understand better the following sections where we will talk about the Fisher matrix of the galaxy clustering and weak lensing observables, let us specify the Fisher matrix defined in Eqn. 3.16 in a more concrete way. In this section we follow closely the argumentations of **cite Dodelson** and **cite Amendola**

We are interested in finding the distribution of matter in the Universe, and as we have seen before, for large sufficient scales, the overdensity of matter at a specific scale k is given by $\delta(k)$. For a galaxy survey, covering a volume $V(z+\Delta z)$ in three-dimensional space and providing m Fourier modes k_i in a redshift bin Δz, we can compute (the full details of the calculation can be found in **cite Dodelson**) that the data covariance matrix between mode k_i and mode k_j is given by:

$$V\langle\delta_{k_i}\delta^*_{k_j}\rangle = C_{k_i k_j} = \frac{\delta_{ij}}{V}\left(P(k_i, z) + \frac{1}{\bar{n}(z+\Delta z)}\right) \tag{3.18}$$

The data covariance $C_{k_i k_j}$ is always a sum of the signal covariance and the noise covariance, which in this case is simply given by the Poisson noise of a discrete random distribution, where $\bar{n}$ is the average number density of galaxies. Again, let us emphasize that we are at relatively large scales, where then we can take our data covariance matrix as being diagonal. If we now assume that the galaxy distribution is well approximated by a Gaussian distribution (and we know it has to be since we are at linear and homogeneous scales), then we can write the likelihood function as:

$$L = \frac{1}{(2\pi)^{m/2}\det(C_{k_i k_j})}\exp\left[-\frac{1}{2}\sum_i^m \frac{\delta^2_{k_i}}{C_{k_i k_j}}\right] \tag{3.19}$$

Then the log-likelihood can be written in the following way

$$\mathcal{L} = -\ln L = \frac{m}{2}\ln(2\pi) + \sum_i^m \ln(C_{k_i k_i}) + \sum_i^m \frac{\delta^2_{k_i}}{2C_{k_i k_i}} \tag{3.20}$$

where we have used the matrix identity: $\ln\det C = \operatorname{tr}\ln C$, which in this case is anyway trivial since C is a diagonal matrix. Now, using definition 3.16, we can

[1] The relation between information matrices and geometry is not only qualitative, it is a formal field of study called Information Geometry **cite Springer book**

evaluate the expectation value of the curvature of the log-likelihood (in model-parameter space θ_α):

$$F_{\alpha\beta} = \sum_i^m \left[\frac{\partial^2 \ln(C_{k_i k_i})}{\partial\theta_\alpha \partial\theta_\beta} + \langle \delta_{k_i}^2 \rangle \frac{\partial}{\partial\theta_\alpha \partial\theta_\beta} \left(\frac{1}{2C_{k_i k_i}} \right) \right] \tag{3.21}$$

The expectation value operator only affects the data $\delta_{k_i}^2$, because the data covariance matrix C is already itself formed by expectation values. On the other hand, the data $\delta_{k_i}^2$ cannot be affected by derivations with respect to the model parameters, because pure measurements should be by definition model independent. Then substituting $\langle \delta_{k_i}^2 \rangle = C_{k_i k_i}$ and doing some algebra, we find:

$$F_{\alpha\beta} = \sum_i^m \left[-\frac{1}{(C_{k_i k_i})^2} \frac{\partial C_{k_i k_i}}{\partial\theta_\alpha} \frac{\partial C_{k_i k_i}}{\partial\theta_\beta} + \frac{3}{(C_{k_i k_i})^2} \frac{\partial C_{k_i k_i}}{\partial\theta_\alpha} \frac{\partial C_{k_i k_i}}{\partial\theta_\beta} \right] \quad . \tag{3.22}$$

Now we can use the data covariance of Eqn. 3.18 to write the Fisher matrix in terms of the matter power spectrum:

$$F_{\alpha\beta} = \frac{1}{2} \sum_i^m \frac{\partial \ln P(k_i)}{\partial\theta_\alpha} \frac{\partial \ln P(k_i)}{\partial\theta_\beta} \left(\frac{\bar{n} P(k_i)}{1 + \bar{n} P(k_i)} \right)^2 \quad , \tag{3.23}$$

where we have assumed that the noise does not depend on the cosmological parameters and we have dropped the redshift dependence for simplicity. We must remember that for a tomographic redshift survey, this Fisher matrix has to be computed at each redshift bin.

Having done this example explicitly we can also write down the more general expression for Gaussian likelihoods, such as the one specified in Eqn.3.15, in which the likelihood contains covariance matrices $\boldsymbol{C}$ which are not diagonal and the average of the data measurements $\boldsymbol{\mu}$ is not zero. (see **cite Luca**; **Dodelson** among other works, for more details). In this case the Fisher matrix can be written as:

$$F_{\alpha\beta} = \frac{1}{2} \mathrm{Tr} \left[\boldsymbol{C}^{-1} \boldsymbol{C}_{,\alpha} \boldsymbol{C}^{-1} \boldsymbol{C}_{,\beta} + \boldsymbol{C}^{-1} \langle \boldsymbol{D}_{,\alpha\beta} \rangle \right] \tag{3.24}$$

This last term, the expected value of the derivatives of the data matrix $\boldsymbol{D}$, is what usually is zero in cosmology, since we study quantities which are fluctuations around the mean and their mean is therefore identically vanishing. In the following sections we will write down the more specific and exact expressions for the Galaxy Clustering and Weak Lensing analysis.

3.4 Fisher Matrix forecasts

The Fisher matrix formalism, which was mainly developed for cosmological observations by **tegmark_measuring_1998** and **seo_baryonic_2005** is one of the most popular tools to forecast the outcome of an experiment, because of its speed and its versatility when the likelihood is approximately Gaussian. The Fisher matrix method is used now ubiquitously in the cosmology literature, with about 200 papers published so far [2], ranging from CMB observations, to Redshift Space Distortions, Supernovae, Lyman-α observations and many more.

[2] According to a full-text search at https://www.arxiv.org

We will present in the following sections, the specific formulas for Galaxy Clustering and Weak Lensing, under some simplifying assumptions (Gaussian likelihood, diagonal data covariance matrices, no redshift-bin correlations). In a paper by **sellentin_breaking_2014** the Gaussian approximation has been dropped and higher order "Fisher tensors" are used to approximate the true underlying likelihood **cite Sellentin more** Also recently, there has been some work in extending the simplifying assumptions of diagonal covariance, redshift-bin correlations and k-space correlations **bailoni_improving_2016**; **Durrer**; **Amendola with Blot**

In **Khedekar**; **Majumdar** the authors also compared directly the Fisher matrix formalism with a direct MCMC approach and have encountered considerable differences in the case of CMB experiments, especially when departing from the standard cosmological model.

3.4.1 Fisher matrix for Galaxy Clustering

how is this different from 3.3.1, I would merge Fisher formalism and fisher forecast sections Using the previously found equation 3.23 for the Fisher matrix of the power spectrum, we can write that expression in a more compact form, where we express the sum over the k-modes as an integral and take into account the number of observed modes into the volume. We end up with the following formula: form (**seo_improved_2007**; **amendola book**; **dodelson**):

$$F_{ij} = \frac{1}{8\pi^2} \int_{-1}^{+1} \mathrm{d}\mu \int_{k_{\rm min}}^{k_{\rm max}} \mathrm{d}k\, k^2 \frac{\partial \ln P_{\rm obs}(k,\mu,z)}{\partial \theta_i} \frac{\partial \ln P_{\rm obs}(k,\mu,z)}{\partial \theta_j} V_{\rm eff}(k,\mu,z) \quad , \tag{3.25}$$

where the effective volume $V_{\rm eff}$ is given by

$$V_{\rm eff} = V_{\rm survey} \left[\frac{n(z) P_{\rm obs}(k,\mu,z)}{n(z) P_{\rm obs}(k,\mu,z) + 1} \right]^2 \quad . \tag{3.26}$$

Here $V_{\rm survey}$ is the volume covered by the survey and contained in a redshift slice Δz, while $n(z)$ is the galaxy number density as a function of redshift. The galaxy number density has to be taken from specifications for each survey, which are obtained by direct measurements in the sky and end-to-end simulations **Dida**; **Guzzo**; **etc** $P_{obs}(k,\mu,z)$ is the observed galaxy power spectrum as a function of redshift z, the wavenumber k and of $\mu \equiv \cos\alpha$, where α is the angle between the line of sight and the 3D-wavevector $\vec{k}$. The derivatives in eq. (3.25) are taken with respect to a vector of cosmological parameters, θ_i that can contain redshift-independent parameters such as h and Ω_m or redshift-dependent parameters such as the growth rate function $f(z)$ or the Hubble function $H(z)$. The details of the observed power spectrum will be treated in the next section 3.4.2.

The minimum and maximum values for the k-integral also depend on the survey specifications and on how much we can trust and rely on non-linear scales. Throughout this work we will use for the smallest wavenumber a value of approximately $k_{\rm min} = 0.008$h/Mpc, while the maximum wavenumber will depend on the specific application and methods used to study the non-linear regime. In general, $k_{\rm max} = 0.10 - 0.15$h/Mpc for linear forecasts and for non-linear forecasts the smallest scales will lie at about $k_{\rm max} = 0.5 - 1.0$h/Mpc.

3.4.2 The observed galaxy power spectrum

The distribution of galaxies in space is not perfectly uniform. Instead it follows, up to a bias, the underlying matter power spectrum so that the observed power spectrum P_{obs} is closely linked to the dark matter power spectrum $P(k)$. The observed power spectrum is the Fourier transform of the real-space two point correlation function (see eq. (3.3)) of the galaxy number overdensity.

The observed power spectrum is then built from the theoretical matter power spectrum $P(k, z)$ (which in turn depends on the fundamental cosmological parameters θ_i) together by the following contributions which modify its signal:

1. The geometrical factor coming from the change of the Baryon Acoustic Peak (BAO) peak
   marked orange in eq. (3.27). It can be shown **cite some BAO review** that the shift of the BAO peak due to geometrical distortions when considering a different cosmology as the reference one is proportional to $H(z)/D_A^2(z)$.

2. The galaxy bias $b(z)$ marked brown in the formula below. This term has to be estimated from observations and simulations for different types of galaxy populations. In this case, we take into account just a linear local bias which means that we assume that the galaxy overdensity δ_g differs from the underlying matter overdensity by a function which just depends on redshift: $\delta_g = b(z)\delta_m$. This assumption breaks at very large and very small scales **cite some bias review**

3. The contribution from Redshift Space Distortions (RSD)
   marked green in eq. (3.27), where $\beta(z) = f(z)/b(z)$, where $b(z)$ is the galaxy bias and $f(z)$ the logarithmic growth rate of perturbations $f(z) = \mathrm{d}\ln(D_+)/\mathrm{d}\ln(a)/$. This term is related to distortions in redshift space caused by peculiar velocity divergences. This term was derived first in (**kaiser_clustering_1987**) and it is commonly known as the Kaiser formula. It represents the distortion caused by peculiar velocities when going from redshift to coordinate space.

4. The geometrical effect of the change of cosmological parameters on the determination of the angles μ and the scale k, which is called the Alcock-Paczynski effect (**alcock1979anevolution**; **ballinger_measuring_1996**; **feldman_power_1994**) and it is marked red in eq. (3.27). In section 3.5.5 we will detail the corresponding formulas.

5. The damping due to spectroscopic redshift errors σ_z and the non-linear pairwise peculiar velocity dispersion $\sigma_v(z)$, which corresponds to a first order correction term to the Kaiser formula and it is also known as the "Fingers of God" effect. These terms damp the power spectrum at small scales, where due to these redshift uncertainties the signal cannot be reproduced accurately. These terms are shown in magenta in eq. (3.27).

6. Extra shot noise due to observational effects which cannot be removed. These can be included in general with the term $P_s(z)$. It is marked blue in the formula below.

$$P_{\rm obs}(z, k, \mu; \theta) = P_{\rm s}(z) + \frac{D_A^2(z)_{ref} H(z)}{D_A^2(z) H(z)_{ref}} b^2(z) \left(1 + \beta(z)\mu^2\right)^2 P(k, z) e^{-k^2\mu^2(\sigma_z^2/H(z) + \sigma_v^2(z))} \tag{3.27}$$

In the previous formula we have neglected relativistic contributions and further more detailed non-linear effects. For a quick review on those topics see **durrer**; **bonvin**; **taruya**; **scoccimarro** In general for the forecasts presented in this work, we will marginalize over the bias $b(z)$, which we will take as a different nuisance parameter at each redshift. We usually fix the spectroscopic redshift error to a value given by the specifications of the instrument, and it is in general a very small number $\sigma_z = 0.001$. We also marginalize over $\sigma_v(z)$, since we don't have a proper way of computing it for each model and relating it to fundamental cosmological parameters. We will take as a fiducial value $\sigma_v = 300$km/s, compatible with the estimates by **de_la_torre_modelling_2012**

Despite the apparent simplicity of eq. (3.25) and eq. (3.27) there are several details that need to be considered when dealing with forecasts in practical applications. One of the main results of this work is the production of a Fisher Matrix code capable of performing forecasts for Galaxy Clustering and Weak Lensing with different options for methods and assumptions. In section 3.5 we will explain more in detail the equations and the structure of the code used to perform the forecasts in this work.

3.4.3 Weak Gravitational Lensing

Light propagating through the universe is deflected by variations in the Weyl potential $\Phi_{\text{Weyl}} = \Phi + \Psi$, leading to distortions in the images of galaxies. The main idea behind weak gravitational lensing is to measure the ellipticity of galaxies in a wide field sky survey and correlate the ellipticities to find a statistical effect caused by the evolution of the lensing potential in the Universe.

The transformation of a light bundle in the case of small angles and small gravitational potentials can be described by a distortion tensor which contains in its diagonal the convergence component and in its off-diagonal elements the shear component (see the comprehensive review by **cite Bartelmann Schneider**). It can be shown that for linear theory, the only component observable in cosmology is the convergence field (see **Amendola**; **Tegmark**; **Hu**; **Bartelmann**) and its power spectrum can be calculated once the matter power spectrum and the evolution of the structures in the Universe in time are known.

In this regime we can write the power spectrum of the convergence field as

$$C_{ij}(\ell) = \frac{9}{4}\int_0^\infty \mathrm{d}z\, \frac{W_i(z)W_j(z)H^3(z)\Omega_m^2(z)}{(1+z)^4}\left[\Sigma(\ell/r(z),z)\right]P_m(\ell/r(z))\,. \tag{3.28}$$

In this expression we are already considered extensions of GR, where we have used Eqn. (4.4) to relate the Weyl potential to Σ and to the matter power spectrum P_m. Furthermore we use the Limber approximation to write down the conversion $k = \ell/r(z)$, where $r(z)$ is the comoving distance given by

$$r(z) = c\int_0^z \frac{d\tilde{z}}{H(\tilde{z})}\,. \tag{3.29}$$

The indices i, j stand for each of the $\mathcal{N}_{bin}$ redshift bins, such that C_{ij} is a matrix of dimensions $\mathcal{N}_{bin} \times \mathcal{N}_{bin}$. The window functions W_i are given by

$$W(z) = \int_z^\infty \mathrm{d}\tilde{z}\left(1 - \frac{r(z)}{r(\tilde{z})}\right)n(\tilde{z}) \tag{3.30}$$

where the normalized galaxy distribution function (in the case of a survey like Euclid 3.1) is

$$n(z) \propto z^2 \exp\left(-(z/z_0)^{3/2}\right) . \tag{3.31}$$

Here the median redshift $z_{\rm med}$ and z_0 are related by $z_{\rm med} = \sqrt{2}z_0$. We integrate this quantity in our redshift range to find the total amount of galaxies and create the redshift bins, such that each of them contains the same number of galaxies. These bins are then called equi-populated bins. The Weak Lensing Fisher matrix is then given by a sum over all possible correlations at different redshift bins (**tegmark_measuring_1998**),

$$F_{\alpha\beta} = f_{\rm sky} \sum_{\ell}^{\ell_{\rm max}} \sum_{i,j,k,l} \frac{(2\ell+1)\Delta\ell}{2} \frac{\partial C_{ij}(\ell)}{\partial\theta_\alpha} {\rm Cov}^{-1}_{jk} \frac{\partial C_{kl}(\ell)}{\partial\theta_\beta} {\rm Cov}^{-1}_{li} . \tag{3.32}$$

The prefactor $f_{\rm sky}$ is the fraction of the sky covered by the survey. The upper limit of the sum, $\ell_{\rm max}$, is a high-multipole cutoff due to our ignorance of clustering and baryon physics on small scales, similar to the role of $k_{\rm max}$ in Galaxy Clustering. In this work we choose $\ell_{\rm max} = 1000$ for the linear forecasts and $\ell_{\rm max} = 5000$ for the non-linear forecasts (this cutoff is not necessarily reached at all multipoles ℓ, as what matters is the minimum scale between $\ell_{\rm max}$ and $k_{\rm max}$, as we discuss below; see also **casas_fitting_2015**). In Eqn. (3.32), ${\rm Cov}_{ij}$ is the corresponding covariance matrix of the convergence power spectrum, which is the sum of the signal and the noise covariance matrix and it has the following form:

$${\rm Cov}_{ij}(\ell) = C_{ij}(\ell) + \delta_{ij}\gamma^2_{\rm int} n_i^{-1} + K_{ij}(\ell) \tag{3.33}$$

where $\gamma_{\rm int}$ is the intrinsic galaxy ellipticity. In Table 3.1 we cite some typical numbers for different surveys. The shot noise term n_i^{-1} is expressed as

$$n_i = 3600 \left(\frac{180}{\pi}\right)^2 n_\theta / \mathcal{N}_{bin} \tag{3.34}$$

with n_θ the total number of galaxies per arcmin^2 and the index i standing for each redshift bin. Since we have equi-populated redshift bins, the shot noise term is equal for each bin. The matrix $K_{ij}(\ell)$ is a diagonal "cutoff" matrix, discussed for the first time in **casas_fitting_2015** whose entries increase to very high values at the scale where the power spectrum $P(k)$ has to be cut to avoid the inclusion of uncertain or unresolved non-linear scales. We choose to add this matrix to have further control on the inclusion of non-linearities. Without this matrix, due to the redshift-dependent relation between k and ℓ, a very high $\ell_{\rm max}$ would correspond at low redshifts, to a very high $k_{\rm max}$ where we do not longer trust the accuracy of the non-linear power spectrum. Therefore, the sum in Eqn. (3.32) is limited by the minimum scale imposed either by $\ell_{\rm max}$ or by $k_{\rm max}$, which is the maximum wavenumber considered in the matter power spectrum $P(k, z)$. As we did for Galaxy Clustering, we use for linear forecasts $k_{\rm max} = 0.15$ and for non-linear forecasts $k_{\rm max} = 0.5$.

3.4.4 Future large scale galaxy redshift surveys

In this work we choose to present results on some of the future galaxy redshift surveys, which are planned to be started and analyzed within the next decade. Our baseline survey will be the Euclid satellite **amendola_cosmology_2013; laureijs_euclid_2011**

Euclid[3] is a European Space Agency medium-class mission scheduled for launch in 2020. Its main goal is to explore the expansion history of the Universe and the evolution of large scale cosmic structures by measuring shapes and redshifts of galaxies, covering 15000deg^2 of the sky, up to redshifts of about $z \sim 2$. It will be able to measure up to 100 million spectroscopic redshifts which can be used for Galaxy Clustering measurements and 2 billion photometric galaxy images, which can be used for Weak Lensing observations (for more details, see **amendola_cosmology_2013**; **laureijs_euclid_2011**). We will use in this work the Euclid Redbook specifications for Galaxy Clustering and Weak Lensing forecasts **laureijs_euclid_2011** some of which are listed in Tables 5.5 and 3.1 and the rest can be found in the above cited references.

Parameter	Euclid	SKA1	SKA2	Description
$f_{\rm sky}$	0.364	0.121	0.75	Fraction of the sky covered
σ_z	0.05	0.05	0.05	Photometric redshift error
n_θ	30	10	2.7	Number of galaxies per arcmin
$\gamma_{\rm int}$	0.22	0.3	0.3	Intrinsic galaxy ellipticity
z_0	0.9	1.0	1.6	Median redshift over $\sqrt{2}$
$\mathcal{N}_{bin}$	12	12	12	Total number of tomographic redshift bins

TABLE 3.1: Specifications for the Weak Lensing surveys Euclid, SKA1 and SKA2 used in this work. Other needed quantities can be found in the references cited in section 3.4.4. For all WL surveys we use a redshift range between $z = 0.5$ and $z = 3.0$, using 6 equi-populated redshift bins.

Another important future survey will be the Square Kilometer Array (SKA)[4], which is planned to become the world's largest radio-telescope. It will be built in two phases, phase 1 split into SKA1-SUR in Australia and SKA1-MID in South Africa and SKA2 which will be at least 10 times as sensitive. The first stage is due to finish observations around the year 2023 and the second phase is scheduled for 2030 (for more details, see **yahya_cosmological_2015**; **santos_hi_2015**; **raccanelli_measuring_2015**; **bull_measuring_2015**). The first phase SKA1, will be able to measure in an area of 5000deg^2 of the sky and a redshift of up to $z \sim 0.8$ an estimated number of about 5×10^6 galaxies; SKA2 is expected to cover a much larger fraction of the sky ($\sim$30000deg^2), will yield much deeper redshifts (up to $z \sim 2.5$) and is expected to detect about 10^9 galaxies with spectroscopic redshifts **santos_hi_2015** SKA1 and SKA2 will also be capable of performing radio Weak Lensing experiments, which are very promising, since they are expected to be less sensitive to systematic effects in the instruments, related to residual point spread function (PSF) anisotropies **harrison_ska_2016** In this work we will use for our forecasts of SKA1 and SKA2, the specifications computed by **santos_hi_2015** for GC and by **harrison_ska_2016** for WL. The numerical survey parameters are listed in Tables 5.5 and 3.1, while the galaxy bias $b(z)$ and the number density of galaxies $n(z)$, can be found in the references mentioned above.

We will also forecast the results from DESI[5], a stage IV, ground-based dark energy experiment, that will study large scale structure formation in the Universe through baryon acoustic oscillations (BAO) and redshift space distortions (RSD),

[3]http://www.euclid-ec.org/
[4]https://www.skatelescope.org/
[5]http://desi.lbl.gov/

using redshifts and positions from galaxies and quasars **desi_collaboration_desi_2016-1**; **desi_collaboration_desi_2016**; **levi_desi_2013** It is scheduled to start in 2018 and will cover an estimated area in the sky of about 14000deg^2. It will measure spectroscopic redshifts for four different classes of objects, luminous red galaxies (LRGs) up to a redshift of $z = 1.0$, bright [O II] emission line galaxies (ELGs) up to $z = 1.7$, quasars (QSOs) up to $z \sim 3.5$ and at low redshifts ($z \sim 0.2$) magnitude-limited bright galaxies (BLGs). In total, DESI will be able to measure more than 30 million spectroscopic redshifts. In this paper we will use for our forecasts only the specifications for the ELGs, as found in **desi_collaboration_desi_2016-1** since this observation provides the largest number density of galaxies in the redshift range of our interest. We cite the geometry and redshift binning specifications in Table 5.5, while the galaxy number density and bias can be found in **desi_collaboration_desi_2016-1**

Parameter	Euclid	DESI-ELG	SKA1-SUR	SKA2	Description
$A_{\rm survey}$	15000 deg^2	14000 deg^2	5000 deg^2	30000 deg^2	Survey area
σ_z	0.001	0.001	0.0001	0.0001	Spectroscopic error
$z_{\rm min,max}$	{0.65, 2.05}	{0.65, 1.65}	{0.05, 0.85}	{0.15, 2.05}	Min. and max. z
Δz	0.1	0.1	0.1	0.1	Δz in bin

TABLE 3.2: Specifications for the spectroscopic galaxy redshift surveys used in this work. The number density of tracers $n(z)$ and the galaxy bias $b(z)$, can be found for SKA in **santos_hi_2015** and for DESI in reference **desi_collaboration_desi_2016-1**

3.4.5 Covariance and correlation matrix and the Figure of Merit

Previously we have defined the data covariance matrix in eq. (3.5), now let us define in more generality the parameter covariance matrix for a *d-dimensional* vector p of model parameters as

$$\mathbf{C} = \langle \Delta p \Delta p^T \rangle \tag{3.35}$$

with $\Delta p = p - \langle p \rangle$ and the angular brackets $\langle\,\rangle$ representing an expectation value. For our Fisher matrix analysis we will assume that $\langle p \rangle$ is the value of the parameter p that maximizes the likelihood. The matrix $\mathbf{C}$, with all its off-diagonal elements set to zero, is called the variance matrix V and contains the square of the errors σ_i for each parameter p_i

$$\mathbf{V} \equiv diag(\sigma_1^2, ..., \sigma_d^2) \ . \tag{3.36}$$

The Fisher matrix $\mathbf{F}$ is the inverse of the parameter covariance matrix

$$\mathbf{F} = \mathbf{C}^{-1} \ . \tag{3.37}$$

The covariance matrix tells us not only the errors on each of the parameters p_i, but also how the errors are correlated among each other. This is more clearly seen by defining the correlation matrix $\mathbf{P}$, which is obtained from the covariance matrix $\mathbf{C}$, in the following way

$$P_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \ . \tag{3.38}$$

If the covariance matrix is non-diagonal, then there are correlations among some elements of p. We can observe this also by plotting the ellipsoidal contours corresponding to the confidence regions. The orientation of the ellipses can tell us if two variables p_i and p_j are correlated ($P_{ij} > 0$), corresponding to ellipses with 45 degree orientation to the right of the vertical line or if they are anti-correlated ($P_{ij} < 0$), corresponding to ellipses oriented 45 degrees to the left of the vertical line.

To summarize the information contained in the Fisher/covariance matrices we can define a Figure of Merit (FoM). The square-root $\sqrt{\det(\mathbf{C})}$ of the determinant of the covariance matrix is proportional to the volume of the error ellipsoid. We can see this if we rotate our coordinate system so that the covariance matrix is diagonal, $\mathbf{C} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots \sigma_d^2)$, then $\det(\mathbf{C}) = \prod_i \sigma_i^2$ and $(1/2)\ln(\det(\mathbf{C})) = \ln \prod_i \sigma_i$ would indeed represent the logarithm of an error volume. Thus, the smaller the determinant (and therefore also $\ln(\det(\mathbf{C}))$), the smaller is the ellipse and the stronger are the constraints on the parameters. We define

$$\mathrm{FoM} = -\frac{1}{2}\ln(\det(\mathbf{C}))\,, \tag{3.39}$$

with a negative sign in front such that stronger constraints lead to a higher Figure of Merit. The FoM allows us to compare not only the constraining power of different probes but also of the different experiments. As the absolute value depends on the details of the setup, we can define the relative figure of merit between probe a and probe b: $\mathrm{FoM}_{a,b} = -1/2\ln(\det(\mathbf{C_a})/\det(\mathbf{C_b})) = \mathrm{FoM}_a - \mathrm{FoM}_b$. The FoM has units of 'nits', since we are using the natural logarithm. These are similar to 'bits', but 'nits' are counted in base e instead of base 2.

An analogous construction allows us to study quantitatively the strength of the correlations encoded by the correlation matrix $\mathbf{P}$. We define the 'Figure of Correlation' (FoC) as:

$$\mathrm{FoC} = -\frac{1}{2}\ln(\det(\mathbf{P}))\,. \tag{3.40}$$

If the parameters are independent, i.e. fully decorrelated, then $\mathbf{P}$ is just the unit matrix and $\ln(\det(\mathbf{P})) = 0$. Off-diagonal correlations will decrease the logarithm of the determinant, therefore making the FoC larger. From a geometrical point of view, the determinant expresses a volume spanned by the vector of (normalized) variables. If these variables are independent, the volume would be maximal and equal to one, while if they are strongly linearly dependent, the volume would be squeezed and in the limit where all variables are effectively the same, the volume would be reduced to zero. Hence, a more positive FoC indicates a stronger correlation of the parameters.

3.4.6 The Kullback-Leibler divergence

In the previous section 3.4.5 we exploited the determinant of the covariance matrix C as a Figure of Merit for our forecasts. Here we summarize a possible alternative to measure the constraining power of a specific forecast, namely the Kullback-Leibler divergence **kullback_information_1951** also called relative entropy or information gain. It has been used in the field of cosmology for model selection, experiment design and forecasting, see among others **kunz_constraining_2006**; **raveri_cosmicfish_2016**; **seehars_information_2014**; **verde_planck_2013**; **Zhao2017** The KL-divergence $\mathcal{D}(p_2||p_1)$ measures for a continuous, d-dimensional random variable θ, the relative entropy between two probability density functions $p_1(\theta)$ and $p_2(\theta)$ and it is given by

$$\mathcal{D}(p_2||p_1) \equiv \int p_2(\theta)\ln\left(\frac{p_2(\theta)}{p_1(\theta)}\right)\mathrm{d}\theta. \tag{3.41}$$

Although it is not symmetric in p_1 and p_2 it can be interpreted as a distance between the two distributions and measures the information gain since it is non-negative ($\mathcal{D}(p_2||p_1) \geq 0$), non-degenerate ($\mathcal{D}(p_2||p_1) = 0$ if and only if $p_1 = p_2$) and it is invariant under re-parameterizations of the distributions as $p_1(\theta)\mathrm{d}\theta = p_1(\tilde{\theta})\mathrm{d}\tilde{\theta}$. In the form given here the information gain is measured in nits as in section 3.4.5, to convert nits to bits it is enough to divide the result by $\ln(2)$.

For the special case of $p_1(\theta)$ and $p_2(\theta)$ being multivariate Gaussian distributions, with the same mean values and covariance matrices $\mathcal{A}$ and $\mathcal{B}$ respectively, we obtain

$$\mathcal{D}(p_2||p_1) = -\frac{1}{2}\left[\ln\left(\frac{\det(\mathcal{A})}{\det(\mathcal{B})}\right) + \mathrm{Tr}\left[\mathbb{1} - \mathcal{B}^{-1}\mathcal{A}\right]\right]. \tag{3.42}$$

We can then define a Kullback-Leibler matrix $\mathcal{K}$, introduced in (**casas_mg_forec** appendix F) composed of the KL-divergence measure among our observables, defined as:

$$\mathcal{K}_{ij} = \mathcal{D}(p_j||p_i). \tag{3.43}$$

where p_i and p_j represent our observables. By looking at the rows of this matrix (one row for each observable) and plotting them in a corresponding matrix plot, one can see graphically which combination of observables yield more information gain. In section 4.6.1 we will show the visual representation of the KL-matrix for an explicit parameterization of modified gravity and for different Galaxy Clustering and Weak Lensing combinations of observables. This will give an intuitive idea of which observable contains more information gain than the other, compared to a reference observable.

3.4.7 Systematic bias on cosmological parameters

As we have seen in previous sections, our forecasts for future observations depend strongly on the knowledge of the theoretical matter power spectrum. For linear scales, this can be calculated with great precision, but for non-linear scales, we have to rely on N-body simulations, semi-analytic methods or perturbation theories. Each of these methods only have a certain range of validity and introduce many numerical and theoretical errors into the calculations. We will see more details on this topic in chapter 6.

In this section we will quantify the effect of the systematic errors on our forecasted statistical error estimations due to the uncertainties on the non-linear power spectrum. We will show how big this systematic bias would be, if we used for our forecasts a power spectrum which is not the "correct" one. The following discussion was introduced in **casas_fitting_2015** in reference to forecasts using non-linear power spectra from N-body simulations and is based mostly on the expressions derived in Appendix B of **taylor_probing_2007**

The linear bias on a cosmological parameter $\delta\theta_i$ due to the bias $\delta\psi_i$ in a parameter of the model which we assume fixed and cannot be measured is given by:

$$\delta\theta_i = -\left[F^{\theta\theta}\right]^{-1}_{ik} F^{\theta\psi}_{kj}\delta\psi_j \tag{3.44}$$

In our case we will have only one systematic parameter ψ, which controls the difference between the "true" power spectrum P_{true} and and our simulated power spectrum P_{num}:

$$P_\psi = \psi P_{num} + (1-\psi)P_{true}$$

ψ can vary continuously so that for $\psi = 1$ we recover P_{num}, while for $\psi = 0$ we obtain P_{true}. We can define the relative difference between P_{true} and P_{num} as:

$$\sigma_p(k,z) \equiv \frac{P_{num}(k,z) - P_{true}(k,z)}{P_{true}(k,z)} \tag{3.45}$$

The $F^{\theta\theta}$ in eqn.3.44 above is simply the usual Fisher matrix:

$$F^{\theta\theta} = \frac{1}{2}\mathrm{tr}\left(C^{-1}\partial_i^\theta C C^{-1}\partial_j^\theta C\right)$$

while the pseudo-Fisher matrix between measured and assumed parameters $F^{\theta\psi}$ is:

$$F_{ij}^{\theta\psi} = \frac{1}{2}\mathrm{tr}\left(C^{-1}\partial_i^\theta C C^{-1}\partial_j^\psi C\right) \tag{3.46}$$

which for one systematic parameter only, is just a column vector.

In the case of galaxy clustering we will compute $F^{\theta\psi}$ in the following way, using the fact that for $\psi = 1$, $C = P_{num}(k,z) + n^{-1}(z)$ and $P_\psi|_{\psi=1} = P_{num}$:

$$F_i^{\theta\psi} \propto \int \mathrm{d}k\, k^2 \left(\frac{n_{eff}(z)P_{num}(k,z)}{n_{eff}(z)P_{num}(k,z)+1}\right)^2 \left(\frac{1}{P_{num}(k,z)}\right)^2 \left.\frac{\partial P_\psi}{\partial \psi}\right|_{\psi=1} \frac{\partial P_{num}}{\partial \theta_i} \tag{3.47}$$

in this step we have assumed that we have no systematic parameters affecting the galaxy number density $n(z)$ and therefore, its derivative w.r.t ψ vanishes. Also, for notational simplicity we left out the integral over μ and the complete form of the observed power spectrum.

In section 5.6.0.2 we will show a concrete example of the application of these formulas and we will see that in the non-linear regime, where the uncertainty in the theoretical power spectrum is of about 10-20%, the systematic errors can be of the same order or larger than the forecasted statistical errors for a future survey like Euclid.

3.5 The equations and structure of the FISHERTOOLS code

In this section we will explain a bit more in detail the implementation of the FISHERTOOLS code, which has been used in the projects and papers explained throughout this work. This code was created by the author and consists on a set of packages for the `Mathematica Wolfram` language.

3.5.1 The structure of FISHERTOOLS

The code contains of 4 basic modules:

1. CosmologyTools: A package defining many background quantities in cosmology, such as the Hubble function $H(z)$, distances, volumes, density fractions $\Omega(z)$, the growth and growth rate functions D_+ and $f(z)$, conversions between different units and time conventions (scale factor, redshift, e-fold time, conformal time) and functions related to observational effects, like RSD and the AP effect (see section 3.4.2), among others. This package also interacts with the package COSMOMATHICA (see section 3.5.2 below), which serves as

an interface and a wrapper to many other useful codes in the community, especially Boltzmann codes.

2. FisherTools: A package containing all functions needed to perform and analyze a Fisher forecast. Many of them are too technical to be included here and belong to a comprehensive manual, but in the next subsection we will specify some of them. The most important ones are the routines for derivation and integration, together with routines for matrix operations and visualization.

3. WeakLensingTools: Some special functions related to Weak Lensing analysis, like window functions, multipoles, computation of bins and bin correlations.

4. UsefulTools: Auxiliary tools used in the code that are not provided by `Mathematica`, such as file exporting and importing, string parsing and matrix operations.

3.5.2 The interface COSMOMATHICA

The COSMOMATHICA code was created originally by Dr. Adrian Vollmer and used for several projects in the field of Fisher matrix forecasts and perturbation theory **papers by Adrian** It is now maintained by the author of this work and it can be found and copied from its repository under the URL: https://github.com/santiagocasas/cosmomathica.

Its main function is to interface many codes and routines used in the cosmology community to a simple `Mathematica` notebook or package. The supported codes are the Boltzmann codes: `CAMB` **lewis_efficient_2000** and `CLASS` **lesgourges** The fitting functions from Eisenstein & Hu **cite Eisenstein** `Halofit` **smith** and the `Cosmic Emulator` **heitmann_coyote_2010**; **heitmann_coyote_2014** and a code that calculates higher order perturbation theory for large scale structure, called `COPTER` **find citation**

This interface (which uses the `MathLink` technology from the `Wolfram` language) allows a very efficient and fast communication between the notebooks, the packages and the codes. Therefore, there is no need to create large quantities of files that have to be read and written in order to deal with cosmological quantities like the power spectrum or the transfer functions.

3.5.3 Derivatives of the observed power spectrum

The main calculation done in this code for galaxy clustering is based on eq. (3.25), which involves derivatives of the observed power spectrum and an integral over wavevectors k and angles μ. For each redshift bin n, the derivatives are evaluated at the center of each bin $\bar{z}_n$ and at the fiducial values of the parameters. The total number of redshift bins is N_b.

Let us first consider the derivatives of the observed power spectrum (see eq. (3.27)) with respect to a number N_p of redshift-independent parameters, which we will denote θ_i, where the index i runs from 1 to N_p. These are parameters of the model which are fixed numbers like n_s, or are defined at redshift $z = 0$, like Ω_m. In this case there are two possible options in terms of code implementation.

1. Full Numerical Derivative method:

$$\left.\frac{\mathrm{d}\ln P_{\mathrm{obs}}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\mathrm{d}\theta_i}\right|_{fid} = \frac{P_{\mathrm{obs}}\left(\bar{z}_n, k, \mu; \theta_i^{+}\right) - P_{\mathrm{obs}}\left(\bar{z}_n, k, \mu; \theta_i^{-}\right)}{2\varepsilon\theta_i^{fid} \times P_{\mathrm{obs}}\left(\bar{z}_n, k, \mu; \theta_i^{fid}\right)} \tag{3.48}$$

where θ_i^+, θ_i^- represent the parameter θ_i evaluated at $\pm\varepsilon$ around the fiducial value θ_i^{fid}:

$$\theta_i^{\pm} = \theta_i^{fid}(1 \pm \varepsilon) \tag{3.49}$$

Almost all functions inside $P_{\rm obs}(\bar{z}, k, \mu; \theta_i)$ can be evaluated at $\theta_i^{\pm}$, except for the bias function $b(z)$, which we cannot relate to the cosmological parameters. Therefore, we need to consider for each redshift bin, an independent parameter $b_n = b(\bar{z}_n)$. This can be considered a redshift-dependent parameter.

2. Chain Rule method:

The observed power spectrum (eq. (3.27)) depends on 5 functions of the redshift, which we can call redshift-dependent variables. However, just three of them, $H(z)$, $D_A(z)$ and $f(z)$ depend on the cosmological parameters θ_i. Then, with the help of the chain rule one can write:

$$\begin{aligned} \left.\frac{\mathrm{d}\ln P_{\rm obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\mathrm{d}\theta_i}\right|_{fid} &= \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln f(\bar{z}_n)} \frac{\partial \ln f(\bar{z}_n)}{\partial \theta_i} \\ &+ \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln H(\bar{z}_n)} \frac{\partial \ln H(\bar{z}_n)}{\partial \theta_i} \\ &+ \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln D_A(\bar{z}_n)} \frac{\partial \ln D_A(\bar{z}_n)}{\partial \theta_i} \\ &+ \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln P(k, \bar{z}_n)} \frac{\partial \ln P(k, \bar{z}_n)}{\partial \theta_i} \\ &+ \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial k} \frac{\partial k}{\partial \theta_i} \\ &+ \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \mu} \frac{\partial \mu}{\partial \theta_i} \end{aligned} \tag{3.50}$$

The last two terms, which consider derivatives of k and μ with respect to θ_i are non-vanishing if one takes the Alcock-Paczynski effect into account for k and μ, since they are affected by geometrical terms, we will detail their expressions in section 3.5.5 below. The redshift dependent functions $b(z)$ and $P_s(z)$ are not known as a function of the fundamental cosmological parameters, and their functional form as a function of redshift is also generally unknown. Therefore, the best we can do is to discretize them in redshift bins and assign some fiducial values for them. We have to assume that each value at each bin is independent of the other. We we will therefore have $2 \times N_b$ unknown parameters in the observed power spectrum, namely $b_n = b(\bar{z}_n)$ and $P_{s,n} = P_s(\bar{z}_n)$, the values of the bias and the extra shot noise at $\bar{z}_n$. In order to simplify the resulting equations, we will use the natural logarithm of these quantities. Then, these derivatives are:

$$\begin{aligned} \left.\frac{\mathrm{d}\ln P_{\rm obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\mathrm{d}P_s(\bar{z}_n)}\right|_{fid} &= \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln P_{s,n}} \\ &= \frac{1}{P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)} \end{aligned} \tag{3.51}$$

where we have assumed the fiducial extra shot noise value to be $P_{s,n} = 0$ at all bins, and

$$\left.\frac{\mathrm{d}\ln P_{\mathrm{obs}}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\mathrm{d}\ln b(\bar{z}_n)}\right|_{fid} = \frac{\partial \ln P_{obs}\left(\bar{z}_n, k, \mu; \theta_i\right)}{\partial \ln b_n} = \frac{2}{1+\beta(\bar{z}_n)\mu^2} \tag{3.52}$$

Therefore, our set of parameters extends from θ_i to $\Theta_i = \{\theta_i, \ln b_n, \ln P_{s,n}\}$. The Fisher matrix eq. (3.25) will depend now on all the combinations of the first derivatives of the power spectrum with respect to Θ and will be of the dimensions $(N_p + 2N_b) \times (N_p + 2N_b)$, since we have 2 redshift-dependent parameters at N_b redshift bins and N_p cosmological redshift-independent parameters.

The intermediate derivatives of $\ln P_{obs}(z, k, \mu; \theta_i)$ with respect to $D_A(z)$, $H(z)$ and $f(z)$ can be calculated analytically from eq. (3.27). We use the logarithm of these quantities, because it simplifies considerably the formulas. As usual, the derivatives are calculated at the *fiducial* value of the cosmological parameters.

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln f(\bar{z})} = \frac{2\beta(\bar{z})\mu^2}{1+\beta(\bar{z})\mu^2} \tag{3.53a}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln H(\bar{z})} = 1 + \mathrm{AP_H} \tag{3.53b}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln D_A(\bar{z})} = -2 + \mathrm{AP_D} \tag{3.53c}$$

$$\frac{\partial \ln P_{obs}\left(\bar{z}, k, \mu; \theta_i\right)}{\partial \ln P(k, \bar{z})} = 1 \tag{3.53d}$$

Here, the $\mathrm{AP_H}$ and $\mathrm{AP_D}$ represent the extra terms appearing from the Alcock-Paczynski effect, where the observed k and μ are corrected by geometrical terms. These formulas will be specified in section 3.5.5 below.

3. The BAO method:

The third option would be to consider first the redshift-dependent functions to be independent of the cosmological parameters θ_i and constrain them independently of the cosmological model. This is the preferred option for observational cosmologists, since it is more connected to the actual observations and it is more model independent. In this method, one simply calculates eq. (3.50) ignoring the derivatives of the redshift-dependent functions with respect to θ_i. Using these three redshift-dependent variables, together with the other two mentioned previously, $b(z)$ and $P_s(z)$, the space of parameters grows to

$$\Theta_j = \{\theta_i, \ln b_n, \ln P_{s,n}, H_n, D_{A,n}, f_n\} \quad , \tag{3.54}$$

where the subscript n, corresponds to the function evaluated at $\bar{z}_n$, the center of the redshift bin. In this case the Fisher matrix would be much bigger, with $(N_p+5N_b) \times (N_p + 5N_b)$ elements, with N_b the number of redshift bins and N_p the number of redshift-independent parameters. If one wishes to project the redshift-dependent parameters into the fundamental cosmological parameters, one simply calculates a Jacobian of the "old" variables with respect to the new variables, where the new fundamental cosmological parameters $\tilde{\theta}_i$ can differ from the "old" variables and leaves open the opportunity of changing the model parameter basis:

$$J_{ab} = \frac{\partial \Theta_a}{\partial \tilde{\theta}_b} \quad . \tag{3.55}$$

Notice that if $\tilde{\theta}_i = \theta_i$, the first $(N_p) \times (N_p)$ elements of the Jacobian matrix will be a unit matrix. This Jacobian is a non-square matrix since Θ_a contains $(N_p + 5N_b)$ components and $\tilde{\theta}_b$ contains just (N_p) components, therefore it cannot be inverted. The new Fisher matrix would change accordingly to $\tilde{F}$ and it would be given by:

$$\tilde{F} = J^T \, F \, J \tag{3.56}$$

This method is called the "BAO" method **seo_improved_2007; seo_improved_2005** since in BAO observations the redshift independent functions $D_A(z)$, $H(z)$ and $f(z)$ are the main observables and one obtains cosmological parameter constraints from them by first constraining their values and then projecting these constraints onto a cosmological model. However, since the bias (which also enters $\beta(z)$) is degenerate with the overall amplitude of the power spectrum as a function of redshift, $\sigma_8(z)$, observers prefer to define an observed power spectrum which depends on $f\sigma_8(z)$ and $b\sigma_8(z)$, so that instead of eq. (3.27), the observed power spectrum would be expressed as (ignoring the exponential damping term for simplicity, which remains unaltered anyway):

$$P_{\rm obs}\left(z, k, \mu; \theta\right) = P_{\rm s}(z) + \frac{D_A^2(z)_{ref} H(z)}{D_A^2(z) H(z)_{ref}} \left(b\sigma_8(z) + f\sigma_8(z)\mu^2\right)^2 \left(\frac{P(k,z)}{\sigma_8^2(z)}\right) \quad . \tag{3.57}$$

3.5.4 Equations for the power spectrum

Let us review here some of the formulas related to the power spectrum, in the way they are implemented in the FISHERTOOLS code.

The normalization σ_R of the power spectrum at $z = 0$ (which is equivalent to the variance smoothed over a scale R) is given by:

$$\sigma_R^2 = \frac{1}{2\pi^2} \int_{k_{min}}^{k_{max}} k^2 P(k, z = 0) W_R^2(kR) \quad . \tag{3.58}$$

The window function smooths $P(k,z)$ over the scale R and has the form: $W(x) = 3(\sin(x) - x\cos(x))/x^3$, where x is a dimensionless variable. This term comes from assuming spherical symmetry and performing the angle integration of the 3-D power spectrum $P(\vec{k})$. If one chooses $R = 8\text{h}^{-1}$Mpc and $[k] = $ h/Mpc, then we can define the parameter

$$\sigma_8^2 \equiv \sigma_R^2 \quad (R = 8\text{Mpc/h}) \quad . \tag{3.59}$$

The quantity σ_8 is a parameter defined always at $z = 0$ and only in linear theory, which means that in eq. (3.58) either $P(k, z = 0)$ is the linear power spectrum or the integration limits are chosen in such a way that only linear scales are taken into account, therefore one should take $k_{max} \approx 0.1$.

The linear power spectrum can be obtained either from `CAMB`, `CLASS` or from the transfer functions of Eisenstein & Hu. In the latter case, we can express the power spectrum $P(k,z)$ in terms of the primordial amplitude $\mathcal{A}_s$, the linear growth

$G(z)$ (equivalent to $D_+(a)$ in alternative notations) and the transfer functions $T(k)$:

$$P(k,z) = G^2(z)T^2(k)k^{n_s}\mathcal{A}_s \quad . \tag{3.60}$$

Since the parameters σ_8^2 and $\mathcal{A}_s$ are fully degenerate in linear theory, one has to be careful to include only one of them into the set of cosmological parameters. The quantity $\sigma_8(z) \equiv \sigma_8^2 G^2(z)$ can be defined accordingly and can be used as an independent and redshift-dependent cosmological parameter, as it was the case in eq. (3.57). We have to emphasize again that these quantities and expressions are only formally valid in the linear regime and in standard Einstein GR; in the deeply non-linear regime or under the effect of modified gravity forces, the growth function might acquire some scale-dependence.

3.5.5 The Alcock-Paczynski Effect

The Alcock-Paczynski effect (AP, see (**alcock1979anevolution**)) relies on the fact that the scales k and the angle cosines μ are changed by geometrical factors of distance, when the cosmological parameters are changed. This means that k and μ depend on $H(z)$ and $D_A(z)$ and therefore are also indirectly functions of the cosmological parameters θ_i. The AP formulas relating the fiducial values of k_{fid} and μ_{fid} to their transformed values are:

$$k_{AP} = R_{AP}(\mu;\theta)k_{\text{fid}} \tag{3.61}$$

$$\mu_{AP} = \frac{H(z;\theta)}{H(z;\theta_{\text{fid}})}\frac{\mu_{\text{fid}}}{R_{AP}(\mu;\theta)} \tag{3.62}$$

where the geometrical function R_{AP} is defined as:

$$R_{AP}(\mu;\theta) = \sqrt{\frac{[D_A(z;\theta)H(z;\theta)\mu_{fid}]^2 - [(D_A(z;\theta_{\text{fid}})H(z;\theta_{\text{fid}})]^2(\mu_{\text{fid}} - 1)}{[D_A(z;\theta)H(z;\theta_{\text{fid}})]^2}} \quad . \tag{3.63}$$

The last two terms of eq. (3.50) come from the fact that now $k = k(H(z;\theta), D_A(z;\theta))$ and $\mu = k(H(z;\theta), D_A(z;\theta))$, so that one has to use the chain rule for the derivative of $P_{obs}(k,\mu,z;\theta)$:

$$\begin{aligned}\frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;H(z);\theta_i)}{\partial \ln H(z)} &= \frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;\theta_i)}{\partial \ln H(z)} \\ &+ \frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;\theta_i)}{\partial \mu}\frac{\partial \mu}{\partial \ln H(z)} \\ &+ \frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;\theta_i)}{\partial k}\frac{\partial k}{\partial \ln H(z)} \quad ,\end{aligned} \tag{3.64}$$

and similarly for the dependence on $D_A(z)$.

Now we can write down explicitly the intermediate derivatives (evaluated at the fiducial value as usual):

$$\frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;\theta_i)}{\partial k} = \frac{\partial \ln P(\bar{z}_n,k;\theta_i)}{\partial k} \tag{3.65}$$

$$\frac{\partial \ln P_{obs}(\bar{z}_n,k,\mu;\theta_i)}{\partial \mu} = \frac{4\beta(z)\mu}{1+\beta(z)\mu^2} \tag{3.66}$$

and the derivatives of k and μ with respect to $\ln D_A$ and $\ln H$ are:

$$\frac{\partial \mu}{\partial \ln D_A(z)} = -\mu(\mu^2 - 1) \tag{3.67a}$$

$$\frac{\partial \mu}{\partial \ln H(z)} = -\mu(\mu^2 - 1) \tag{3.67b}$$

$$\frac{\partial k}{\partial \ln D_A(z)} = k(\mu^2 - 1) \tag{3.67c}$$

$$\frac{\partial k}{\partial \ln H(z)} = k\mu^2 \tag{3.67d}$$

The interesting thing about eq. (3.67) is that one can see that when evaluating the derivative of $P_{\rm obs}$ w.r.t a cosmological parameter, the scales k and the direction cosine angles μ get mixed, giving a very powerful probe of cosmology.

3.5.6 Fisher matrix operations

Another important part of the code is the handling of the Fisher matrix itself for post-analysis. We are interested in the constraints on the model parameters and therefore we need to calculate the covariance matrix $\mathbf{C} = F^{-1}$, whose diagonal contains the square of the fully marginalized 1σ errors of the model parameters and its off-diagonal entries encode the level of correlation among parameters.

Maximization: If we have a Fisher matrix corresponding to an N number of parameters, and we want to *maximize* over the parameter θ_m, i.e. we want to obtain the errors on all other parameters, when the parameter θ_m is fixed, then we just need to remove from the Fisher matrix, the rows and the columns corresponding to that parameter at position m:

$$\tilde{F}_{(N-1)\times(N-1)} = M^T_{(N-1)\times N} F_{N\times N} M_{N\times(N-1)} \tag{3.68}$$

where the matrix M can be defined in terms of the Kronecker delta $\delta_{i,j}$ as:

$$M_{ij} = \delta_{\mho(i),j} \quad , \tag{3.69}$$

where:

$$\mho(i) \equiv \begin{cases} i, & i < m \\ N, & i = m \\ i-1, & i > m \end{cases} \quad , \tag{3.70}$$

and the indices run as $i = \{1, \ldots, N\}$ and $j = \{1, \ldots, (N-1)\}$. The new Fisher matrix $\tilde{F}$ will be then of dimensions $(N-1) \times (N-1)$.

Marginalization: If we want to *marginalize* over the parameter m, then we have to invert the matrix F to obtain the parameter covariance matrix $C = F^{-1}$ and remove from C the columns and rows corresponding to the index m:

$$\tilde{C}_{(N-1)\times(N-1)} = M^T_{(N-1)\times N} C_{N\times N} M_{N\times(N-1)} \quad , \tag{3.71}$$

where M is given by eq. (3.69) and eq. (3.70). The operation of multiplying by matrix M and its transpose can be done more efficiently for a large number of parameters to be removed, by using appropriate functions in the `Wolfram Mathematica` language.

Combination: If we want to combine the Fisher matrix from an observation (or experiment) A with another Fisher matrix from experiment B, where both F_A and F_B have the same fiducial parameters and the same ordering, we just need to add the Fisher matrices together: $F_{A+B} = F_A + F_B$ and from F_{A+B} we can obtain the combined contours and errors.

Changing of parameter basis: If our Fisher matrix F was performed in the parameter basis X^i and we want to change to a different parameter basis Y^i; for example in ΛCDM when we want to go from $\{\Omega_m, \Omega_b, h, \sigma_8\}$ to $\{\omega_c, \omega_b, h, \ln(10^1 0 A_s)\}$, we can use a simple Jacobian operation. This can be done since, as we already mentioned, the Fisher matrix is composed of derivatives of the likelihood, which in turn are in our case first derivatives of the data covariance matrix $\mathbf{C}$:

$$F_{\alpha\beta} \propto \frac{\partial \mathbf{C}}{\partial X^\alpha} \frac{\partial \mathbf{C}}{\partial X^\beta} \tag{3.72}$$

Therefore we can write the transformed Fisher matrix $\tilde{F}$ in the new basis as:

$$\begin{aligned} \tilde{F}_{\mu\nu} &\propto \frac{\partial X^\alpha}{\partial Y^\mu} \frac{\partial \mathbf{C}}{\partial X^\alpha} \frac{\partial \mathbf{C}}{\partial X^\beta} \frac{\partial X^\beta}{\partial Y^\nu} \\ \tilde{F}_{\mu\nu} &= J^T F J \end{aligned} \tag{3.73}$$

with $J_{\alpha\nu} = \partial X^\alpha / \partial Y^\nu$. In some cases, if the "old" parameter basis X^i cannot be solved in terms of the "new" parameter basis Y^i, then one can use the identity:

$$J^{-1}_{ab} = \frac{\partial Y^a}{\partial X^b} \quad . \tag{3.74}$$

Ellipsoidal confidence contour regions: Since we are assuming a Gaussian likelihood at the maximum, the confidence contours for a 2-dimensional slice of the likelihood will be ellipsoidal, simply from the fact that the exponent of the two-dimensional Gaussian probability distribution for a set of two parameters $X_i = \{x_1, x_2\}$ and a symmetric covariance matrix $C_{ij} \equiv c_{ij} = c_{ji}$ contains the term:

$$X_i C^{-1}_{ij} X_j = \left(\frac{c_{22} x_1^2}{\det C} + \frac{c_{11} x_2^2}{\det C} - \frac{2 c_{12} x_1 x_2}{\det C} \right) \quad . \tag{3.75}$$

Then the ellipse will be oriented along the eigenvectors of C_{ij} with an anlge α with respect to the coordinate axes:

$$\tan(2\alpha) = \frac{2 c_{12}}{c_{11}^2 - c_{22}^2} \tag{3.76}$$

By integrating a bi-variate Gaussian eq. (3.15) with the exponent eq. (3.75) to find the confidence regions as in eq. (3.14), we obtain that the 1, 2 and 3σ confidence regions with major semi-axes a and minor semi-axes b are given by:

$$\begin{aligned} 1\sigma: &\quad a = 1.51\sqrt{\lambda_1}; \quad b = 1.51\sqrt{\lambda_2} \\ 2\sigma: &\quad a = 2.49\sqrt{\lambda_1}; \quad b = 2.49\sqrt{\lambda_2} \\ 3\sigma: &\quad a = 3.44\sqrt{\lambda_1}; \quad b = 3.44\sqrt{\lambda_2} \quad , \end{aligned}$$

where $\lambda_1,\ \lambda_2$ are the largest and the smallest eigenvalues, respectively. The 2×2 matrix from where this ellipses can be visualized can be either obtained by marginalizing or maximizing over all the other $(N-2)$ parameters. The FISHERTOOLS code contains these options and many other tools for plotting and visualization.

3.5.7 Extensions of the Fisher matrix approach

The Fisher matrix a

Chapter 4

Linear and non-linear Modified Gravity forecasts

4.1 Introduction

Future large scale structure surveys will be able to measure with percent precision the parameters governing the evolution of matter perturbations. While we have the tools to investigate the standard model, the next challenge is to be able to compare those data with cosmologies that go beyond General Relativity, in order to test whether a fluid component like Dark Energy or similarly a Modified Gravity scenario can better fit the data. On the theoretical side, while many Modified Gravity models are still allowed by type Ia supernova (SNIa) and Cosmic Microwave Background (CMB) data **planck_collaboration_planck_2016** structure formation can help us to distinguish among them and the standard scenario, thanks to their signatures on the matter power spectrum, in the linear and mildly non-linear regimes (for some examples of forecasts, see **amendola_cosmology_2013**; **casas_fitting_2015**; **bielefeld_cosmological_2014**).

The evolution of matter perturbations can be fully described by two generic functions of time and space **kunz_phenomenological_2012**; **amendola_observables_2013** which can be measured via Galaxy Clustering and Weak Lensing surveys. In this work we want to forecast how well we can measure those functions, in different redshift bins.

While any two independent functions of the gravitational potentials would do, we follow the notation of **planck_collaboration_planck_2016** and consider μ and η: the first modifies the Poisson equation for Ψ while the second is equal to the ratio of the gravitational potentials (and is therefore also a direct observable **amendola_observables_2013**). We will consider forecasts for the planned surveys Euclid, SKA1 and SKA2 and a subset of DESI, DESI-ELG, using as priors the constraints from recent *Planck* data (see also **Alonso2016**; **hojjati_cosmological_2012**; **baker_observational_2015**; **bull_extending_2015**; **Gleyzes2016** for previous works that address forecasts in Modified Gravity.

In section 4.2 we define μ and η and parameterize them in three different ways. First, in a general manner, we let these functions vary freely in different redshift bins. Complementarily, we also consider two specific parameterizations of the time evolution proposed in **planck_collaboration_planck_2016** Here, we also specify the fiducial values of our cosmology for each of the parameterizations considered. Section 4.3 discusses our treatment for the linear and mildly non-linear regime. Linear spectra are obtained from a modified Boltzmann code **hojjati_testing_2011** the mild non-linear regime (up to k $\sim$ 0.5 h/Mpc) compares two methods to emulate the non-linear power spectrum: the commonly used Halofit **smith_stable_2003**; **takahashi_revising_2012** and a semi-analytic prescription to model the screening mechanisms present in Modified Gravity models **hu_parameterized_2007** In section **??** we explain the method used to produce the Fisher forecasts both for Weak Lensing and Galaxy Clustering. We explain how we compute and add the CMB *Planck* priors to our Fisher matrices. Section 4.4 discusses the results obtained for the redshift binned parameterization both for Galaxy Clustering and for Weak Lensing in the linear and non-linear cases. We describe our method to decorrelate the errors in section 4.4.4. The results for the other two time parameterizations are instead discussed in sections 4.7.1 and 4.7.2, both for Weak Lensing and Galaxy Clustering in the linear and mildly non-linear regimes. To test the effect of our non-linear prescription, we show in section 4.7.3 the impact of different choices of the non-linear prescription parameters on the cosmological parameter estimation.

4.2 Parameterizing Modified Gravity

In linear perturbation theory, scalar, vector and tensor perturbations do not mix, which allows us to consider only the scalar perturbations in this paper. We work in the conformal Newtonian gauge, with the line element given by

$$ds^2 = -(1+2\Psi)dt^2 + a^2(1-2\Phi)dx^2 \quad . \tag{4.1}$$

Here Φ and Ψ are two functions of time and scale that coincide with the gauge-invariant Bardeen potentials in the Newtonian gauge.

In theories with extra degrees of freedom (Dark Energy, DE) or modifications of General Relativity (MG) the normal linear perturbation equations are no longer valid, so that for a given matter source the values of Φ and Ψ will differ from their usual values. We can parameterize this change generally with the help of two new functions that encode the modifications. Many different choices are possible and have been adopted in the literature, see e.g. **planck_collaboration_planck_2016** for a limited overview. In this paper we introduce the two functions through a gravitational slip (leading to $\Phi \neq \Psi$ also at linear order and for pure cold dark matter) and as a modification of the Poisson equation for Ψ,

$$\begin{aligned} -k^2\Psi(a,k) &\equiv 4\pi G a^2 \mu(a,k)\rho(a)\delta(a,k) \quad ; & (4.2)\\ \eta(a,k) &\equiv \Phi(a,k)/\Psi(a,k) \quad . & (4.3)\end{aligned}$$

These expressions define μ and η. Here $\rho(a)$ is the average dark matter density and $\delta(a,k)$ the comoving matter density contrast – we will neglect relativistic particles and radiation as we are only interested in modeling the perturbation behaviour at late times. In that situation, η, which is effectively an observable **amendola_observables_2013** is closely related to modifications of GR **saltas_anisotropic_2014**; **sawicki_non-standard_2016** while μ encodes for example deviations in gravitational clustering, especially in redshift-space distortions as non-relativistic particles are accelerated by the gradient of Ψ.

When considering Weak Lensing observations then it is also natural to parameterize deviations in the lensing or Weyl potential $\Phi + \Psi$, since it is this combination that affects null-geodesics (relativistic particles). To this end we introduce a function $\Sigma(t,k)$ so that

$$-k^2(\Phi(a,k) + \Psi(a,k)) \equiv 8\pi G a^2 \Sigma(a,k)\rho(a)\delta(a,k) \quad . \tag{4.4}$$

Since metric perturbations are fully specified by two functions of time and scale, Σ is not independent from μ and η, and can be obtained from the latter as follows:

$$\Sigma(a,k) = (\mu(a,k)/2)(1+\eta(a,k)) \ . \tag{4.5}$$

Throughout this work, we will denote the standard Lambda-Cold-Dark-Matter (ΛCDM) model, defined through the Einstein-Hilbert action with a cosmological constant, simply as GR. For this case we have that $\mu = \eta = \Sigma = 1$. All other cases in which these functions are not unity will be labeled as Modified Gravity (MG) models.

Using effective quantities like μ and η has the advantage that they are able to model *any* deviations of the perturbation behaviour from ΛCDM expectations, they are relatively close to observations, and they can also be related to other commonly used parameterization **pogosian_how_2010** On the other hand, they are not easy to map to an action (as opposed to approaches like effective field theories that are based on an explicit action) and in addition they contain so much freedom that we normally restrict their parameterisation to a subset of possible functions.

This has however the disadvantage of loosing generality and making our constraints on μ and η parameterization-dependent. In this paper, we prefer to complement specific choices of parameterizations adopted in the literature (we will use the choice made in **planck_collaboration_planck_2016**) with a more general approach: we will bin the functions $\mu(a)$ and $\eta(a)$ in redshift bins with index i and we will treat each μ_i and η_i as independent parameters in our forecast; we will then apply a variation of Principal Component Analysis (PCA), called Zero-phase Component Analysis (ZCA). This approach has been taken previously in the literature by **hojjati_cosmological_2012** where they bin μ and η in several redshift and k-$scale$ bins together with a binning of $w(z)$ and cross correlate large scale structure observations with CMB temperature, E-modes and polarization data together with Integrated Sachs-Wolfe (ISW) observations to forecast the sensitivity of future surveys to modifications in μ and η. In the present work, we will neglect a possible k-dependence, we will focus on Galaxy Clustering (GC) and weak lensing (WL) surveys and we will show that there are important differences between the linear and non-linear cases; including the non-linear regime generally reduces correlations among the cosmological parameters. In the remainder of this section we will introduce the parameterizations that we will use.

4.2.1 Parameterizing gravitational potentials in discrete redshift bins

As a first approach we neglect scale dependence and bin the time evolution of the functions μ and η without specifying any parameterized evolution. To this purpose we divide the redshift range $0 \leq z \leq 3$ in 6 redshift bins and we consider the values $\mu(z_i)$ and $\eta(z_i)$ at the right limiting redshift z_i of each bin as free parameters, thus with the i index spanning the values $\{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$. The first bin is assumed to have a constant value, coinciding with the one at $z_1 = 0.5$, i.e. $\mu(z < 0.5) = \mu(z_1)$ and $\eta(z < 0.5) = \eta(z_1)$. The $\mu(z)$ function (and analogously $\eta(z)$) is then reconstructed as

$$\mu(z) = \mu(z_1) + \sum_{i=1}^{N-1} \frac{\mu(z_{i+1}) - \mu(z_i)}{2} \left[1 + \tanh\left(s\frac{z - z_{i+1}}{z_{i+1} - z_i}\right)\right], \tag{4.6}$$

where $s = 10$ is a smoothing parameter and N is the number of binned values. We assume that both μ and η reach the GR limit at high redshifts: to realize this, the last $\mu(z_6)$ and $\eta(z_6)$ values assume the standard ΛCDM value $\mu = \eta = 1$ and both functions are kept constant at higher redshifts $z > 3$.

Similarly, the derivatives of these functions are obtained by computing

$$\mu'(\bar{z}_j) = \frac{\mu(z_{i+1}) - \mu(z_i)}{z_{i+1} - z_i}, \tag{4.7}$$

with $\bar{z}_j = (z_{i+1} + z_i)/2$, using the same $\tanh(x)$ smoothing function:

$$\frac{d\mu(z)}{dz} = \mu'(\bar{z}_1) + \sum_{j=1}^{N-2} \frac{\mu'(\bar{z}_{j+1}) - \mu'(\bar{z}_j)}{2} \left[1 + \tanh\left(s\frac{z - \bar{z}_{j+1}}{\bar{z}_{j+1} - \bar{z}_j}\right)\right] . \tag{4.8}$$

In particular we assume $\mu' = \eta' = 0$ for $z < 0.5$ and for $z > 3$. We set the first five amplitudes of μ_i and η_i as free parameters, thus the set we consider is: $\theta = \{\Omega_m, \Omega_b, h, \ln 10^{10} A_s, n_s, \{\mu_i\}, \{\eta_i\}\}$, with i an index going from 1 to 5. We take as fiducial cosmology the values shown in Tab. 4.1 columns 5 and 6. We only modify the evolution of perturbations and assume that the background expansion is well described by the standard ΛCDM expansion law for a flat universe with given values of Ω_m, Ω_b and h.

4.2.2 Parameterizing gravitational potentials with simple smooth functions of the scale factor

As an alternative approach, we assume simple specific, time parameterizations for the μ and η MG functions, adopting the ones used in the *Planck* analysis **planck_collaboration_planck_2016** We neglect here as well any scale dependence:

- a parameterization in which the time evolution is related to the dark energy density fraction, to which we refer as 'late-time' parameterization:

$$\mu(a,k) \equiv 1 + E_{11}\Omega_{\rm DE}(a) \ , \tag{4.9}$$
$$\eta(a,k) \equiv 1 + E_{22}\Omega_{\rm DE}(a) \ ; \tag{4.10}$$

- a parameterization in which the time evolution is the simplest first order Taylor expansion of a general function of the scale factor a (and closely resembles the $w_0 - w_a$ parametrization for the equation of state of DE), referred to as 'early-time' parameterization, because it allows departures from GR also at high redshifts [1]:

$$\mu(a,k) \equiv 1 + E_{11} + E_{12}(1-a) \ , \tag{4.11}$$
$$\eta(a,k) \equiv 1 + E_{21} + E_{22}(1-a) \ . \tag{4.12}$$

The late-time parameterization is forced to behave as GR ($\mu = \eta = 1$) at high redshift when $\Omega_{\rm DE}(a)$ becomes negligible; the early time one allows more freedom as the amplitude of the deviations from GR do not necessarily reduce to zero at high redshifts. Both parameterizations have been used in **planck_collaboration_planck_2016** In **bull_extending_2015**; **Gleyzes2016**; **Alonso2016** the authors used a similar time

[1] Notice that our early-time parametrization is called 'time-related' in **planck_collaboration_planck_2016**

parameterization in which the Modified Gravity parameters depend on the time evolution of the dark energy fraction. In **bull_extending_2015** an extra parameter accounts for a scale-dependent μ: their treatment keeps η (called γ in their paper) fixed and equal to 1; it uses linear power spectra up to $k_{\max}(z)$ with $k_{\max}(z=0) = 0.14/\text{Mpc} \approx 0.2$ h/Mpc. In **Gleyzes2016** the authors also use a combination of Galaxy Clustering, Weak Lensing and ISW cross-correlation to constrain Modified Gravity in the Effective Field Theory formalism **Gubitosi2013** In **Bellini2016** and **Alonso2016** a similar parameterization was used to constrain the Horndeski functions **Bellini2014** with present data and future forecasts respectively, in the linear regime.

For the late-time parameterization, the set of free parameters we consider is: $\theta = \{\Omega_m, \Omega_b, h, \ln 10^{10} A_s, n_s, E_{11}, E_{22}\}$, where E_{11} and E_{22} determine the amplitude of the variation with respect to ΛCDM. As fiducial cosmology we use the values shown in Table 4.1, columns 1 and 2, i.e. the marginalized parameter values obtained fitting these models with recent *Planck* data; notice that these results differ slightly from the *Planck* analysis in **planck_collaboration_planck_2016** for the same parameterization, because we don't consider here the effect of massive neutrinos.

For the early time parameterization we have E_{11} and E_{21} which determine the amplitude of the deviation from GR at present time ($a = 0$) and 2 additional parameters ($E_{12},\ E_{22}$), which determine the time dependence of the $\mu(a)$ and $\eta(a)$ functions for earlier times. The fiducial values for this model, obtained from the *Planck*+BSH best fit is given in columns 3 and 4 of Table 4.1.

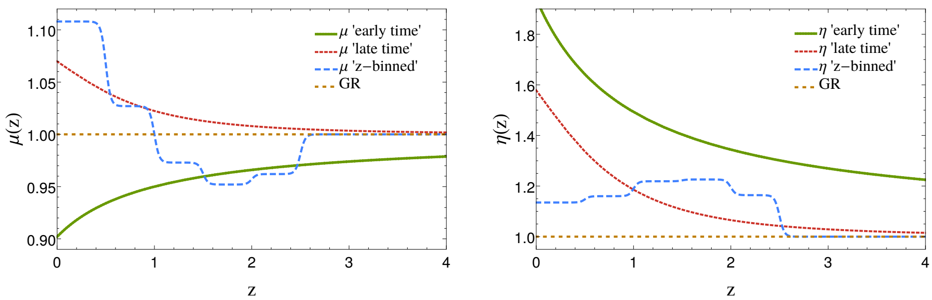


FIGURE 4.1: The Modified Gravity functions μ and η as a function of redshift z for each of the models considered in this work, evaluated at the fiducials specified inTable 4.1. In light long-dashed blue lines, the 'redshift binned model' (Eqns. 4.6-4.8). In short-dashed red lines the late-time parameterization (Eqns. 4.9-4.10) and in green solid lines the early-time parameterization (Eqns. 4.11-4.12). Finally the medium-dashed orange line represents the standard ΛCDM model (GR) for reference.

4.3 The power spectrum in Modified Gravity

4.3.1 The linear power spectrum

In this work we will use linear power spectra calculated with MGCAMB **zhao_searching_2009**; **hojjati_testing_2011** a modified version of the Boltzmann code CAMB **lewis_efficient_2000** We do so, as MGCAMB offers the possibility to input directly any parameterization of μ and η without requiring further assumptions: MGCAMB uses our Eqns. (4.2)

Late time		Early time		Redshift Binned	
Parameter	Fiducial	Parameter	Fiducial	Parameter	Fiducial
Ω_c	0.254	Ω_c	0.256	Ω_c	0.254
Ω_b	0.048	Ω_b	0.048	Ω_b	0.048
n_s	0.969	n_s	0.969	n_s	0.969
$\ln 10^{10} A_s$	3.063	$\ln 10^{10} A_s$	3.091	$\ln 10^{10} A_s$	3.057
h	0.682	h	0.682	h	0.682
E_{11}	0.100	E_{11}	-0.098	μ_1	1.108
E_{22}	0.829	E_{12}	0.096	μ_2	1.027
		E_{21}	0.940	μ_3	0.973
		E_{22}	-0.894	μ_4	0.952
				μ_5	0.962
				η_1	1.135
				η_2	1.160
				η_3	1.219
				η_4	1.226
				η_5	1.164

TABLE 4.1: Fiducial values for the Modified Gravity parameterizations and the redshift-binned model of μ and η used in this work. The DE related parameterization contains two extra parameters E_{11} and E_{22} with respect to GR; the early-time parametrization depends on 4 extra parameters E_{11}, E_{12}, E_{22} and E_{21} with respect to GR; the redshift-binned model contains 10 extra parameters, corresponding to the amplitudes μ_i and η_i in five redshift bins. In this work we will use alternatively and for simplicity the notation $\ell\mathcal{A}_s \equiv \ln(10^{10}A_s)$. The fiducial values are obtained performing a Monte Carlo analysis of *Planck*+BAO+SNe+H_0 (BSH) data **planck_collaboration_planck_2016**

and (4.3) in the Einstein-Boltzmann system of equations, providing the modified evolution of matter perturbations, corresponding to our choice of the gravitational potential functions. Non-relativistic particles like cold dark matter are accelerated by the gradient of Ψ, so that especially the redshift space distortions are sensitive to the modification given by $\mu(a, k)$. For relativistic particles like photons and neutrinos on the other hand, the combination of $\Phi + \Psi$ (and therefore Σ) enters the equations of motion. The impact on the matter power spectrum is more complicated, as the dark matter density contrast is linked via the relativistic Poisson equation to Φ. In addition, an early-time modification of Φ and Ψ can also affect the baryon distribution through their coupling to radiation during that period. As already mentioned above, we will not consider the $k-$dependence of μ and η in this work and our modifications with respect to standard GR will be only functions of the scale factor a.

4.3.2 Non-linear power spectra

As the Universe evolves, matter density fluctuations (δ_m) on small scales (k > 0.1 h/Mpc) become larger than unity and shell-crossing will eventually occur. The usual continuity and Euler equations, together with the Poisson equation become singular **bernardeau_large-scale_2001**; **bernardeau_evolution_2013** therefore making a computation of the matter power spectrum in the highly non-linear regime practically impossible under the standard perturbation theory approach. However, at intermediate scales of around $k \approx 0.1 - 0.2$ h/Mpc, there is the possibility of calculating semi-analytically the effects of non-linearities on the oscillation patterns of the baryon acoustic oscillations (BAO). There are many approaches in this direction, from renormalized perturbation theories **crocce_renormalized_2006; blas_time-sliced_2016; taruya_closure_2008** to time flow equations **pietroni_flowing_2008; anselmi_nonlinear_2012** effective or coarse grained theories **carrasco_effective_2012; baumann_cosmological_2012; pietroni_coarse-grained_2011; manzotti_coarse_2014** and many others. This direction of research is important, since from the amplitude and widths of the first few BAO peaks, one can extract more information from data and break degeneracies among parameters, especially in Modified Gravity.

Computing the non-linear power spectrum in standard GR is still an open question, and even more so when the Poisson equations are modified, as it is in the case in Modified Gravity theories. A solution to this problem is to calculate the evolution of matter perturbations in an N-body simulation **springel_cosmological_2005; fosalba_mice_2013; takahashi_revising_2012; lawrence_coyote_2010; heitmann_coyote_2014** however, this procedure is time-consuming and computationally expensive.

Because of these issues, several previous analyses have been done with a conservative removal of the information at small scales (see for example the *Planck* Dark Energy paper **planck_collaboration_planck_2016** several CFHTLenS analysis **heymans_cfhtlens_2013; kitching_3d_2014** or the previous PCA analysis by **hojjati_cosmological_2012**). However, future surveys will probe an extended range of scales, therefore removing non-linear scales from the analysis would strongly reduce the constraining power of these surveys. Moreover, at small scales we also expect to find means of discriminating between different Modified Gravity models, such as the onset of screening mechanisms needed to recover GR at small scales where experiments strongly constrain deviations from it. For these reasons it is crucial to find methods which will allow us to investigate, at least approximately, the non-linear power spectrum.

Attempts to model the non-linear power spectrum semi-analytically in Modified Gravity have been investigated for f(R) theories in **zhao_modeling_2014; taruya_regularized_201** for coupled dark energy in **casas_fitting_2015; saracco_non-linear_2010; vollmer_efficient_2014** and for growing neutrino models in **brouzakis_nonlinear_2011** Typically they rely on non-linear expansions of the perturbations using resummation techniques based on **pietroni_flowing_2008; taruya_closure_2008** or on fitting formulae based on N-Body simulations **casas_fitting_2015; takahashi_revising_2012; bird_massive_2011** A similar analysis is not available for the model-independent approach considered in this paper. In order to give at least a qualitative estimate of what the importance of non-linearities would be for constraining these Modified Gravity models, we will adopt in the rest of the paper a method which interpolates between the standard approach to non linear scales in GR and the same applied to MG theories.

4.3.2.1 Halofit

We describe here the effect of applying the standard approach to non linearities to MG theories. This is done using the revised Halofit **takahashi_revising_2012** based on **smith_stable_2003** which is a fitting function of tens of numerical parameters that reproduces the output of a certain set of ΛCDM N-body simulations in a specific range in parameter space as a function of the linear power spectrum. This fitting function is reliable with an accuracy of better than 10% at scales larger than $k \lesssim 1$h/Mpc and redshifts in between $0 \leq z \leq 10$ (see **takahashi_revising_2012** for more details). This fitting function can be used within Boltzmann codes to estimate the non-linear contribution which corrects the linear power spectrum as a function of scale and time. We will use the Halofit fitting function as a way of approximating the non-linear power spectrum in our models even though it is really only valid for ΛCDM. In Fig. 4.2, the left panel shows a comparison between the linear and non-linear power spectra calculated by MGCAMB in two different models, our fiducial late-time model (as from Table 4.1) and GR, both sharing the same ΛCDM parameters. At small length scales (large k), the non-linear deviation is clearly visible at scales $k \gtrsim 0.3$ h/Mpc and both MG and GR seem to overlap due to the logarithmic scale used. In the right panel, we can see the ratio between MG and GR for both linear and non-linear power spectra, using the same 5 ΛCDM parameters $\{\Omega_m, \Omega_b, h,\ A_s, n_s\}$. We can see clearly that MG in the non-linear regime, using the standard Halofit, shows a distinctive feature at scales in between $0.2 \lesssim k \lesssim 2$. This feature however, does not come from higher order perturbations induced by the modified Poisson equations (4.2, 4.4), because Halofit, as explained above, is calibrated with simulations within the ΛCDM model and does not contain any information from Modified Gravity. The feature seen here is caused by the different growth rate of perturbations in Modified Gravity, that yields then a different evolution of non-linear structures.

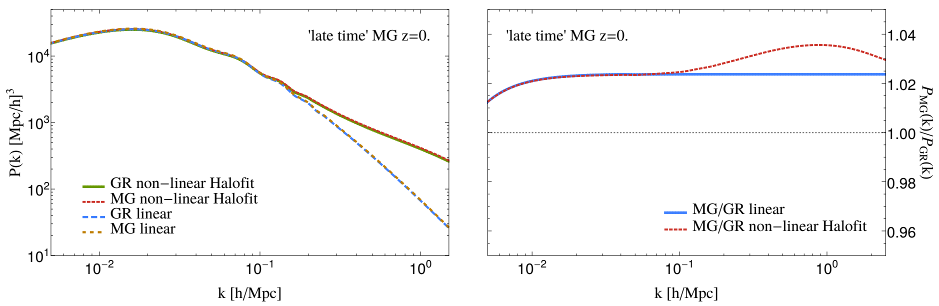


FIGURE 4.2: **Left:** matter power spectra computed with MGCAMB (linear) and MGCAMB+Halofit (non-linear), illustrating the impact of non-linearities at different scales. As an illustrative example, MG in this plot corresponds to the fiducial model in the late-time parametrization defined in Eq. (4.9). All curves are computed at $z = 0$. The green solid line is the GR fiducial in the non-linear case, the blue long-dashed line is also GR but in the linear case. The short-dashed red line is the MG fiducial in the non-linear case and the medium-dashed brown line the MG fiducial in the linear case. **Right:** in order to have a closer look at small scales, we plot here the ratio of the MG power spectrum to the GR power spectrum for the linear (blue solid) and non-linear (red short-dashed) cases separately. The blue solid line compared to the horizontal grey dashed line, shows the effect of Modified Gravity when taking only linear spectra into account. While the red dashed line, which represents the non-linear case, shows that the ratio to GR presents clearly a bump that peaks around $k \approx 1.0$ h/Mpc, meaning that the power spectrum in MG differs at most 4% from the non-linear power spectrum in GR. We will see later that we are able to discriminate between these two models using future surveys, especially when non-linear scales ($k \gtrsim 0.1$ h/Mpc) are included.

4.3.2.2 Prescription for mildly non-linear scales including screening

As discussed above, modifications to the Φ and Ψ potentials make the use of Halofit to compute the evolution at non linear scales unreliable. In order to take into account the non-linear contribution to the power spectrum in Modified Gravity, we investigate here a different method, which starts from the consideration that whenever we modify μ and η with respect to GR, we modify the strength of gravitational attraction in a way universal to all species: this means that, similarly to the case of scalar-tensor or $f(R)$ theories, we need to assume the existence of a non-perturbative screening mechanism, acting at small scales, that guarantees agreement with solar system experiments. In other words, it is reasonable to think that the non-linear power spectrum will have to match GR at sufficiently small scales, while at large scales it is modified. Of course, without having a specific model in mind, it remains arbitrary how the interpolation between the small scale regime and the large scale regime is done. In this paper, we adopt the Hu & Sawicki (HS) Parametrized Post-Friedmann prescription proposed in **hu_parameterized_2007** which was used for the case of $f(R)$ theories previously by **zhao_modeling_2014** Given a MG model, this prescription interpolates between the non-linear power spectrum in Modified Gravity (which is in our case just the linear MG power spectrum corrected with standard Halofit, $P_{\rm HMG}$) and the non-linear power spectrum in GR calculated with Halofit ($P_{\rm HGR}$). The resulting power spectrum will be denoted as $P_{\rm nlHS}$

$$P_{\rm nlHS}(k,z) = \frac{P_{\rm HMG}(k,z) + c_{\rm nl} S_{\rm L}^2(k,z) P_{\rm HGR}(k,z)}{1 + c_{\rm nl} S_{\rm L}^2(k,z)}\,, \tag{4.13}$$

with

$$S_{\rm L}^2(k,z) = \left[\frac{k^3}{2\pi^2} P_{\rm LMG}(k,z)\right]^s\,. \tag{4.14}$$

The weighting function $S_{\rm L}$ used in the interpolation quantifies the onset of non-linear clustering and it is constructed using the linear power spectrum in Modified Gravity ($P_{\rm LMG}$). The constant $c_{\rm nl}$ and the constant exponent s are free parameters. In Figure 4.3 we show the ratio $P_{\rm nlHS}/P_{\rm HGR}$, which illustrates the relative difference between the non-linear HS prescription in MG and the Halofit non-linear power spectrum in GR, for different values of $c_{\rm nl}$ (left panel) and different values of s (right panel). The parameter $c_{\rm nl}$ controls at which scale there is a transition into a non-linear regime in which standard GR is valid (this can be the case when a screening mechanism is activated); s controls the smoothness of the transition and is in principle a model and redshift dependent quantity. When $c_{\rm nl} = 0$ we recover the Modified Gravity power spectrum with Halofit $P_{\rm HMG}$; when $c_{\rm nl} \to \infty$ we recover the non-linear power spectrum in GR calculated with Halofit $P_{\rm HGR}$. In **zhao_modeling_2014**; **zhao_n-body_2011**; **koyama_non-linear_2009** the $c_{\rm nl}$ and s constants were obtained fitting expression (4.13) to N-Body simulations or to a semi-analytic perturbative approach. In the case of $f(R)$, $s = 1/3$ seems to match very well the result from simulations up to a scale of $k = 0.5{\rm h/Mpc}$ **koyama_non-linear_2009** A relatively good agreement up to such small scales is enough for our purposes. In the absence of N-Body simulations or semi-analytic methods available for the models investigated in this work, we will assume unity for both parameters, which is a natural choice, and we will test in Section 4.7.3 how our results vary for different values of these parameters, namely $c_{\rm nl} = \{0.1, 0.5,\ 1, 3\}$ and $s = \{0,\ 1/3,\ 2/3,\ 1\}$. This will give a qualitative estimate of the impact of non-linearities on the determination of cosmological parameters.

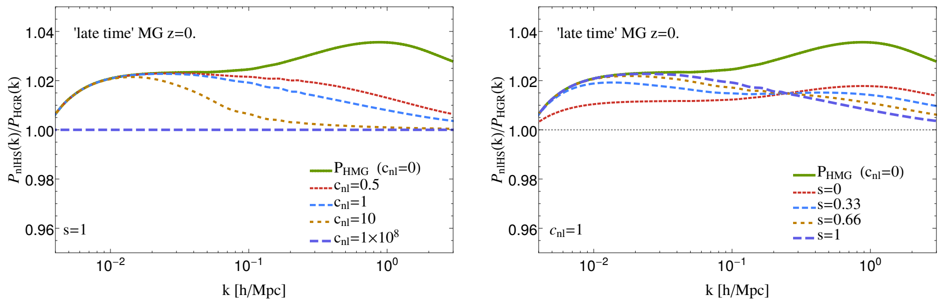


FIGURE 4.3: The ratio of the Modified Gravity non-linear power spectrum using the HS prescription by **hu_parameterized_2007** ($P_{\rm nlHS}$) with respect to the GR+Halofit fiducial non-linear power spectrum $P_{\rm HGR}$, for different values of $c_{\rm nl}$ (left panel) and s (right panel), illustrated in Eqns.(4.13, 4.14). The value $c_{\rm nl} = 0$ (green solid line) corresponds to MG+Halofit $P_{\rm HMG}$. All curves are calculated at $z = 0$. **Left:** We show the ratio for $c_{\rm nl} = \{0.5, 1.0, 10, 10^8\}$, plotted as short-dashed red, medium-dashed blue, short-dashed brown and medium-dashed purple lines respectively. When $c_{\rm nl} \to \infty$, Eqn. (4.13) corresponds to the limit of $P_{\rm HGR}$ and therefore the ratio is just 1. The effect of the HS prescription is to grasp some of the features of the non-linear power spectrum at mildly non-linear scales induced by Modified Gravity, taking into account that at very small scales, a screening mechanism might yield again just a purely GR non-linear power spectrum. The parameter $c_{\rm nl}$ interpolates between these two cases. **Right:** in this panel we show the effect of the parameter s, for $s = \{0, 0.33, 0.66, 1\}$ (short-dashed red, medium-dashed blue, short-dashed brown and long-dashed purple, respectively). Both parameters need to be fitted with simulations in order to yield a reliable match with the shape of the non-linear power spectrum in Modified Gravity, as it was done in **zhao_n-body_2011** and references therein. The grey dashed line marks the constant value of 1.

4.3.3 CMB *Planck* priors

Alongside the information brought by LSS probes, we also include CMB priors on the parameterizations considered. In order to obtain these, we analyze the binned and parameterized approaches described in Section 4.2 with the *Planck*+BSH combination of CMB and background (BAO+SN-Ia+H_0) datasets discussed in the *Planck* Dark Energy and Modified Gravity paper **planck_collaboration_planck_2016** We use a Markov Chain Monte Carlo (MCMC) approach, using the publicly available code COSMOMC **lewis_cosmological_2002**; **lewis_efficient_2013** interfaced with our modified version of MGCAMB. The MCMC chains sample the parameter vector Θ which contains the standard cosmological parameters $\{\omega_b \equiv \Omega_b h^2,\ \omega_c \equiv \Omega_c h^2,\ \theta_{\rm MC},\ \tau, n_s, \ln 10^{10} A_s\}$ to which we add the E_{ij} parameters when we parameterize the time evolution of μ and η with continuous functions of the scale factor, and the $\mu_i,\ \eta_i$ parameters in the binned case. On top of these, also the 17 nuisance parameters of the *Planck* analysis are included. From the MCMC analysis of the *Planck* likelihood we obtain a covariance matrix in terms of the parameters Θ. We marginalize over the nuisance parameters and over the optical depth τ since this parameter does not enter into the physics of large scale structure formation.

θ is usually the ratio of sound horizon to the angular diameter distance at the time of decoupling. Since calculating the decoupling time $z_{\rm CMB}$ is relatively time consuming, as it involves the minimization of the optical transfer function, COSMOMC uses instead an approximate quantity $\theta_{\rm MC}$ based on the following fitting formula from **hu_small_1996**

$$\begin{aligned} z_{\rm CMB} &= 1048 \times (1 + 0.00124\omega_b^{-0.738}) \\ &\times \left(1 + 0.0783\omega_b^{-0.238}/(1 + 39.5\omega_b^{0.763}\right) \\ &\times \left(\omega_d + \omega_b)^{0.560/(1+21.1\omega_b^{1.81})}\right) \end{aligned} \tag{4.15}$$

where $\omega_d \equiv (\Omega_c + \Omega_\nu)h^2$. The sound horizon is defined as

$$r_s(z_{\rm CMB}) = cH_0^{-1} \int_{z_{\rm CMB}}^{\infty} {\rm d}z \frac{c_s}{E(z)} \tag{4.16}$$

where the sound speed is $c_s = 1/\sqrt{3(1 + \overline{R}_b a)}$with the baryon-radiation ratio being $\overline{R}_b a = 3\rho_b/4\rho_\gamma$. $\overline{R}_b = 31500\Omega_b h^2(T_{\rm CMB}/2.7{\rm K})^{-4}$. However, CAMB approximates it as $\overline{R}_b a = 30000a\Omega_b h^2$.

Therefore we first marginalize the covariance matrix over the nuisance parameters and the parameter τ, which cannot be constrained by LSS observations. Then, we invert the resulting matrix, to obtain a *Planck* prior Fisher matrix and then use a Jacobian to convert between the MCMC parameter basis Θ_i and the GC-WL parameter basis θ_i. We use the formulas above for the sound horizon r_s and the angular diameter distance d_A to calculate the derivatives of $\theta_{\rm MC}$ with respect to the parameters of interest. Our Jacobian is then simply (see Appendix **??** for details)

$$J_{ij} = \frac{\partial \Theta_i}{\partial \theta_j} \, . \tag{4.17}$$

4.4 Results: Euclid forecasts for redshift binned parameters

In this section we analyze the Modified Gravity functions $\mu(a)$ and $\eta(a)$, described in Section 4.2, when they are allowed to vary freely in five redshift bins.

For this purpose, we calculate a Fisher matrix of fifteen parameters: five for the standard ΛCDM parameters $\{\Omega_m, \Omega_b, h, \ln 10^{10}A_s, n_s\}$, five for μ (one for each bin amplitude μ_i) and five for η (one for each bin amplitude η_i), corresponding to the 5 redshift bins z={0-0.5, 0.5-1.0, 1.0-1.5, 1.5-2.0, 2.0-2.5}. The fiducial values for all fifteen parameters were calculated running a Markov-Chain-Monte-Carlo with *Planck* likelihood data and can be found in Table 4.1.

We first show the constraints on our 15 parameters for Galaxy Clustering (GC) forecasts in subsection 4.4.1, while in subsection 4.4.2 we report results for Weak Lensing (WL). In subsection 4.4.3, we comment on the combination of forecasts for GC+WL together with *Planck* data. All forecasts are performed using Euclid Redbook specifications. Other surveys will be considered for the other two time parameterizations in section 4.7.1.6 and 4.7.2.3. For each case, we show the correlation matrix obtained from the covariance matrix and argue that the redshift-binned parameters show a strong correlation, therefore we illustrate the decorrelation procedure for the covariance matrix in subsection 4.4.4 where we also include combined GC+WL and GC+WL+*Planck* cases.

4.4.1 Euclid Galaxy Clustering Survey

For the Galaxy Clustering survey, we give results for two cases, one using only linear power spectra up to a maximum wavevector of $k_{\text{max}} = 0.15$ h/Mpc and another one using non-linear power spectra up to $k_{\text{max}} = 0.5$ h/Mpc, as obtained by using the HS parameterization of Eqn. (4.13). For the redshift-binned case, we will report forecasts only for a Euclid survey, using Euclid Redbook specifications which are detailed in section 3.4.1.

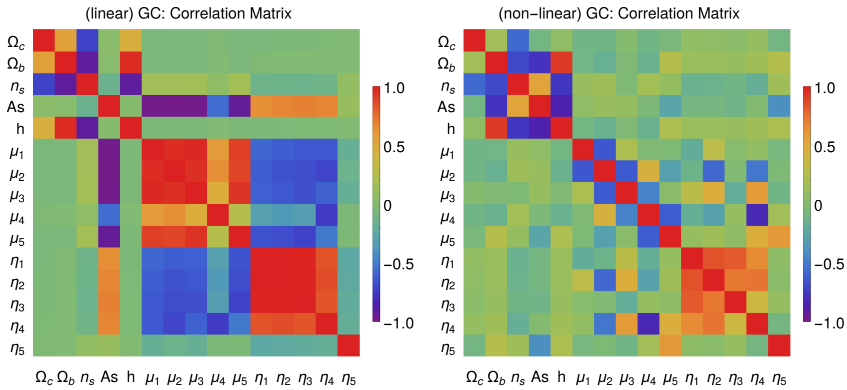


FIGURE 4.4: Correlation matrix $\mathbf{P}$ defined in (3.38) obtained from the covariance matrix in the MG-binning case, for a Galaxy Clustering Fisher forecast using Euclid Redbook specifications. **Left panel:** Linear forecasts. Here there are strong positive correlations among the μ_i and η_i parameters and anti-correlations between $\ln 10^{10} A_s$ and the μ_i parameters, as well as between μ_i and η_i. The FoC in this case is ≈ 65. (see Eqn. (3.40) for its definition). **Right panel:** Non-linear forecasts using the HS prescription. Interestingly, the anti-correlations between $\ln 10^{10} A_s$ and μ_i have disappeared, as well as the correlations among the μ_i parameters. The FoC is in this case ≈ 32, meaning that the variables are much less correlated than in the linear case. This is due to the fact that taking into account non-linear structure formation breaks degeneracies between the primordial amplitude parameter and the modifications to the Poisson equation.

We calculate the Fisher matrix for the 15 parameters $\theta = \{\Omega_m, \Omega_b, h, \ln 10^{10} A_s, n_s, \mu_i, \eta_i\}$ where η_i and μ_i represent ten independent parameters, one for each function at each of the 5 redshift bins corresponding to the redshifts z={0-0.5, 0.5-1.0, 1.0-1.5, 1.5-2.0, 2.0-2.5}. As a standard procedure, we marginalize over the unknown bias parameters. From the covariance matrix, defined previously in Eqn. (3.35), we obtain the correlation matrix P_{ij} defined in Eqn. (3.38) for the set of parameters θ_i. In Figure 4.4 we show the matrix P_{ij} in the linear (left panel) and non-linear-HS (right panel) cases. Redder (bluer) colors signal stronger correlations (anti-correlations).

A covariance matrix that contains strong correlations among parameter A and B, means that the experimental or observational setting has difficulties distinguishing between A and B for the assumed theoretical model, i.e. this represents a parameter degeneracy. Therefore if for example parameter A is poorly constrained, then parameter B will be badly constrained as well. The appearance of correlations among parameters is linked to the non-diagonal elements of the covariance matrix. Subsequently, this means that the fully marginalized errors on a single parameter, will be larger if there are strong correlations and will be smaller (closer to the value of the fully maximized errors) if the correlations are negligible.

In the linear case, μ_i and η_i parameters show correlations among each other, while the primordial amplitude parameter $\ln 10^{10} A_s$ exhibits a strong anti-correlation with all the μ_i. This can be explained considering that a larger growth of structures in linear theory can also be mimicked with a larger initial amplitude of density fluctuations.

Interestingly, including non-linear scales in the analysis (right panel of Fig. 4.4) leads to a strong suppression of the correlations among the μ_i. Also the correlation between these and $\ln 10^{10} A_s$ is suppressed as a change in the initial amplitude of the power spectrum is not able to compensate for a modified Poisson equation when non-linear evolution is considered.

As discussed in Section 3.4.5, we can also express the difference between the correlation matrix of the linear forecast and the non-linear forecast in a more quantitative way, by computing the determinant of the correlation matrix, or equivalently the FoC (3.40). If the correlations were negligible, this determinant would be equal to one (and therefore its FoC would be 0), while if the correlations were strong, the determinant would be closer to zero with a corresponding large positive value of the FoC. For the linear forecast, the FoC is about 62, while for the non-linear forecast, it is much smaller at approximately 35. In Table 4.2 we show the 1σ constraints obtained on $\ln(10^{10} A_s)$ and on the μ_i and η_i parameters, both in the linear and non-linear cases for a Euclid Redbook GC survey (top rows). While linear GC alone ($k_{\max}$ = 0.15 h/Mpc) is not very constraining in any bin, the inclusion of non-linear scales ($k_{\max}$= 0.5 h/Mpc) drastically reduces errors on the μ_i parameters: the first three bins in μ_i (0. < z < 1.5) are the best constrained, to less than 10%, with the corresponding η_i constrained at 20% by non-linear GC alone. This is also visible in the FoM which increases by 19 nits ('natural units', similar to bits but using base e instead of base 2), nearly 4 nits per redshift bin on average, when including the non-linear scales. The fact that the error on $\ln 10^{10} A_s$ improves from 90% to 0.68% shows that the decorrelation induced by the non-linearities breaks the degeneracy with the amplitude and therefore improves considerably the determination of cosmological parameters. This shows that it is important to include non-linear scales in GC surveys (and not only in Weak Lensing ones, which is usually more expected and will be shown in the next subsection).

4.4.2 Euclid Weak Lensing Survey

In the case of Weak Lensing, the linear forecast is performed with linear power spectra up to a maximum multipole $\ell_{\max} = 1000$, while the non-linear forecast is performed with non-linear spectra up to a maximum multipole of $\ell_{\max} = 5000$, as explained in section **??**. Since we limit our power spectrum to a maximum in k-space, as explained in section **??**, these multipole values are not reached at every redshift. Like for GC, also for WL it is very important to include information from the non-linear power spectrum, since in that range will lie most of the constraining power of next-generation surveys like Euclid. In Figure 4.5 we show the correlation matrices for the linear (left panel) and non-linear (right panel) Fisher Matrix forecasts. In this case, as opposed to the GC correlation matrices, it is not visually clear which case is more correlated than the other. At a closer look, in the linear case we can observe strong anti-correlations between the μ_j and η_j parameters for $j = 2, 3, 4, 5$ and an anti-correlation between η_1 and the primordial amplitude $\ln 10^{10} A_s$. In the non-linear case, the primordial amplitude parameter is effectively decorrelated from the Modified Gravity parameters, and the anti correlation between μ_j and the η_j affects the first three bins (effectively increasing degeneracies in the first bin). The anti-correlation between these two sets of parameters is expected, since WL is sensitive to Σ, which is a product of η and μ, c.f. Eqn. (4.5). The decrease of correlation when going from the linear to the non-linear case is confirmed, also quantitatively, by the FoC: the one for the linear case is of ≈ 69, larger than the one for the non-linear case ≈ 32. Once again, the inclusion of non-linear scales breaks degeneracies, especially among the linear amplitude of the power spectrum and the MG parameters.

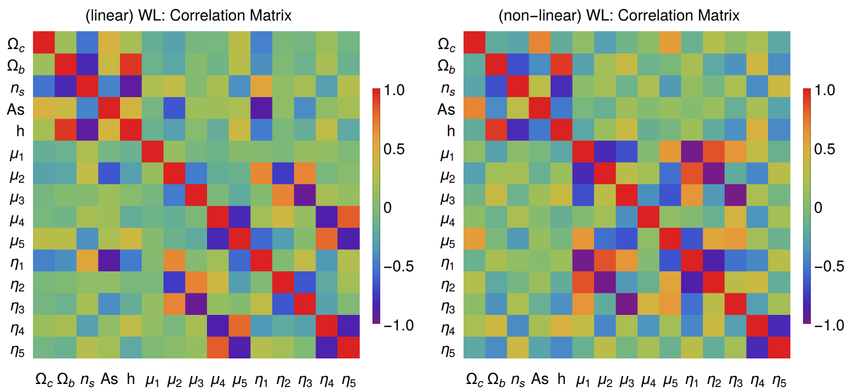


FIGURE 4.5: Correlation matrix obtained from the covariance matrix in the MG-binning case, for a Weak Lensing Fisher Matrix forecast using Euclid Redbook specifications. **Left panel:** linear forecasts. Strong anti-correlations are present between the μ_i and the η_i parameters for the same value of i in 2,3,4,5. The amplitude $\ln(10^{10}A_s)$ parameter is mostly uncorrelated except with η_1. The FoC (3.40) in this case is approximately 69. **Right panel:** non-linear forecasts using the HS prescription. Here the same trend as in the linear case is present, with just subtle changes. The FoC in this case is about 32, meaning that the variables are indeed less correlated than in the linear case. The parameter $\ln(10^{10}A_s)$ is effectively not correlated to other parameters, and the anti-correlation of μ_i and the η_i for the same value of the index i is present in the first three bins. The anti-correlation between these two sets of parameters is expected, since WL is sensitive to the Weyl potential Σ, which is a product of μ and η.

Table 4.2 shows the corresponding 1σ marginalized errors on $\ln\left(10^{10}A_s\right)$ and on the μ_i and η_i, both in the linear and non-linear cases for a Euclid Redbook WL survey. As in the case of GC, linear WL cannot constrain alone any of the amplitudes of the Modified Gravity parameters μ_i and η_i for any redshift bin. Being able to include non-linear scales improves constraints on the amplitude $\ln(10^{10}A_s)$ by a factor 100. The 1σ errors on μ_i and η_i improve up to one order of magnitude, with the FoM increasing by 17 nits, although remaining quite unconstraining for Modified Gravity parameters in all redshift bins. Notice, however, that the 1σ error on μ_1 from WL in the linear case is slightly smaller than in the non-linear case. This can be attributed to the fact that in the linear case, μ_1 is uncorrelated to any other parameter, as shown in Figure 4.5 and on that specific bin (0 < z < 0.5) non-linearities don't seem to improve the constraints on this parameter.

4.4.3 Combining Euclid Galaxy Clustering and Weak Lensing, with *Planck* data

The combination of Galaxy Clustering and Weak Lensing is expected to be very powerful for Modified Gravity parameters as they measure two different combinations of μ and η, thus breaking their degeneracy as illustrated in Fig. 4.17. This is shown in Table 4.2, where the sensitivity drastically increases with respect to the two separate probes, especially in the low redshifts bins ($0. < z < 1.5$), where the lensing signal is dominant. Adding non-linearities further doubles the FoM. The *Planck* data constrains mostly the standard ΛCDM parameters and has only a limited ability to constrain the MG sector. However, the additional information breaks parameter degeneracies and in this way significantly decreases the uncertainties on all parameters, so that the linear GC+WL+*Planck* is comparable to the non-linear GC+WL. Also, quantitatively, the correlation among parameters is reduced by combining GC+WL with *Planck* data. The FoC in this case is of ≈ 22. We also see in Table 4.2 that the differences between the non-linear prescription adopted here (nl-HS) and a straightforward application of Halofit to the MG case (nl-Halofit) is not very large. We will investigate further the impact of the parameters used in the non-linear prescription in section 4.7.3.

4.4.4 Decorrelation of covariance matrices and the Zero-phase Component Analysis

In the previous subsections we highlighted how the MG parameters and the amplitude of the primordial power spectrum exhibit significant correlations and showed how including non-linearities helps to decorrelate them. Even without including non-linearities, however, it is interesting to investigate how we can completely decorrelate the parameters, identifying in this way those parameter combinations which are best constrained by data.

Given again a *d-dimensional* vector p of random variables (our originally correlated parameters), we can calculate its covariance matrix $\mathbf{C}$ defined in Eqn. (3.35). The process of decorrelation is the process of making the matrix $\mathbf{C}$ a diagonal matrix.

Let us define some important identities. The covariance matrix can be decomposed in its eigenvalues (the elements of a diagonal matrix Λ) and eigenvectors (the rows of an orthogonal matrix U).

$$\mathbf{C} = U\Lambda U^T \quad \Leftrightarrow \quad \mathbf{F} = U\Lambda^{-1}U^T \ , \tag{4.18}$$

where $\mathbf{F}$ is the Fisher Matrix.

It is possible to show that applying a transformation matrix W to the p parameter vector, thus obtaining a new vector of variables $q = Wp$, the covariance matrix of the transformed vector q is whitened (i.e. it is the identity matrix, and whitening is defined as the process of converting the covariance matrix into an identity matrix)

$$\begin{aligned} \tilde{\mathbf{C}} &= W\langle\Delta\hat{p}\Delta\hat{p}^T\rangle W^T = \langle\Delta\hat{q}\Delta\hat{q}^T\rangle \\ &= \mathbb{1} \ . \end{aligned} \tag{4.19}$$

This means that the transformed q parameters are decorrelated, since their correlation matrix is diagonal. The choice of W is not unique as several possibilities exist; we focus on a particular choice in the rest of the paper referred to as Zero-phase

Component Analysis (ZCA, first introduced by **Bell19973327** in the context of image processing), but we show two other possible choices and their effect on the analysis in Appendix 4.6.

Zero-phase Component Analysis (sometimes also called Mahalanobis transformation **kessy_optimal_2015**) is a specific choice of decorrelation method that minimizes the squared norm of the difference between the q_i and the p_i vector $\|\vec{p} - \vec{q}\|$, under the constraint that the vector $\vec{q}$ should be uncorrelated **kessy_optimal_2015** In this way the uncorrelated variables q will be as close as possible to the original variables p in a least squares sense. This is achieved by using the W matrix:

$$W \equiv F^{1/2} \quad . \tag{4.20}$$

Then in this case, the covariance matrix is whitened, by following Eqn. (4.19):

$$\tilde{C} = F^{1/2} F^{-1} F^{1/2} = \mathbb{1} \quad . \tag{4.21}$$

In our case, since we do not want to whiten, but just decorrelate the covariance matrix, we renormalize the W matrix by multiplying with a diagonal matrix $N_{jj} \equiv \sum_j (W_{ij}^{-2})$, such that the sum of the square of the elements on each row of the new weighting matrix $\tilde{W} \equiv NW$, is equal to unity; therefore the final transformed covariance is still diagonal but is not the identity matrix:

$$\tilde{C} = \tilde{W} C \tilde{W} = N^2 \mathbb{1} \tag{4.22}$$

and at the same time we ensure that the vector of new variables q_i will have the same norm as the old vector of variables p_i.

4.4.4.1 ZCA for Galaxy Clustering

From Fig. 4.4 we can see that the correlations are present in sub-blocks, one for the standard ΛCDM parameters and another one for the Modified Gravity parameters. The exception lies in the linear case where $\ell\mathcal{A}_s \equiv \ln(10^{10} A_s)$ is strongly anti-correlated with all the μ_i and positively correlated with the η_i. To use a more objective criterion, we choose the 10×10 block of MG parameters μ_i and η_i with parameter indices 6 to 15, and only add to this block a ΛCDM parameter with index a if the following condition is satisfied:

$$\sum_{i=6}^{i=15} (P_{ai})^2 \geq 1$$

where the index a corresponds to one of the first five standard parameters. For Galaxy Clustering, the only index satisfying this condition is $a = 4$ in the linear case, corresponding to $\ell\mathcal{A}_s \equiv \ln(10^{10} A_s)$: i.e. the standard parameter corresponding to the amplitude is, as said, degenerate with Modified Gravity parameters μ_i and η_i. In the non-linear case no parameter satisfies this condition (because, as we have seen, non-linearities are able to eliminate correlation with the amplitude), but for consistency we will use the same vector of 11 parameters $p_i = \{\ell\mathcal{A}_s, \mu_i, \eta_i\}$ for our decorrelation procedure. Therefore we will also have 11 transformed uncorrelated q_i parameters, function of the original p_i parameters, in all the cases presented below.

Figure 4.6 shows the coefficients that relate the q_i parameters to the original p_i ones, in the linear (left panel) and the non-linear (right panel) cases, also shown explicitly in Tables **??** and **??** of Appendix 4.5. We plot in Figure 4.7 a comparison between the 1σ errors on the primary parameters p_i (represented by circles connected with dark green dashed lines) and the decorrelated parameters q_i (represented by squares connected with orange solid lines). In the linear case (left panel), we can see that the errors on the q_i parameters are 2 orders of magnitude better than the errors on the p_i parameters. In the non-linear case (right panel) the improvement is of at most 1 order of magnitude and that for a completely decorrelated parameter like $\ell\mathcal{A}_s$, the error on its corresponding q_i is exactly the same. This shows that a decorrelation procedure is still worth to do, even when including the non-linear regime, even if the degeneracy with the amplitude is already completely broken thanks to the non-linear prescription. The fact that the curve of 1σ errors for the q_i follows the same pattern as the curve for the p_i errors, is due to the fact that we have used a ZCA decomposition (see Section 4.4.4) and therefore the q_i are as similar as possible to the p_i.

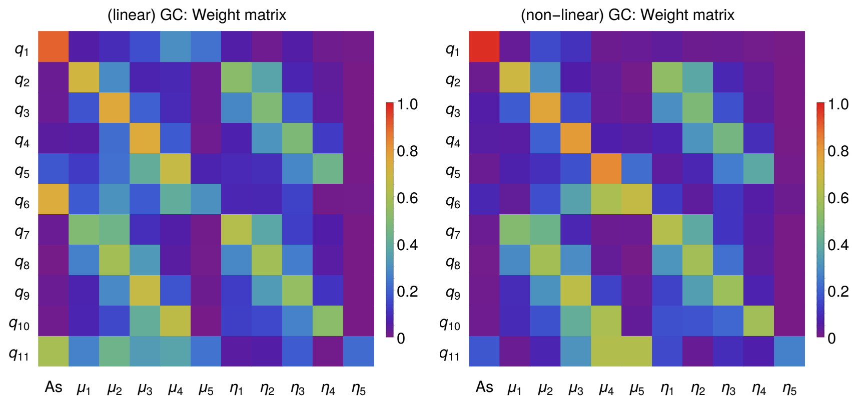


FIGURE 4.6: Entries of the matrix W that relates the q_i parameters to the original p_i ones, after applying the ZCA decorrelation of the covariance matrix in the linear and non-linear GC cases. This matrix shows for each new variable q_i on the vertical axis, the coefficients of the linear combination of parameters μ_i, η_i and A_s that give rise to that variable q_i. The red (blue) colors, indicate a large (small) contribution of the respective variable on the horizontal axis. **Left panel:** linear forecast for Euclid Redbook specifications. **Right panel:** non-linear forecast for Euclid Redbook specifications, using the HS prescription. In both cases one can observe that most q_i parameters have only small or negligible contributions from μ_5 and η_5, which are found to be the less constrained bins.

We are interested in finding the best combination of primary parameters p_i giving rise to the best constrained uncorrelated parameters q_i. In order to find the errors on the parameters q_i, we need to look at the diagonal of the decorrelated covariance matrix $\tilde{\mathbf{C}}$ expressed in Eqn. (4.22) and identify the q_i parameters with the smallest relative errors (σ_{q_i}/q_i): we find than in the linear GC case, the best constrained combinations of primary parameters (ordered from most to least constrained) are given approximately by:

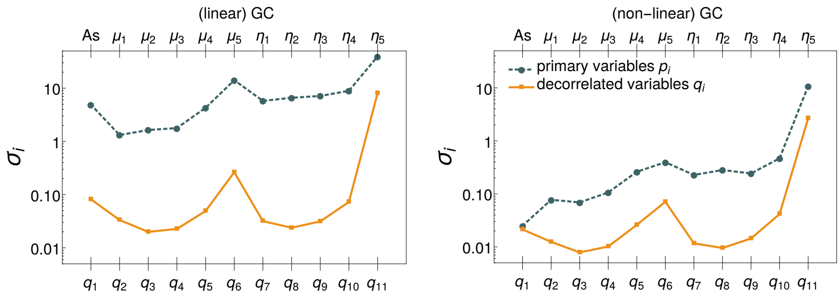


FIGURE 4.7: Results for a Euclid Redbook GC survey, with redshift-binned parameters, before and after applying the ZCA decorrelation. Each panel shows the 1σ fully marginalized errors on the primary parameters p_i (green dashed lines), and the 1σ errors on the decorrelated parameters q_i (orange solid lines). **Left:** linear forecasts, performed using linear power spectra up to a maximum wavenumber $k_{\max} = 0.15\text{h/Mpc}$. **Right:** non-linear forecasts using non-linear spectra with the HS prescription up to a maximum wavenumber $k_{\max} = 0.5\text{h/Mpc}$. In the linear case, the errors on the decorrelated q_i parameters are about 2 orders of magnitude smaller than for the primary parameters, while in the non-linear HS case, the improvement in the errors is of one order of magnitude. This means that applying a decorrelation procedure is worth even when non-linearities are considered. In both cases for GC, the least constrained parameters are μ_5 and η_5, corresponding to $2.0 < z < 2.5$.

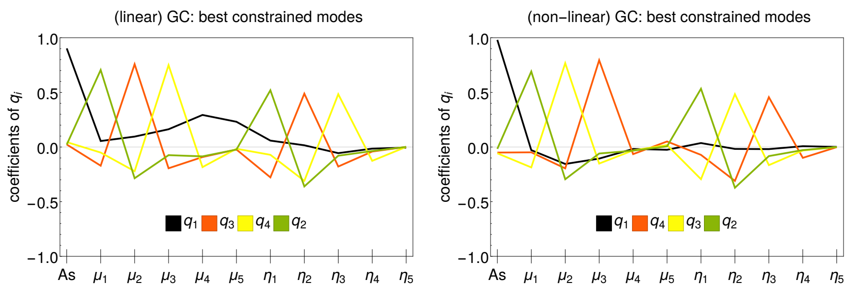


FIGURE 4.8: Best constrained modes for a Euclid Redbook GC survey, with μ and η binned in redshift, after transforming into uncorrelated q parameters via ZCA. Each of the four best constrained parameters q_i, shown in the panels, is a linear combination of the primary parameters p_i. The q_i in the legends are ordered from left to right, from the best constrained to the least constrained.

$$
\begin{aligned}
q_1 &= +0.9\ell\mathcal{A}_s + 0.32\mu_4 \\
q_3 &= +0.75\mu_2 - 0.29\eta_1 + 0.50\eta_2 \\
q_4 &= -0.25\mu_2 + 0.74\mu_3 - 0.32\eta_2 + 0.49\eta_3 \\
q_2 &= +0.70\mu_1 - 0.30\mu_2 + 0.52\eta_1 - 0.36\eta_2 \quad .
\end{aligned}
\tag{4.23}
$$

In contrast, for the non-linear GC case, the parameter $\ell\mathcal{A}_s \equiv \ln(10^{10}A_s)$ is not correlated to any other, and therefore it is well constrained on its own. The best 4 constrained parameters (ordered from most to least constrained) in the non-linear case, are:

$$
\begin{aligned}
q_1 &= +0.99\ell\mathcal{A}_s \\
q_4 &= -0.28\mu_2 + 0.76\mu_3 - 0.33\eta_2 + 0.47\eta_3 \\
q_3 &= +0.73\mu_2 - 0.32\eta_1 + 0.49\eta_2 \\
q_2 &= +0.68\mu_1 - 0.35\mu_2 + 0.52\eta_1 - 0.37\eta_2 \ .
\end{aligned}
\tag{4.24}
$$

The best constrained decorrelated parameters q_i for a Euclid GC survey, expressed in the set of Equations (4.23) (linear) and (4.24) (non-linear HS), can be seen graphically in Fig. 4.8 for the linear (left panel) and non-linear HS (right panel) cases respectively. From these combinations we see that a survey like Euclid, using GC only, will be sensitive to Modified Gravity parameters μ and η mainly in the first three redshift bins, corresponding to a range $0. < z < 1.5$. The complete matrix W of coefficients relating the q_i to the p_i parameters, can be found in Tables **??** and **??** of Appendix 4.5.

4.4.4.2 ZCA for Weak Lensing

We apply the same decorrelation procedure to the WL case, obtaining the q vectors shown in the weight matrix of Figure 4.9 again reported explicitly in Tables **??** and **??** in Appendix 4.5.

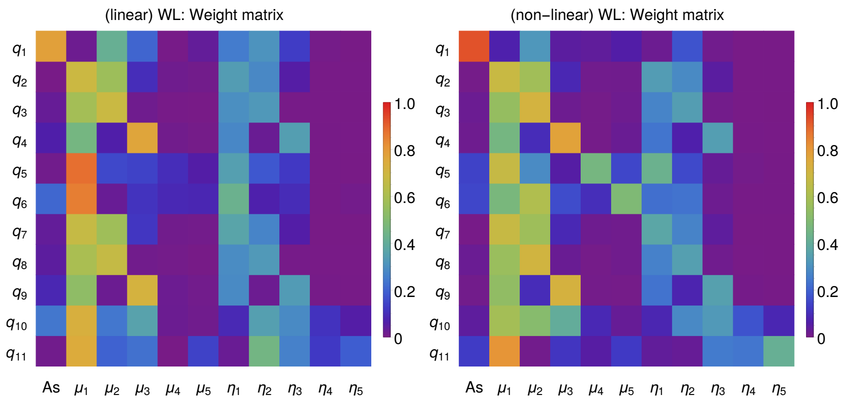


FIGURE 4.9: Entries of the matrix W that relates the q_i parameters to the original p_i ones, after applying the ZCA decorrelation of the covariance matrix in the linear and non-linear WL cases. This matrix shows for each new variable q_i on the vertical axis, the coefficients of the linear combination of parameters μ_i, η_i and A_s that give rise to that variable q_i. The red (blue) colors, indicate a large (small) contribution of the respective variable on the horizontal axis. **Left panel:** linear forecast for Weak Lensing Euclid Redbook specifications. **Right panel:** non-linear forecast for Weak Lensing Euclid Redbook specifications, using the HS prescription. As for GC, most q_i parameters have only small or negligible contributions from μ_5 and η_5, which are found to be the less constrained bins.

In Figure 4.10 we show the comparison between the errors on the primary parameters p_i and the de-correlated ones q_i. As for the GC case, the errors in the linear case improve by 2 orders of magnitude after applying the decorrelation procedure (left panel). In the non-linear case (right panel) the improvement is smaller, but still worth to do, especially to constrain q_2, q_3, q_7, q_8.

More generally, as we did for the GC case in the previous section, we look for the q_i parameters with the smallest relative errors (σ_{q_i}/q_i) and find in the linear

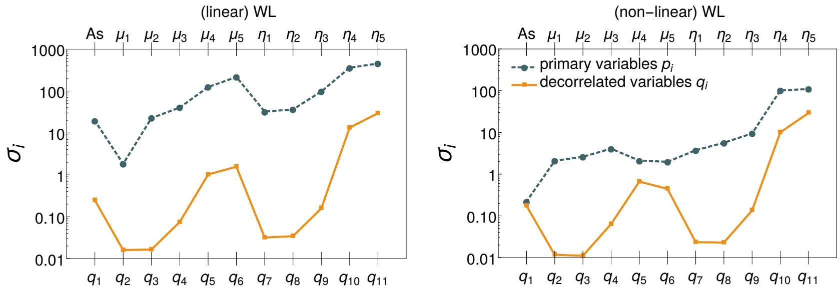


FIGURE 4.10: Results for a Euclid Redbook WL survey, with redshift-binned parameters, before and after applying the ZCA decorrelation. Each panel shows the 1σ fully marginalized errors on the primary parameters p_i (green dashed lines), and the 1σ errors on the decorrelated parameters q_i (orange solid lines). **Left:** Linear forecasts, performed with an $\ell_{\rm max} = 1000$ and linear matter power spectra. **Right:** Non-linear forecasts using the non-linear spectra with the HS prescription, up to an $\ell_{\rm max} = 5000$. The errors in the non-linear HS case, are about 1 order of magnitude smaller than in the linear case. For the best constrained q_i parameters, the decorrelated errors are up to 2 orders of magnitude smaller than the corresponding fully marginalized parameters on the parameters p_i.

WL case, that the best constrained combinations (ordered from most to least constrained) of primary parameters are given approximately by:

$$\begin{aligned} q_1 &= +0.76\ell\mathcal{A}_s + 0.48\mu_2 + 0.33\eta_2 \\ q_3 &= -0.59\mu_1 + 0.67\mu_2 - 0.30\eta_1 + 0.32\eta_2 \\ q_7 &= +0.65\mu_1 - 0.60\mu_2 + 0.36\eta_1 - 0.28\eta_2 \\ q_2 &= +0.67\mu_1 - 0.59\mu_2 + 0.33\eta_1 - 0.29\eta_2 \ . \end{aligned} \tag{4.25}$$

This means that WL in the linear case will only be able to constrain combinations of the first two redshift bins in μ and η (corresponding to $0. < z < 1.0$). This can also be observed graphically in the left panel of Figure 4.11. For the non-linear WL case, the combinations remain practically the same, except for q_1, which will depend much more strongly on the parameter $\ell\mathcal{A}_s$. The best 4 constrained parameters in this case, are (ordered from most to least constrained):

$$\begin{aligned} q_3 &= -0.55\mu_1 + 0.71\mu_2 + -0.27\eta_1 + 0.34\eta_2 \\ q_1 &= +0.93\ell\mathcal{A}_s - 0.32\mu_2 \\ q_2 &= +0.67\mu_1 - 0.60\mu_2 + 0.33\eta_1 - 0.29\eta_2 \\ q_4 &= -0.46\mu_1 + 0.29\mu_2 + 0.73\mu_3 + 0.31\eta_3 \ . \end{aligned} \tag{4.26}$$

These combinations can also be visualized in the right panel of Figure 4.11. The complete matrix W of coefficients relating the q_i to the p_i parameters, can be found in Tables **??** and **??** of Appendix 4.5.

4.4.4.3 ZCA for Weak Lensing + Galaxy Clustering + CMB *Planck* priors

As mentioned earlier, Galaxy Clustering and Weak Lensing are particularly important, combined together, to constrain Modified Gravity parameters, as they probe

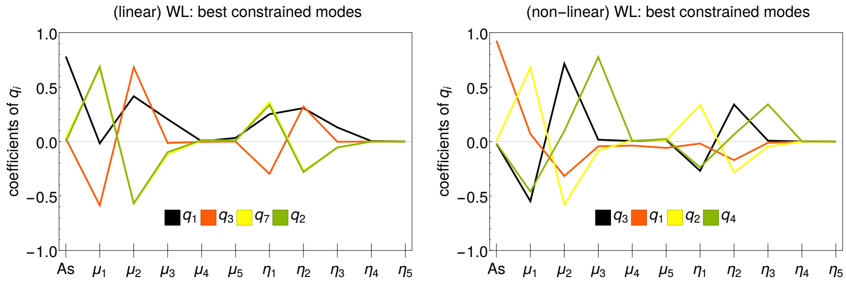


FIGURE 4.11: Best constrained modes for a Euclid Redbook WL survey, with μ and η binned in redshift, after transforming into uncorrelated q parameters via ZCA. Each of the four best constrained parameters q_i, shown in the panels, is a linear combination of the primary parameters p_i. q_i in the label are ordered from the best constrained to the least constrained.

two independent combinations of the gravitational potentials. We now show results for their combination, using for both the non-linear HS prescription, together with a *Planck* prior (which was obtained by performing an MCMC analysis on *Planck*+BSH background data, as specified in Section 4.3.3). Notice that we neglect here any information coming from the cross correlation of the two probes; we therefore assume that these two observables are independent of each other and we simply add the GC and WL Fisher matrices to obtain our combined results; this appears to be a conservative (pessimistic) choice **Lacasa2016** In Table 4.2, we can see that the inclusion of the *Planck* prior improves considerably certain parameters, especially the less constrained ones by GC+WL, namely $\mu_{4,5}$ and $\eta_{4,5}$. In terms of correlations, we can observe in the left panel of Fig. 4.12, that the structure of the correlation matrix resembles the one of the linear WL case (Fig. 4.5), except that the block of standard ΛCDM parameters is much less correlated and that the anti-correlation among μ_i and η_i is much stronger now. On the other hand, applying the decorrelation procedure (Section 4.4.4), the weight matrix W (right panel of Fig. 4.12), resembles the W matrix observed in the non-linear GC case, illustrated in Figure 4.6). Notice that now the variables q_i depend quite strongly on only one of the μ_i for $i = \{1, 2, 3, 4\}$, while the q_i for $i = \{6, 7, 8, 9, 10\}$ depend on a balanced sum of μ_i and η_i for $i = \{6, 7, 8, 9, 10\}$.

Finally, in this combined case the best constrained q_i variables, are $q_1,\ q_2,\ q_3,$, q_4 approximately given by:

$$\begin{aligned} q_1 &= +0.93\ell\mathcal{A}_s \\ q_2 &= +0.84\mu_1 + 0.48\eta_1 \\ q_3 &= +0.80\mu_2 - 0.26\eta_1 + 0.45\eta_2 \\ q_4 &= +0.28\ell\mathcal{A}_s + 0.79\mu_3 - 0.29\eta_2 + 0.39\eta_3 \ . \end{aligned} \tag{4.27}$$

These combinations of primary parameters are illustrated in the right panel of Figure 4.13. The combination q_2 is similar to the combination $2\mu + \eta$ that was also identified in **planck_collaboration_planck_2016** as being well-constrained. The best constrained modes q_2, q_3 and q_4 all contain terms of the form $a\mu_i + b\eta_i$ for $i = \{1, 2, 3\}$, with positive coefficients a and b, where $a \approx 2b$.

All errors are shown in the left panel of Figure 4.13. Notice how in this case, the improvement on the errors of the q_i variables is less than an order of magnitude, thus smaller than what found in GC and WL separately; this is due to combination

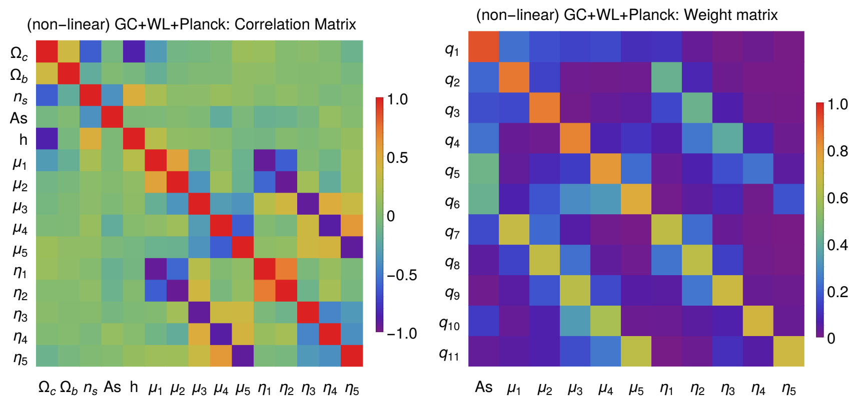


FIGURE 4.12: Results for the combined forecasts of Euclid Redbook GC+WL using the non-linear HS prescription together with the addition of *Planck* CMB priors. **Left panel:** correlation matrix obtained from the covariance matrix in the MG-binning case. Red (purple blue) colors represent strong positive (negative) correlations. The structure of this matrix is considerably diagonal, except for the strong anti-correlations of the pair (μ_i, η_i) for $i = \{1, 2, 3, 4, 5\}$, which resembles the correlations found for the WL case alone (see Figure 4.5). However, the sub-block of standard cosmological parameters is now much more diagonal and shows less correlations than in the GC (Fig. 4.4) or WL cases. The natural FoC (defined in 3.40) in this case is ≈ 22, showing that the variables are much less correlated than in the two previous cases. **Right panel:** entries of the matrix W for the ZCA decorrelation of the covariance matrix. This matrix shows for each new variable q_i on the vertical axis, the coefficients of the linear combination of parameters μ_i and η_i that give rise to that variable q_i. The red (blue) colors, indicate a large (small) contribution of the respective variable on the horizontal axis.

of GC and WL which, together with the inclusion of the CMB prior, lead to smaller correlations among the parameters. When combining GC+WL in the non-linear HS case, the FoC (defined in 3.40) is ≈ 31, showing that there is not much gain in decorrelation, compared to GC or WL alone, where this quantity was approximately 32. However, combining GC+WL (non-linear HS) with *Planck* priors yields FoC ≈ 22, showing that correlations among parameters are drastically reduced. The fact that the curve of 1σ errors for the q_i (orange line, marked with circles) follows the same pattern as the curve for the p_i errors (green dashed line, marled with circles), is due to the fact that we have used a ZCA decomposition and therefore the q_i are as close as possible to the p_i. The complete matrix of coefficients W can be found in Table **??** in appendix 4.5.

4.5 Weight matrices coefficients

Here we explicitly show the coefficient of the W matrices illustrated in Figures 4.6, 4.9 and 4.12. Tables **??** and **??** contain the coefficient matrices obtained for Galaxy Clustering in the linear and non-linear HS prescription cases respectively, while Tables **??** and **??** contain the coefficients obtained for WL. Finally, Table **??** reports the coefficients for the combination of GC, WL and *Planck* priors, in the non-linear HS case.

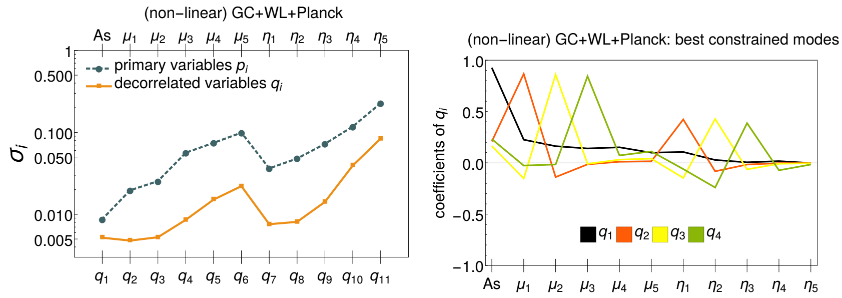


FIGURE 4.13: **Left:** the 1σ fully marginalized errors on the primary parameters p_i (green dashed lines), and the 1σ errors on the decorrelated derived parameters q_i (yellow solid lines). As opposed to the GC or WL cases (figs.4.7,4.10), here the decorrelated errors are much more similar to the standard errors. This is due to the fact that in the GC+WL+*Planck* combination, the cosmological parameters are not so strongly correlated. **Right:** best constrained modes for a Euclid Redbook GC+WL case using the non-linear HS prescription and adding a CMB *Planck* prior. Each panel shows the four best constrained parameters q_i. Each of them is a linear combination of the primary parameters p_i. The best constrained modes are sums $a\mu_i + b\eta_i$ for $i = \{1, 2, 3\}$ and positive values a and b.

4.6 Other Decorrelation Methods

In section 4.4.3 we have worked with a special decorrelation method, ZCA (Zero-phase component analysis, first introduced by **Bell19973327** in the context of image processing), which allows us to find a new vector of decorrelated variables q that is as similar as possible to the original vector of variables p. Other decorrelation methods do not share this property, so in this section we want to illustrate their difference with respect to ZCA. In the next subsections we show for the Principal Component Analysis and Cholesky decomposition methods, a subset of our previous results, namely the Galaxy Clustering non-linear case, using the HS prescription for a Euclid survey with Redbook specifications.

4.6.0.1 Principal Component Analysis

Principal Component Analysis (PCA) **Friedman-PCA** is a well known method, which rotates the vector of variables p into a new basis, using the eigenmatrix of the covariance matrix C. At the same time, it is the method that maximizes the compression of all components of p into the components of q using as measure the cross-covariance between q and p (see **kessy_optimal_2015** for more details and references). This method is useful for dimensional reduction or data compression, since the information is stored in as few components as possible. This is achieved by using the W matrix

$$W = \Lambda^{-1/2}U^T \tag{4.28}$$

where Λ and U represent the eigensystem of C (defined in Eqn. 4.18).

Euclid (Redbook)	$\ell\mathcal{A}_s$	μ_1	μ_2	μ_3	μ_4	μ_5	η_1	η_2	η_3	η_4	η_5	MG FoM
Fiducial	3.057	1.108	1.027	0.973	0.952	0.962	1.135	1.160	1.219	1.226	1.164	relative
GC (lin)	160%	119%	159%	183%	450%	1470%	509%	570%	586%	728%	3390%	0
GC (nl-HS)	0.8%	7.0%	6.7%	10.9%	27.4%	41.1%	20%	24.3%	19.9%	38.2%	930%	19
WL (lin)	640%	165%	2210%	4150%	13100%	22500%	2840%	3140%	8020%	29300%	39000%	-27
WL (nl-HS)	7.3%	188%	255%	419%	222%	206%	330%	488%	775%	8300%	9380%	-10
GC+WL (lin)	11.3%	5.8%	10%	19.2%	282%	469%	7.9%	9.6%	16.1%	276%	2520%	12
GC+WL+*Planck* (lin)	1.1%	3.4%	4.8%	7.8%	9.3%	13.1%	6.2%	7.7%	9.1%	12.7%	23.6%	27
GC+WL (nl-HS)	0.8%	2.2%	3.3%	8.2%	24.8%	34.1%	3.6%	5.1%	8.1%	25.4%	812%	24
GC+WL+*Planck* (nl-HS)	0.3%	1.8%	2.5%	5.8%	7.8%	10.3%	3.2%	4.1%	5.9%	9.6%	19.5%	33
GC+WL+*Planck* (nl-Halofit)	0.4%	2.0%	2.4%	5.1%	7.4%	10.2%	3.5%	4.1%	5.8%	9.2%	18.9%	33

TABLE 4.2: 1σ fully marginalized errors (as a percentage of the corresponding fiducial) on cosmological parameters for Euclid (Redbook) Galaxy Clustering and Weak Lensing surveys, alone and combining the two probes. We compare forecasts using linear spectra (lin) and forecasts using the nonlinear HS prescription (nl-HS). In Galaxy Clustering, the cutoff is set to $k_{\rm max} = 0.15$ h/Mpc in the linear case and $k_{\rm max} = 0.5$ h/Mpc in the non-linear case. For WL, the maximum cutoff in the linear case is at $\ell_{\rm max} = 1000$, while in the nl-HS case it is $\ell_{\rm max} = 5000$. At the bottom, we add on top a *Planck* prior (see section 4.3.3). For comparison, we also show in the last row the combined GC+WL+*Planck*, using just Halofit power spectra. The last column indicates the relative Figure of Merit (FoM$_{a,b}$) of the MG parameters in nits ('natural units', i.e. using the natural logarithm), with respect to our reference GC linear case, see (3.39) and surrounding text. A larger FoM indicates a more constraining probe. We notice a considerable improvement, in both GC and WL, when non-linearities are included. The combination GC+WL in the linear case constrains the MG parameters in the first two bins (z < 1.0) to less than 10% and including *Planck* priors allows to access higher redshifts with the same accuracy. A significant improvement in the constraints is obtained when adding the non-linear regime, in agreement with the observed reduction in correlation seen in Figs. 4.4 and 4.5. This is especially well exemplified by the error on $\ell\mathcal{A}_s \equiv \ln(10^{10}A_s)$, which reduces from 160% to 0.82%, from the linear to the non-linear forecast in the GC case and from 640% to 7.3% in the WL case. Finally, we note that since we are showing errors on μ and η, WL seems to be unfairly poor at constraining parameters. However, when converting this errors into errors on Σ, which is directly measured by WL, the constrains on $\Sigma_{1,2,3}$ are slightly better, of the order of 40% for WL(nl-HS) as can be guess from the degeneracy directions shown in Fig. 4.17. The FoM itself is nearly unaffected by the choice of $\{\mu, \eta\}$ vs $\{\mu, \Sigma\}$ as it is rotationally invariant.

Then, it follows that the transformed covariance matrix $\tilde{C}$ is whitened:

$$\tilde{C} = \Lambda^{-1/2}U^T C \Lambda^{-1/2}U^T \tag{4.29}$$

$$= \Lambda^{-1/2}U^T U\Lambda U^T \Lambda^{-1/2}U^T \tag{4.30}$$

$$= \mathbb{1} \tag{4.31}$$

As done previously, we renormalize W in such a way that the sum of the square of the elements of each row sum up to unity. In Figure 4.14 we show in the left panel the weight matrix W and in the right panel the 1σ errors for the original and decorrelated variables. We see that the most constrained q variables are the last components q_8-q_{11}, where most of the information has been compressed into. These

4 variables are complicated linear combinations of A_s, $\mu_{1,2,3}$ and $\eta_{1,2,3}$. This is the same we found for ZCA (cf. Fig. 4.8 and Eqns. (4.8)). However, the interpretation in terms of the old variables p is not so simple anymore. We can see from the weight matrix and the 1σ errors, that the most unconstrained parameter is q_1, which is basically equivalent to η_5. This means that this parameter can be eliminated from the analysis if one wants to do dimensional reduction.

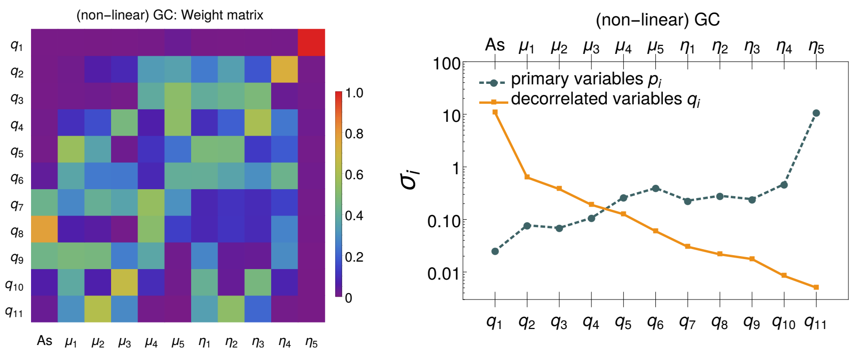


FIGURE 4.14: **Left**: weight matrix W for PCA. **Right**: 1σ fully maximized errors on the primary parameters p (blue lines) and the errors on the uncorrelated derived parameters q (orange lines). Notice how all the important information is constrained in as few variables as possible, namely the last elements of q_i.

4.6.0.2 Cholesky decomposition

Cholesky decomposition of the Fisher matrix $F = LL^T$, allows us to define a decorrelation method that compresses all components of p into an upper triangular matrix of components of q. This is achieved by using the W matrix:

$$W = L^T \tag{4.32}$$

Then the covariance matrix will be whitened:

$$\tilde{C} = L^T(LL^T)^{-1}L \tag{4.33}$$
$$= L^T(L^T)^{-1}L^{-1}L \tag{4.34}$$
$$= \mathbb{1} \tag{4.35}$$

As done previously, we renormalize W in such a way that the sum of the square of the elements of each row sum up to unity. In Cholesky decomposition, since we are constructing it via an upper triangular matrix (see left panel of Fig. 4.15), the new parameter q_{11} will be identical to the parameter η_5, which is, as we have seen before (section 4.4.4.1), the less constrained parameter. This decorrelation method is useful if one wants to have an ordering of the variables (see **kessy_optimal_2015** and references therein). On the other hand q_1 is almost identical to A_s, since as we have seen in section 4.4.4.1, it becomes decorrelated from the MG parameters in the non-linear case.

4.6.1 Kullback-Leibler divergence measure

In section 3.4.6 we introduced the Kullback-Leibler divergence (see eq. (3.42)) as another way of measuring the constraining power of a survey and we defined

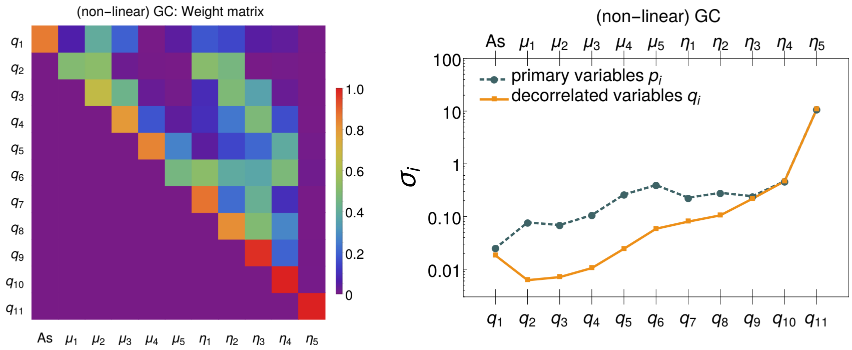


FIGURE 4.15: **Left:** weight matrix W for the Cholesky decorrelation. **Right:** 1σ fully maximized errors on the primary parameters p (blue lines) and the errors on the uncorrelated derived parameters q (orange lines). Notice how in this case, because of the upper triangular construction, the new variables $q_{9,10,11}$ are equivalent to $\eta_{3,4,5}$. The parameter A_s is decorrelated in the non-linear HS case for GC, therefore it is almost equivalent to q_1 as was also the case in ZCA (compare with Fig. 4.6)

the Kullback-Leibler matrices (eq. (3.43)) (first introduced in **casas-MGforec**) as a way of visualizing the information gain between different probes. Here we will visualize the KL-matrix for the probes: GC(lin), GC(nl-HS), WL(lin), WL(nl-HS), GC+WL(lin) and GC+WL(nl-HS) in the redshift binned parameterization of section 4.4. In this way we can quantitatively see how much information is gained when going from one probe like Galaxy Clustering to another probe like Weak Lensing. In Figure 4.16 we plot in the left panel the KL matrix $\mathcal{K}_{ij}$ when no *Planck* priors are added and in the right panel the KL matrix when *Planck* priors are added to all probes. For visualization purposes, we plot the logarithm of the KL matrix: $\log_{10}\mathcal{K}_{ij}$. Blue (red) colours represent small (large) information gain, while black represents no information gain at all, which is by construction the case on the diagonal, $\mathcal{K}_{ii} = 0$. The rows of the matrix represent the reference observable of p_1 and the columns, the new observable p_2. The information gain when going from a linear WL observable to a non-linear GC observable or to a combined GC+WL (linear and non-linear) observable is considerably high at about $\approx 10^6$-10^7. However, when adding a *Planck* prior, this gets reduced by at least 2 orders of magnitude, since the prior is strong and then there is not as much new information gained in the new observables. On the opposite case, we can see that the row corresponding to GC+WL(nl-HS) is mostly blue, meaning that there is little gain of information when going to one of the other observables.

4.7 Modified gravity with simple smooth functions of the scale factor

As discussed in Sec. 4.2, μ and η (or an equivalent pair of functions of the gravitational potentials) depend in general on time and space. We will now investigate the time dependence further, starting from the two parameterizations proposed in **planck_collaboration_planck_2016** and recalled in Eqns. (4.9-4.12) in this paper. We extend the analysis of the *Planck* paper by testing different prescriptions for

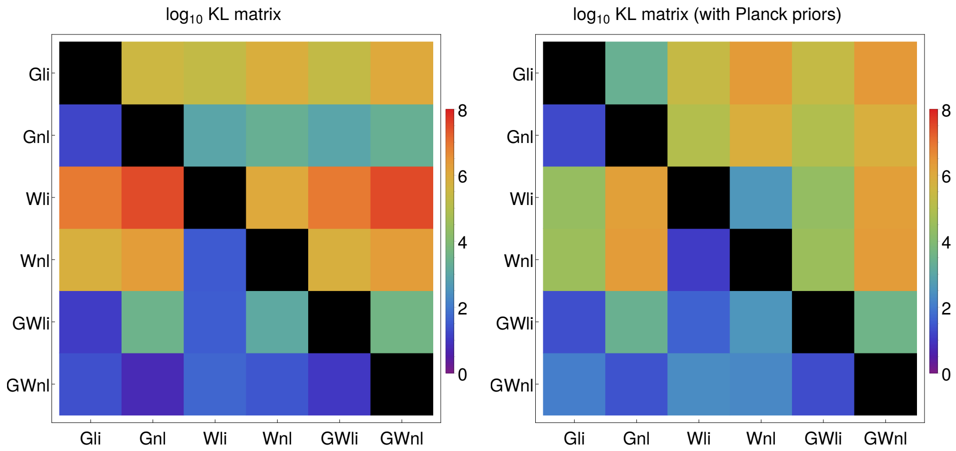


FIGURE 4.16: Kullback-Leibler divergence matrices $\mathcal{K}_{ij}$. This matrix represents graphically the information gain between all possible observables in the redshift binned parameterization of section 4.4. We have plotted here the logarithm of the KL-divergence matrix, for illustrative purposes. Therefore the diagonal is $-\infty$ and it is represented by a black color. **Left:** KL matrix without *Planck* priors. The maximum gain is about 10^7 when going from WL(lin) to GC+WL(non-linear) and we can observe that GC+WL does not gain extra information when complemented with the other observables, which is expected. **Right:** In this case we compare the observables, when a *Planck* prior is added beforehand. The overall information gain is now smaller, with a maximum of about 10^6. The maximum gain comes when comparing WL (linear and non-linear) to GC and GC+WL (non-linear).

the non-linear regime in Modified Gravity (as illustrated in Section 4.3) and further investigate forecasts for future experiments like Euclid, SKA, DESI. In the following subsections we first give results for the late-time parameterization of Eqns. (4.9,4.10) and then for the early time parameterization of Eqns. (4.11,4.12). In both cases we consider Galaxy Clustering and Weak Lensing, neglecting, as in Section 4.4 any information coming from the cross correlation of the two probes.

4.7.1 Modified Gravity in the late-time parameterization

The late-time parameterization is defined in Eqns. (4.9,4.10). We now calculate forecasts for Galaxy Clustering and Weak Lensing, with future surveys, in the linear and mildly non-linear regimes. We also include prior information obtained from the analysis of the *Planck*+BSH datasets (where we recall that BSH stands for BAO + SNe + H_0 prior), as discussed in Section 4.3.3.

4.7.1.1 Galaxy Clustering in the linear and mildly non-linear regime

In Table 4.3 we show forecasts for the Euclid survey **laureijs_euclid_2011** for Galaxy Clustering (top part of the table) and three different cases: using only linear scales with a cutoff at $k_{\rm max} = 0.15{\rm h/Mpc}$, labeled GC(lin); extending forecasts in the mildly non-linear regime, obtained using the prescription described in Sec. 4.3.2.2, with a cutoff at $k_{\rm max} = 0.5$ h/Mpc, labeled GC(nl-HS); combining the mildly non-linear case with *Planck* priors, as described in Sec. 4.3.3. We take into account the BAO features, redshift space distortions and the full shape of the power spectrum, as well as the survey specifications of the Euclid Redbook, recalled in Section 3.4.1 for convenience. The columns correspond to the marginalized errors on five standard cosmological parameters $\{\Omega_c, \Omega_b, n_s, \ln(10^{10}A_s), h\}$ and three combinations of the gravitational potentials $\{\mu, \eta, \Sigma\}$: the latter are reconstructed in time, according to the late-time parameterization, as defined in Eqns. (4.9,4.10). We recall that only two of these three functions are independent and fully determine cosmological linear perturbations.

In the late-time scenario, for a Galaxy Clustering survey, neither η nor Σ are actually constrained by a linear forecast, while μ is mildly constrained. Adding the non-linear regime improves constraints on μ, while the other parameters remain unconstrained, unless we also include *Planck* priors, which yields an improvement in the FoM of 6.3 nits. In general the observable power spectrum may depend on both μ (explicitely appearing in the last term of the equation for the density perturbation (cf. Eqn.(4.52)) and on η (implicitly contained in the derivatives of the gravitational potential in the same equation). The contribution of the derivative of the potentials is larger in the early-time parameterization, with respect to the late-time one by construction. This is due to the fact that in the late-time case deviations from GR go to zero at large redshifts. In this sense, in the specific case of the late time parameterization, the observed power spectrum mainly depends on μ only, which explains why this is the only quantity (mildly) constrained by GC alone. In the early-time parameterization, though, modifications can appear also at earlier times, so that both η and μ effectively affect the power spectrum, which explains why they can both be constrained with a smaller uncertainty, as we will discuss in Section 4.7.2.

In Appendix 4.7.1.5 we review the equation governing the evolution of cold dark matter density fluctuations, as a function of the Modified Gravity functions $\mu(a)$ and $\eta(a)$ as they are implemented in the code MGCAMB **hojjati_testing_2011** The inability of GC to constrain η in this parametrization, is also visible in Fig. 4.17 which shows that the GC constraints are degenerate along the η or Σ directions. Therefore, we show how with this parameterization choice Euclid GC will be extremely sensitive to modifications of the Poisson equation for Ψ, while it would require additional information to constrain departures from the standard Weyl potential.

4.7.1.2 Late time parameterization

Our main observables are parameterized in terms of the primary variables E_{11} and E_{22} from Equation 4.11; we are however interested in forecasting the constraints on the pair of secondary variables $\{\mu, \eta\}$ or on the pair $\{\mu, \Sigma\}$.

Therefore we need to transform the variables using a Jacobian $J_{ij} = \frac{\partial\theta_i}{\partial\tilde{\theta}_j}$, where θ_i is the set of primary variables and $\tilde{\theta}_i$ is the vector of secondary variables. Eqns. (4.5,4.9,4.10) allow us to express the $\tilde{\theta}_i$ as a function of the variables θ_i and to obtain the non-vanishing derivatives of μ and η w.r.t to all cosmological parameters:

$$\frac{\partial\mu}{\partial\Omega_c} = -E_{11}, \qquad \frac{\partial\eta}{\partial\Omega_c} = -E_{22} \tag{4.36}$$

$$\frac{\partial\mu}{\partial\Omega_b} = -E_{11}, \qquad \frac{\partial\eta}{\partial\Omega_b} = -E_{22} \tag{4.37}$$

$$\frac{\partial\mu}{\partial E_{11}} = 1 - \Omega_b - \Omega_c - \Omega_\nu, \qquad \frac{\partial\eta}{\partial E_{22}} = 1 - \Omega_b - \Omega_c - \Omega_\nu \ . \tag{4.38}$$

Euclid (Redbook)	Ω_c	Ω_b	n_s	$\ell\mathcal{A}_s$	h	μ	η	Σ	MG FoM
Fiducial	0.254	0.048	0.969	3.060	0.682	1.042	1.719	1.416	relative
GC(lin)	1.9%	6.4%	3%	2.8%	4.5%	17.1%	1030%	641%	0
GC(nl-HS)	0.9%	2.5%	1.3%	0.8%	1.7%	1.7%	475%	291%	2.9
GC(nl-HS)+*Planck*	0.7%	0.6%	0.3%	0.2%	0.3%	1.7%	16.8%	10.3%	6.3
WL(lin)	7.8%	25.7%	9.9%	10.3%	19.1%	58.2%	106%	9.3%	3.2
WL(nl-HS)	6.3%	20.7%	4.6%	5.8%	13.8%	23.3%	40.9%	4.6%	4.5
WL(nl-HS)+*Planck*	2.1%	1.1%	0.4%	0.7%	0.7%	11.8%	21.8%	2.8%	5.7
GC+WL(lin)	1.8%	5.9%	2.8%	2.3%	4.2%	7.1%	10.6%	2%	6.6
GC+WL(lin)+*Planck*	1.0%	0.7%	0.4%	0.4%	0.4%	6.2%	9.8%	1.5%	7.0
GC+WL(nl-HS)	0.8%	2.2%	0.8%	0.7%	1.5%	1.6%	2.4%	1.0%	8.8
GC+WL(nl-HS)+*Planck*	0.7%	0.6%	0.2%	0.2%	0.3%	1.6%	2.4%	0.9%	8.9
GC+WL(nl-Halofit)+*Planck*	0.6%	0.5%	0.2%	0.2%	0.2%	0.8%	1.7%	0.8%	9.6

TABLE 4.3: 1σ fully marginalized errors on the cosmological parameters in the late-time parameterization of Modified Gravity for a Euclid Galaxy Clustering forecast (top), a Weak Lensing forecast (middle) and the combination of both probes (bottom): Modified Gravity is encoded in two of the three functions μ, η, Σ, which are reconstructed in the late-time parameterization defined in Eqns. (4.9,4.10). For each case, we also list the forecasted errors using a *Planck*+BSH prior. Linear forecasts are labeled by "lin", and correspond to a cutoff $k_{\rm max} = 0.15$h/Mpc for GC and $\ell_{\rm max} = 5000$ for WL; non-linear forecasts use the prescription described in sec. 4.3.2.2, are labeled by "nl-HS" and correspond to a cutoff of $k_{\rm max} = 0.5$h/Mpc for GC and $\ell_{\rm max} = 5000$ for WL. in both cases, power spectra have been computed using the MGCAMB Boltzmann code. For completeness, we also show the GC+WL+*Planck* case using the non-linear power spectra computed using Halofit only (nl-Halofit). In the last column, we show for each observation the Figure of Merit (FoM) relative to our base observable in this parametrization, which is GC(linear). GC(linear) has an absolute FoM of -0.94 in 'nits'. We can see that in both GC and WL there is a considerable gain when including non-linear scales and *Planck* priors. The difference in the FoM between the non-linear HS prescription and the standard Halofit approach is quite small.

Euclid (Redbook)	Ω_c	Ω_b	n_s	$\ell\mathcal{A}_s$	h	μ	η	Σ	MG FoM
Fiducial	0.256	0.048	0.969	3.091	0.682	0.902	1.939	1.326	relative
GC(lin)	3.7%	19.2%	9.1%	16.6%	16.5%	16.7%	758%	489%	0
GC(nl-HS)	1.1%	2.3%	1.3%	0.7%	1.6%	1.8%	7.9%	4.8%	6.6
GC(nl-HS)+*Planck*	0.8%	0.7%	0.3%	0.3%	0.3%	1.7%	7.6%	4.6%	6.7
WL(lin)	12.1%	28.9%	11.3%	13.3%	24%	6.8%	11.1%	11.9%	4.9
WL(nl-HS)	6.5%	21.9%	6.6%	5.9%	15.8%	2.8%	8.0%	3.4%	6.6
WL(nl-HS)+*Planck*	2.1%	1.3%	0.5%	0.9%	0.7%	2.2%	7.2%	2.9%	7.2
GC+WL(lin)	1.8%	6.6%	3.4%	5.6%	5.2%	3.0%	6.8%	3.4%	6.4
GC+WL(lin)+*Planck*	1.2%	0.9%	0.6%	2.3%	0.4%	2.4%	6.5%	2.8%	6.9
GC+WL(nl-HS)	1.0%	2.2%	1.2%	0.7%	1.6%	1.3%	4.4%	1.9%	8.0
GC+WL(nl-HS)+*Planck*	0.8%	0.7%	0.3%	0.3%	0.3%	1.3%	4.4%	1.9%	8.2
GC+WL(nl-Halofit)+*Planck*	0.7%	0.7%	0.3%	1.3%	0.3%	0.9%	2.3%	1%	8.9

TABLE 4.4: Same as Table (4.3) but for the early-time parameterization. Note that the last column (MG FoM) cannot be compared to the one for a different parameterization (Table 4.3) since the reference value is different (GC(lin)) and the two parameterizations have a different number of primary parameters). In this case the absolute MG FoM of GC(lin) is ≈ -0.47 nits.

	Ω_c	Ω_b	n_s	$\ell\mathcal{A}_s$	h	μ	η	Σ	MG FoM
Fiducial	0.254	0.048	0.969	3.060	0.682	1.042	1.719	1.416	relative
GC(nl-HS)									
Euclid	0.9%	2.5%	1.3%	0.8%	1.7%	1.7%	475%	291%	2.9
SKA1-SUR	5%	15.3%	8.7%	3.8%	10.8%	18.1%	165%	108%	1.7
SKA2	0.5%	1.3%	0.4%	0.4%	0.8%	0.7%	86.8%	53.2%	5.5
DESI-ELG	1.6%	4.1%	2.3%	1.3%	2.9%	3.3%	899%	552%	1.8
WL(nl-HS)									
Euclid	6.3%	20.7%	4.6%	5.8%	13.8%	23.3%	40.9%	4.6%	4.5
SKA1	30.8%	109%	35%	36.5%	77.6%	220%	405%	36.8%	0.5
SKA2	6%	22.5%	5.9%	6.8%	15.9%	19%	33.2%	3.7%	4.9
GC+WL(lin)									
Euclid	1.8%	5.9%	2.8%	2.3%	4.2%	7.1%	10.6%	2%	6.6
SKA1	10.1%	47.6%	25.4%	21.7%	40.4%	26.4%	28.8%	13.6%	3.7
SKA2	1.2%	4.5%	2.2%	1.9%	3.3%	4.1%	5.5%	1.6%	7.5
GC+WL(lin)+*Planck*									
Euclid	1.0%	0.7%	0.4%	0.4%	0.4%	6.2%	9.8%	1.5%	6.9
SKA1	2.4%	1.2%	0.4%	1.2%	0.7%	12%	19.8%	3.8%	5.3
SKA2	0.7%	0.6%	0.3%	0.4%	0.3%	3.6%	5.2%	1.2%	7.8
GC+WL(nl-HS)									
Euclid	0.8%	2.2%	0.8%	0.7%	1.5%	1.6%	2.4%	1.0%	8.7
SKA1	4.7%	14.3%	6.2%	3.6%	9.6%	12.8%	11%	7.3%	5.5
SKA2	0.4%	1.3%	0.3%	0.4%	0.8%	0.7%	0.9%	0.6%	10.3
GC+WL(nl-HS)+*Planck*									
Euclid	0.7%	0.6%	0.2%	0.2%	0.3%	1.6%	2.4%	0.9%	8.9
SKA1	2.0%	1.0%	0.4%	0.8%	0.6%	3.5%	6%	2.7%	6.9
SKA2	0.4%	0.5%	0.2%	0.1%	0.2%	0.6%	0.9%	0.5%	10.3

TABLE 4.5: 1σ fully marginalized errors on the cosmological parameters $\{\Omega_m, \Omega_b, h, \ell\mathcal{A}_s, n_s, \mu, \eta, \Sigma\}$ in the late-time parameterization comparing different surveys in the linear and non-linear case. In the last column, we show for each observation the Modified Gravity Figure of Merit (MG FoM) relative to our base observable, which is the Euclid Redbook GC(linear), see Table 4.3. We can see that in general terms, SKA2 is the most powerful survey, followed by Euclid and SKA1. In the case of GC alone, DESI-ELG is more constraining than SKA1-SUR. Notice that in this parameterization, a GC survey would only constrain μ with a high accuracy, while a WL survey would constrain Σ with a very good accuracy. The combination of both is much more powerful than the single probes. Adding *Planck* priors (last row) improves considerably the constraints on the ΛCDM parameters but has an almost negligible effect on the MG parameters (MG FoM remain almost constant when adding *Planck* priors to the GC+WL (non-linear HS) case. The marginalized contours for the μ-η plane , comparing these surveys, can be seen in the left panel ofFig.4.21.

With these derivatives we can construct the inverse of the Jacobian $J^{-1}_{ij} = \frac{\partial\tilde{\theta}_j}{\partial\theta_i}$. The Fisher matrix in the secondary variables $\tilde{F}_{ij}$ is then given by

$$\tilde{F} = J^T F\, J \tag{4.39}$$

For the parameter set containing $\tilde{\theta}_i = \{\mu, \Sigma\}$ we obtain the following non-vanishing derivatives

	Ω_c	Ω_b	n_s	$\ell\mathcal{A}_s$	h	μ	η	Σ	MG FoM
Fiducial	0.256	0.0485	0.969	3.091	0.682	0.902	1.939	1.326	relative
GC(nl-HS)									
Euclid	1.1%	2.3%	1.3%	0.7%	1.6%	1.8%	7.9%	4.8%	6.6
SKA1-SUR	7.9%	14.2%	13.4%	4.2%	11%	12.6%	82.7%	52.6%	2.2
SKA2	0.6%	1.3%	0.7%	0.4%	0.9%	0.9%	3.4%	1.8%	8.3
DESI-ELG	2.0%	4.3%	2.7%	1.4%	3.0%	8.2%	32%	28.6%	4.3
WL(nl-HS)									
Euclid	6.5%	21.9%	6.6%	5.9%	15.8%	2.8%	8.0%	3.4%	6.6
SKA1	32%	106%	37.2%	33%	79.3%	13.1%	37.1%	16.4%	3.4
SKA2	5.9%	22.1%	6.7%	6.1%	16.1%	2.4%	7.0%	2.9%	6.9
GC+WL(lin)									
Euclid	1.8%	6.6%	3.4%	5.6%	5.2%	3.0%	6.8%	3.4%	6.4
SKA1	10.3%	46.4%	24.2%	33.6%	40.2%	14.4%	29.6%	15.5%	3.3
SKA2	1.3%	4.9%	2.5%	4.2%	3.9%	2.5%	5.7%	2.7%	6.8
GC+WL(lin)+*Planck*									
Euclid	1.2%	0.9%	0.6%	2.3%	0.4%	2.4%	6.5%	2.8%	6.8
SKA1	2.5%	1.5%	0.8%	2.9%	0.8%	8.8%	22.2%	8.5%	4.5
SKA2	0.9%	0.7%	0.6%	2.1%	0.3%	2.1%	5.4%	2.3%	7.2
GC+WL(nl-HS)									
Euclid	1.0%	2.2%	1.2%	0.7%	1.6%	1.3%	4.4%	1.9%	8.1
SKA1	7.1%	13.4%	10.7%	4%	10%	8.2%	24.4%	10.5%	4.4
SKA2	0.6%	1.3%	0.7%	0.4%	0.9%	0.8%	2.7%	1.3%	8.8
GC+WL(nl-HS)+*Planck*									
Euclid	0.8%	0.7%	0.3%	0.3%	0.3%	1.3%	4.4%	1.9%	8.1
SKA1	2.1%	1.3%	0.5%	0.9%	0.7%	7%	20.8%	8.2%	4.9
SKA2	0.5%	0.5%	0.3%	0.2%	0.2%	0.8%	2.7%	1.3%	8.8

TABLE 4.6: Same as Table 4.5 but for the early-time parameterization. The last 4 columns correspond to the projection of the errors on E_{11} and E_{22} onto μ, η and Σ, respectively. We have marginalized over E_{12} and E_{21} since at $z = 0$ they don't contribute to the Modified Gravity parameters. Notice that in this parameterization, a GC survey alone is able to constrain both μ and Σ to a good level for all surveys, better than with the late time parameterization, more often used in literature. The combination of GC+WL is however less constraining in the early time parametrization than in late time parameterization one. The reference case for the MG FoM is the Euclid (Redbook) GC linear forecast (Table 4.4). The non-linear forecast for GC+WL+*Planck* would yield, for Euclid and SKA2, contraints at the 1-2% accuracy on μ, Σ, while for SKA1 the contraints would be at the 8% level. The marginalized contours for the μ-η plane , comparing these surveys, can be seen in the right panel ofFig.4.21.

$$\frac{\partial \mu}{\partial \Omega_c} = -E_{11}, \qquad \frac{\partial \Sigma}{\partial \Omega_c} = -\frac{1}{2}E_{22}(1 + E_{11}(1 - \Omega_b - \Omega_c - \Omega_\nu)) \tag{4.40}$$

$$- \frac{1}{2}E_{11}(2 + E_{22}(1 - \Omega_b - \Omega_c - \Omega_\nu)) \tag{4.41}$$

$$\frac{\partial \mu}{\partial \Omega_b} = -E_{11}, \qquad \frac{\partial \Sigma}{\partial \Omega_b} = -\frac{1}{2}E_{22}(1 + E_{11}(1 - \Omega_b - \Omega_c - \Omega_\nu)) \tag{4.42}$$

$$- \frac{1}{2}E_{11}(2 + E_{22}(1 - \Omega_b - \Omega_c - \Omega_\nu)) \tag{4.43}$$

$$\frac{\partial \mu}{\partial E_{11}} = 1 - \Omega_b - \Omega_c - \Omega_\nu, \qquad \frac{\partial \Sigma}{\partial E_{11}} = \frac{1}{2}(2 + E_{22}(1 - \Omega_b - \Omega_c - \Omega_\nu))(1 - \Omega_b - \Omega_c - \Omega_\nu) \tag{4.44}$$

$$\frac{\partial \Sigma}{\partial E_{22}} = \frac{1}{2}(1 + E_{11}(1 - \Omega_b - \Omega_c - \Omega_\nu))(1 - \Omega_b - \Omega_c - \Omega_\nu) \tag{4.45}$$

4.7.1.3 Weak Lensing in the linear and mildly non-linear regime

Using the Euclid Weak Lensing specifications described in Section sub:Fisher-Weak-Lensing, we obtain the results displayed in the middle panel of Table 4.3. Also in this case, we use the late-time parameterization, in three different cases: the first uses only linear quantities, with a maximum multipole of $\ell_{\max} = 1000$; the second case uses the non-linear HS prescription of section 4.3.2.2 up to a maximum multipole of $\ell_{\max} = 5000$; the third case combines Weak Lensing with *Planck* priors, as described in 4.3.3.

In the linear case, WL forecast yields constraints on the standard ΛCDM parameters at around 10% of accuracy, with the exception of Ω_b which is poorly constrained at around 26%, and the expansion rate h (19%). This is likely due to the fact that WL is only directly sensitive to the total matter distribution in the Universe and cannot differentiate baryons from dark matter. All constraints improve when adding non-linear information, with Ω_b and h still constrained only at the level of about 20%. When the *Planck* priors are included, though, constraints shrink down to about 1% for all cosmological parameters. The Modified Gravity parameters show the expected trend; in the linear case, only Σ is constrained, at 11%, as this parameter is directly defined in terms of the lensing potential $\Phi + \Psi$; a Weak Lensing probe is however not directly sensitive to μ and η separately, as can be seen in Fig. 4.17. The linear FoM is only slightly weaker than the one from GC for this parameterization, probably because both probes have an effectively unconstrained degeneracy direction. When adding non-linear information, errors on μ and η improve, though still remaining in a poorly constrained interval (25%-44%). Already on its own, however, Weak Lensing could rule out many models of Modified Gravity that change Σ at more than 5% (or even 2.9%, if we include *Planck* priors). The combination GC+*Planck* and WL+*Planck* has a comparable overall constraining power, with GC+*Planck* being about 1 nit stronger and providing smaller errors on μ.

4.7.1.4 Combining Weak Lensing and Galaxy Clustering

After using the two primary probes from Euclid separately, we discuss here the constraints obtained combining Weak Lensing and Galaxy Clustering. The combination between GC and WL can be seen in the bottom panel of Table 4.3. In the late-time parameterization, in the linear case, Weak Lensing combined with a galaxy clustering for a Euclid survey (Redbook specifications) constrains the standard ΛCDM parameters in the range $2\% - 6\%$, and below 1% when *Planck* priors are included. Modified Gravity parameters μ and η are now also constrained below 10%, reaching 1% when adding non-linear scales. The remarkable improvement can be attributed to the fact that the combination of GC and WL Fisher matrices breaks many degeneracies in the parameter space. This is shown in Figure 4.17, where it is possible to notice how the two probes are almost orthogonal both in the μ-η- and μ-Σ-plane. Weak Lensing measures the changes in the Weyl potential, parametrized by Σ, while μ is related to the Poisson equation, and therefore to the potential Ψ, modified by peculiar velocities and sensitive to Galaxy Clustering; η can also be written as a combination of μ and Σ (see Eqn. 4.5).

Further improvement is brought by the sensitivity of Galaxy Clustering to standard ΛCDM parameters; even though GC constraints on Σ and η are not as good as the ones for Weak Lensing, the better measurement of standard parameters provided by Galaxy Clustering breaks degeneracies in the Modified Gravity sector of

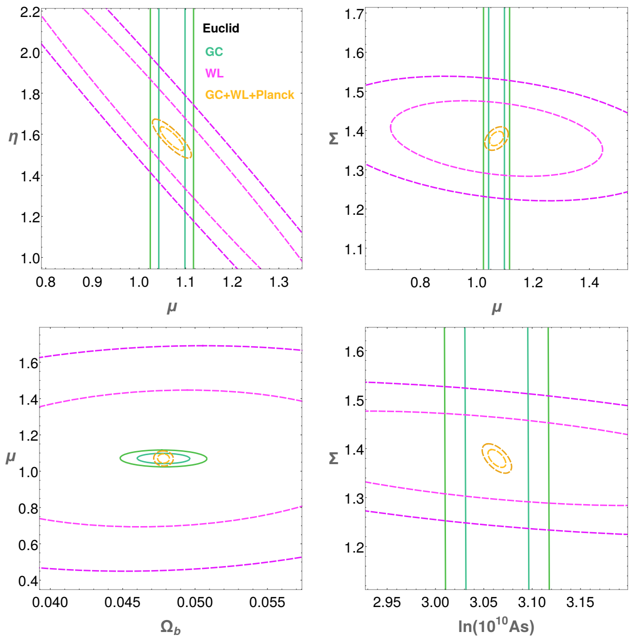


FIGURE 4.17: Fisher Matrix marginalized contours (1, 2 σ) for the Euclid space mission in the late-time parameterization using mildly non-linear scales and the HS prescription. Green lines represent constraints from a Galaxy Clustering survey, pink lines stand for the Weak Lensing observables, and orange lines represent the GC+WL+*Planck* combined confidence regions. **Upper Left:** contours for the fully marginalized errors on η and μ. **Upper Right:** contours for the fully marginalized errors on Σ and μ. **Lower Left:** contours for the fully marginalized errors on μ and Ω_b. **Lower Right:** contours for the fully marginalized errors on Σ and $\ln(10^10 A_s)$. The fact that the combination of GC, WL and *Planck* breaks many degeneracies in the 7-dimensional parameter space, explains why the combined contours (yellow) have a much smaller area. Notice that in this parametrization, GC measures mostly μ and WL constrains mostly just Σ.

the parameter space, leading to narrower bounds for η and Σ with respect to both probes taken separately. WL is instead not sensitive to modifications of the Poisson equation for matter and this explains why constraints on μ are not improved by the combination of the two probes, but are rather dominated by GC. The correlation among parameters can also help us explain the observed results. The Figure of Correlation, defined in Eq. (3.40), for GC (non-linear HS) alone is 4.9, while for WL (non-linear HS) the correlation is higher, with FoC $= 16.9$. When combining both probes (GC+WL (non-linear HS)) the FoC goes to an intermediate point of 7.6.

Given the constraining power of the GC+WL combination on MG functions, adding the *Planck* priors does not lead to significant improvements on the dark energy related parameters. On the other hand, standard parameters significantly benefit from the inclusion of CMB and background priors and we can expect this to be a relevant factor for MG models with degeneracies with ΛCDM parameters, e.g. models affecting also the expansion history of the universe. An overview of the constraints on Modified Gravity described in this section is shown in Fig.4.17, with Euclid GC, Euclid WL and Euclid GC+WL combined with *Planck* priors.

4.7.1.5 Derivatives of the Power Spectrum with respect to μ and η

In this section we investigate the derivatives of the power spectrum with respect to the MG parameters. Using the Jacobians from the previous Appendix **??**, we can convert our fundamental derivatives $\partial P(k,z)/\partial E_{ij}$ to derivatives of $\partial P(k,z)/\partial(\mu,\eta)$ evaluated at $z=0$.

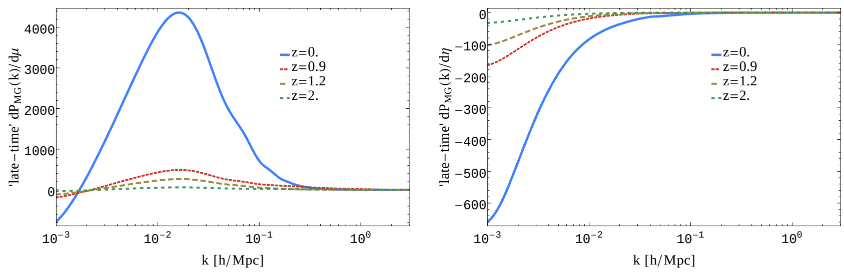


FIGURE 4.18: Derivatives of the matter power spectrum $P(k,z)$ w.r.t. the MG parameters μ and η in the late-time parametrization.

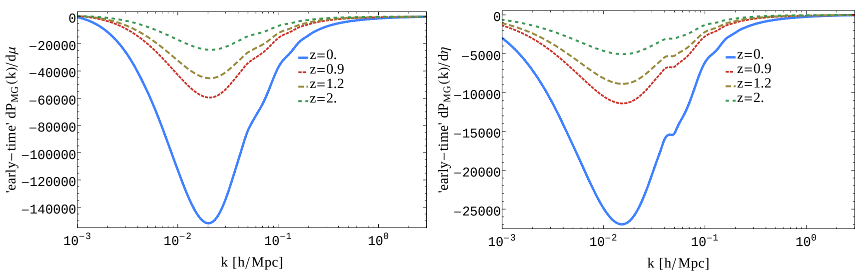


FIGURE 4.19: Derivatives of the matter power spectrum $P(k,z)$ w.r.t. the MG parameters μ and η in the early-time parametrization.

We can see from the previous figures that in the early-time parametrization the derivative of the power spectrum with respect to η has a similar shape as the derivative with respect to μ, making η detectable by a Galaxy Clustering survey. To explain this, we derive the equation governing the evolution of density fluctuations for a cold dark matter (CDM) species, based on the equations implemented on the code MGCAMB presented in **hojjati_testing_2011** expressed here in the conformal Newtonian gauge. In the following, a dot represents derivative with respect to conformal time τ:

$$\dot{\delta} = -(1+w)(\theta - 3\dot{\Phi}) - 3\mathcal{H}(\frac{\delta P}{\delta \rho} - w)\delta \tag{4.46}$$

$$\dot{\theta} = -\mathcal{H}(1-3w)\theta - \frac{\dot{w}}{1+w}\theta + \frac{\delta P/\delta\rho}{1+w}k^2\delta - k^2\sigma + k^2\Psi \tag{4.47}$$

For CDM we have $\sigma = w = c_s^2 = \delta P/\delta\rho = 0$, then:

$$\dot{\delta} = -(\theta - 3\dot{\Phi}) \tag{4.48}$$

$$\dot{\theta} = -\mathcal{H}\theta + k^2\Psi \tag{4.49}$$

We have parameterized the solution to Ψ as:

$$k^2\Psi = -4\pi G a^2 \mu(\tau)\rho(\tau)\delta(\tau) \quad , \tag{4.50}$$

and since we are also requiring gravitational slip $\eta = \Phi/\Psi$, we then have the Poisson equation for Φ:

$$k^2\Phi = -4\pi G a^2 \mu(\tau)\eta(\tau)\rho(\tau)\delta(\tau) \tag{4.51}$$

Taking the time derivative of (4.48) and using (4.49) and (4.48) to replace $\dot{\theta}$ and θ, and substituting Ψ from (4.50), we obtain:

$$\ddot{\delta} + \mathcal{H}\dot{\delta} = 3\mathcal{H}\dot{\Phi} + 3\ddot{\Phi} + 4\pi G a^2 \mu\rho\delta \tag{4.52}$$

In general, derivatives of Φ appearing on the right hand side will depend on both μ and η. Their contribution is larger in the early-time parameterization with respect to the late-time one.

4.7.1.6 Forecasts in Modified Gravity for SKA1, SKA2 and DESI

For the SKA1 and SKA2 surveys (whose specifications are explained in detail in Section 3.4.4), previous work on forecasting cosmological parameters has been done, among others, by **baker_observational_2015** and **bull_extending_2015** In the latter work, the author parameterizes the evolution of $\mu(a)$ using the late-time parameterization, but also adds an extra parameter allowing for a scale dependence in $\mu(a)$ and including a *Planck* prior. For a fixed scale, the 1σ errors on the amplitude of μ lie between 0.045 and 0.095, depending on the details of the SKA1 specifications, while for SKA2, this error is of about 0.017. This setting would correspond to our GC+WL(linear) + *Planck* case (see Table 4.5) where we find for SKA1 a 1σ error on μ of 0.12 and for SKA2 the forecasted error is 0.036. Our errors are somewhat larger, but we also have extended our analysis to let the gravitational slip η be different from 1 at present time, our departure from $\mu = 1$ at present time is larger by a factor 4 and our linear forecast is conservative in the sense that it includes less wavenumbers k at higher redshifts, compared to theirs.

In Figure 4.20 we show the 1σ fully marginalized forecasted errors on the parameters $\{\Omega_m, \Omega_b, h, \ell\mathcal{A}_s, n_s, \mu, \Sigma\}$ for different Weak Lensing (left panel) and Galaxy Clustering (right panel) surveys in the late-time parameterization. In the GC case, the surveys considered are DESI-ELG (yellow), SKA2 (green), SKA1-SUR (orange) and Euclid (blue). For the WL forecast, we considered Euclid (blue), SKA1 (orange) and SKA2 (green). These constraints correspond to the ones listed in Table 4.5. The marginalized confidence contours for the μ-η plane , comparing all these surveys, can be seen in the left panel ofFig.4.21. The 1σ fully marginalized constraints on the parameters are weaker for WL than for GC, which may be a consequence of the higher correlation among variables for the Weak Lensing observable, described in the previous section 4.7.1.4. Comparing the different surveys, the general trend is that Euclid and SKA2 perform at a similar level for WL at both linear and non-linear level; for GC and when combining both probes, SKA2 gives the strongest constraints, followed by Euclid, SKA1 and DESI (for GC). Notice that in this parameterization, a SKA1-SUR GC survey constrains the Σ parameter alone better than a Euclid Galaxy Clustering survey (although Euclid is overall much stronger as can be seen with the FoM). This is due to the fact that SKA1-SUR probes much lower redshifts (from $z = 0.05 - 0.85$) than Euclid and is therefore suitable to better constrain those parameterizations in which the effect of the Modified Gravity parameters is stronger at lower redshifts; this is the case of the late-time parameterization, which is proportional to the dark energy density, dominating at low redshifts only. This result is reversed in the early time parameterization, in which Modified Gravity can play a role also at earlier redshifts.

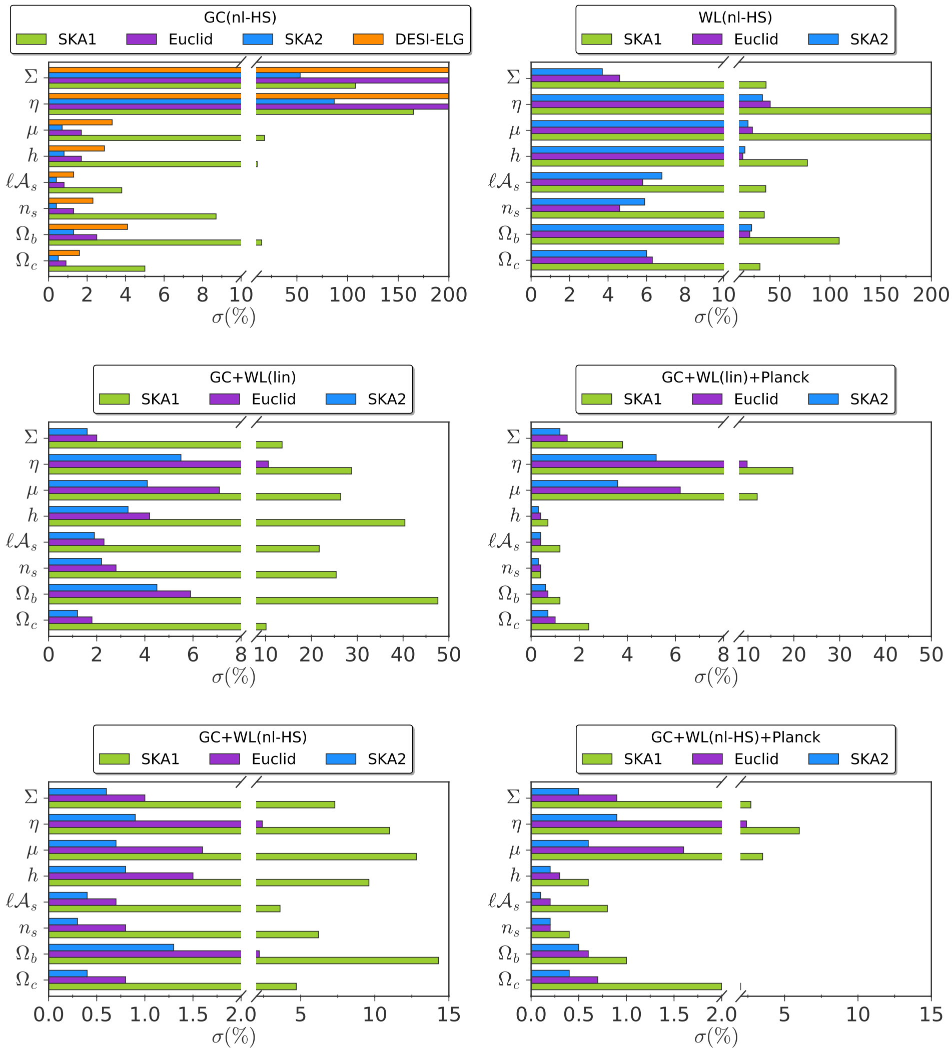


FIGURE 4.20: 1σ fully marginalized errors on the parameters $\{\Omega_m, \Omega_b, h, \ell\mathcal{A}_s, n_s, \mu, \eta, \Sigma\}$ for the late-time parameterization of MG obtained by forecasts on Galaxy Clustering (non-linear HS) (top left panel), Weak Lensing (non-linear HS) (top right panel), the combinations GC+WL (linear) (middle left) and GC+WL+*Planck* (linear) (middle right) and the combinations GC+WL (non-linear HS) (bottom left) and GC+WL+*Planck* (non-linear) (bottom right). In the GC case, the surveys considered are SKA2 (blue), SKA1-SUR (green), Euclid Redbook (purple) and DESI-ELG (orange). For forecasts including WL, only Euclid, SKA1 and SKA2 are included. Although the 1σ constraints on the standard parameters are overall weaker for WL than for GC, Weak Lensing surveys perform better on Modified Gravity parameters. Comparing the different surveys, Euclid and SKA2 perform similarly well for the WL observable alone, if non-linearities are included. Notice that SKA1-SUR performs better than Euclid on the η and Σ parameters, because it can measure better at lower redshifts. Including the *Planck* prior, the GC+WL combination for Euclid and SKA2 constrains all parameters at much better than percent accuracy. Detailed specifications of the different surveys are explained in the text.

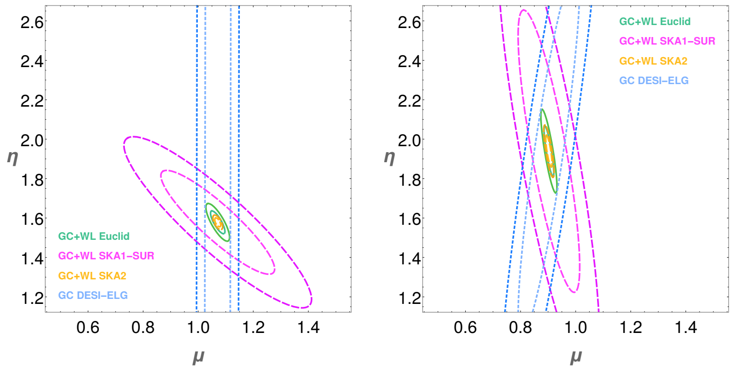


FIGURE 4.21: 1σ and 2σ fully marginalized confidence contours on the parameters μ and η, for 3 different surveys combining Galaxy Clustering (GC) and Weak Lensing (WL): Euclid, SKA1-SUR and SKA2 and for GC only: DESI-ELG, all in the late-time (left panel) and early-time (right panel) parameterizations of sections 4.7.1 and 4.7.2, respectively. As explained in the main text, the constraints are parameterization-dependent, especially on η, where in the late-time scenario GC alone is not able to constrain it, while in the early-time scenario GC can constrain both μ and η.

4.7.2 Modified Gravity in the early-time parameterization

4.7.2.1 Galaxy Clustering, Weak Lensing and its combination

We extend our analysis now to an alternative choice, the early time parameterization specified in Eqns. (4.11) and (4.12). As before, we use Euclid Redbook specifications for WL and GC and the cut in scales discussed previously for the two observables, i.e. a maximum wavelength cutoff at $k_{\max} = 0.15$ for GC and a maximum multipole of $\ell_{\max} = 1000$ for WL in the linear case, and a cutoff $k_{\max} = 0.5\text{h/Mpc}$ and a maximum multipole of $\ell_{\max} = 5000$ for WL in the non-linear regime, which is analyzed using the prescription described in Sec. 4.3.2.2. We use the two observables both separately and in combination, without accounting for cross correlation of the two (as discussed in section 4.4 this seems to correspond to a conservative choice), with and without *Planck* priors.

Results are shown in Table 4.4 and Figure 4.22. The general behaviour of the constraints is similar to the one in the late-time parameterization, with the combination of GC and WL able to break the degeneracies with standard cosmological parameters, leading to a significant improvement of the constraints on MG parameters, constraining μ and Σ at the 1-2% level. There are some other interesting differences with the late-time scenario. First, the addition of *Planck* priors does not really improve much the constraints obtained by GC or WL alone, which was not the case in the late-time parametrization. This is related to the fact that in the early-time parameterization, GC and WL (non-linear) on their own are already good at constraining both μ (at 2-3 %) and η (at around 8%), with consequently small errors on Σ ($\approx 3\%$). In Appendix 4.7.1.5 we show the derivatives of the matter power spectrum with respect to the MG parameters μ and η in both parameterizations. We can observe that in the early-time scenario, the derivative $dP(k)/d\eta$ is larger than in the late-time parameterization, leading therefore to better constraints. Another difference lies in the correlation among parameters, which for WL and GC+WL is considerably smaller than in the late-time scenario. The Figure of Correlation (defined in Eqn. 3.40) for GC (non-linear HS) alone is 4.7, while for WL (non-linear HS) the correlation is somewhat higher, with FoC $= 7.3$. When combining both probes (GC+WL (non-linear HS)) the FoC goes to an intermediate point of 5.2.

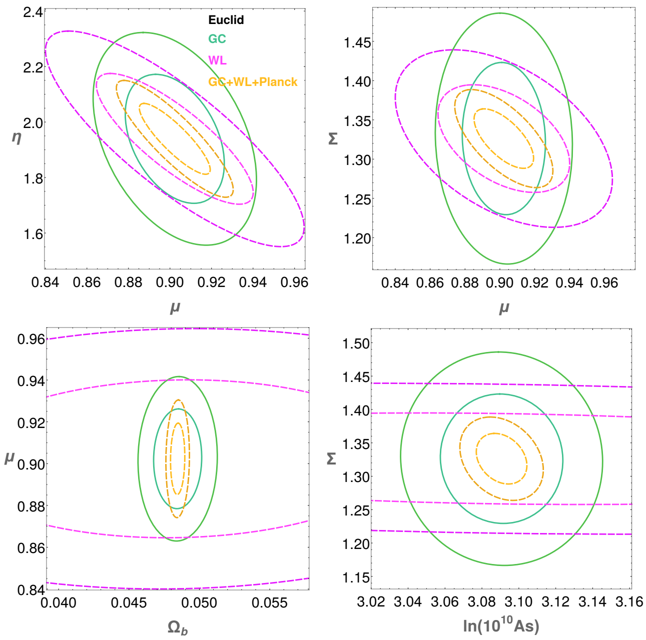


FIGURE 4.22: Fisher Matrix marginalized forecasted contours (1σ, 2σ) for the Euclid Redbook satellite in the early-time parameterization using mildly non-linear scales and the HS prescription. Green lines represent constraints from the Galaxy Clustering survey, pink lines stand for the Weak Lensing observables, and orange lines represent the GC+WL+*Planck* combined confidence regions. **Upper left:** contours for the fully marginalized errors on η and μ. **Upper right:** contours for the fully marginalized errors on Σ and μ. **Lower left:** contours for the fully marginalized errors on μ and Ω_b. **Lower right:** contours for the fully marginalized errors on Σ and $\ln 10^{10} A_s$. Notice that in this parameterization, GC and WL are able to constrain both μ and η or Σ on their own.

4.7.2.2 Early time parameterization

In the early time case, we have two parameters for each $\mu(a)$ and $\eta(a)$ function as in Eq. (4.11). If we are interested in the parameters μ, η and Σ today ($a = 1$), then the parameters E_{12} and E_{21} are not important anymore and we can simply marginalize over them in our Fisher matrix. Then the relation between μ and η is very simple $\{\mu, \eta\} = 1 + \{E_{11}, E_{22}\}$. The corresponding Jacobian is simply a 7×7 identity matrix and we can apply it to the Fisher matrix after we marginalize over the two unimportant parameters.

For the transformation to the pair μ-Σ we use the definition of Σ of Eq. (4.5) and then find the derivatives with respect to E_{ij}. We obtain

$$\frac{\partial \Sigma}{\partial E_{11}} = \frac{1}{2}(2 + E_{22}) \tag{4.53}$$

$$\frac{\partial \Sigma}{\partial E_{22}} = \frac{1}{2}(2 + E_{11}) \quad , \tag{4.54}$$

while all other derivatives remain the same.

4.7.2.3 Other Surveys: DESI-ELG, SKA1 and SKA2

Also in the early time parameterization we obtain the 1-σ fully marginalized errors for Galaxy Clustering (top left panel of Figure 4.23) considering DESI-ELG (yellow), SKA2 (green), SKA1-SUR (orange) and Euclid (blue), and for Weak Lensing (top right panel of Figure 4.23) using Euclid (blue), SKA1 (orange) and SKA2 (green). We also report the results in Table 4.6, where it is possible to notice how the conclusions drawn on the hierarchy of the considered experiments do not change with respect to the late-time parameterization. The main difference is that in this case the full SKA1-SUR GC survey does not constrain the Σ parameter better than the Euclid survey; this is due to the fact that in the early time parametrization, deviations from ΛCDM are present also at high redshift, therefore we do expect the information present at small redshift to be as relevant as the one coming from higher z, where Euclid performs significantly better than SKA1-SUR. The marginalized contours for the μ-η plane, comparing all these surveys, can be seen in the right panel of Fig.4.21.

4.7.3 Testing the effect of the Hu-Sawicki non-linear prescription on parameter estimation

In this section we show the effect of changing the parameters $c_{\rm nl}$ and s used in the HS prescription (specified in Eq. 4.13) for the mildly non-linear matter power spectrum. As mentioned already in Section 4.3.2.2, previous works (see **zhao_modeling_2014**; **zhao_n-body_2011**; **koyama_non-linear_2009**), have fitted the values of $c_{\rm nl}$ and s to match N-body simulations in specific Modified Gravity models. In all these cases the HS parameters $c_{\rm nl}$ and s have been found to be of order unity, with $c_{\rm nl}$ ranging usually from 0.1 to 3 and s from about $1/3$ to a value of around 2. In the absence of N-body simulations for our models, we selected our benchmark HS parameters to be $c_{\rm nl} = 1$ and $s = 1$, as discussed in Section 4.3.2.2, which we used for all the analysis presented above. However, in order to test the effect of a change of c_{nl} and s non-linear parameters on our estimation of errors on the cosmological parameters, we perform our GC and WL forecasts on the MG late-time model (Section 4.7.1) also changing one at a time the values of both HS parameters. We use the following values for our test: $c_{\rm nl} = \{0.1, 0.5,\ 1,\ 3\}$ and $s = \{0,\ 1/3,\ 2/3,\ 1\}$.

In Figure 4.24 we show the percentage discrepancy between the 1σ marginalized error obtained on parameters in the GC non-linear HS benchmark case (see Table 4.3 for the exact values) and the corresponding error obtained by performing the same forecast with a different value of the $c_{\rm nl}$-s parameters. In general terms we see that for the MG parameter μ (left panel of Figure 4.24) the relative difference in the estimation of the 1σ forecasted errors can lie between 90% (at $c_{\rm nl} = 3$, $s = 0.33$) and -30% for $c_{\rm nl} = 0.1$ and $s = 1.0$. The behavior of the contour lines shows us that for a fixed value of $c_{\rm nl}$, the forecasted error on the parameter μ remains unaffected. For η (right panel) the relative discrepancy lies between 40% and -2%. Here, to get the same 1σ errors on η, one would have to vary both $c_{\rm nl}$ and s. We also tested the effect on the standard ΛCDM parameters and found it to be smaller than 4% for all choices of $c_{\rm nl}$ and s. In the case of WL forecasts, we perform the same tests, which are shown in Figure 4.25. We can observe that the relative discrepancies in the case of the μ parameter lie between $\sim$60% and $\sim$-15%, while for Σ the discrepancy is considerably smaller. The 1σ error on Σ varies only between $\pm$6%. This is however

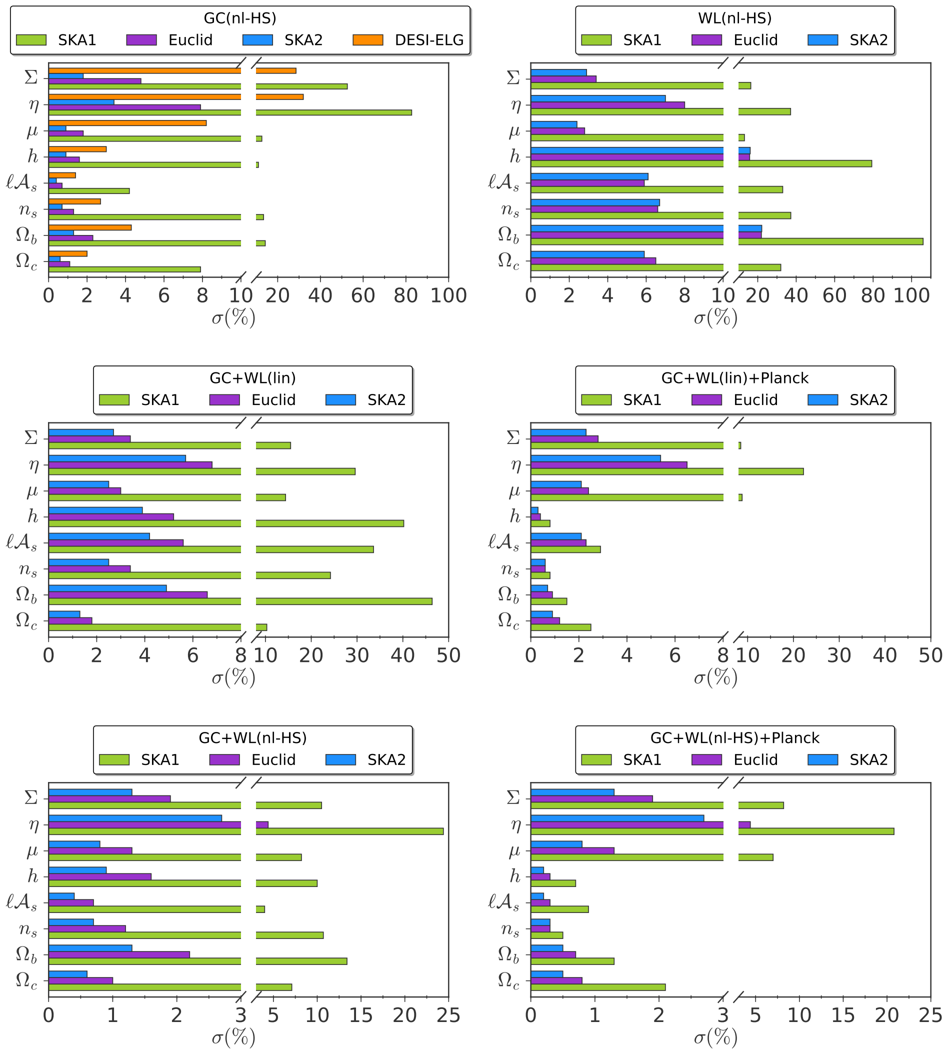


FIGURE 4.23: Same as Fig.4.20 but for the early-time parameterization (Eqns. 4.11,4.12) The 1σ fully marginalized constraints on the parameters are weaker for WL than for GC, which is a consequence of the higher correlation among variables for the Weak Lensing observable.

a particular effect of choosing Σ, which is the true WL observable. If we perform this test on η using WL, we will find a stronger discrepancy, that lies in between $\sim$50% and $\sim$-15%, similar to the one found for μ.

We can also test the effect of adding these two HS parameters as extra nuisance parameters to our model of the non-linear power spectrum; therefore, by taking derivatives of the observed power spectrum with respect to these parameters, we can forecast what would be the estimated error on $c_{\rm nl}$ and s. Then, marginalizing over $c_{\rm nl}$ and s would yield more realistic constraints on our cosmological parameters, and take into account our ignorance on the correct parameters of the HS prescription. In Table 4.7 we list the 1σ marginalized constraints on $c_{\rm nl}$ and s for our benchmark fiducial ($c_{\rm nl} = 1$, $s = 1$) and using the standard fiducial for the cosmological parameters used previously for the MG late time parametrization.

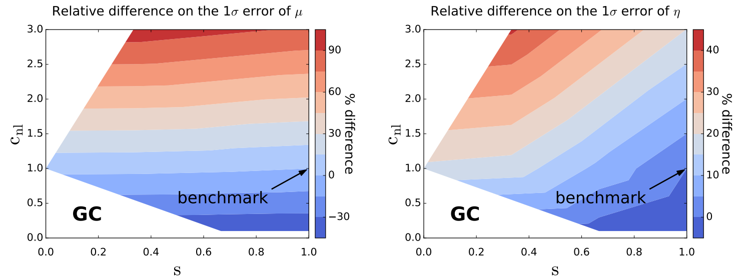


FIGURE 4.24: Effect of the $c_{\rm nl}$ and s parameters on the 1σ marginalized error of the μ (left panel) and the η (right panel) parameters in the MG late-time parametrization for a Euclid Galaxy Clustering forecast with Redbook specifications. The colored contours show the percentage discrepancy when departing from the benchmark case $c_{\rm nl} = 1$ and $s = 1$ (marked with a black arrow) to all other points in the $c_{\rm nl}$-s space. The red (blue) contours signal the regions of maximum positive (negative) discrepancy. For example in the left panel, choosing $c_{\rm nl}$ and s in the red region, will yield a 1σ marginalized error on μ which is 90% larger than in the benchmark case (see Table 4.3 for the benchmark forecast). For the standard ΛCDM cosmological parameters (not shown here) the discrepancy is smaller than 4% for all choices of $c_{\rm nl}$ and s.

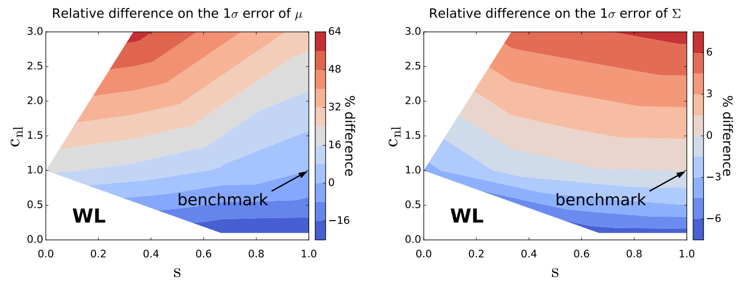


FIGURE 4.25: Same as Fig.4.24 but for a Weak Lensing Euclid forecast using Redbook specifications. In the left panel, choosing $c_{\rm nl}$ and s in the red region, will yield a 1σ marginalized error on μ which is 60% larger than in the benchmark case (see Table 4.3 for the benchmark forecast). In the right panel we see that for the MG parameter Σ the maximum positive and negative discrepancy is only of about 6%. The maximum and minimun discrepancy for the MG parameter η is of -15% and 50%. This means that the parameter Σ (defined as the lensing Weyl potential, and therefore directly constrained by Weak Lensing) is much less sensitive to changes in the non-linear prescription parameters.

Taking these HS parameters into account in our Fisher forecast, will automatically change the constraints on the other parameters. For our method to be consistent, we would like this effect to remain as small as possible. In Table 4.7, we list the same constraints reported in Table 4.3, but obtained marginalizing over $c_{\rm nl}$ and s at the benchmark fiducial $c_{\rm nl} = 1$, $s = 1$. Comparing the two tables, we can see that for GC the errors on the cosmological parameters remain quite stable, except for η and Σ; this is understandable since those parameters are not well constrained by GC alone. For WL, we see a difference of 4 to 7 percent points in the errors on h, n_s and Ω_b, while all other errors remain stable, with Σ varying less than 2 percent points. Remarkably, the combined constraints from GC+WL are even less affected by the two nuisance parameters. Comparing the Modified Gravity FoM between the two tables, shows the expected behavior, the MG FoM is reduced when adding two extra parameters, but the change is very small, of just 0.3-0.5 nits.

Euclid (Redbook)	Ω_c	Ω_b	n_s	$\ell\mathcal{A}_s$	h	μ	η	Σ	$c_{\rm nl}$	s	MG FoM
Fiducial	0.254	0.048	0.969	3.060	0.682	1.042	1.719	1.416	1	1	relative
GC(nl-HS)	1.0%	2.8%	1.3%	1.1%	2.0%	1.7%	784%	480%	372%	236%	2.4
WL(nl-HS)	6.5%	25%	8.3%	9.1%	19%	25%	46%	6.0%	1680%	899%	4.2
GC+WL(nl-HS)	1%	2.8%	1.2%	1%	1.9%	1.6%	2.6%	1.2%	333%	166%	8.5

TABLE 4.7: 1σ fully marginalized errors on the cosmological parameters and the two HS parameters $c_{\rm nl}$ and s for a Euclid Galaxy Clustering forecast, a Weak Lensing forecast and the combination of both in the late-time parameterization of Modified Gravity using non-linear scales and the HS prescription. In contrast to Table 4.3 (where c_{nl} and s had been fixed to the benchmark value) here we include $c_{\rm nl}$ and s as free parameters and marginalize over them. The MG FoM is computed relative to the same Euclid Redbook GC linear case, used previously. The errors and the MG FoM show the expected behavior of adding two nuisance parameters, and remain quite stable. All other naming conventions are the same as for Table 4.3. Remarkably, the combination of GC and WL is still able to constrain all Modified Gravity parameters at the level of 1-2 % after marginalizing over the non-linear parameters.

4.8 Conclusions of Modified Gravity forecasts

We study in this paper the constraining power of upcoming large scale surveys on Modified Gravity theories, choosing a phenomenological approach that does not require to specify any particular model. To this purpose we consider the two functions μ and η that encode general modifications to the Poisson equation and the anisotropic stress. We study three different approaches to MG: redshift-binning, where we discretize the functions $\mu(z)$ and $\eta(z)$ in 5 redshift bins and let the values of μ and η in each of the bins vary independently with respect to the others; an early-time parameterization, where μ and η are allowed to vary at early times and their amplitude can be different from unity today; a late-time parameterization where μ and η are linked to the energy density of dark energy and therefore they are very close to unity in the past, but they can vary considerably at small redshifts. For convenience, we summarize all results in Table **??**, with a direct link to the section and tables related to each scenario.

We use the predictions of linear perturbation theory to compute the linear power spectrum in Modified Gravity and then use a prescription to add the non-linearities, by interpolating between Halofit non-linear corrections computed for the linear power spectrum for the MG model and for the corresponding GR model ($\eta = \mu = 1$). We find that the non-linear power spectrum is sensitive to changes in μ and η; limiting the analysis to linear scales significantly reduces the constraining power on the anisotropic stress. Using this prescription, we perform Fisher forecasts for Galaxy Clustering and Weak Lensing, taking into account linear and non-linear scales. We use the specifications for Euclid (using Redbook specifications), SKA1 & SKA2 and DESI (only ELG). In addition to these surveys we also include *Planck* priors obtained by performing an MCMC analysis with *Planck* data for the MG parametrizations considered here.

In the redshift-binned case, we find that in the linear case the μ_i and η_i parameters are strongly correlated, while including the information coming from non linear scales reduces this correlation. We compute a figure of merit (FoM), given by the determinant of the μ-η part of the Fisher matrix, for the cases examined, finding that the combination of Galaxy Clustering and Weak Lensing is able to break the degeneracies among Modified Gravity parameters; as an example the error obtained with the non-linear prescription on μ (η) in the first redshift bin changes from 7% (20%) for Galaxy Clustering to 2.2% (3.6%) when this is combined with Weak Lensing, even if Weak Lensing alone is not very constraining for the same parameters. Overall, constraints are stronger at low redshifts, with the first two bins (0 < z < 1) being constrained at better than 5% for both μ and η if non-linearities are included (while the constraints are half as good, $< 10\%$, if we only consider linear scales).

Given the significant correlation between the μ_i and η_i parameters, we apply the ZCA decorrelation method, in order to find a set of uncorrelated variables, which gives us information on which redshift dependence of μ and η will be best constrained by future surveys. If one combines GC+WL (Euclid Redbook)+*Planck*, the best constrained combinations of parameters (effectively $2\mu + \eta$ in the lowest redshift bin) will be measured with a precision of better than 1%. In the linear case, the errors on the decorrelated q_i parameters are about 2 orders of magnitude smaller than for the primary parameters, while in the non-linear HS case, the improvement in the errors is of one order of magnitude. This also shows that applying a decorrelation procedure is worth even when non-linearities are considered.

In addition to binning Modified Gravity functions in redshift, we also forecast the constraining power of the same probes in the case where we assume a specific time evolution for the μ and η functions. We choose two different and complementary time evolutions, used in **planck_collaboration_planck_2016** and to which we refer as late-time and early-time evolution. We investigate also in this case the impact of the non-linear prescription interpolating between Halofit and the MG power spectrum. For these parameterizations we extract constraints on the present reconstructed value of μ, η and Σ, where the latter is the parameter actually measured directly by Weak Lensing. In the late time parameterization, in the linear case, μ is mainly constrained by GC (although poorly, at the level of 17%) while WL constrains directly Σ, the modification of the lensing potential, at the level of 9%. Adding non-linear scales allows to significantly improve constraints down to less than 2% for GC (on μ) and to less than 5% on Σ for those two probes. Combining probes allows to reach $1-2\%$ on all Modified Gravity values of the μ, η and Σ functions at z = 0.

In the early-time parameterization, we find that including non-linearities allows to constrain also the η and Σ functions with GC alone at the level of 8% and 5%, respectively. This is related to the early time deviations from GR allowed by this parameterization, which are not present in the late time case: a variation in η can yield a variation of the amplitude of the power spectrum, which can then be measured in the mildly non-linear regime. Overall, also in this case the combination of Weak Lensing and Galaxy Clustering leads to errors of the order of 1% on the values of these functions at present.

Finally, we test the impact on the forecasts given by uncertainties appearing in the non linear HS prescription, related in particular to the parameters c_{nl} and s. We find that the errors on the parameters μ and η can vary by up to 90% for the Galaxy Clustering case and up to 65% for Weak Lensing when we change the fiducial values of the HS parameters in a region between 0 and 3 for c_{nl} and 0 and 1 for s. The effect on Σ is quite small, with a discrepancy of $\pm 6\%$ compared to the benchmark case. Interestingly, when we include these two parameters as extra nuisance parameters in our forecast formalism and marginalize over them, the effect is very small and the errors found previously remain stable both for GC, WL and their combination.

It is clear that limiting the analysis to linear scales discards important information encoded in structure formation. On the other hand, a realistic analysis of non-linear scales would have to include several further effects (baryonic effects, higher order RSD's, damping of BAO peaks, corrections to peculiar velocity perturbations, higher order perturbation theory in Modified Gravity, just to name a few), which make our non-linear case an optimistic limit. Therefore, the quantitive true constraints given by a survey like Euclid will probably lie in between these two limiting cases.

Chapter 5

Fitting and forecasting coupled dark energy in the non-linear regime

5.1 Introduction

Cosmological data have reached a level of precision that forces us to address effects which may have been safely neglected in the past. Among them, predicting the behaviour of theoretical models at non-linear scales gives a more concrete chance to disentangle ΛCDM from alternative scenarios. The non-linear power spectrum contains useful information encoded in the clustering of galaxies and the weak lensing correlations, which can help to constrain more tightly the cosmological parameters and remove certain degeneracies between them. The drawback relies in the fact that the full estimate of the non-linear power spectrum requires time-demanding and computationally expensive N-body simulations. As a consequence, there exist simulations only for a handful of models beyond ΛCDM on a very limited subset of the parameter space of dark energy theories. This makes it hard to use the non-linear power spectrum in forecasts (see e.g.), both when doing cosmological parameter estimation via Monte Carlo simulations and with the simpler approximation of the Fisher matrix formalism. The alternative is to employ fitting functions for the non-linear power spectrum, like Halofit (**smith_stable_2003**; **takahashi_revising_2012**) or machine learning estimators like the cosmological emulator (CosmicEmu) (**heitmann_coyote_2010**; **lawrence_coyote_2010**; **heitmann_coyote_2014**) and the interpolator PkANN **agarwal_pkann_2012**; **agarwal_pkann_2014** based on artificial neural networks. All these methods are calibrated or trained using large sets of N-body simulations and typically permit an exploration of the parameter space just around a fiducial ΛCDM cosmology or wCDM cosmology with a constant w. An extension of Halofit to account for massive neutrinos was realised in (**bird_massive_2011**), which was then used in (**audren_neutrino_2013**) to perform a parameter estimation using a Markov Chain Monte Carlo technique. Other authors have recently used polynomial fits to take into account systematic effects in the non-linear regime like baryonic effects such as in **bielefeld_cosmological_2014** or to parametrize the cosmological dependence of non-linear clustering, beyond the Zeldovich approximation (see **mohammed_analytic_2014**). These fitting functions have, however, an intrinsic error that usually increases when scales become highly non-linear; therefore one has to be aware of the range in scales and redshifts they were designed to work on, in order to keep the errors induced on the power spectrum under control.

As an alternative to fitting formulae and N-body simulations, there has been in recent years a substantial progress in semi-analytical methods to calculate the non-linear power spectrum at mildly non-linear scales, such as for example renormalized perturbation theory (**crocce_renormalized_2006**) and all other resummation methods derived from it (for example **Anselmi2012**), the time renormalisation group (TRG) (**Pietroni_2008**) and effective field theories of large scale structure **baumann_cosmological_2010** Most of these approaches claim to reach percent accuracy at the baryon acoustic oscillation (BAO) scale and a bit beyond the second BAO peak, but they are still not able to predict what happens at highly non-linear scales, when the single stream approximation does not hold any longer.

In this paper, we will use N-Body simulations to find fitting functions for a class of models beyond ΛCDM usually refered to as coupled Dark Energy (CDE). These models, widely discussed in literature (**Wetterich_1995**; **Amendola_2000**; **Amendola_2004**; **pettorino_baccigalupi_2008**), involve an extra degree of freedom, associated to a scalar field that provides acceleration and mediates a fifth force, in addition to gravity, which is felt by dark matter particles only. Semi-analytical non-linear analysis **wintergerst_clarifying_2010**; **Saracco_etal_2010** and cosmological N-body simulations within coupled Dark Energy have been performed by many different groups (**baldi_etal_2010**; **Li_Barrow_2011**; **maccio_coupled_2004**; **carlesi_hydrodynamical_2014-** **baldi_codecs_2012**; **2015arXiv150407243P**) and their effects on large scale structure formation have been identified and characterised. The power spectrum, halo mass functions and concentration, halo spin and sphericity, voids and amount of substructures show noticeable differences compared to a simple ΛCDM model (see for example (**mainini_mass_2006**; **maccio_concentration_2008**; **baldi_clarifying_2011**; **cui_halo_2012**; **baldi_effect_2011**; **sutter_observability_2014**) and the review article (**Baldi_2012b**)). For a constant coupling, constraints have been found for a variety of probes (**pettorino_constraints_2012**; **pettorino_testing_2013**; **amendola_skewness_2004**; **Xia_2009**), with the latest ones discussed by the Planck collaboration in (**planckcollaboration_planck2015_2** Recently there have been attempts to constrain more general couplings between dark matter (DM) and dark energy (DE) using large-scale structure (**creminelli_single-field_2014**), CMB (**morris_cosmicmicrowave_2014**) or laboratory experiments (**hamilton_atominterferometry_2015**). Forecasts on coupled dark energy using galaxy clustering (GC) and weak lensing (WL) measurements for future surveys like Euclid have been discussed in (**amendola2012testing**), but have been performed using only linear power spectra for the CDE models. The TRG method has been extended to coupled Dark Energy (**saracco_non-linear_2010**) and to massive neutrinos (**lesgourgues_non-linear_2009**). However the TRG method does not produce a reliable estimation of the power spectrum for scales larger than $k \approx 0.3h/\mathrm{Mpc}$, which makes them less suitable for forecasts which attempt to extract information on highly non-linear scales.

The CoDECS (Coupled Dark Energy Cosmological Simulations) set of N-body simulations (**baldi_codecs_2012**) has shown that CDE models have characteristic and measurable features in the morphology and history of non-linear structures, such as halos, subhalos and voids, and therefore in the non-linear power spectrum.

The aim of this paper is to create fitting functions which are valid in the observable regime of non-linear perturbations at all interesting redshifts and reproduce the subtle effects of coupled dark energy on the non-linear power spectrum while allowing us to vary the different parameters of the model. We use this to perform forecasts of cosmological parameters assuming coupled dark energy as the fiducial model, using galaxy clustering and weak lensing as observational tools, as expected for future surveys like Euclid **laureijs_euclid_2011**; **amendola_cosmology_2012-short** We do a careful treatment of errors and systematics, so that we take into account all errors induced by our fitting functions, the cosmic emulators and the extraction of the power spectrum from the N-body simulation into the analysis, systematics related to non-linear effects that affect the redshift space distortions and the lensing signals. In this way we obtain a conservative estimate on how well a probe like Euclid, will be able to measure a DM-DE coupling.

5.2 Coupled Dark Energy

In this section we briefly review the main equations governing the coupled DE models that will be investigated in the present work. For a more detailed and extended discussion and for the full derivation of the equations we refer the interested reader to some of the original publications on the subject (**Amendola_2000**; **Amendola_2004**; **pettorino_baccigalupi_2008**; **baldi_etal_2010**).

The background evolution for the coupled DE scenario model is described by the following equations, in which the subscripts r, b, c and ϕ, indicate radiation, baryons, cold dark matter (CDM) and the dark energy scalar field, respectively:

$$
\begin{aligned}
\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} &= \sqrt{\frac{2}{3}}\beta(\phi)\frac{\rho_c}{M_{Pl}} \ , && (5.1)\\
\dot{\rho}_c + 3H\rho_c &= -\sqrt{\frac{2}{3}}\beta(\phi)\frac{\rho_c\dot{\phi}}{M_{Pl}} \ , && (5.2)\\
\dot{\rho}_b + 3H\rho_b &= 0 \ , && (5.3)\\
\dot{\rho}_r + 4H\rho_r &= 0 \ , && (5.4)\\
3H^2 &= \frac{1}{M_{Pl}^2}(\rho_b + \rho_c + \rho_r + \rho_\phi) \ . && (5.5)
\end{aligned}
$$

We express from now on the scalar field ϕ in units of the Planck mass $M_{pl} \equiv 1/\sqrt{8\pi G}$, and choose as potential $V(\phi)$ an exponential $V(\phi) = Ae^{-\alpha\phi}$ (**Lucchin_Matarrese_1984**; **Wetterich_1988**). The coupling function $\beta(\phi)$ defines the strength of the interaction between the DE fluid and CDM particles and in the present work we will restrict our analysis to the simplified case of a constant coupling $\beta(\phi) = \beta$, although in general it could be a field-dependent quantity **Amendola_2004**; **Baldi_2011a**

The current constraints on a coupling to ordinary matter are very tight. The "post-Einstein" coupling parameter $\bar{\gamma}$ that measures the local admixture of a spin-0 field to gravity is constrained in Solar System experiments (see e.g. the PDG review **Agashe:2014kda** and also (**Will_2005**; **Bertotti_Iess_Tortora_2003**)) roughly to $|\bar{\gamma}| \leq 4 \cdot 10^{-5}$. This parameter enters the modification of the effective Newton constant as $G_{eff} = G_N(1 - \bar{\gamma}/2)$; in our notation, this is $G_{eff} = G_N(1 + 4\beta^2/3)$ (see below) and therefore $\beta^2 = -3\bar{\gamma}/8$. A coupling $\beta^2_{baryons}$ appears then constrained to be smaller than 10^{-5} roughly, and we assume therefore that is completely negligible. As a consequence, baryons follow the usual geodesics of a FLRW cosmology, which allows coupled DE to pass the stringent local gravity constraints without the need to employ any screening mechanism (**2015arXiv150203888H**).

Due to the exchange of energy between DE and CDM, the energy density of the latter will no longer scale as the cosmic volume, and by assuming the conservation of the CDM particle number one can derive the time evolution of the CDM particle mass by integrating Eq.(5.2) between the present time ($z = 0$) and any other redshift z:

$$
m_c(z) = m_{c,0}e^{-\beta(\phi(z)-\phi(0))} \ . \qquad (5.6)
$$

Parameter	Explanation	Value	Reference ΛCDM
A	Potential normalization	0.0218	–
α	Potential slope	0.08	–
$\phi(z=0)$	Scalar field normalization	0	–
β	Coupling parameter	$\{0.05,\ 0.10,\ 0.15\}$	–
$w_\phi(z=0)$	DE equation of state	$\{-0.997,\ -0.995,\ -0.992\}$	-1
$\sigma_8(z=0)$	Power spectrum amplitude	$\{0.825,\ 0.875,\ 0.967\}$	0.809

TABLE 5.1: The main parameters and normalizations of the three coupled DE models of the Codecs suite considered in the present work. The last column displays the corresponding valuem for the reference ΛCDM cosmology.

5.3 The CoDECS simulations

The CoDECS suite[1] includes simulated periodic volumes of the universe at different scales and with different physical ingredients (as e.g. simulations with and without hydrodynamics) in the context of a series of coupled DE cosmologies characterised by various choices of the self-interaction potential $V(\phi)$ and of the coupling functions $\beta(\phi)$. The simulations have been performed with a suitably modified version (see (**baldi_codecs_2012**) for more details on the numerical implementation) of the widely-used TreePM N-body code GADGET (**springel_cosmological_2005**). Such modified version includes all the relevant effects that characterise coupled DE cosmologies, from the modified background evolution to the CDM particle mass variation, the "fifth-force" and the velocity-dependent acceleration appearing in equation (**??**).

As already stated above, in this work we will consider – besides the reference ΛCDM simulation – the subset of CoDECS runs characterised by an exponential potential and by a constant coupling function. This consists of three different coupled DE models with the same potential slope α and with three values of the coupling $\beta = \{0.05\,, 0, 1\,, 0.15\}$. The short names for these simulations are respectively EXP001, EXP002 and EXP003. All the models have been built in order to have the same cosmological parameters at $z = 0$ consistent with the WMAP7 results (**komatsu_seven-year_2010**), see Table 5.2, with the obvious exception of the value of the equation of state parameter w_0, that changes from model to model due to the different dynamics of the DE scalar field. The present observational constraints on the cosmological parameters have only slightly changed with the latest updated release of Planck data (**Planck:2015xua**) with respect to the assumed WMAP7 values, and are still good enough for the purposes of this paper, being in good agreement with large scale structure observations. For what concerns linear perturbations, all cosmologies have been normalised to have the same statistics (i.e. the same power spectrum shape and amplitude) of density fluctuations at the redshift of the Cosmic Microwave Background $z_{\rm CMB} \approx 1100$. As a consequence of this choice and of the different growth associated with the various coupling values, all the models will have a different normalisation σ_8 of the linear perturbations amplitude at $z = 0$. The main features of these models are summarised in Table ??. We refer to (**baldi_codecs_2012**) for further details.

For the purposes of the present work, we will employ the matter power spectra extracted from the CoDECS runs of these cosmologies at different redshifts in order to find a fitting formula that captures with high accuracy the deviations of the coupled DE nonlinear power spectra from the reference ΛCDM case.

5.4 Extracting the power spectrum at small scales

The power spectrum $P(k)$, is defined as the ensemble average of the density contrast in Fourier space $\langle\delta(k)\delta(k')\rangle \equiv (2\pi)^3 P(k)\delta_D(k+k')$. When trying to extract this quantity from an N-body simulation, one has to take into account several technicalities related to *sampling* effects which appear as a consequence of treating a *discrete*

[1]see also the public CoDECS database at www.marcobaldi.it/CoDECS/

Parameter	Value
H_0	$70.3\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$
Ω_{CDM}	0.226
Ω_{DE}	0.729
$\mathcal{A}_s$	2.42×10^{-9}
Ω_b	0.0451
n_s	0.966

TABLE 5.2: The set of cosmological parameters used in all CoDECS simulations, consistent with the WMAP7 results.

distribution of particles. For more details related to this problem and different solution methods, see (**cui_ideal_2008**).

At large scales (or equivalently small k) the power spectrum suffers from uncertainties due to the finite size of the simulation box, since there are only few independent modes to sample the signal from. On the other hand, at high k, one is limited by the resolution of the simulation, since one cannot sample wave modes smaller than the typical grid size L/N where L is the length of one side of the simulation box and N is the number of particles. This maximum frequency is called the Nyquist frequency $k_{Ny} \equiv 2\pi N/L$ and modes smaller than this cannot be reliably measured (this corresponds to the so-called aliasing effect). For the CoDECS simulations used in this work, the Nyquist frequency has the value $k_{Ny} \approx 2.2$ h/Mpc at present time.

The power spectrum computation embedded in GADGET-3 that was adopted in the CODECS simulations employs the so-called folding method developed by **colombi_accurate_2009** – which is based on **jenkins_galaxy_1999** – to calculate the matter power spectrum for smaller scales than the Nyquist frequency. Following this method one ends up with two separate power spectra, P_{top} which is calculated using the simulation particle mesh (PM) at $k \lesssim k_{Ny}$ and P_{fold} which is the folded power valid for $k \gtrsim k_{Ny}$. In order to provide a single sampling of the power spectrum across k_{Ny} the CoDECS project employed a simple interpolation procedure around k_{Ny} by averaging the two power spectra in the region of overlap. However, this might introduce some spurious features that appear only when considering the ratio of two power spectra $P(k)$ at highly non-linear scales. Although such features are harmless for most practical purposes, for the aims of the present work it is very important to have accurate ratios of power spectra, since we want to calculate fitting functions that can capture the effect of the dark energy - dark matter coupling compared to ΛCDM and therefore we need to correct for these spurious effects.

We then developed a PYTHON code that finds the optimal interpolation and matching between the P_{top} and P_{fold} power spectra. By evaluating the first and second derivatives of the ratios and minimizing abrupt changes, it finds the optimal number of points to be removed from P_{top} due to the aliasing error and the number of points to be removed from the P_{fold} due to the low sampling error; at the same time it looks for the optimal linear interpolation weights between them. Moreover, it cuts off the power spectrum when the shot noise error ($P_{shot} = 1/N$) reaches 10% of the estimated power spectrum. This method improves considerably the convergence of the fitting functions at non-linear scales, allowing us to reach our accuracy goal of 1-2% (see section 5.5 on fitting functions). In figure (5.4), the uncertainties on $P(k)$ are plotted, and the shaded region represents the error on the power spectrum due to the finite number of modes available. The clear jump present in this shaded region occurs at the scale in which the folded and top level power spectra have been matched together, which corresponds roughly to k_{Ny}.

5.5 Fitting functions

5.5.1 The net effect of the DM-DE coupling at non-linear scales

We now model the net effect of a coupled DE model with an exponential potential and a constant coupling on top of a fiducial ΛCDM non-linear power spectrum, by evaluating the ratio of the non-linear power spectrum of the coupled DE model, with respect to the one in the ΛCDM model, both extracted from the CoDECS simulations. In figure (5.1) we show the ratio $R(k;\beta,z) \equiv P_{Exp}(k;\beta,z)/P_{LCDM}(k;\beta,z)$ as a function of the scale k, for four different redshifts $z = \{0,\, 0.55,\, 1.0,\, 1.61\}$; each panel contains the curves corresponding to the three available constant couplings in the CoDECS simulations: $\beta = \{0.05,\, 0.10,\, 0.15\}$. Since all coupled DE models have the same amplitude of perturbations at recombination, an increasing coupling has the effect of inducing a higher linear normalization σ_8 of the power spectrum at $z = 0$ (**baldi_codecs_2012**). Therefore, in order to see the net effect of the coupling at non-linear scales, each power spectrum ratio has been re-normalized by dividing each model by its respective σ_8^2, so that at linear scales $k \lesssim 0.1$h/Mpc all the ratios are unity. The net effect of the fifth force is a "bump" at non-linear scales, whose amplitude increases with higher couplings and whose maximum is shifted into higher wavenumbers k for higher redshifts. This extra information imprinted into the non-linear power spectrum is what we want to use to improve the estimation of parameters using future surveys. To achieve this, we will fit these curves which are functions of redshift, physical scale and coupling, using the minimal number of numerical parameters possible, while keeping the accuracy goal at the 1% level.

5.5.2 Generating the fitting functions

The fact that one can observe a clear trend that relates the amplitude of the signal to an increase of the coupling, together with a shift of the peak towards larger length scales when looking at smaller redshifts, is an indication that we should be able to find a relatively simple fitting formula describing this behaviour, which will be then a function of z, k and the coupling constant β only.

To perform the fit, we use a least-squares-minimizing technique, using the conjugate gradient method **weisstein_least_????** Taking into account the particular form in k-space of the ratios $R(k;\beta,z) \equiv P_{EXP}(k;\beta,z)/P_{LCDM}(k;\beta,z)$ that we need to fit, we use as an Ansatz different sigmoid functions to reproduce the particular form of the peak. For each fitting model, we keep the same functional form for the k-dependence of $R(k;\beta,z)$ at all redshifts and for all couplings. We tried 7 different sigmoid functions as fitting models, but we only show the best two models M2 and M7 in table 5.3. All models contain 5 coefficients, which are dependent on the coupling β and the redshift z: ($a_i = a_0,\, a_1,\, c,\, b,\, k_0$, with $i = 0, ..., 5$). The coefficients k_0 and b determine qualitatively the form of the peak to be fitted, while the others control mostly the shifting and the flattening of the peak.

Each coefficient a_i is then fitted using a polynomial in β and z, up to a maximum of third order in powers of β and z. Polynomials of order 4 and 5 were also examined, but the gain in goodness of fit was minimal compared to the increase in the number of free parameters. Therefore, third order polynomials were the best compromise between complexity and goodness of fit.

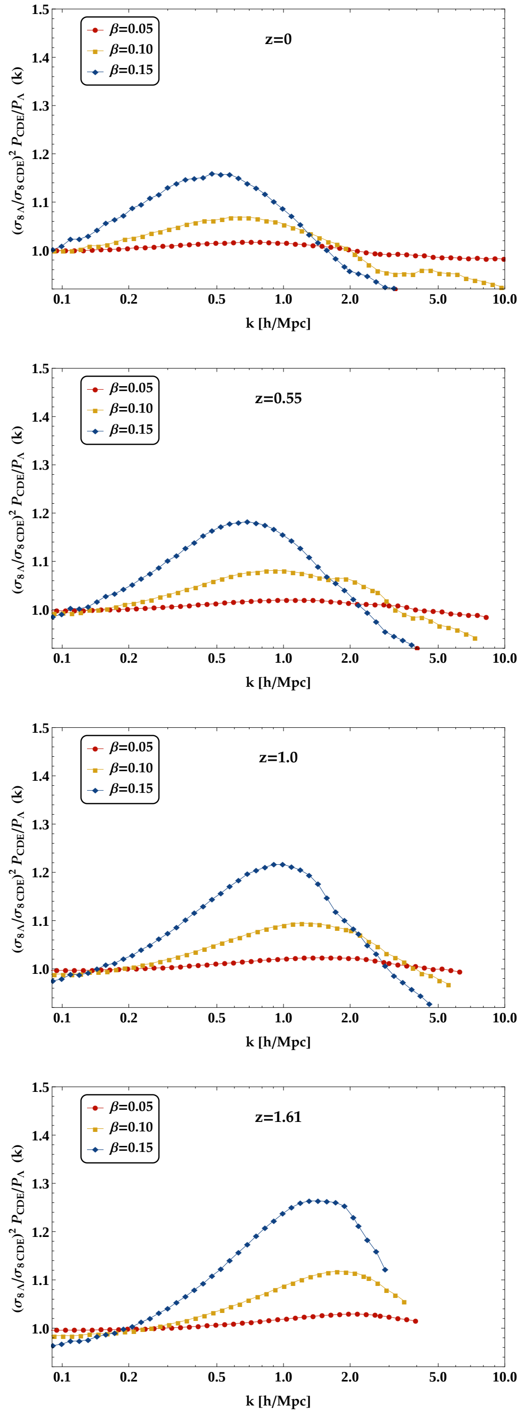


FIGURE 5.1: Ratios of the non-linear matter power spectra of the CoDECS CDE models with an exponential potential, with respect to the CoDECS ΛCDM power spectra, normalized by their respective σ_8^2, evaluated at four different redshifts and three different coupling constants β. Upper left panel: $z = 0$, upper right panel: $z = 0.55$, lower left panel: $z = 1.0$, lower right panel: $z = 1.61$. The blue line represents the model with strongest coupling $\beta = 0.15$, the yellow line $\beta = 0.10$ and the red line the smallest available coupling $\beta = 0.05$. In order to be able to observe the net effect of the β coupling at non-linear scales we have divided each power spectrum by its respective σ_8^2.

Model Name	Functional form	# of coefficients	R^2-value
M2	$f(k) = 1 + a_0 + a_1 \cdot k + c \cdot k \cdot \arctan((k - k_0) \cdot b)$	5	0.99996
M7	$f(k) = 1 + a_0 + a_1 \cdot k + c \cdot k \cdot \frac{b \cdot (k-k_0)}{\sqrt{1+b^2 \cdot (k-k_0)^2}}$	5	0.999989

TABLE 5.3: Fitting models M2 and M7 with their corresponding number of fitting coefficients and their R^2-value. Each coefficient a_0, a_1, c, b and k_0 is fitted as a polynomial in the coupling parameter β and the redshift z (see Appendix **??** for further details). The R^2-value is a measure of the goodness of fit: a value of 1 corresponds to a perfect fit, while 0 means that the model does not fit the data.

The best fitting models were chosen according to their coefficient of determination (also known as R^2-value), which is a statistical measure for the goodness-of-fit **weisstein_correlation_????**. It can be simply defined as $R^2 = 1 - S_{res}/S_{tot}$, where S_{res} is the residual sum of squares (the residual between the data points and the fitting function) and S_{tot}, which is the total sum of squares and is proportional to the variance of the data. An R^2-value of 1 corresponds to a perfect fit. The analytical expressions for the best models M2 and M7 are shown in table 5.3, together with their R^2-value. We performed the whole Fisher analysis (see section **??** below), for both of the models and the results on the parameter estimation are basically the same (less than half of a percent relative difference in the estimated final errors).

5.5.3 Fitting functions and cosmological parameters

In order to use the fitting formulae obtained before to forecast cosmological parameters of the model, we need to assume that the shape of the non-linear coupled DE signal does not change dramatically if the other cosmological parameters, apart from β and σ_8, are modified by small amounts. This is justified since in the deeply non-linear regime, the evolution of perturbations is ruled by mode-mode coupling between high k wavevectors (non-linear k modes are not independent of each other anymore), which erases most of the information about initial conditions and makes the shape of the non-linear power spectrum at large k practically independent of cosmological background parameters, such as n_s, Ω_b and h. This has been shown to be the case when calculating perturbatively non-linear corrections to the power spectrum, see for example (**crocce_memory_2006**; **pietroni_coarse-grained_2011**), but also analyzing the covariance matrix of non-linear power spectra using a large suite of N-body simulations as was investigated in (**takahashi_non-gaussian_2011**). Furthermore, the decoupling of virialized structures in the small scale regime from the background dynamics of the universe, is one of the cornerstones of the recently developed effective field theory of large-scale structure **baumann_cosmological_2012-2** and was also shown to be approximately true using a coarse grained cosmological perturbation approach **manzotti2014acoarse** Since β and σ_8 are quantities that affect directly the linear perturbations and the virialization dynamics and we are only looking at a particular signal at very small scales, they should be the main parameters determining the shape of the non-linear "bump". An investigation of how robust these fitting formulas are, with respect to a change of cosmological parameters, would need either more high-resolution N-body simulation for coupled DE scenarios or the development of a consistent perturbation theory for modified gravity and scalar-tensor theories that reaches highly non-linear scales.

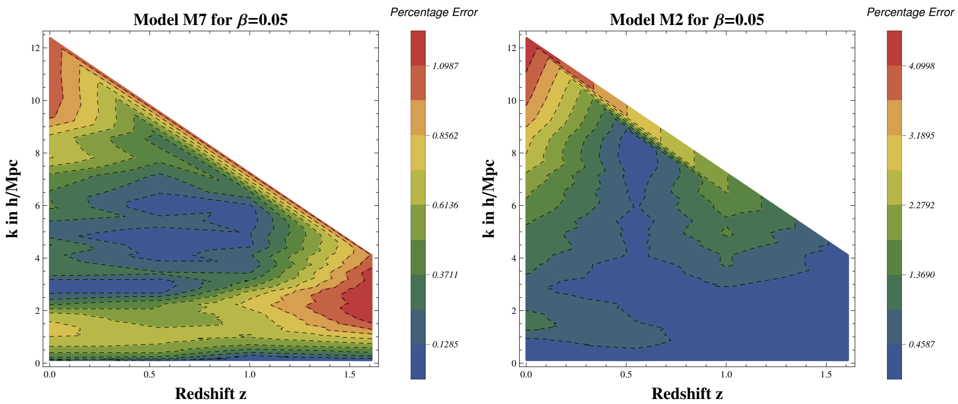


FIGURE 5.2: Error contour plot for the fitting functions of model M7 (left) and M2 (right) applied compared directly to the N-body simulations. We show that for the smallest coupling constant available in the simulations, $\beta = 0.05$, the error remains below 1% for the scales and redshifts we are interested in.

5.6 Non-linear power spectrum and error estimation

The accuracy of the fitting functions when compared to the original N-body simulation power spectra is shown in figure 5.2, where the absolute value of the percentage error between fitting function and the original power spectra is plotted as a function of redshift and scale. In this case we show that for the model M7 the error remains well below 1% for a coupling constant of $\beta = 0.05$ and for the scales and redshifts we are interested in. For higher couplings the error goes up to a maximum of 1.5%. For the model M2 the same test gives similar results, yielding <1% error in all interesting scales and redshifts. We can then expect that when applying these fitting functions on top of ΛCDM power spectra, the intrinsic error is less than 1% and the extra sources of error are all given by the N-body simulation spectra and by the estimators of the non-linear ΛCDM power spectrum.

Since the fitting functions shown in the previous section are useful when applied on top of a ΛCDM non-linear power spectrum, we need to choose an estimator for the non-linear ΛCDM $P(k)$. Our tests show that *FrankenEmu* (**heitmann_coyote_2014**), an improved version of the original Coyote Cosmic Emulator, performs better than the revised version of Halofit by (**takahashi_revising_2012**), which is included in recent versions of CAMB (**lewis_efficient_1999**). While at $z = 0$ both estimators work similarly well with an accuracy at the 5% level in the BAO range, Halofit performs much worse with increasing redshift and increasing k, as illustrated in figure 5.3. At $z = 0$ the Cosmic Emulator shows a flat error curve for all scales up to the Nyquist frequency, below the error estimated for Halofit. A comparison between power spectrum estimators and N-body simulations has been performed also in (**fosalba_mice_2013**), where they found similar results: Halofit and the Cosmic Emulator perform similarly in the BAO range, but the errors introduced by Halofit is above the 10% level at scales of around $k \approx 1\text{Mpc}/h$ and $z \gtrsim 1$. Therefore, we use the *FrankenEmu* as the preferred ΛCDM non-linear power spectrum estimator for our forecasting purposes. They claim to be accurate at the 1-3% percent level around the scales of interest and they have performed very careful resolution tests using hundreds of realizations.

Another source of error is the sample variance error of the power spectrum when extracted from the N-body simulation. This depends on the number on the

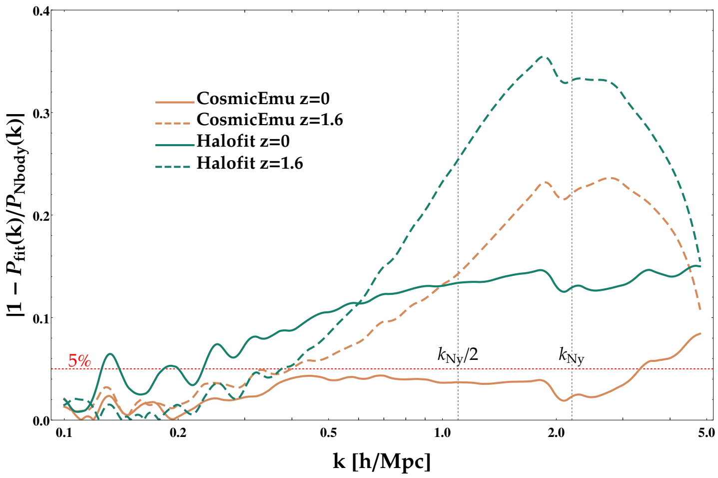


FIGURE 5.3: Relative error of the fitting functions Halofit (green) and CosmicEmulator (orange) (FrankenEmu) with respect to the CoDECS ΛCDM N-body power spectra at two different redshifts (solid lines: $z = 0$, dashed lines: $z = 1.6$). While the relative error with respect to the CosmicEmulator remains below the 5% limit (horizontal red dashed line) for all interesting k-values, the error compared to Halofit is already bigger than 10% for $k \gtrsim 0.5$h/Mpc. Both the error of our simulation with respect to Halofit and to the CosmicEmulator increase as a function of redshift. The change of trend after the Nyquist frequency (marked with the vertical grey line) can be attributed to the use of the folding method for the CoDECS nonlinear power spectra.

number n_{mod} of independent modes available at each wavevector bin in k and its given by **feldman_power_1994** $\sigma_P = \sqrt{\frac{2}{n_{mod}}} P(k)$. In figure 5.4 it can be seen as a blue shaded region marked by a red and blue solid line. For large scales this error is considerable, but there one usually uses just linear power spectra computed by Boltzmann codes, like CAMB or CLASS. At $k = 0.1h/\mathrm{Mpc}$ the binning error is at the percent level and decreases rapidly to negligible values, but then it increases again quickly at $k \sim 2h/\mathrm{Mpc}$ since there the folded mesh for the power spectrum has again only few modes to sample from, as explained in section 5.4.

The intrinsic error of the fitting function is shown in figure 5.4 as a dashed green line and it remains well below the 1% level at all the scales of interest.

We include all errors discussed here in our Fisher forecast analysis. It turns out that they do not affect considerably the results for a survey like Euclid, which will measure such a high number of galaxies, that the sampling of the clustering signal will be not affected by small sources of noise in the data.

5.6.0.1 Including sources of error in the Fisher formalism

In section 5.6 we discussed several sources of error affecting the non-linear power spectrum, the intrinsic errors of the coupled DE fitting functions, mode-binning errors in the N-body power spectrum and the estimation and interpolation error of the ΛCDM non-linear power spectrum obtained from the CosmicEmulator. We will take these sources of error into account in our Fisher forecast analysis by including them as extra noise affecting the observed power spectrum. The term in square brackets in eq. (3.25) corresponds to the inverse of the covariance $C = P(k, z) +$

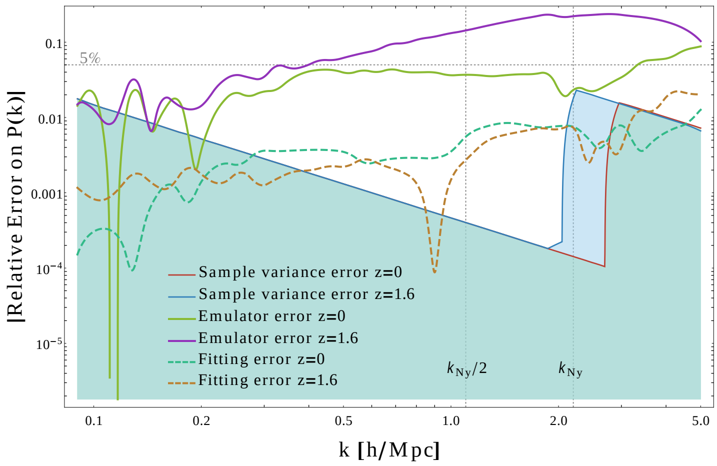


FIGURE 5.4: Different sources of error affecting the nonlinear power spectrum. Each source of error is shown at two different redshifts, $z = 0$ and $z = 1.6$. The sample variance error (red solid line and blue solid line, together with their shaded regions) corresponds to the error induced by the limited number of available k-modes when extracting the power spectrum; it has a sharp increase at the Nyquist frequency, since there the folded mesh has fewer modes available than the top level mesh, see section 5.4 for more details. The fitting error (green and orange dashed lines) corresponds to the intrinsic error of the fitting functions with respect to an N-body simulation using the same parameters, calculated in section 5.5. The emulator error (green and purple thick lines) is the error of the Cosmic Emulator compared to a ΛCDM N-body simulation from CoDECS (these two lines correspond to the orange lines of figure 5.3). The relative error increases with redshift and scale and reaches more than 15% at the Nyquist frequency for the highest redshift. The vertical grey-dashed lines mark the scales $k_{Ny}/2$ and k_{Ny}.

$n(z)^{-1}$. The "noise term" $n(z)^{-1}$ is the number density of sampling points for the matter power spectrum (galaxies in a survey), which gives us an estimate of the signal-to-noise ratio we can expect from the forecast: for a higher number density, the power spectrum is better sampled and more information can be extracted from it. In order to take into account the theoretical and numerical errors on $P(k, z)$, we decrease $n(z)$ by a factor that contains the relative errors on $P(k, z)$. In eq. (3.25), instead of $n(z)$, we then have an "effective" number density:

$$n_{eff}(k, z) = n(z)/(1 + n(z)\tilde{\sigma}_p(k, z)) \tag{5.7}$$

The term $\tilde{\sigma}_p(k, z)$ is a scale- and redshift-dependent term which is the square root of the sum of the relative errors squared. We take into account all error sources which affect $P(k, z)$ due to different reasons, as explained in section 5.6. One of them is the difference between our power spectrum estimator and the N-body simulation: $\sigma_p(k, z) = (P_{numerical}(k, z) - P_{true}(k, z))/P_{true}(k, z)$. If $\tilde{\sigma}_p(k, z) = 0$ or it is negligible, the effective number density will be the observed one $n_{eff}(z) = n(z)$; otherwise, $n_{eff}(z) < n(z)$, the effective number of sampling points being reduced, together with the amount of information one can extract from the power spectrum. As long as $n(z)P(k, z) \gg 1$ for all z and k, the sampling will be always good enough to extract cosmological information from the power spectrum even in the presence of noise. For the specifications used in this work (see table 5.5 below), $n(z)P(k,z)$ is larger than 1 in all scales of interest and therefore the theoretical and numerical error on $P(k,z)$ does not have such a considerable effect on the parameter estimation, as one would expect naively. We test the inclusion of the effective number density $n_{eff}(z)$ on the Fisher forecast analysis and we find that the relative 1-σ marginalised errors on each cosmological parameter are between 8% and 15% higher when using the estimated uncertainties on the power spectrum.

5.6.0.2 Systematic bias due to uncertainty in the power spectrum

In section 3.4.7 we have detailed the theory on how to calculate the systematic error bias on our cosmological parameters, when taking into account theoretical uncertainties on the determination of the non-linear power spectrum.

There we have seen in eq. (3.47) that the systematic-bias Fisher matrix is proportional to the derivative of the power spectrum with respect to the systematic parameter ψ. For our case, we have then:

$$\partial_\psi P_\psi = -P_{true} + P_{num} = P_{true}\sigma_p(k,z) \tag{5.8}$$

and we just need to perform eqn. 3.47 with eqn. 3.44 in order to obtain the systematic biases on the cosmological parameters.

In the following table 5.4 we present the results on the systematic bias, for a standard ΛCDM forecast, for different maximum k coverages, up to a maximum k of 1.1h/Mpc. We will regard as a "true" power spectrum P_{true}, the one obtained by the Cosmic (Franken) Emulator **heitmann_coyote_2014** since they have performed a careful analysis of resolution effects using a large set of simulations and claim to be accurate for $k < 1.0\,$h/Mpc at the 1% percent level compared to state of the art N-body simulations. On the other hand, the numerical "biased" power spectrum P_{num}, is the one obtained from the CoDECS ΛCDM run, which consists on only one realization. We have left out the CDE coupling parameter β, since in that case we do not have any other prediction in the non-linear regime to compare with.

Parameter	h	$\ln \mathcal{A}_s$	n_s	ω_b	ω_c
fiducial	0.7036	2.42	0.966	0.04503	0.2256
$\mathbf{k_{max} = k_{Ny}/2 = 1.1h/Mpc}$ syst. bias with $\delta\psi = 1$	-0.0016	-0.15	0.061	0.0028	-0.0031
$\mathbf{k_{max} = 0.35h/Mpc}$					
syst. bias with $\delta\psi = 1$	-0.0094	-0.11	0.045	0.0021	-0.0039
$\mathbf{k_{max} = 0.15h/Mpc}$					
syst. bias with $\delta\psi = 1$	-0.0032	-0.04	0.018	0.00026	-0.0024

TABLE 5.4: Systematic bias on the ΛCDM cosmological parameters evaluated at our fiducial model. The systematic bias is higher when probing smaller scales, where uncertainties in the non-linear power spectrum are higher. This highlights the importance of modelling correctly the non-linear power spectrum in order to analyze data. Using the wrong non-linear power spectrum produces statistical errors which are larger or of the same order as the statistical results.

The results of table 5.4 show how important it is to model accurately the non-linear power spectrum in order to make forecasts and to analyze the upcoming data of large scale structure surveys like Euclid **amendola_cosmology_2012-short**

The systematic errors on the cosmological parameters can be bigger than the statistical errors (compare to table 5.6 and figure 5.5 in the results section below). This is the case if, as in our example scenario, we would use a non-linear power spectrum that is inaccurate by about 10-15% in the non-linear regime at higher redshifts (which was shown in figure 5.3). Therefore, it is well justified to choose for our Fisher forecasts the Cosmic Emulator as the "true" ΛCDM non-linear power spectrum estimator, since this is the most accurate predictor up to date. It would still be interesting to know how robust is the signal of the coupling parameter β with respect to changes in the other parameters or in the ΛCDM non-linear prediction, but as long as we do not have better and faster semianalytic methods applicable to general models of dark energy, the estimation of systematic biases of extra parameters is an impossible task.

5.6.0.3 Choice of the k_{max} integration limit in the Fisher formalism

There are at least three ways of setting the maximum k mode used in the Fisher forecast integration (eqn. 3.25). One common choice is to set a hard cut in k at all redshifts, and depending if one wants to include or not non-linear effects, this cut can be taken at linear scales, smaller than $k = 0.1h/\text{Mpc}$, or at non-linear scales $k > 0.1h/\text{Mpc}$. In the latter case if one is using a power spectrum calculated in linear theory, one needs to to use some Lagrangian damping correction, as introduced originally in **seo2007improved**; **seo_nonlinear_2008** in order to take into account broadening effects on the BAO peak induced by the non-linear evolution of densities and velocities. Another option to cut off the power spectrum is to demand that the variance $\sigma_8(z; k)$ stays below a specified value at each redshift, therefore implicitly changing the cutoff in k as a function of z. An usual choice for this is $\sigma_8(z; k_{cut}) = 0.25$, as was done previously in **amendola_testing_2012** Since we are assuming to have a knowledge of the non-linear power spectrum up to the Nyquist frequency, we implement for our forecasts a hard-cut method, at k_{Ny} and at $k_{Ny}/2$, without the need of any Lagragian damping correction. However, to be conservative, we cite as our main results the ones in which the cut is performed at $k_{Ny}/2$ in order to eliminate any possible unknown contribution from the numerical high-frequency noise entering the estimation of $P(k)$ at the Nyquist frequency (see e.g. **colombi_accurate_2009**; **Fosalba2013** for similar prescriptions on where to cut the non-linear power spectrum).

5.6.1 Adding non-linear corrections to the power spectrum

We are interested in a Fisher forecast that includes information from non-linearities. In this case we cannot separate the power spectrum, into a power spectrum at redshift zero multiplied by the square of the normalized growth factor, but we need to evaluate directly the non-linear power spectra $P(k, z)$ at different redshifts. A full non-linear correction for the redshift space distortions would be also desirable, but since the modelling of that effect is yet to be understood in general cases, we use as a first approximation an exponential damping of the form $\exp(-k^2\mu^2\sigma_v^2)$, where σ_v is the pairwise peculiar velocity of galaxies induced by non-linearities in the matter and velocity power spectrum. This is the first term of a set of corrections that can be applied to the Kaiser formula (**kaiser_clustering_1987**) for the clustering in redshift space (see e.g. **de_la_torre_modelling_2012 scoccimarro_redshift-space_2004 taruya_baryon_2010wang_toward_2012**). We use a value of $\sigma_v = 300\text{km/s}$ which is an approximate and conservative value based on the estimations by (**de_la_torre_modelling_2012**), in which the authors test a variety of redshift-space distortion models.

As already mentioned, the damping term, introduced by (**seo2007improved**), which should correct the linear $P(k)$, especially the position of the BAO peaks, for non-linearities, is not included here, since we assume that we have a complete model of the non-linear power spectrum and therefore all possible corrections are already included in our fitting functions and power spectrum emulators.

In order to implement our model of coupled Dark Energy, we use for $H(z)$ and $D(z)$ tables precomputed using a modified version of CAMB, that includes the exponential and inverse-power law potentials for coupled dark energy **amendola_testing_2012**; **pettorino_testing_2013** We calculate these tables for each of the parameters $\theta_i \pm \epsilon$, where $\epsilon = 0.03$. We do the same for the linear perturbation quantity $G(z)$, whose logarithmic derivative with respect to $\ln(a)$, known as the growth rate $f(z)$ enters the redshift space distortion term in **??**. The background quantities $H(z)$ and $D(z)$ are important in the Fisher forecast for the Alcock-Paczynski geometrical term, which we take into account in the observed power spectrum.

The full non-linear power spectrum we use in our method is then obtained as follows.

- At linear scales $k \lesssim 0.1h/\text{Mpc}$ the linear power spectrum is obtained from our modified version of CAMB which includes the effect of the DM-DE coupling.
- At non-linear scales, $k \gtrsim 0.1h/\text{Mpc}$, we use a combination of the power spectrum calculated with the cosmic emulator *FrankenEmu* and our fitting formulas for CDE that account for the non-linear dynamics in presence of a fifth force.
- The matching at $k \approx 0.1h/\text{Mpc}$ is performed using different interpolation methods, either averaging on both sides and smoothing out or allowing for a small discontinuity. The matching point is left out of the Fisher integral and we check for this effect, showing that the different methods have a negligible effect (less than 2 percent) on the final absolute errors on the parameters.
- In order to be conservative in terms of numerical errors and noise, we cut the power spectra at half of the Nyquist frequency (we also test and compare with a cutoff at k_{Ny}) and we include all sources of errors specified in section 5.6.0.1 , as effective noise terms into our Fisher estimation.

5.7 Results

We now present results for galaxy clustering and weak lensing, using the specifications for a Euclid-like survey as described in tables 5.5 and 5.7 below. We use the Fisher formalism described in section **??** to forecast the errors in the cosmological parameters, using information from the non-linear power spectrum for coupled Dark Energy models, as obtained with the procedure described in section 5.4 together with the fitting functions obtained in section 5.5. To make our estimation more realistic, we also take into account all sources of error and systematics for the power spectrum which were discussed in section 5.6. As a way of testing our improvement on the parameter estimation, we also perform two extra Fisher forecasts, first using only linear power spectra for the CDE model and then correcting these linear $P(k)$ with the latest version of Halofit **takahashi_revising_2012** which was designed for ΛCDM only.

Parameter	Value	Description
A_{survey}	15000 deg^2	Survey area in the sky
σ_z	0.001	Photometric redshift error
σ_v	300 km/s	Fiducial pairwise velocity dispersion
Δz	0.2	Redshift bin width
$\{z_{min}, z_{max}\}$	{0.6, 2.0}	Min. and max. limits for redshift bins

TABLE 5.5: Specifications for the Fisher Matrix of an Euclid-like galaxy clustering survey.

The fiducial parameters are $\omega_c = 0.1117$, $\omega_b = 0.0223$, $n_s = 0.966$, $\log \mathcal{A}_s = -19.8395$, $h = 0.7036$, $\beta^2 = 0.0025$, which are consistent with WMAP7 results (see table 5.2). The fiducial values for the galaxy bias $b(z)$ used for the GC probe are taken from the Euclid specifications (see **amendola_cosmology_2012-short**; **laureijs_euclid_2011**; **orsi2010probing**) For our final results, we convert these parameters into the set $p_i = \{\beta^2, h, 10^9\mathcal{A}_s, n_s, \Omega_c, \Omega_b\}$ marginalizing over the bias $b(z)$ (for the GC case) and using a Jacobian transformation to convert into the new set of parameters, which is allowed by the Fisher matrix formalism. We choose to forecast the error on the square of the coupling parameter β^2, because this is the quantity entering the modified gravitational Newton constant G_{eff} in the limit of linear perturbations (cf. eq. **??**), therefore giving the strength of the "fifth force". The corresponding fiducial value for the coupling, $\beta = 0.05$, is still compatible with recent limits set by analyzing the data from the Planck Satellite (see **pettorino_testing_2013**; **planckcollaboration_planck2015_2015**).

5.7.1 Galaxy clustering

Table 5.5 shows the specifications of a Euclid-like survey, which are used in our Fisher forecast. While Euclid specifications use 14 redshift bins of a width $\Delta z = 0.1$ (see **amendola_cosmology_2012-short** table 3 in that work), we use only 6 bins of a width of 0.2. We check in the case of ΛCDM, that this re-binning (done using the specified number of galaxies and the corresponding comoving volumes in our cosmology) has a very small effect (of a few percent) on the estimation of the 1-σ errors on the parameters.

In table 5.6 we show the fully marginalized 1-σ errors on the final cosmological parameters $p_i = \{\beta^2, h, 10^9\mathcal{A}_s, n_s, \Omega_c, \Omega_b\}$, performing the non-linear power spectrum cut-off at two different scales: $k_{max} = k_{Ny}/2$ and $k_{max} = k_{Ny}$. As explained above, we take as a reference result the one corresponding to a cutoff at $k_{Ny}/2$. The gain in constraining the β^2 parameter when going from $k_{Ny}/2$ to k_{Ny} is of two percent points in the relative errors, while for the other cosmological parameters the improvement is negligible.

5.7.1.1 Variation of the k_{max} integration limit

We now test the gain in information obtained by including progressively more non-linear wavemodes k into the Fisher integration. We perform the same Fisher forecast, each time increasing the maximum mode k_{max} at which the integration is cut off. In figure 5.5 we show how the 1-σ fully marginalized error on the cosmological parameters p_i changes with an increase of k_{max}. The error decreases steadily with an increase of k_{max}, where the biggest gain is achieved when going from linear ($k \approx 0.1h$/Mpc) to mildly non-linear ($k \approx 0.3h$/Mpc) scales. For parameters

Parameter	β^2	h	$10^9\mathcal{A}_s$	n_s	Ω_b	Ω_c
fiducial	0.0025	0.7036	2.42	0.966	0.04503	0.2256
$\mathbf{k_{max} = k_{Ny}/2}$						
abs. error	0.000346	0.00160	0.01855	0.00267	0.00084	0.00088
relative error	14%	0.23%	0.77%	0.28%	1.9%	0.39%
$\mathbf{k_{max} = k_{Ny}}$						
abs. error	0.000305	0.00151	0.01808	0.00249	0.00081	0.00085
relative error	12%	0.22%	0.75%	0.26%	1.8%	0.38%

TABLE 5.6: 1-σ fully marginalized errors on the cosmological parameters of the model for a galaxy clustering Fisher forecast. Two cases are presented, a hard cutoff of the power spectrum at $k_{max} = k_{Ny}/2$, and a hard cut at k_{Ny}.

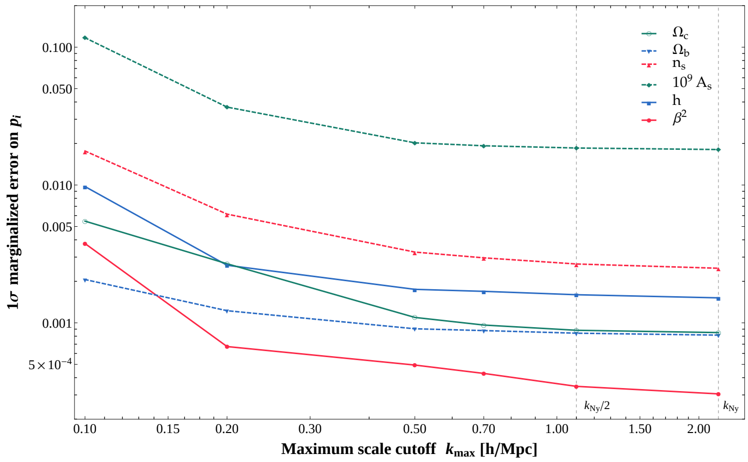


FIGURE 5.5: Change of the 1-σ fully marginalized error on the set of cosmological parameters p_i for a galaxy clustering Fisher forecast, as a function of the maximum mode k_{max} used as a cutoff in the Fisher matrix integration. When increasing the maximum k_{max} the errors on the parameters get steadily smaller, especially when going from linear ($k \approx 0.1h$/Mpc) to mildly non-linear ($k \approx 0.3h$/Mpc) scales. The vertical dashed grey lines, mark the half and the full Nyquist frequencies.

like h and Ω_b, an approximate plateau is reached already before $k_{max} \approx 1.0$, while for β^2 there is still a considerable gain in parameter constraints when going into smaller scales, even beyond $k_{Ny}/2$ (consistent with table 5.6 above). This happens, qualitatively, because at small scales we have a well-defined characteristic signal coming from non-linear interactions including the fifth force which itself involves the β coupling, while the information on the backgorund cosmological parameters gets washed out (c.f. section 5.5.3 above).

5.7.2 Weak lensing

Table 5.7 show the specifications for a weak lensing probe in an Euclid-like survey. The redshift bins are chosen in such a way that they contain an equal number of galaxies (equipopulated bins). The bins are then given by:

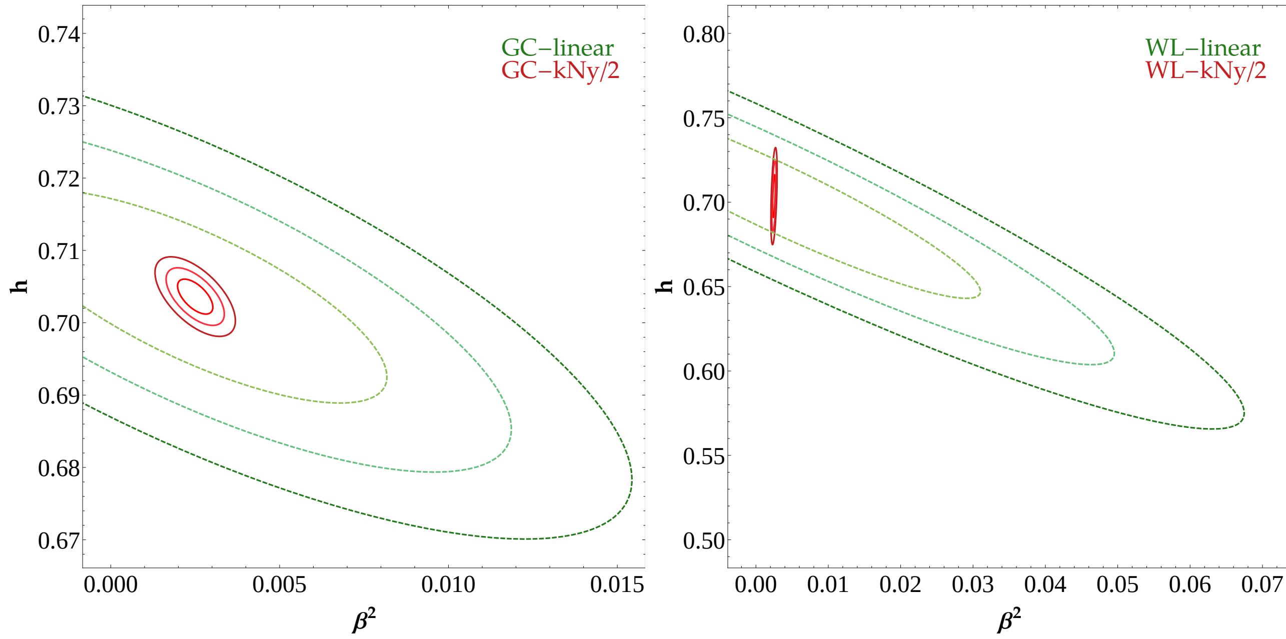


FIGURE 5.6: Marginalized confidence contour regions 1,2,3-σ for β^2 and h. The plots correspond to a comparison between linear (green-dashed lines) and non-linear (red-solid lines) scales for GC (left) and WL (right) using the scale cutoff at $k_{Ny}/2$.

Parameter	Value	Description
f_{sky}	0.364	Observed sky fraction
γ_{int}	0.22	Intrinsic alignment
n_θ	30	Galaxy density per $arcmin^2$
σ_{pz}	0.05	Photometric redshift error

TABLE 5.7: Specifications for the Fisher Matrix of an Euclid-like weak lensing survey.

$$n_i(z) = \frac{1}{2} n(z) \left[\mathrm{Erf}\left(\frac{\tilde{z}_{i+1} - z}{\sigma_{pz}\sqrt{2}} \right) - \mathrm{Erf}\left(\frac{\tilde{z}_i - z}{\sigma_{pz}\sqrt{2}} \right) \right] \tag{5.9}$$

where $\tilde{z}_i$ are the values of the bin intervals in the range $z_{range} = 0.5 \le z \le 3$ chosen such that for each interval the integral over the galaxy distribution function $n(z)$ (eqn. 3.31) is equal. The resulting peaks of the bins and their full width at half maximum are specified in table 5.8.

Analogously to the galaxy clustering case, we show in table 5.9 the fully marginalised 1-σ errors on the parameters p_i. The sum over multipoles ℓ in eq. (3.32) is performed from $\ell_{min} = 5$ up to a maximum of $\ell_{max} = 20000$, but as explained in section 4.4.2, we perform a cutoff at each redshift bin, so that no scales in the non-linear power spectrum beyond the half of the Nyquist (for our reference case) or beyond the Nyquist frequency (for our second case) contribute to the WL signal. The values of these cutoffs, ℓ_{cut} for the two different cases $k_1 = k_{Ny}/2$ and $k_2 = k_{Ny}$ are listed in table 5.8. In contrast to the GC case, going from $k_{Ny}/2$ to k_{Ny} in a WL survey does bring a noticeable improvement on the estimation of parameters.

The dependence of the 1-σ fully marginalized error on each parameter p_i with

Parameter	Value	Description
$\mathcal{N}$	6	Number of redshift bins
z_{peak}	$\{0.59, 0.75, 0.90, 1.06, 1.28, 1.57\}$	z-position of peak of the bin
w_z	$\{0.22, 0.23, 0.25, 0.27, 0.32, 0.51\}$	full width at half maximum of the peak
$\ell_{cut,\,k_1}$	$\{1686, 2070, 2410, 2753, 3155, 4344\}$	cutoff in multipole ℓ at the center of each bin for k_{max} =
$\ell_{cut,\,k_2}$	$\{3372, 4141, 4820, 5506, 6311, 8689\}$	cutoff in multipole ℓ at the center of each bin for k_{max}
z_{range}	$0.5 \leq z \leq 3$	Total range in redshift of each bin

TABLE 5.8: Redshift bins specifications for an Euclid-like weak lensing survey using equipopulated redshift bins in the range $0 \leq z \leq 3$ and the corresponding values for the cutoff applied in the multipoles ℓ at each redshift bin, for two different cases: scales larger than $k_1 = k_{Ny}/2$ and scales larger than $k_2 = k_{Ny}$.

Parameter	β^2	h	$10^9 A_s$	n_s	Ω_b	Ω_c
fiducial	0.0025	0.7036	2.42	0.966	0.04503	0.2256
$\boldsymbol{\ell_{cut,\,k_{Ny}/2}}$						
abs. error	0.000125	0.00835	0.112	0.0105	0.0032	0.0046
relative errror	5.0%	1.2%	4.6%	1.1%	7.1%	2.0%
$\boldsymbol{\ell_{cut,\,k_{Ny}}}$						
abs. error	0.000097	0.0068	0.058	0.0085	0.0022	0.0032
relative error	3.9%	0.97%	2.4%	0.88%	4.9%	1.4%

TABLE 5.9: 1σ fully marginalized errors on the cosmological parameters of the model for a weak lensing Fisher forecast. Two cases are presented, a redshift-dependent cutoff $\ell_{cut,\,k_{Ny}/2}$ and a cutoff $\ell_{cut,\,k_{Ny}}$, corresponding to cutting off the non-linear power spectrum at the half and at the full Nyquist frequency (analogously to the GC case) as explained in the main text. As opposed to the GC case, going into smaller scales in a WL survey does bring a noticeable improvement.

respect to ℓ_{max} is shown in figure 5.7. When using just linear power spectra information, the error on the parameters does not improve if one increases the scale ℓ_{max}, while in the case where non-linear information is used, increasing the maximum multipole ℓ_{max} improves considerably the 1-σ error on the parameters, especially on the coupling β^2 and on the initial amplitude of scalar fluctuations $10^9 A_s$. This is due to the fact that the extra signal on the coupling coming from the non-linear part of the power spectrum, the so called "bump", greatly enhances the constraints on the parameter estimation. The double lines corresponding to each parameter in figure 5.7, show how the error estimation is changed if scales up to k_{Ny} are included (lower line) compared to the upper line where only scales up to $k_{Ny}/2$ contribute. At small ℓ both lines are on top of each other and only start diverging at around $\ell = 2000$, when the extra amount of information contained in highly non-linear scales starts becoming important. The most significant gains occur again on the parameters β^2 and $10^9 A_s$.

5.7.3 Combined results

In figure 5.8 we show the 1-, 2- and 3-σ confidence contours from the Fisher forecast for WL and GC. These confidence regions for each pair of parameters are obtained

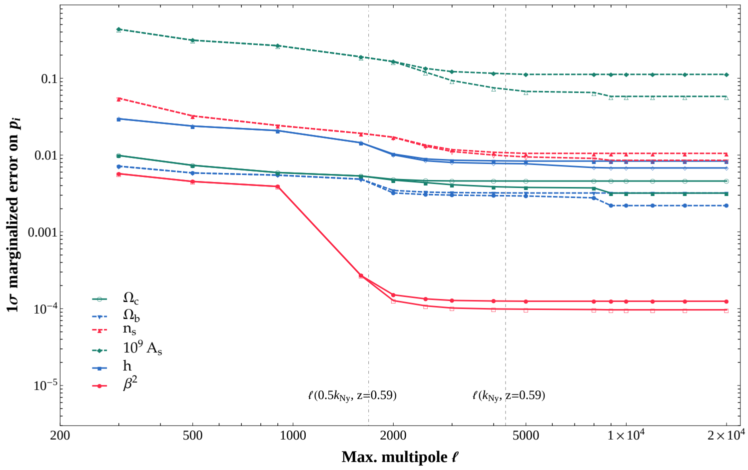
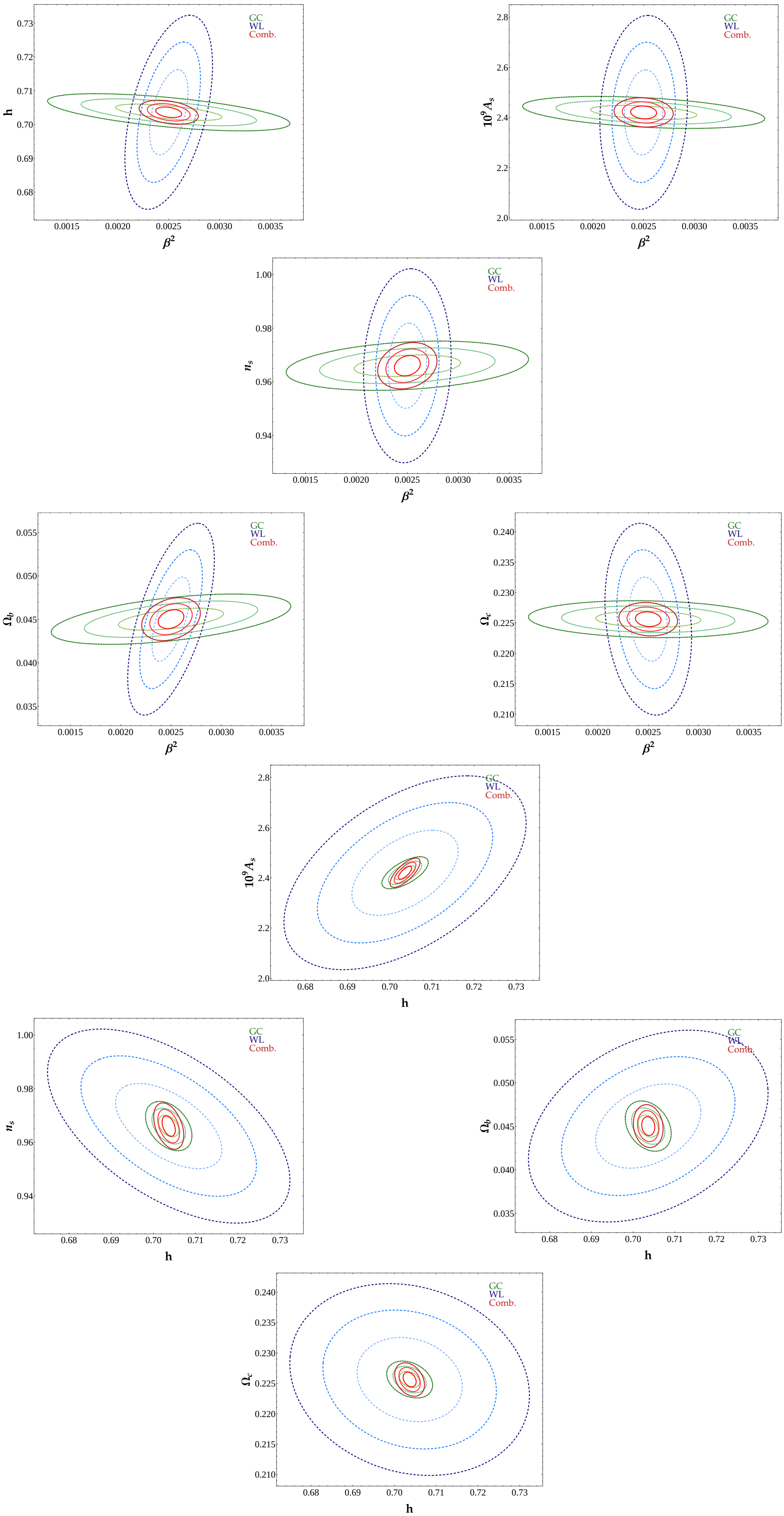


FIGURE 5.7: Variation in the 1-σ fully marginalized errors for each cosmological parameter p_i as a function of the maximum multipole used in the weak lensing Fisher forecast. Increasing the observed multipole range improves considerably the determination of a parameter, especially on the coupling β^2 (red solid line) and the initial amplitude of perturbations $10^9 A_s$ (dashed green line). The double lines corresponding to each parameter, show how the error changes if a cut in the matter power spectrum $P(k)$ is performed at k_{Ny} (lower line) instead of a cut at $k_{Ny}/2$ (upper line, respectively). The vertical dashed grey lines mark the ℓ_{cut} at the peak of the first redshift bins for the cases $\ell_{cut,\, k_{Ny}/2}$ and $\ell_{cut,\, k_{Ny}}$.

after marginalizing over all the other parameters. As it can be seen, some degeneracies are broken when combining the confidence ellipses from two different observations, for example in the case of the plane Ω_b, β^2. Other parameter combinations, as n_s, β^2, show the same orientation of the ellipses for WL and GC, so that the combination of both probes does not help to disentangle the degeneracies. While GC constraints much better the usual parameters $\{h,\, 10^9\mathcal{A}_s,\, n_s,\, \Omega_c,\, \Omega_b\}$, WL constrains the coupling parameter β^2 much better which can be seen in the vertical orientation of the ellipses that correspond to β^2. Therefore combining the observations on GC and WL, as a future survey like Euclid will do, is a powerful way of constraining degenerate parameters in cosmology.

In table 5.10, we cite the 1-σ fully-marginalised errors on the parameters p_i for three different cases: a) using only linear CDE power spectra computed from our modified version of CAMB; b) applying a non-linear correction to these linear CDE power spectra using the latest version of Halofit from **takahashi_revising_2012** which was designed for ΛCDM-only; c) using the full coupled DE non-linear power spectra computed with our fitting functions following the procedure explained in section 5.6.1.

This shows the value of using the N-body-calibrated fitting functions on the coupling β^2. Using the proper β-dependent non-linear correction instead of the standard Halofit correction, the constraints on β^2 improve by more than an order of magnitude for WL and by a factor of order three for GC. This makes very clear the importance of applying non-linear corrections that depend on the parameter to be tested.


GC
WL
Comb.
h
β^2
$10^9 A_s$
n_s
Ω_b
Ω_c

Parameter	β^2	h	$10^9\mathcal{A}_s$	n_s	Ω_b	Ω_c
fiducial	0.0025	0.7036	2.42	0.966	0.04503	0.2256
WL: 1-σ abs. error, using:						
linear CDE	0.0189	0.040	0.221	0.0139	0.0062	0.0127
linear CDE+Halofit	0.0184	0.044	0.256	0.0109	0.0066	0.0079
non-linear CDE fitting functions	0.000125	0.00835	0.112	0.0105	0.0032	0.0046
GC: 1-σ abs. error, using:						
linear CDE	0.0038	0.0097	0.117	0.0176	0.0021	0.0055
linear CDE+Halofit	0.0011	0.0029	0.024	0.0023	0.0007	0.0006
non-linear CDE fitting functions	0.00035	0.0016	0.018	0.0027	0.0008	0.0009
comb. WL+GC: 1-σ error, using:						
non-linear CDE fitting functions (abs.)	0.000084	0.0010	0.0169	0.00251	0.00072	0.00080
non-linear CDE fitting functions (rel.)	3.4%	0.16%	0.7%	0.26%	1.6%	0.35%

TABLE 5.10: 1-σ fully marginalized errors on the cosmological parameters for WL, GC and the combined Fisher matrix WL+GC, using three different power spectra. **Linear CDE**: Using only information from the linear power spectrum for the CDE model up to a scale of $k = 0.1h/\text{Mpc}$. **Linear CDE+Halofit:** Using the linear power spectrum for CDE plus a non-linear correction using the latest Halofit from **takahashi_revising_2012 Non-linear CDE fitting functions:** Using the fully non-linear power spectra for CDE obtained from the fitting functions and the emulator as explained in 5.5, which we regard as the most reliable description of the model in this range of scales. In all these cases we are using our reference Fisher forecasts corresponding to the cutoff at $k_{Ny}/2$.

5.8 Conclusions Fitting Functions

The goal of this paper is to exploit the cosmological information contained in the non-linear regime in order to improve parameter estimation from future large-scale observations. The main obstacle along this road is that we have accurate non-linear corrections for the matter power spectrum only for ΛCDM and a few other relatively simple variants, but not for the large variety of modifed gravity models that have been proposed in recent years.

The first part of this paper has then been devoted to the task of finding corrections to the linear power spectrum in the range of $k \approx 0.1 - 1h/\text{Mpc}$ for a selected class of modified gravity models, namely coupled dark energy. This model is indeed one of the simplest possible extensions of Einstein's gravity and depends entirely on a single parameter, the coupling constant β (in addition to the standard ones). Employing the CoDECS suite of simulations (**baldi_codecs_2012**) we build different fitting function models, such that when multiplied by the ΛCDM non-linear power spectrum (we use the estimator provided in ref. (**heitmann_coyote_2014**)) they reproduce the N-body results to an accuracy of 1%, for scales k between 0.1 and 5 h/Mpc and a range in z, between 0 and 1.8. To achieve this accuracy in the fitting functions we need to perform a careful extraction and interpolation of the power spectrum from the simulation mesh.

The accurate fitting functions have been then employed to extend the regime of validity of the forecasts for future experiments. We focused on a Euclid-like survey that includes weak lensing and redshift-space distortions (galaxy clustering) and predicted the constraints in the cosmological parameters, with particular emphasis on the dark matter-dark energy coupling β. We find that β is better constrained by weak lensing than by galaxy clustering (contrary to all the other standard parameters). We find that the extension into non-linear scales improves the constraints by more than an order of magnitude compared to previous results using only linear power spectra, but also by more than an order of magnitude in WL and a factor of three in GC compared to using a wrong ΛCDM Halofit non-linear correction. We also show that using the wrong non-linear power spectrum, can bias systematically the estimation of errors on the cosmological parameters, yielding systematic errors of the same order of magnitude as the statistical ones. This makes very clear that it is important to include the proper non-linear corrections to the parameter to be tested, especially for models beyond $\Lambda\mathrm{CDM}$ in which the small-scale gravitational dynamics are modified.

To make our forecast more realistic, we take into account all known sources of error entering the estimation of the power spectra and the fitting functions in the way of a reduced effective number density of galaxies and then perform a conservative cut of the power spectrum at half of the simulation Nyquist frequency, to avoid other sources of unknown numerical noise affecting the results. In the case of GC we include also a first approximation to the correction to redshift space distortions, caused by peculiar pairwise velocities at non-linear scales. We find that a space probe like Euclid will be able to constrain the coupling parameter β^2 around the fiducial value 0.0025 at 1-σ with a relative accuracy of 14% when using weak lensing alone, 5% when using only galaxy clustering and at 3.4% when combining both probes.

It is interesting to note that the most stringest constraint we obtain amounts to $\Delta\beta^2 \approx 8 \cdot 10^{-5}$; this level of precision on the dark matter-dark energy coupling is not far from the current best limits reached with Solar System observations on a coupling to baryons **Agashe:2014kda** which can be translated in our notation as $\beta^2 \leq 2 \cdot 10^{-5}$ at 1-σ.

Chapter 6

The non-linear evolution of matter perturbations

6.1 Standard perturbation theory and higher orders

- Vlasov-Poisson system
- System of equations in Fourier space
- Velocity divergence and vorticity
- Stress tensor and higher order hierarchy terms
- Field variable doublet and 1-loop expansion

6.2 The non-linear standard equations in Fourier space

For studying large scale structure formation we will treat the Dark Matter distribution as a perfect fluid coupled to gravity, described by the continuity, Euler and Poisson's equations. This amounts to neglecting higher moments of the Vlasov equation, such as the stress tensor and velocity dispersions **bernardeau_large-scale_2001** This is equivalent to the single stream approximation and is one of the main limitations of Eulerian perturbation theory, since the theory breaks down as soon as shell crossing and multi-streaming start being important.

In the Einstein frame we have the following fluid equations for a general modification of gravity or a coupling to other scalar degrees of freedom **pietroni_flowing_2008**

$$\frac{\partial \delta_c}{\partial \tau} + \nabla \cdot [(1+\delta_c)\,\mathbf{v}] = 0 \tag{6.1}$$

$$\frac{\partial \mathbf{v}}{\partial \tau} + \mathcal{H}(\mathbf{v} + [\mathcal{A}\mathbf{v}]) + (\mathbf{v}\cdot\nabla)\,\mathbf{v} = -\nabla\Psi \tag{6.2}$$

$$\nabla^2\Psi = \frac{3}{2}\mathcal{H}^2\Omega_c(\tau)\,(\delta_c + [\mathcal{B}\delta_c]) \tag{6.3}$$

Here τ is the conformal time,$\delta_c(\mathbf{x},\tau)$ and $\mathbf{v}(\mathbf{x},\tau)$ are respectively the matter density contrast and the peculiar velocity, $\mathcal{H} = \mathrm{d}\log a/\mathrm{d}\tau$ is the conformal Hubble parameter, $\Psi(\mathbf{x},\tau)$ the gravitational potential and the functions$\mathcal{A}(\mathbf{x},\tau)$ and $\mathcal{B}(\mathbf{x},\tau)$ are general functions of space and time that parametrize different cosmologies, when particles' geodesics are modified, i.e. the Poisson and Euler equations. This can happen due to couplings with a scalar field or more general modifications of gravity (MG), such as any Horndeski theory or massive bigravity theories **amendola_observables_2012** However, in the Jordan frame (where most MG theories are formulated) the Euler equation is not modified and there is only a space-time dependent modification to the Poisson equation, which can be expressed as in eqn. 2.6 once we go to Fourier space. In this work $\Omega_m(\tau)\equiv\Omega_c(\tau)$ are the time functions representing the cold dark matter (CDM) or the full matter density of the Universe, since we are not considering baryonic matter nor non-universal couplings.

The last term in the left hand side of eqns. 6.1 and 6.2 is what causes the nonliniearities in the evolution of the fluid. This can be seen in Fourier space as a mode-mode coupling or expressed as a nonlocality of the equations. This is the property that describes how a specific Fourier mode is able to depend on another mode at a different wave vector. We now go to Fourier space and will express equations 6.1-6.3 in a compact form introduced by **scoccimarro_new_2000**

The density is a scalar degree of freedom and the velocity can be decomposed such that:

$$\mathbf{v}(\mathbf{k}) = \mathbf{v}_\theta(\mathbf{k}) + \mathbf{v}_\omega(\mathbf{k}) \tag{6.4}$$

where:

$$\mathbf{k}\cdot\mathbf{v}_\omega(\mathbf{k}) = 0$$
$$\mathbf{k}\times\mathbf{v}_\theta(\mathbf{k}) = 0$$

According to linear theory the vorticity component $\mathbf{v}_\omega(\mathbf{k})$ decays with the expansion of the Universe as a^{-1}, so if we assume non-vortical initial conditions, we can neglect this term and look only at the velocity divergence $\theta = i\mathbf{k}\cdot\mathbf{v}$, which is now a scalar function. Inverting this last relation to get: $\mathbf{v} = -\frac{i\mathbf{k}}{k^2}\theta$, allows us to perform the Fourier transforms explicitly:

$$\begin{aligned} FT\left\{\nabla\cdot(\delta_m\mathbf{v})\right\} &= +i\mathbf{k}\int \mathrm{d}^3q\mathrm{d}^3p\,\delta_D(\mathbf{p}+\mathbf{q}-\mathbf{k})\frac{-i\mathbf{p}}{p^2}\delta_m(\mathbf{q},\tau)\theta(\mathbf{p},\tau) \\ &= \int \mathrm{d}^3q\mathrm{d}^3p\,\delta_D(\mathbf{k}-\mathbf{p}-\mathbf{q})\underbrace{\frac{(\mathbf{p}+\mathbf{q})\cdot\mathbf{p}}{p^2}}_{\alpha(\mathbf{p},\mathbf{q})}\delta_m(\mathbf{q},\tau)\theta(\mathbf{p},\tau) \end{aligned}$$

$$\begin{aligned} FT\left\{\nabla\cdot[(\mathbf{v}\cdot\nabla)\cdot\mathbf{v}]\right\} &= i\mathbf{k}\cdot\int \mathrm{d}^3q\mathrm{d}^3p\,\delta_D(\mathbf{q}+\mathbf{p}-\mathbf{k})\left(\frac{-i\mathbf{q}}{q^2}\cdot i\mathbf{p}\right)\frac{-i\mathbf{p}}{p^2}\theta(\mathbf{q},\tau)\theta(\mathbf{p},\tau) \\ &= \mathbf{k}\cdot\int \mathrm{d}^3q\mathrm{d}^3p\,\delta_D(\mathbf{k}-\mathbf{q}-\mathbf{p})\left(\frac{\mathbf{p}\cdot\mathbf{q}}{p^2q^2}\right)\mathbf{p}\,\theta(\mathbf{q},\tau)\theta(\mathbf{p},\tau) \\ &= \int \mathrm{d}^3q\mathrm{d}^3p\,\delta_D(\mathbf{k}-\mathbf{q}-\mathbf{p})\underbrace{\frac{(\mathbf{p}\cdot\mathbf{q})^2\mathbf{p}\cdot\mathbf{q}}{2p^2q^2}}_{\beta(\mathbf{q},\mathbf{p})}\theta(\mathbf{q},\tau)\theta(\mathbf{p},\tau) \end{aligned}$$

where in the last step we used the symmetry between $\mathbf{p}$ and $\mathbf{q}$. The terms marked with an underbrace, are the ones responsible for the mode-mode coupling:

$$\alpha(\mathbf{q},\mathbf{p}) = \frac{(\mathbf{p}+\mathbf{q})\cdot\mathbf{q}}{q^2} = \alpha(-\mathbf{q},-\mathbf{p});\;\; \beta(\mathbf{q},\mathbf{p}) = \frac{(\mathbf{p}+\mathbf{q})^2\mathbf{p}\cdot\mathbf{q}}{2p^2q^2} = \beta(-\mathbf{q},-\mathbf{p}) \tag{6.5}$$

6.2.0.1 The field notation

In the perturbation theory literature **bernardeau_large-scale_2001; crocce_renormalized_2005; pietroni_flowing_2008** eqns. 6.1-6.3 are written in a more compact way (also called field notation):

$$\partial_\eta \varphi_a(\mathbf{k},\eta) = -\Omega_{ab}(\mathbf{k},\eta)\varphi_b(\mathbf{k},\eta) + e^{\eta}\gamma_{abc}(\mathbf{k},-\mathbf{p},-\mathbf{q})\varphi_b(\mathbf{p},\eta)\varphi_c(\mathbf{q},\eta) \tag{6.6}$$

Where the field doublet is defined as:

$$\varphi_a(\mathbf{k},\eta) = e^{-\eta}\begin{pmatrix} \delta_m(\mathbf{k},\eta) \\ -\theta(\mathbf{k},\eta)/\mathcal{H} \end{pmatrix} \tag{6.7}$$

and we define a new time variable which will prove to be very convenient in this notation

$$\eta \equiv \log \frac{a}{a_{in}}$$

On the l.h.s. of eqn. 6.6 the second term represents all the nonlinearities and non-localities, if we neglect this term we go back to linear perturbation theory. The γ_{abc} functions in this formalism can be understood as interaction vertices and its only non-vanishing components are precisely given by the mode-mode coupling functions 6.5 :

$$\gamma_{121}(\mathbf{k},\mathbf{p},\mathbf{q}) = \frac{1}{2}\delta_D(\mathbf{k}+\mathbf{p}+\mathbf{q})\alpha(\mathbf{p},\mathbf{q}) \qquad \gamma_{121}(\mathbf{k},\mathbf{q},\mathbf{p}) = \gamma_{112}(\mathbf{k},\mathbf{p},\mathbf{q}) \tag{6.8}$$
$$\gamma_{222}(\mathbf{k},\mathbf{p},\mathbf{q}) = \delta_D(\mathbf{k}+\mathbf{p}+\mathbf{q})\beta(\mathbf{p},\mathbf{q})$$

in this notation and throughout this work, an integration over $\mathbf{p},\mathbf{q}$ is understood and $\mathbf{k}$ is always the "external" momentum. The integral signs will be added only when there could be a confusion.

Due to mode-mode coupling, there is a loss of information about the initial conditions in the limit of high-k, since density perturbations get mixed **crocce_memory_2006** This has as a consequence that highly-nonlinear corrections to the power spectrum become cosmology-independent in a rough way.

The information about the cosmological background is contained in:

$$\Omega_{ab}(\mathbf{k},\eta) = \begin{pmatrix} 1 & -1 \\ -\frac{3}{2}\Omega_m(\eta)(1+\mathcal{B}(\mathbf{k},\eta)) & 2+\frac{\mathcal{H}'}{\mathcal{H}}+\mathcal{A}(\mathbf{k},\eta) \end{pmatrix} \tag{6.9}$$

Where we have the usual definitions: $\mathcal{H} = aH,\ a\mathrm{d}\tau = \mathrm{d}t,\ a' = \frac{\mathrm{d}a}{\mathrm{d}\ln a} = a,\ \frac{\mathcal{H}'}{\mathcal{H}} = \frac{H'}{H}+1$, where $' := \frac{\mathrm{d}}{\mathrm{d}\ln a}$ and $N \equiv \ln a$ is what is typically called the e-foldings time.

For details on how to go back to the usual fluid equations 6.1-6.3 in Fourier space starting from the field equation 6.6, see Appendix **??**.

Equation 6.6 is the starting point for all renormalization and resummation methods **crocce_renormalized_2005; bernardeau_evolution_2013; bernardeau_constructing_2012; valageas_matter_2013; anselmi_nonlinear_2012; anselmi_next–leading_2010** The linear part can be easily solved and the function relating the initial primordial density perturbations to the final one, is called the linear propagator and will be studied in section 6.2.1. The nonlinear part, cannot be solved exactly analytically, nor numerically. But a perturbative approach using all tools from Quantum Field Theory can be used to regularize its divergences which are caused by the fact that the density perturbations (which is at the same time the perturbation variable) grows with time and increasing wave vector. For this part we will use the resummation technique of **anselmi_nonlinear_2012** and this will be explained in detail in section 6.3.

Renormalized perturbation theory (RPT) can help in finding the evolution of the nonlinear power spectrum at small scales and late times, but even if we could calculate exactly its result at all loop orders, there are still intrinsic limitations given by the starting equations 6.1-6.3. Apart from neglecting vorticity in the later stages of evolution, the initial equations are derived in the single-stream-approximation. This means that at a single point in space, there can be only one velocity direction. This clearly breaks down in the virialization regime and even before. The RPT approach can be extended and improved by including these other sources of density power into the equations in an effective way, see the discussion in **manzotti_coarse_2014**; **pietroni_coarse-grained_2011** and recent results in the effective field theory of large scale structures **baumann_cosmological_2012**; **pajer_renormalization_2013**; **senatore_ir-resummed_2014**; **carrasco_effective_2012**

6.2.1 Solving for the general scale dependent linear propagator

As was explained above, in eqn. 6.6 the non-linearity of the Vlasov-Poisson system of equations is fully encoded in the vertex γ_{abc} which represents the mode-mode coupling. Without this term, we recover the linear equation:

$$\partial_\eta \varphi_a(\mathbf{k},\eta) = -\Omega_{ab}(\mathbf{k},\eta)\varphi_b(\mathbf{k},\eta) \tag{6.10}$$

which is valid for a fully scale and time dependent Ω_{ab}.

The linear propagator is just the function that connects the initial density perturbations with the final ones, or in other words solves the above equation **crocce_renormalized_2005** If this equation has solutions of the form **pietroni_flowing_2008**

$$\varphi_{sol}(\mathbf{k},\eta) = \begin{pmatrix} 1 \\ f(\mathbf{k},\eta) \end{pmatrix} \varphi(\mathbf{k},\eta)$$

then we can find an equation that describes the evolution of the growth of perturbations.

For the index $a = 1$:

$$\partial_\eta \varphi = -\Omega_{11}\varphi - \Omega_{12} f\varphi = -(\Omega_{11} + \Omega_{12} f)\varphi \tag{6.11}$$

For the index $a = 2$:

$$\begin{aligned} \partial_\eta(f\varphi) &= f\partial_\eta\varphi + \varphi\partial_\eta f = -\Omega_{21}\varphi - \Omega_{22} f\varphi \\ \Rightarrow \partial_\eta f &= \Omega_{12} f^2 + (\Omega_{11} - \Omega_{22}) f - \Omega_{21} \qquad (6.12) \\ \Rightarrow \partial_\eta f &= \Omega_{12}(f - \bar{f}_+)(f - \bar{f}_-) \qquad (6.13) \end{aligned}$$

Equation 6.12 is what we usually know as the growth rate equation for $f = \mathrm{d}\ln D/\mathrm{d}\ln a$, being D the growth factor of density perturbations and in this general case it can have a time and scale dependent solution.

The zeros of eqn. 6.13 are given by:

$$\bar{f}_\pm(\mathbf{k},\eta) = \frac{(\Omega_{22} - \Omega_{11}) \mp \sqrt{(\Omega_{22} - \Omega_{11})^2 + 4\Omega_{21}\Omega_{12}}}{2\Omega_{12}} \tag{6.14}$$

The solution of eqns. 6.11 and 6.13 is then given by:

$$\begin{aligned}
\varphi(\eta) &= e^{-\int_{\eta'}^{\eta}(\Omega_{11}+\Omega_{12}f)\mathrm{d}x}\varphi(\eta') \\
f(\eta)\varphi(\eta) &= e^{-\int_{\eta'}^{\eta}(\Omega_{21}+\Omega_{22}f)\mathrm{d}x}f(\eta')\varphi(\eta') \\
&= e^{-\int_{\eta'}^{\eta}(\Omega_{11}+\Omega_{12}f)\mathrm{d}x}\varphi(\eta')f(\eta')\frac{f(\eta)}{f(\eta')}
\end{aligned}$$

One can identify the basis solutions by setting their initial conditions as:

$$f_{\pm}^{in} = \bar{f}_{\pm}(\mathbf{k}, \eta_i)$$

where η_i is an initial time that can be set at high redshift where the Universe is approximately Einstein-deSitter (E-dS) or strongly matter dominated.

For E-dS, $\Omega_m = 1$ we have very simple background quantities: $\mathcal{H}'\backslash\mathcal{H} = -1/2 - 3w_{eff}/2 = -1/2$, so that the Ω_{ab} from equation 6.9 is simply:

$$\Omega_{ab} = \begin{pmatrix} 1 & -1 \\ -\frac{3}{2} & \frac{3}{2} \end{pmatrix}$$

which gives the following initial conditions for the growing u and decaying v modes:

$$u_a = \begin{pmatrix} 1 \\ f_+^{in} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$v_a = \begin{pmatrix} 1 \\ f_-^{in} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{3}{2} \end{pmatrix}$$

The growing mode will be the mode of interest that we will use in section 6.3, when we want to calculate the evolution of the power spectrum.

The instantaneous projectors on the two basis solutions are defined as:

$$\begin{aligned}
\mathbf{M}^+(\mathbf{k},\eta)\begin{pmatrix} 1 \\ f_+(\mathbf{k},\eta) \end{pmatrix} &= \begin{pmatrix} 1 \\ f_+(\mathbf{k},\eta) \end{pmatrix} \\
\mathbf{M}^+(\mathbf{k},\eta)\begin{pmatrix} 1 \\ f_-(\mathbf{k},\eta) \end{pmatrix} &= 0 \\
\mathbf{M}^-(\mathbf{k},\eta)\begin{pmatrix} 1 \\ f_-(\mathbf{k},\eta) \end{pmatrix} &= \begin{pmatrix} 1 \\ f_-(\mathbf{k},\eta) \end{pmatrix} \\
\mathbf{M}^-(\mathbf{k},\eta)\begin{pmatrix} 1 \\ f_+(\mathbf{k},\eta) \end{pmatrix} &= 0
\end{aligned}$$

The growing projector can be written explicitly by subtracting the decaying projector from the unity matrix:

$$\mathbf{M}^+(\mathbf{k},\eta) = \mathbb{1} - \mathbf{M}^-(\mathbf{k},\eta) = \frac{1}{f_- - f_+}\begin{pmatrix} f_- & -1 \\ f_- f_+ & -f_+ \end{pmatrix}$$

For the Einstein-deSitter case, the projectors do not evolve in time, since u_a and v_a are constant, and they are given by:

$$\mathbf{M}^+ = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix}$$
$$\mathbf{M}^- = \frac{1}{5}\begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}$$

The linear propagator is defined as the operator giving the linear evolution of the field φ_a, or in other words, solves eqn.6.10 :

$$\varphi_a(\mathbf{k},\eta) = g_{ab}(\mathbf{k},\eta,\eta')\varphi_b(\mathbf{k},\eta)$$

and it has to fulfill following properties:

$$\partial_\eta g_{ab}(\mathbf{k},\eta,\eta') = -\Omega_{ac}(\mathbf{k},\eta)\cdot g_{cb}(\mathbf{k},\eta,\eta')$$
$$\lim_{\eta'\to\eta} g_{ab}(\mathbf{k},\eta,\eta') = \mathbb{1}_{ab}$$
$$g_{ab}(\mathbf{k},\eta,\eta')\cdot g_{bc}(\mathbf{k},\eta',\eta'') = g_{ac}(\mathbf{k},\eta,\eta'')$$

Using these properties and the projectors, the propagator can be written in general as:

$$g(\mathbf{k},\eta,\eta') = \Theta(\eta-\eta')\Bigg[e^{-\int_{\eta'}^{\eta}(\Omega_{11}+\Omega_{12}f_+)\mathrm{d}x}\begin{pmatrix} 1 & 0 \\ 0 & \frac{f_+(\mathbf{k},\eta)}{f_+(\mathbf{k},\eta')} \end{pmatrix}\mathbf{M}^+(\mathbf{k},\eta') + e^{-\int_{\eta'}^{\eta}(\Omega_{11}+\Omega_{12}f_-)\mathrm{d}x}\begin{pmatrix} 1 & 0 \\ 0 & \frac{f_-(\mathbf{k},\eta)}{f_-(\mathbf{k},\eta')} \end{pmatrix}\mathbf{M}^-(\mathbf{k},\eta')\Bigg] \tag{6.15}$$

In the E-dS case, this would reduce to simply (since there is no k-dependence in any quantity and the growing mode is constant):

$$g(\eta,\eta') = \Theta(\eta-\eta')\left[\frac{1}{5}\begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}e^{-5/2(\eta-\eta')}\right]$$

Using the same procedure, we will calculate the linear propagator for the Horndeski case, in which the growth factor D is now scale and time dependent. Then the linear propagator can be used to calculate its fully nonlinear renormalized version, which then is a crucial ingredient of the evolution equation in section 6.3.

6.2.2 The linear propagator in the Horndeski case

Using eqn. 2.6 as the modification of the Poisson equation we can write the general Ω_{ab} matrix as:

$$\Omega_{ab}(\mathbf{k},\eta) = \begin{pmatrix} 1 & -1 \\ -\frac{3\mu(k)\Omega_m(\eta)}{2} & \frac{\mathcal{H}'}{\mathcal{H}}+2 \end{pmatrix} \tag{6.16}$$

in this case, the initial conditions from eqn.6.14 are:

$$\begin{aligned} f_-^{in} &= -\frac{1}{2}(\Sigma+\omega) \\ f_+^{in} &= \frac{1}{2}(\Sigma-\omega) \end{aligned} \tag{6.17}$$

where $\omega = 1+\frac{\mathcal{H}'}{\mathcal{H}} = \frac{1}{2}-\frac{3}{2}w_{eff}$ and $\Sigma = \sqrt{6Y\Omega_m+\omega^2}$ are quite general functions of scale and time. Inserting this into eqn.6.15, one can find the most general form of the propagator for the Horndeski theory. However, this is not so practical, since we would have to solve for a scale and time dependent growth rate f and evaluate the $\mathcal{H}$ function at each step.

For models which are close to ΛCDM, it is more convenient to change the time variable from $\eta = \ln\frac{a}{a_{in}}$ to $\mathcal{X} = \ln\frac{D(\tau)}{D(\tau=\tau_{in})}$, where $D(\tau)$ is the growth function usually written as the growing solution of the linear density perturbation equation : $\delta_c(\tau) = D_+(\tau)\delta_c^{in}$. In our case, however, the time variable $\mathcal{X}$ would itself be scale dependent, since the growth rate equation in the Horndeski case is scale dependent, so that $\mathcal{X}(k) = \ln\frac{D(\tau,k)}{D(\tau=\tau_{in},k)}$. Using $\frac{\partial}{\partial\eta} = \frac{\mathrm{d}\ln D}{\mathrm{d}\ln a}\frac{\partial}{\partial\ln D} = f\frac{\partial}{\partial\mathcal{X}}$, we can rewrite eqns.6.11-6.12 as (where $' := \frac{\partial}{\partial\mathcal{X}}$):

$$f(\mathbf{k},\mathcal{X})\delta_c'(\mathbf{k},\mathcal{X}) + \frac{\theta(\mathbf{k},\mathcal{X})}{\mathcal{H}} + \frac{\mathbf{k}\cdot\mathbf{q}}{q^2}\delta_c(\mathbf{q},\mathcal{X})\frac{\theta(\mathbf{p},\mathcal{X})}{\mathcal{H}} = 0 \tag{6.18}$$

$$f(\mathbf{k},\mathcal{X})\frac{\theta(\mathbf{k},\mathcal{X})'}{\mathcal{H}} + \frac{\theta(\mathbf{k},\mathcal{X})}{\mathcal{H}} + \frac{1}{2}\frac{\mathbf{k}^2\mathbf{q}\cdot\mathbf{p}}{q^2p^2}\frac{\theta(\mathbf{q},\mathcal{X})}{\mathcal{H}}\frac{\theta(\mathbf{p},\mathcal{X})}{\mathcal{H}} = -\frac{3}{2}\Omega_m(\mathcal{X})\mu(\mathbf{k})\mathcal{H}^2\delta_c(\mathbf{k},\mathcal{X}) \tag{6.19}$$

The doublet 6.7 can now be redefined as:

$$\tilde{\varphi}_a = \begin{pmatrix} \tilde{\varphi}_1 \\ \tilde{\varphi}_2 \end{pmatrix} = \begin{pmatrix} e^{-\mathcal{X}}\delta_c \\ -e^{-\mathcal{X}}\frac{\theta}{\mathcal{H}f} \end{pmatrix} \tag{6.20}$$

Substituting in the previous equations the values for δ_c, θ and $\delta_c' = e^{\mathcal{X}}(\varphi_1'+\varphi_1)$, $\theta' = -e^{\mathcal{X}}f(\mathcal{X})\mathcal{H}\left(\tilde{\varphi}_2\left(1+\frac{f'(\mathcal{X})}{f(\mathcal{X})}+\frac{\mathcal{H}'}{\mathcal{H}}\right)+\tilde{\varphi}_2'\right)$ we get:

$$\tilde{\varphi}_1' + \tilde{\varphi}_1 - \tilde{\varphi}_2 - \alpha e^{\mathcal{X}}\tilde{\varphi}_1\tilde{\varphi}_2 = 0 \tag{6.21}$$

$$\begin{aligned} -\tilde{\varphi}_2' - \tilde{\varphi}_2\left(1+\frac{f'}{f}+\frac{1}{f}+\frac{\mathcal{H}'}{\mathcal{H}}\right) + \beta e^{\mathcal{X}}\tilde{\varphi}_2\tilde{\varphi}_2 &= -\frac{3}{2}\Omega_m\mu\frac{\tilde{\varphi}_1}{f^2} \\ \Rightarrow -\tilde{\varphi}_2' - \frac{3}{2}\frac{\Omega_m\mu}{f^2}\tilde{\varphi}_2 + \frac{3}{2}\frac{\Omega_m\mu}{f^2}\tilde{\varphi}_1 + \beta e^{\mathcal{X}}\tilde{\varphi}_2\tilde{\varphi}_2 &= 0 \end{aligned} \tag{6.22}$$

where we have omitted for notational simplicity the momentum and time dependence. In the last step we used the growth rate equation $\ddot{\delta}_m + \dot{\delta}_m\left(1+\frac{\dot{\mathcal{H}}}{\mathcal{H}}\right) = \frac{3}{2}\Omega_m\delta_m\mu$, where in this case an overdot represents a derivative with respect to $\eta = \ln\frac{a}{a_{in}}$, since the $'$-symbol is now reserved for the $\mathcal{X}$ time variable.

Comparing eqns. 6.21-6.22 with 6.6, we get the following Ω matrix:

$$\tilde{\Omega}(\mathbf{k},\mathcal{X}) = \begin{pmatrix} 1 & -1 \\ -\frac{3}{2}\frac{\Omega_m(\mathcal{X})}{f_+^2(\mathcal{X})}\mu(k) & \frac{3}{2}\frac{\Omega_m(\mathcal{X})}{f_+^2(\mathcal{X})}\mu(k) \end{pmatrix} \tag{6.23}$$

It is of great importance to note that in the Horndeski case, the quantity $\frac{\Omega_m(\mathcal{X})}{f_+^2(\mathcal{X})}$ cannot be approximated to 1 as it is done in LCDM for all times, since in this case it can have a much greater or lower value during structure formation, but more importantly it is scale-dependent through the scale dependence of $\mathcal{X}(k)$. This can be seen clearly in plot **??**.

6.2.3 Assuming a constant $\mu \neq 1$ for the 1-loop quantities

Assuming a constant μ different from 1 and the approximation that $\frac{\Omega_m(\mathcal{X})}{f_+^2(\mathcal{X})} = 1$ throughout the history of the Universe (which is valid only for E-dS but turns out to be a good approximation (much more accurately than 1%) for the nonlinear ΛCDM power spectrum), we can follow the steps given above in section 6.2.1 and obtain the initial growing and decaying modes:

$$u_a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$v_a = \frac{-2}{3Y}\begin{pmatrix} 1 \\ -\frac{3Y}{2} \end{pmatrix}$$

and with this the projectors:

$$\mathbf{M}^+ = \frac{1}{2+3Y}\begin{pmatrix} 3Y & 2 \\ 3Y & 2 \end{pmatrix}$$
$$\mathbf{M}^- = \frac{1}{2+3Y}\begin{pmatrix} 2 & -2 \\ -3Y & 3Y \end{pmatrix}$$

The propagator would then look like:

$$g(\mathcal{X},\mathcal{X}') = \Theta(\mathcal{X}-\mathcal{X}')\left[\frac{1}{2+3Y}\begin{pmatrix} 3Y & 2 \\ 3Y & 2 \end{pmatrix} + \frac{1}{2+3Y}\begin{pmatrix} 2 & -2 \\ -3Y & 3Y \end{pmatrix} e^{-\frac{(2+3Y)}{2}(\mathcal{X}-\mathcal{X}')}\right] \tag{6.24}$$

For the rest of this work, eqn. 6.24 will be the linear propagator we will use, even if we are treating more general cases where $\mu(k)$ is an arbitrary function and $\Omega_m/f^2 \neq 1$. This approximation can be justified better in the next section, when we will see that g only enters in the evolution equation of the power spectrum inside the integral of the self-energy and mode-coupling 1-loop quantities, which should contribute only sub-dominantly in the final power spectrum.

6.3 The Evolution Equation for the Power Spectrum

We are interested in solving the general scale and time dependent evolution equation from the resummation method presented in **anselmi_nonlinear_2012** (eqn. 53):

$$\begin{aligned}\partial_{\mathcal{X}} \tilde{P}_{ab}(k;\mathcal{X}) &= -\tilde{\Omega}_{ac}(\mathbf{k};\mathcal{X})\tilde{P}_{cb}(\mathbf{k};\mathcal{X}) - \tilde{\Omega}_{bc}(\mathbf{k};\mathcal{X})\tilde{P}_{ac}(\mathbf{k};\mathcal{X}) \\ &+ H_{\mathbf{a}}(k;\mathcal{X},\mathcal{X}_{in})\tilde{P}_{\mathbf{a}b}(\mathbf{k};\mathcal{X}) + H_{\mathbf{b}}(k;\mathcal{X},\mathcal{X}_{in})\tilde{P}_{a\mathbf{b}}(\mathbf{k};\mathcal{X}) \\ &+ \int \mathrm{d}s \left[\tilde{\Phi}_{ad}(k;\mathcal{X},s)G^{eik}_{bd}(k;\mathcal{X},s) + G^{eik}_{ad}(k;\mathcal{X},s)\tilde{\Phi}_{db}(k;\mathcal{X},s)\right]\end{aligned} \tag{6.25}$$

where $\mathcal{X} = \ln(D(a)/D(a_{in}))$ (notice that in ref. **anselmi_nonlinear_2012** η is the time variable and represents the growth factor). The first line of this equation corresponds to the linear evolution equation of the power spectrum already discussed before. The second and third lines contain the 1PI (one-particle-irreducible) functions, the self-energy Σ_{ab} and the mode-coupling term $\tilde{\Phi}G^{AB}_{ab}$ accounting for the contributions at the large- and small-k limit at 1-loop order.

Massimo: Can you check from eqns. 33-36?

If we transform this equation to $\eta = \ln \frac{a}{a_{in}}$, using the variable transformation $\partial\mathcal{X}/\partial\eta = \frac{\mathrm{d}\ln D(a)}{\mathrm{d}\ln a} = f(\eta)$, where $f(\eta) \equiv f(N(\eta))$ and the relation between N and η is given by $N = \eta + \ln(a_{in})$, we have to transform also the power spectrum since the field has been redefined (see 6.20):

$$\tilde{P}_{ab} = e^{-2\mathcal{X}(\eta)}e^{2\eta}(\delta_{a1} + \frac{1}{f(\eta)}\delta_{a2})(\delta_{b1} + \frac{1}{f(\eta)}\delta_{b2})P_{ab}(\eta) \tag{6.26}$$

where we can call this transformation Ξ_{ab}:

$$\tilde{P}_{ab} = \Xi_{ab}P_{ab}(\eta)$$

and its inverse would be:

$$\Xi^{-1}_{ab} = e^{2\mathcal{X}(\eta)}e^{-2\eta}(\delta_{a1} + f(\eta)\delta_{a2})(\delta_{b1} + f(\eta)\delta_{b2}) \tag{6.27}$$

However, eqn. 6.25 is only invariant under the transformation:

$$\Upsilon_{ab}(\eta) = f(\eta)\Xi^{-1}_{ab} = f(\eta)e^{2\mathcal{X}(\eta)}e^{-2\eta}(\delta_{a1} + f(\eta)\delta_{a2})(\delta_{b1} + f(\eta)\delta_{b2}) \tag{6.28}$$

since we also have to transform the derivatives and the Ω_{ab} functions.

Inserting these transformations into eqn. 6.25, and generalizing to the case where the growth rate is k-dependent, $f(\eta,k)$,we obtain:

$$\begin{aligned}\partial_\eta P_{ab}(k;\eta) &= -\Omega_{ac}(\mathbf{k};\eta)P_{cb}(\mathbf{k};\eta) - \Omega_{bc}(\mathbf{k};\eta)P_{ac}(\mathbf{k};\eta) \\ &+ [H_{\mathbf{a}}(k;\mathcal{X}(\eta,k),\mathcal{X}(\eta_{in},k))f(\eta,k)P_{\mathbf{a}b}(k;\eta) \\ &+ H_{\mathbf{b}}(k;\mathcal{X}(\eta,k),\mathcal{X}(\eta_{in},k))f(\eta,k)P_{a\mathbf{b}}(k;\eta)] \\ &+ \Upsilon_{ab}(\eta,k) \times \int \mathrm{d}s \left[\tilde{\Phi}_{ad}(k;\mathcal{X}(\eta,k),s)G^{eik}_{bd}(k;\mathcal{X}(\eta,k),s) + G^{eik}_{ad}(k;\mathcal{X}(\eta,k),s)\tilde{\Phi}_{db}(k;\mathcal{X}(\eta,k),\end{aligned} \tag{6.29}$$

The quantities such as $H_{\mathbf{a}}(k;\eta,\eta_{in})$ and $\tilde{\Phi}_{ad}(k;\eta,s)$, shown in eqns 6.33-6.36, which are actually calculated in the $\mathcal{X}$ variable, are now indirectly a function of η through the relation $\mathcal{X}(\eta) = \ln(D(N(\eta))/D(N(\eta_{in})))$.

The evolution equation as it is written in eqn. 6.25 relies on three different assumptions: the power spectrum is well behaved for $k \to 0$, which is in our cosmology a good assumption, since it behaves just a power law k^n; there is a clear separation of scales between "hard" and "soft" modes, or in other words, the eikonal limit is fulfilled. This means that the modes k we are interesting in are much bigger than the internal coupling modes p, q. The third assumption is of course the single-stream approximation, which is used in all forms of resummed and renormalized perturbation theories in cosmology as was explained already in section 6.2.0.1.

In this work we want to solve eqn. **??** for the simple Horndeski models stated above. We will test two apporaches, first we will include inside the $\Omega_{ac}(\mathbf{k};\eta)$ functions, the full scale and time dependence of the parametrized Horndeski models. These terms will have a dominant effect on the evolution of $P(\mathbf{k})$ and we will calculate the 1-loop integrals using the usual ΛCDM model. Since the $H_{\mathbf{a}}(k;\eta,\eta_{in})$ and $\tilde{\Phi}_{ab}(k;\eta,s)$ functions are calculated for a scale-dependent growth factor, we will have to test if this scale dependence has a big effect on the result.

As a second approach we will calculate the 1PI functions using the linear Horndeski propagator in the case that μ is a constant different from one. Once these integrals are calculated, we can then solve the evolution equation completely. In this case we have to test how much a change in $\mu(k)$ affects the result, i.e. if assuming a constant μ is a good approximation or not. In the next section we will show the explicit expressions for these integrals.

6.3.1 The 1PI functions in the Horndeski constant $\mu \neq 1$ case

The Horndeski variable is Y for compatibility with the code for now, but it can be changed to μ later on.

Now we give the general expression for for the 1PI functions $\Sigma^{(1)}_{ab}(k;\mathcal{X},\mathcal{X}')$ and $\Phi^{(1)}_{ab}(k;\mathcal{X},\mathcal{X}')$ computed at 1-loop in eRPT. The general expression for $\Sigma^{(1)}_{ab}(k;\mathcal{X},\mathcal{X}')$ is given by:

$$\Sigma^{(1)}_{ab}(k;\mathcal{X},\mathcal{X}') = \\ 4e^{\mathcal{X}+\mathcal{X}'}\int \mathrm{d}^3q\gamma_{acd}(\mathbf{k},-\mathbf{q},\mathbf{q}-\mathbf{k})u_cP^0(q)u_e\gamma_{feb}(\mathbf{k}-\mathbf{q},\mathbf{q},-\mathbf{k})g_{df}(\mathcal{X},\mathcal{X}') \tag{6.30}$$

If we insert for g_{ab} the linear propagator from 6.24 and the coupling vertices γ_{abc} from 6.8 we have to perform an the angular integration of d^3q, see Appendix **??** for the explicit terms.

The $H_1(k;\mathcal{X},-\infty)$, $H_2(k;\mathcal{X},-\infty)$ functions are a time integration of the $\Sigma^{(1)}_{ab}(k;\mathcal{X},\mathcal{X}')$ quantities in the internal time from minus infinity to the external time η:

$$H_1(k;\mathcal{X},-\infty) = \int_{-\infty}^{\eta} \mathrm{d}s\Sigma^{(1)}_{1b}(k;\mathcal{X},\mathcal{X}')u_b \tag{6.31}$$

$$H_2(k;\mathcal{X},-\infty) = \int_{-\infty}^{\eta} \mathrm{d}s\Sigma^{(1)}_{2b}(k;\mathcal{X},\mathcal{X}')u_b \tag{6.32}$$

They have the following form after performing the time integration:

$$H_{1Y}(k;\mathcal{X},-\infty) = -\frac{\pi k^3 e^{2\mathcal{X}}}{3(3Y+4)} \int \mathrm{d}r \left[16 + 3Y\left(3r^4 - 8r^2 + 1\right) - \frac{9Y}{2r}\left(r^2-1\right)^3 \log\left|\frac{1+r}{1-r}\right|\right] P^0(kr) \tag{6.33}$$

$$H_{2Y}(k;\mathcal{X},-\infty) = -\frac{\pi k^3 e^{2\mathcal{X}}}{3(3Y+4)} \int \mathrm{d}r \left[-\frac{9Y}{r^2} + 9r^2\mu + 4(9Y+4) - \frac{9Y}{2r^3}\left(r^2-1\right)^3 \log\left|\frac{1+r}{1-r}\right|\right] P^0(kr) \tag{6.34}$$

The third line of the evoution equation 6.25 contains the mode-coupling function, which is obtained by integrating the counter-term 1-loop quantity:

$$\tilde{\Phi}^{(1)}_{ad}(k;\mathcal{X},\mathcal{X}') = 2e^{\mathcal{X}+\mathcal{X}'} \int \mathrm{d}^3 q \gamma_{acd}(\mathbf{k},-\mathbf{q},\mathbf{p}) u_c P^0(q) u_d \tag{6.35}$$

$$\times\, u_e P^0(p) u_f \gamma_{bef}(\mathbf{k},-\mathbf{q},\mathbf{p}) \tag{6.36}$$

where $\mathbf{p} = \mathbf{k}-\mathbf{q}$, together with the renormalized propagator from Crocce-Scoccimarro $\bar{G}^L_{bd}(k;\eta,s)$:

$$\int \mathrm{d}s \left[\tilde{\Phi}_{ad}(k;\mathcal{X},s)\bar{G}^L_{bd}(k;\mathcal{X},s) + \bar{G}^L_{ad}(k;\mathcal{X},s)\tilde{\Phi}_{db}(k;\mathcal{X},s)\right] = \tilde{\Phi}G^A_{ab}(k;\mathcal{X}) + \tilde{\Phi}G^B_{ab}(k;\mathcal{X}) \tag{6.37}$$

where for all a, b, the B terms are:

$$\tilde{\Phi}G^B_{ab}(k;\mathcal{X}) = u_a u_b y^2 \left(\frac{\sqrt{\pi}}{2}\left(2y^2+1\right)\mathrm{Erf}(y) + \left(e^{-y^2} - 2\right) y\right) P^0(k) \tag{6.38}$$

The $\tilde{\Phi}G^B_{ab}(k;\eta)$ expression has to be switched off in the small k-limit since it contains 2-loop expressions valid only at large k, therefore it has to be "filtered" by a momentum-cutoff function:

$$\tilde{\Phi}G^A_{ab}(k;\eta) + \frac{(k/\bar{k})^4}{1+(k/\bar{k})^4}\tilde{\Phi}G^B_{ab}(k;\eta)$$

The $\bar{k}$ quantity can be set to a reasonable scale at which the large scale expression starts to be applicable, usually we can set here $\bar{k} = 0.2h/\mathrm{Mpc}$. The A terms read (a new variable $W = \frac{3}{2}Y$ has been defined for simplicity):

$$\tilde{\Phi}G^A_{11} = y(\Phi^{(1)}_{11} - \Phi^{(1)}_{11})\mathcal{B}(y^2;W) \tag{6.39}$$
$$+ \frac{\sqrt{\pi}\mathrm{Erf}(y)}{W+1}(\Phi^{(1)}_{11}W + \Phi^{(1)}_{12})$$

$$\tilde{\Phi}G^A_{12} = \frac{y\mathcal{B}_{12}(y^2;W)}{(1+W)^2(2+W)}(\Phi^{(1)}_{12} - \Phi^{(1)}_{22} - W(\Phi^{(1)}_{11} - \Phi^{(1)}_{12})) \tag{6.40}$$

$$+ \frac{\sqrt{\pi}\mathrm{Erf}(y)}{2(W+1)}(W(\Phi^{(1)}_{11} + \Phi^{(1)}_{12}) + \Phi^{(1)}_{12} + \Phi^{(1)}_{22}) \tag{6.41}$$

$$\tilde{\Phi}G_{22}^{A} = \frac{yW}{(W+1)^2}(\Phi_{22}^{(1)} - \Phi_{12}^{(1)})\mathcal{B}(y^2;W) + \frac{\sqrt{\pi}}{(W+1)}\text{Erf}(y)(\Phi_{12}^{(1)}W + \Phi_{22}^{(1)}) \tag{6.42}$$

and where we have used the combination of generalized Hypergeometric functions:

$$\begin{aligned}\mathcal{B}(y^2;W) = &\ {}_2F_2\left(\frac{1}{2},1;\frac{W}{2}+1,\frac{W}{2}+\frac{3}{2};-y^2\right)\\ &+\frac{W+1}{W+2}\,{}_2F_2\left(\frac{1}{2},1;\frac{W}{2}+\frac{3}{2},\frac{W}{2}+2;-y^2\right)\\ &-\frac{1}{W+2}\,{}_2F_2\left(1,\frac{3}{2};\frac{W}{2}+\frac{3}{2},\frac{W}{2}+2;-y^2\right)\end{aligned}$$

$$\begin{aligned}\mathcal{B}_{12}(y^2;W) = &\ (2+W)\,{}_2F_2\left(\frac{1}{2},1;\frac{W}{2}+1,\frac{W}{2}+\frac{3}{2};-y^2\right)\\ &-{}_2F_2\left(1,\frac{3}{2};\frac{W}{2}+\frac{3}{2},\frac{W}{2}+2;-y^2\right)\end{aligned}$$

Assuming a general scale-dependent function $\mu(k)$ would turn out impossible an analaytic calculation of the intermediate quantities, which would render a practical implementation of this method to be numerically prohibitive. Therefore, using the fact that the μ function behaves as a step function in $\log k$, having a well defined minimum and maximum value we can use simply a constant and check its effect on the 1-loop and 2-loop corrections. Besides, in any viable model compatible with observations μ can differ from 1 by about 15-20%. So that we can be sure that the overall effect will be just a perturbation around the LCDM value and should reduce to it in the limit $\mu \to 1$.

Issue: In the LCDM case from Massimo's paper, all three $\tilde{\Phi}G_{ab}^{A}$ functions can be written using the same combination of Hypergeometric Functions. In this case, this is only possible for $\tilde{\Phi}G_{11}^{A}$ and $\tilde{\Phi}G_{22}^{A}$.

6.3.2 Method of implementation

Equation **??** is the final and main equation we want to solve and it represents a differential equation in time, for each external momentum k that we are interested in. This means that if we want to compute the power spectrum on a grid with 100 points in k-space, we need to solve 100 times the same differential equation. That is why parallelism plays an important role here to speed up the computational calulation.

First we compute the linear growth function for a specific Horndeski model, using the linear part of 6.19. With this we obtain the growth rate and the growth factor, which are then used to calculate the loop integrals.

The initial power spectrum is obtained from CAMB and its evaluated at 100 points in k-space. This is used to compute analytically eqns. 6.33, 6.34. These integrals can be stored as long as the cosmological parameters are not modified.

Finally we compute the 6.25 for each k mode in parallel, which using 8 cores on a normal machine, takes about 15 seconds until redshift $z = 0$.

6.4 Numerical approaches

6.4.1 N-body simulations

- The Vlasov-Poisson system can be discretized on a grid and the denisty can be represented with particles
- Particle-Mesh method
- Tree method
- Extraction of information: The power spectrum and Halo catalogs

As we have seen in the previous chapters, non-linear dynamics of gravitational interactions, are hard to treat analytically. If we want to study the growth of perturbations in the full non-linear regime, as well as in large scales and at galactic scales, we need to perform numerical simulations. Numerical N-body simulations have many advantages, as well as some drawbacks related to their specific practical implementation and we will try to explain here only some of their most fundamental properties.

First we will review some general techniques to deal with great numbers of particles in gravitational interactions, then we will review some of the most important Dark Universe simulations and finally we will explain some technicalities regarding the extraction of the power spectrum from simulation data.

6.5 Computational Techniques

To perform an exact evolution of the density fluctuations, beyond linear perturbations, the density field has to be represented by a set of fictitious discrete particles with a certain mass. These particles do not represent real galaxies or clusters of galaxies, they just sample the underlying density field. For current cosmological simulations, depending on the desired resolutions, the particle masses are around $m_p \approx 10^9 - 10^{12} M_{\odot}$. (see **kuhlen_numerical_2012**)

The equations of motion for each particle depend on solving for the gravitational field due to all the other particles, finding the change in particle positions and velocities over some small time step. Then particle positions and velocities have to be updated and the gravitational potentials have to be recalculated to start a new iteration. For cosmological simulations, where we suppose the Universe to be isotropic and homogeneous at large scales and described by a smooth FLRW metric, we use the fact that at smaller scales the Universe must tend to locally inertial frames where Newton's laws are valid. Therefore we can just use Newtonian dynamics and the expansion history of the Universe is taken into account by using comoving coordinates, where the expansion rate $a(t)$ is factored out (see **peacock_cosmological_1999**; **dehnen_n-body_2011**). We then obtain for the comoving velocity u the following equations:

$$\frac{d}{dt}\vec{u} = -2\frac{\dot{a}}{a}u - \frac{1}{a^2}\nabla\Phi \tag{6.43}$$

where Φ is the Newtonian gravitational potential due to density perturbations. If we change the time variable from t to a we obtain:

$$\frac{d}{d\ln a}(a^2\vec{u}) = \frac{a}{H}\vec{g} = \frac{G}{aH}\sum_i m_i \frac{\vec{x_i} - \vec{x}}{|\vec{x_i} - \vec{x}|^3} \tag{6.44}$$

We see that in order to get the gravitational acceleration for one single particle, we have to sum the contributions from all other particles, which leads to the exact solution, but it is computationally prohibitive for a large number of particles, since the number of needed computations grows as $\mathcal{O}(N^2)$.

Since we have only finite resources for computation, we need to calculate the particles in a finite box of size L and a finite number of particles N. In this case the walls of the box would break our desired homogeneity and isotropy, therefore we need to introduce periodic boundary conditions, such that the cubic box is actually a 3-torus, where the walls on opposite sides are identified. Then the gravitational potential can be described as (**dehnen_n-body_2011**):

$$\Phi(\mathbf{x}, t) = -G\sum_{\mathbf{n}} \int d\mathbf{x}' \, \frac{\rho(\mathbf{x}' + \mathbf{n}L, t)}{|\mathbf{x} - \mathbf{x}' - \mathbf{n}L|} \tag{6.45}$$

where the sum is performed over $\mathbf{n} = (N_x, \, N_y, \, N_z)$ and accounts for all periodic replica. Usually all the N_i are the same in all directions. Because of the fact, that performing an infinite sum of replicas is in practice not possible, the sum is approximated using Ewald's method (see **ewald_berechnung_1921**), which was originally developed for periodic crystals in solid state physics and was first adapted to this field by **hernquist_application_1991**

The gravitational Newtonian force has a singularity when two particles approach too close to each other, that is why a so-called softening term has to be added to the force equation in order to avoid unphysical accelerations during close encounters. The force can be modified for small distances to something like (see **trenti_gravitational_2008**):

$$\vec{F_i} = -\sum_{j\neq i} Gm_i m_j \frac{\vec{x_i} - \vec{x_j}}{\left(|\vec{x_i} - \vec{x_j}|^2 + \epsilon^2\right)^{3/2}} \tag{6.46}$$

where $\epsilon > 0$ is the softening or smoothing length, which is a typical size below which the gravitational interaction is suppressed. In current cosmological N-body simulations, the softening length is found to be: $1.0h^{-1}$kpc $\leq \epsilon \leq 150h^{-1}$kpc, meaning that most of the simulations cannot resolve halos or subhalos of galaxies and can only focus on large scale structure formation. The choice of gravitational softening length is still a debate among the community, since making it to small increases computational effort (due to smaller time steps), but allows for more realistic gravitational potentials and on the other hand it introduces spurious two-body relaxation effects that can cause artificial fragmentation of structure (see **kuhlen_numerical_2012** and references therein).

Many alternatives to direct summation of forces, have been developed in the last 20 years, also using hybrid approaches. The most relevant ones for our purposes are the following (see reviews by **trenti_gravitational_2008**; **dehnen_n-body_2011**; **kuhlen_numerical_2012**):

- **Particle Mesh PM:** Since the problem is to solve Poisson's equation, a faster approach is to use Fourier methods for discretized systems, such as the FFT (see **??**) and solve directly for Poisson's equation in Fourier space:

$$(\nabla\Phi)_k = -i\Phi_k\mathbf{k} = \frac{-i4\pi Ga^2\overline{\rho}}{k^2}\delta_k\mathbf{k} \tag{6.47}$$

Then we can eliminate the matter density in terms of Ω_m and for a given particle, the equation of motion would be:

$$\frac{d}{d\ln a}(a^2\mathbf{u}) = \sum \mathbf{F}_k \exp(-i\mathbf{k}\cdot\mathbf{x}), \qquad \mathbf{F}_k = -i\mathbf{k}\frac{\Omega_m Ha^2}{2k^2}\delta_k \tag{6.48}$$

By interpolating the density field ρ_k over a finite grid, one can solve Poisson's equation and then use the FFT again to calculate the forces and velocities on the particles. The complicated part of the algorithm is the assignement of the mass of the particles onto the grid cells and then interpolating back the evaluated force onto the particles, for consistency the same procedure has to be used for both these steps. Particles do not interact with each other but only through the mean field, which causes that the maximum force resolution to be limited to about the size of the mesh. The computational advantage of this method is that the number of computations is now of the order $\mathcal{O}(N_g log(N_g))$, for N_g the number of grid cells. This is the scaling of the FFT method also (see **??**).

- **Particle-particle-particle-mesh $\mathbf{P^3M}$** : Since PM codes are gravitationally softened below the mesh size and the resolution is therefore low, this hybrid method uses a coarse grid to calculate forces at larger scales between distant cells and for particles in the same or neighbouring cell, a direct sumation between particles is performed, increasing the force resolution, but also in some way the computational time, if there is a strong clustering of particles.

- **Adaptive Mesh Refinement AMR:** The dynamic range of particle-mesh codes can be increased by using instead of a static grid, an adaptive one, which has more concentrated grid elements in the high density regions, where the forces are also more varying and stronger. This allows to truncate the error to the desired precision level by refining the mesh at specific points. The complicated part of this method is to match the solution at the grid interfaces, where they might change drastically. Since the force softening can change along a particle's orbit, this method can lead to an unphysical violation of energy conservation. To accelerate the calculation of an AMR code, one can evolve the particles asynchronously, leaving the particles in the coarse grid, while evolving with smaller time steps the ones in the finer grid and using the coarser grid potential as a boundary condition.

- **Tree Codes:** If close encounters are not important and the force contributions from distant particles do not have to be calculated at high accuracy, this method is well suited for cosmological simulations. The simulation box is split into eight cubic cells, containing a determined number of particles, each cubic cell containing fewer than n_{max} number of particles is split again into eight cubic child cells of half their parent's particle size. This results in a tree-like binary (oct) hierarchy of cubic nodes, containing the root box (that contains all N particles) at his bottom. For each cubic cell, the total mass and center of mass is calculated and stored as an information on the node. Then at the moment of calculating the force acting on a particle at position $\vec{x}$, one just adds the contributions from different cells with center of mass $\vec{z_a}$, depending on some opening angle: $\sin(\theta) = |\vec{x} - \vec{z}_a| / w_a$, where w_a is the linear size of the box (this is the easiest approximation, for a formal derivation of the tree code, using taylor expansions and multipoles, see **dehnen_n-body_2011**). If the opening angle is bigger than wanted, one applies the algorithm to the daughter cells. In this way one gets the contributions from all possible groups of particles. This method offers a computational time that scales like the depth of the tree, therefore being of order $\mathcal{O}(N \ln(N))$, which is similar to the above mentioned methods. For nearby particles, a standard particle-particle interaction is calculated. The complicated aspect of the tree method is, among others, how to visit each cell only once, in a highly parallelizable way. Really efficient methods have been developed, using Peano-Hilbert curves, such as in the Gadget-2 code, which is a Tree-PM code (**springel_cosmological_2005**). A drawback of this method is that each interaction is only one-sided, meaning that the force of a single particle on the group of particles of a distant cell is not calculated and therefore Newton's third law is violated. This can lead to fluctuations in the total energy of the N-body system, but it has been shows that this spurious effects can be kept quite small (see **dehnen_n-body_2011**). In 6.1 (top), the binary tree hierarchy of the simulation box can be seen and (bottom) a comparison between the computation of the force with direct sumation and the tree algorithm. The computational advantage of the tree algorithm can be really appreciated in this case.

- **Fast Multipole Methods FMM:** This modern method uses the fact that in a standard tree implementation, one does not take advantage of the fact that the force caused by distant cells will be the same for nearby particles. This method not only expands the distant source potentials into multipoles, but also expands the force at the sink position, leading to a gain in speed that is quite considerable, leading to computational efforts of order $\mathcal{O}(N)$. Since the interaction is approximated by cells at both ends of the interaction (each interaction is mutual), Newton's third law is respected and no spurious effects arise. At nearby distances a particle-particle method is employed as usual. The drawback of this method is that its parallelization is not so straighforward and has not been developed yet. For more references, check **dehnen_n-body_2011**

FIGURE 6.1: **Top**: Hierarchical structure of a Tree code, the fundamental box is divided into always smaller cells. A Peano-Hilbert curve method (space filling curve) is used to visit all cells in parallel. Figure taken from: **springel_cosmological_2005** **Bottom**: Comparison between force calculation for 100 particles. Left: Each red curve represents a direct particle-particle interaction. Middle: The tree method, with its cells and groups of particles (green lines represent particle-cells interactions). Only 7 interactions are needed to compute total force. Right: Interactions needed to compute all the forces among particles, requiring 902 cell-particle interactions and 306 particle-particle interactions, instead of 4950 computations in the case of the direct summation. Figure taken from: **dehnen_n-body_2011**

6.5.1 Fitting formulae and prescriptions

Chapter 7

Growing Neutrino Quintessence

7.1 Introduction

The observed accelerated expansion of the Universe is currently well described by the standard ΛCDM scenario, where a cosmological constant leads the background of the Universe to a final de Sitter state. Such a scenario, however, raises a "coincidence" (or "why-now") problem, as it is not understood why dark energy becomes important just recently, marking the present cosmological epoch as a special one within cosmic history. Alternative models of dynamical dark energy or modified gravity should address the "why now problem" and the associated fine tuning of parameters. At the same time they need to explain why dark energy is almost static in the present epoch, such that the present dark energy equation of state w is close to the observed range near -1.

Growing neutrino quintessence **amendola_growing_2008**; **wetterich_growing_2007** explains the end of a cosmological scaling solution (in which dark energy scales as the dominant background) and the subsequent transition to a dark energy dominated era by the growing mass of neutrinos, induced by the change of the value of the cosmon field which is responsible for dynamical dark energy. The dependence of the mass of neutrinos on the cosmon (dark energy) field ϕ,

$$m_\nu = m_\nu(\phi) \propto \hat{m}_\nu e^{-\int \beta(\phi) d\phi} \; , \beta(\phi) = -\frac{\partial \ln m_\nu(\phi)}{\partial \phi} \tag{7.1}$$

involves the cosmon-neutrino coupling $\beta(\phi)$ which measures the strength of the fifth force (additional to gravity). The constant $\hat{m}_\nu$ is a free parameter of the model which determines the size of the neutrino mass. (We take for simplicity all three neutrino masses equal - or equivalently m_ν stands qualitatively for the average over the neutrino species.) The special role of the neutrino masses (as compared to quark and charged lepton masses) is motivated at the particle physics level by the way in which neutrinos get masses **wetterich_growing_2007** Growing neutrino quintessence with a sufficiently large negative value of β successfully relates the present dark energy density and the mass of the neutrinos. The evolution of the cosmon is effectively stopped once neutrinos become non-relativistic. Dark energy becomes important now because neutrinos become non-relativistic in a rather recent past, at typical redshifts of about $z = 5$ **mota_neutrino_2008** In this way, the "why now problem" is resolved in terms of a "cosmic trigger event" induced by the change in the effective neutrino equation of state, rather than by relying on the fine tuning of the scalar potential. This differs from other mass varying neutrino cosmologies (usually known as MaVaN's) **brookfield_cosmology_2007**; **la_vacca_mass-varying_2013**; **bi_cosmological_2005**; **fardon_dark_2004**; **kaplan_neutrino_2004**; **spitzer_stability_2006**; **takahashi_speed_2006** Some of the observational consequences of those models were studied in **la_vacca_mass-varying_2013**; **kaplan_neutrino_2004** and more recently a new scalar field - neutrino coupling that produces viable cosmologies was proposed in **simpson_dark_2016** A viable cosmic background evolution of growing neutrino quintessence offers interesting prospects of a possible observation of the neutrino background.

The case in which the coupling β is constant has been largely investigated in literature at the linear level **mota_neutrino_2008** in semi-analytical non-linear methods **wintergerst_clarifying_2010**; **wintergerst_very_2010**; **brouzakis_nonlinear_2011** joining linear and non-linear information to test the effect of the neutrino lumps on the cosmic microwave background **pettorino_neutrino_2010** and within N-Body simulations **ayaita_neutrino_2013**; **ayaita_structure_2012**; **baldi_oscillating_2011**; **ayaita_nonlinear_2016** For the values of β ($\beta \gtrsim 10^3$) needed for dark energy to dominate today, the cosmic neutrino background is clumping very fast. Large and concentrated neutrino lumps form and induce very substantial backreaction effects. These effects are so strong that the deceleration of the evolution of the cosmon gets too weak, making it difficult to obtain a realistic cosmology **fuhrer_backreaction_2015**

In this paper we instead consider the case in which the neutrino-cosmon coupling $\beta(\phi)$ depends on the value of the cosmon field and increases with time. In a particle physics context this has been motivated **wetterich_growing_2007** by a decrease with ϕ of the heavy mass scale (B-L-violating scale) entering inversely the light neutrino masses. In this scenario $\beta(\phi)$ has not been large in all cosmological epochs - the present epoch corresponds to a crossover where β gets large. A numerical investigation **baldi_oscillating_2011** of this type of model has revealed compatibility with observations for the case of a present neutrino mass $m_{\nu,0} = 0.07$ eV. In the present paper we investigate the dependence of cosmology on the value of the neutrino mass by varying the parameter $\hat{m}_\nu$ in eq. 7.1. For large neutrino masses we find a qualitative behavior similar to the case of a constant neutrino-cosmon coupling β, with difficulties to obtain a realistic cosmology. In contrast, for small neutrino mass, the neutrino lumps form and dissolve, with small influence on the overall cosmological evolution. In this case, the neutrino-induced gravitational potentials are found to be much smaller than the ones induced by dark matter. As we will discuss in this paper, it will not be easy to find observational signals for the neutrino lumps. In-between the regions of small and large neutrino masses we expect a transition region for intermediate neutrino masses where, by continuity, observable effects of the neutrino lumps should show up.

7.2 Growing neutrinos with varying coupling

We consider here cosmologies in which neutrinos have a mass that varies in time, along the framework of "varying growing neutrino models" **wetterich_growing_2007** As long as neutrinos are relativistic, the coupling is inefficient and the dark energy scalar field ϕ rolls down a potential, as in an early dark energy scenario. As the neutrino mass increases with time, neutrinos become non-relativistic, typically at a relatively late redshift $z \approx 4 - 6$ **pettorino_neutrino_2010** This influences the evolution of ϕ, which feels the effect of neutrinos via a coupling to the neutrino mass $m_\nu(\phi)$. The evolution of the scalar field slows down and practically stops, such that the potential energy of the cosmon behaves almost as a cosmological constant at recent times. In other words, in these models the cosmological constant behavior observed today is related to a cosmological trigger event (i.e. neutrinos becoming non-relativistic) and the present dark energy density is directly connected to the value of the neutrino mass. In the following we will detail the formalism and equations used to describe the cosmological evolution of the model.

We start with the linearized Friedman-Lemaitre-Robertson-Walker metric in the Newtonian gauge:

$$ds^2 = -(1+2\Psi)dt^2 + a^2(1-2\Phi)d\mathbf{x}^2 \ . \tag{7.2}$$

Moreover, we use a quasi-static approximation for sub-horizon scales ($H/k \ll 1$) which allows us to neglect time derivatives with respect to spatial ones. Then the quasi-static, first-order perturbed Einstein equations, are the Poisson equation **ma_cosmological_1994**

$$k^2\Phi = 4\pi G a^2 \delta T^0_0 \ , \tag{7.3}$$

and the "stress" equation

$$k^2(\Phi - \Psi) = 12 G a^2(\bar{\rho} + \bar{P})\sigma \ , \tag{7.4}$$

where δT^0_0 is the perturbation of the $0-0$ component of the energy momentum tensor $T_{\mu\nu}$ and σ is the anisotropic stress of the fluid which depends on the traceless component of the spatial part of the energy-momentum tensor, $T^i_j - \delta^i_j T^k_k/3$. This stress tensor is in our case only important for relativistic particles (i.e. the neutrinos). The source term of the Poisson equation 7.3 will contain contributions from all matter species (dark matter & neutrinos) and from the cosmon field. It is proportional to the total density contrast $\delta\rho_t = \delta\rho_\nu + \delta\rho_m + \delta\rho_\phi$.

The cosmon field can be described through a Lagrangian in the standard way

$$-\mathcal{L}_\phi = \frac{1}{2}\partial^\nu\phi\partial_\nu\phi + V(\phi) \tag{7.5}$$

where for this work we choose an exponential potential $V(\phi) \propto e^{-\alpha\phi}$. The field dependent mass (eq.7.1) allows for an energy-momentum transfer between neutrinos and the cosmon, which is proportional to the trace of the energy momentum tensor of neutrinos $T_{(\nu)}$ and to a coupling parameter $\beta(\phi)$

$$\nabla_\eta T^{\mu\eta}_{(\phi)} = +\ \beta(\phi)T_{(\nu)}\partial^\mu\phi \ , \tag{7.6}$$

$$\nabla_\eta T^{\mu\eta}_{(\nu)} = -\ \beta(\phi)T_{(\nu)}\partial^\mu\phi \ . \tag{7.7}$$

The cosmon is the mediator of a fifth force between neutrinos, acting at cosmological scales. Its evolution is described by the Klein-Gordon equation sourced by the trace of the energy-momentum tensor $T_{(\nu)}$ of the neutrinos,

$$\nabla_\mu\nabla^\mu\phi - V'(\phi) = \beta(\phi)T_{(\nu)}. \tag{7.8}$$

As long as the neutrinos are relativistic ($T_{(\nu)} = 0$) the source on the right hand side vanishes. During this time, the coupling has no effect on the evolution of ϕ. While the potential term $\sim V'$ drives ϕ towards larger values, the term $\sim \beta$ has the opposite sign and stops the evolution effectively once $\beta T_{(\nu)}$ equals V'. The trace of the energy momentum tensor T_ν, entering eq.7.8 is equal to:

$$T_\nu = m_\nu(\phi)\tilde{n}(\phi) \tag{7.9}$$

where $\tilde{n}_\nu(\phi) = n_\nu(\phi)/\gamma$ is the ratio of the number density of neutrinos n_ν, divided by the relativistic γ factor. Eq.7.9 is valid for both relativistic and non-relativistic neutrinos. Here we consider a coupling β between neutrino particles and the quintessence scalar field ϕ as a field dependent quantity:

$$\beta(\phi) \equiv -\frac{1}{\phi_c - \phi}\,. \tag{7.10}$$

From eq.7.1 the neutrino mass is then given by:

$$m_\nu(\phi) = \frac{\bar{m}_\nu}{\phi_c - \phi}\,. \tag{7.11}$$

Here ϕ_c denotes the asymptotic value of ϕ for which β and $m_\nu(\phi)$ would formally become infinite. By an additive shift in ϕ it can be set to an arbitrary value, e.g. $\phi_c = 0$. We consider the range $\phi < \phi_c$. The divergence of β for $\phi \to \phi_c$ in eq.7.10 is not crucial for the results of the present paper - β and m_ν never increase to large values, such that the immediate vicinity of ϕ_c plays no role.

The coupling induces a total force acting on neutrinos given by $\nabla(\Phi_\nu + \beta\delta\phi)$ and appearing in the corresponding Euler equation **pettorino_neutrino_2010** as usual in coupled cosmologies **baldi_hydrodynamical_2010** For values $2\beta^2 > 1$ the fifth force induced on neutrinos by the cosmon becomes larger than the gravitational attraction. For the large values of $|\beta| \approx 10^2$ reached during the cosmological evolution, the attraction induced by the cosmon gives rise to the formation of neutrino lumps. As shown in **mota_neutrino_2008**; **pettorino_neutrino_2010** this represents the major difficulty encountered within growing neutrino models and also, simultaneously, one of its clearest predictions with respect to alternative dark energy models: the presence of neutrino lumps at scales of ≈ 10 Mpc or even larger, depending on the details of the model **mota_neutrino_2008** Since the attractive force between neutrinos is 10^4 times bigger than gravity, therefore also the dynamical time scale of the clumping of neutrino inhomogeneities is a factor 10^4 faster than the gravitational time scale. Even the tiny inhomogeneities in the cosmic neutrino background grow very rapidly non-linear. The impact of such structures, has been shown to depend crucially on the strength of backreaction effects **ayaita_structure_2012**; **ayaita_nonlinear_2016** For constant coupling, the effect of backreaction is strong and can lead to neutrino lumps with rapidly growing concentration, reaching values of the gravitational potential which exceed observational constraints. The effect is so strong that it is able to destroy the oscillatory effect first encountered in **baldi_hydrodynamical_2010** in which neutrino lumps were forming and then dissipating. No realistic cosmology has been found in this case **fuhrer_backreaction_2015** With the varying coupling of eq.7.10 a similar behavior will be found for large neutrino masses. For small neutrino masses the oscillatory effects will be dominant and realistic cosmologies seem possible **ayaita_nonlinear_2016**

7.3 Numerical treatment of growing neutrino cosmologies

7.3.1 Modified Boltzmann code

For the early stages of the evolution of the growing neutrino quintessence model, neutrinos behave as standard relativistic particles and the coupling to the cosmon field is suppressed. Therefore the Klein-Gordon equation can be linearized and no important backreaction effects are present. The Einstein-Boltzmann system of equations for the relativistic neutrinos and all other species has been solved using a modified version of the code CAMB **lewis_efficient_2000** (hereafter referred to as nuCAMB), used and developed already in previous papers on mass-varying and growing neutrino cosmologies. We refer the reader to previous publications **mota_neutrino_2008**; **pettorino_neutrino_2010**; **wintergerst_very_2010**; **brookfield_cosmology_2007** for details about its implementation. These equations are valid until neutrinos become non-relativistic and as long as perturbations are still linear. The neutrinos can be seen as a weakly-interacting gas of particles in thermal equilibrium with a phase space distribution $f(p)$ with p denoting the momentum. The statistical description is in the case of neutrinos a Fermi-Dirac distribution, given by

$$f_{FD}(p) = \frac{1}{e^{(E(p)-\mu)/T}+1} \; , \tag{7.12}$$

where μ is the chemical potential and $E(p) = \sqrt{m^2+p^2}$ the particle energy. Then, the number density of neutrinos, the energy density and the pressure are given respectively by

$$n_\nu = \frac{2}{(2\pi)^3}\int \mathrm{d}^3p\, f_{FD}(p) \; , \tag{7.13}$$

$$\rho_\nu = \frac{2}{(2\pi)^3}\int \mathrm{d}^3p\, E(p) f_{FD}(p) \; , \tag{7.14}$$

$$P_\nu = \frac{2}{(2\pi)^3}\int \mathrm{d}^3p\, \frac{p^2}{E(p)} f_{FD}(p) \; . \tag{7.15}$$

The solution of the Boltzmann hierarchy of neutrinos coupled to the perturbed Einstein equations 7.3 and 7.4, together with the solution of the background Klein-Gordon equation 7.8, form the basis of the modification of nuCAMB with respect to the standard code CAMB, which handles dark matter, photons and baryons altogether. We recall that for growing neutrino quintessence, the neutrino mass depends on the cosmon field ϕ and therefore on the scale factor a.

The ratio of the initial mass of the neutrinos to their temperature (given in eV), is calculated in nuCAMB as follows

$$\hat{r}_{\nu eV} \equiv \left(\frac{m}{T}\right)_{\nu,camb} = \frac{(7/8)(\pi^4/15)}{(3/2)\zeta(3)} \times \frac{\rho_{cr}\Omega_{\nu,input}}{\rho_\nu} \; . \tag{7.16}$$

The first fraction comes from the relation $m_\nu \approx \rho_\nu/n_\nu = ((\frac{7}{8}\frac{\pi^4}{15})/\frac{3}{2}\zeta(3))T_\nu$, which is valid in the non-relativistic limit of eqns.7.12-7.15; the critical density is defined as usual: $\rho_{cr} = \frac{3H^2_{0,input}}{8\pi G}$. The second fraction is a re-scaling that corrects the neutrino density in order to match the wanted $\Omega_{\nu,input}$ given as input value. The code performs an iterative routine that varies initial conditions in such a way that the input parameters are obtained at present time. Since this is not exact, the final values of H_0 and Ω_ν might vary slightly with respect to the given input values. The ratio $\hat{r}_{\nu eV}$ depends on the input parameters $H_{0,input}$ (via the critical density) and on $\Omega_{\nu,input}$ [1]. Furthermore, the neutrino energy density ρ_ν and the photon energy density ρ_γ at

[1] $\hat{r}_{\nu eV}$ is also the conversion factor between the mass units in the N-Body code and units in eV.

relativistic times are related as

$$\rho_\nu = N_\nu \times \frac{7}{8} \times \left(\frac{4}{11}\right)^{4/3} \rho_\gamma \tag{7.17}$$

where $N_\nu = 3$ is the number of neutrino species. The use of these formulae is valid if initial conditions are set when neutrinos are non-relativistic, where the linear regime still applies. For initial conditions set at earlier time relativistic corrections have to be taken into account. After solving the Einstein-Boltzmann system, realizations of the fields $\delta_\nu(\mathbf{k})$ and $v_{pec,\nu}(\mathbf{k})$ at an early time are obtained from nuCAMB and are then used as the initial conditions for the neutrino distribution in the growing neutrino quintessence N-body simulation. This will be explained more in detail at the end of the following section.

7.3.2 N-body simulation

For N-body simulations, we use here the code developed in **ayaita_neutrino_2013**; **ayaita_nonlinear_2016**; **ayaita_structure_2012** and then refined in **fuhrer_backreaction_2015** and in the present work, which uses a particle-mesh approach for the neutrino and dark matter particle evolution and a multi-grid approach for solving the non-linear scalar field equations. In table 7.1 we describe the parameters of the models discussed in this paper. We consider 5 models with different neutrino masses.

Our N-body simulation differs from standard Newtonian N-body codes in many ways, the most important one being that we evolve the cosmon ϕ and the gravitational potentials Φ and Ψ separately. While neutrinos, dark matter and the cosmon are non-linear in the N-body simulations, we assume that the gravitational potentials Φ and Ψ are small, which is valid in cosmological applications, even for large deviations of standard ΛCDM and at small scales. The perturbation in the dark energy scalar field $\delta\rho_\phi$ can be calculated from the perturbation of the energy density of the cosmon field

$$\delta\rho_\phi = \frac{\bar{\phi}'\delta\phi}{a^2} + V(\bar{\phi})\delta\phi \ . \tag{7.18}$$

The evolution of the homogeneous potential of the cosmon field can be obtained through its energy density and pressure in the following way

$$V_\phi(a) = \frac{1}{2}(\rho_\phi(a) - p_\phi(a)) \ , \tag{7.19}$$

while the perturbations in the potential can be approximated by

$$\delta V_\phi(a) = -\frac{1}{2}(\delta\rho_\phi(a) + 3\delta p_\phi(a)) \ . \tag{7.20}$$

The cosmon field can cluster and therefore its spatial gradients are non-vanishing, so that after averaging over the volume of the box the energy density of the cosmon field is

$$\bar{\rho}_\phi = \frac{1}{2}\overline{\dot{\phi}^2} + \frac{1}{2a^2}\overline{(1+2\Phi)\partial_i\phi\partial_j\phi\delta^{ij}} + \overline{V(\phi)} \ , \tag{7.21}$$

while its pressure reads

$$\bar{P}_\phi = \frac{1}{2}\overline{\dot{\phi}^2} - \frac{1}{6a^2}\overline{(1+2\Phi)\partial_i\phi\partial_j\phi\delta^{ij}} + \overline{V(\phi)} \ . \tag{7.22}$$

Cosmological parameters	Growing neutrino models					
Linear values	M1	M2	M3	M4	M5	M6
$\Omega_{\nu 0}+\Omega_{\phi 0}$	0.686	0.688	0.692	0.701	0.693	0.697
$\Omega_{\nu 0}$	3.8×10^{-3}	2.6×10^{-2}	1.64×10^{-2}	4.7×10^{-2}	6.1×10^{-2}	9.4×10^{-2}
h	0.671	0.673	0.6818	0.701	0.722	0.740
$m_{\nu 0}$ [eV]	0.060	0.407	0.239	0.730	1.000	1.712
$\langle m_\nu\rangle[0.4:0.6]$[eV]	0.040	0.067	0.134	0.277	0.399	0.701
$\langle m_\nu\rangle[0.8:1.0]$[eV]	0.099	0.152	0.318	0.661	0.907	1.51
$\langle w_{\nu\phi}\rangle[0.9:1.0]$	-0.97	-0.97	-0.95	-0.92	-0.90	-0.85
N-body values						
$\Omega_{\nu 0}+\Omega_{\phi 0}$	0.688	0.690	-	-	-	-
$\Omega_{\nu 0}$	2.5×10^{-2}	1.9×10^{-2}	-	-	-	-
$m_{\nu 0}$ [eV]	0.038	0.078	-	-	-	-
$\langle m_\nu\rangle$ [0.4 : 0.6] [eV]	0.048	0.069	0.1436	0.280	0.401	0.676
$\langle m_\nu\rangle$ [0.8 : 1.0] [eV]	0.120	0.164	-	-	-	-
$\langle w_{\nu\phi}\rangle[0.9:1.0]$	-0.95	-0.96	-	-	-	-
a_{final}	1.0	1.0	0.84	0.70	0.65	0.67

TABLE 7.1: Table of parameters for the six models considered in this work. The top part refers to the output values computed with the linear nuCAMB code. The bottom part refers to values computed within the N-body simulation. Quantities denoted with a subscript 0 are values at present time, $a = 1.0$. The $\langle m_\nu\rangle\,[a_1 : a_2]$ is the root mean squared (RMS) value of the neutrino mass in units of eV computed between $a = a_1$ and $a = a_2$. The same notation is also valid for $\langle w_{\nu\phi}\rangle[a_1, : a_2]$ corresponding to the equation of state of the combined cosmon and neutrino fluid which represents dynamical dark energy. a_{final} is the final time at which simulations were computed accurately. Therefore, for the models M3-M6 we cannot cite values of present time quantities or averages at times beyond a_{final}. The input values for nuCAMB corresponding to all models can be found in table 7.2 of appendix 7.9.

We will use for the following a convention in which bars denote spatial averages, while angular brackets denote time averaged quantities. The evolution of the cosmon field is solved using a multigrid relaxation algorithm, known as the Newton-Gauß-Seidel solver, which was originally developed for $f(R)$ modified gravity simulations **puchwein_modified-gravity-gadget:_2013** and has also been implemented into the growing neutrino N-body simulations in **ayaita_nonlinear_2016** The bottom part of table 7.1 lists the results of the six models computed using the N-body simulations.

In the case of neutrinos, the mass is a time-varying quantity following eq.7.11. Neutrinos obey a modified geodesic equation

$$\frac{du^\mu}{d\tau}+\Gamma^\mu_{\nu\lambda}u^\nu u^\lambda=\beta(\phi)\partial^\mu\phi+\beta(\phi)u^\nu u^\mu\partial_\nu\phi\ , \tag{7.23}$$

in which the right hand side gets a contribution from the coupling.

Simulations start at an initial value of $a_{in} = 0.02$. Until $a \approx 0.30$ the dark matter particles, the cosmon field and the gravitational potentials are evolved on the grid. For dark matter particles we take standard initial conditions from nuCAMB and start the particle-mesh algorithm that solves the Poisson equation 7.3 at an initial redshift of $z = 49$. This is not the most accurate way of setting initial conditions for cosmological dark matter simulations (see for example recent N-body comparisons by **schneider_matter_2016**), but since in this work we are not interested in detailed substructures of dark matter halos or a percent-accurate power spectrum, we find that our approach gives a correct description at the scales of interest. Neutrinos are first treated differently from other particles, as a distribution of relativistic particles in thermal equilibrium and no backreaction effects from neutrino structures are taken into account. Starting from a scale factor of approximately $a_{ini} \approx 0.30$

(depending on the exact parameters of each model), which is when neutrinos become non-relativistic, neutrinos are also projected on the grid: their phase-space distribution is sampled using effective particles. Since their equation of state is non-relativistic, we can approximate the phase-space distribution by

$$f_\nu(\mathbf{x},\mathbf{v}) = \bar{n}_\nu f_{FD}(|\mathbf{v}_\nu - \mathbf{v}_{\mathbf{pec},\nu}(\mathbf{x})|)(1+\delta_v(\mathbf{x})) \ , \tag{7.24}$$

where f_{FD} is the Fermi-Dirac distribution (7.12). The thermal velocities of the neutrinos are the difference between their total velocities and their peculiar velocities $\mathbf{v}_{th,\nu} = \mathbf{v}_\nu - \mathbf{v}_{pec,\nu}$. We obtain $\delta_v(\mathbf{x})$ and $\mathbf{v}_{\mathbf{pec},\nu}(\mathbf{x})$ by Fourier transforming the momentum-space realization of those fields obtained at the time a_{ini} from nuCAMB. Equation 7.24 is solved for $\mathbf{v}_{th,\nu}$ in order to obtain the correct thermal distribution of particles and we duplicate the number of neutrino particles in each grid, assigning to each of them a thermal velocity which is equal in magnitude but opposite in direction, to avoid a distortion of the distribution of peculiar velocities at larger scales than a single grid cell size. For a large enough number of effective neutrino particles (i.e. when there is much more than one particle per cell), the distortion of the peculiar velocities by thermal velocities should be negligible. The correct neutrino density one would obtain from the Fermi-Dirac distribution for a non-relativistic particle reads

$$\langle \rho_\nu(\mathbf{x}) \rangle_{f_\nu} = \int \mathrm{d}^3v \, m_\nu f_\nu(\mathbf{x},\mathbf{v}) = m_\nu \bar{n}_\nu (1+\delta_v(\mathbf{x})) \ . \tag{7.25}$$

Since we need to enforce the right hand side of 7.25 at each grid cell of comoving volume $a^3\Delta V$, where the mass of the neutrinos is given by the scalar field, we have a condition on the number of particles N_{part}, such that

$$\frac{M_\nu \langle N_{part} \rangle}{a^3 \Delta V} = m_\nu \bar{n}_\nu (1+\delta_v(\mathbf{x})) \ , \tag{7.26}$$

is fulfilled (more details of these method can be found in **ayaita_structure_2012**). When neutrinos enter as particles into the N-body simulation and therefore backreaction effects from neutrino structures start becoming important, the calculation of the fields and the potentials becomes computationally demanding, due to the non-linearity of the terms sourcing the continuity 7.7 and Klein-Gordon equations 7.8. Since these equations cannot be linearized due to the large values of the coupling parameter $\beta(\phi)$, the multigrid Newton-Gauß-Seidel solver is of crucial importance. For the parallelization of the code, we use a simple *OpenMP* approach, which calculates in parallel, for the available processing cores, the equations of motion of the particles and the fast Fourier transforms. In Tab.7.3 we describe all parameters related to the N-body simulations, including box and grid size.

7.4 Lump dynamics and the low mass - high mass divide

We find two different regimes for the non-linear evolution of neutrino lumps, depending on the average value of the neutrino mass. For light neutrino masses, during the lump formation process, the neutrinos are accelerated to relativistic velocities. Subsequently, the lumps dissolve and form again periodically, as described in detail in ref. **ayaita_nonlinear_2016**; **baldi_oscillating_2011** We demonstrate this behavior in the left panel of fig. 7.1. The repeated acceleration epochs heat the neutrino fluid to a huge effective temperature, such that neutrinos have again an almost relativistic equation of state during alternating periods of time.

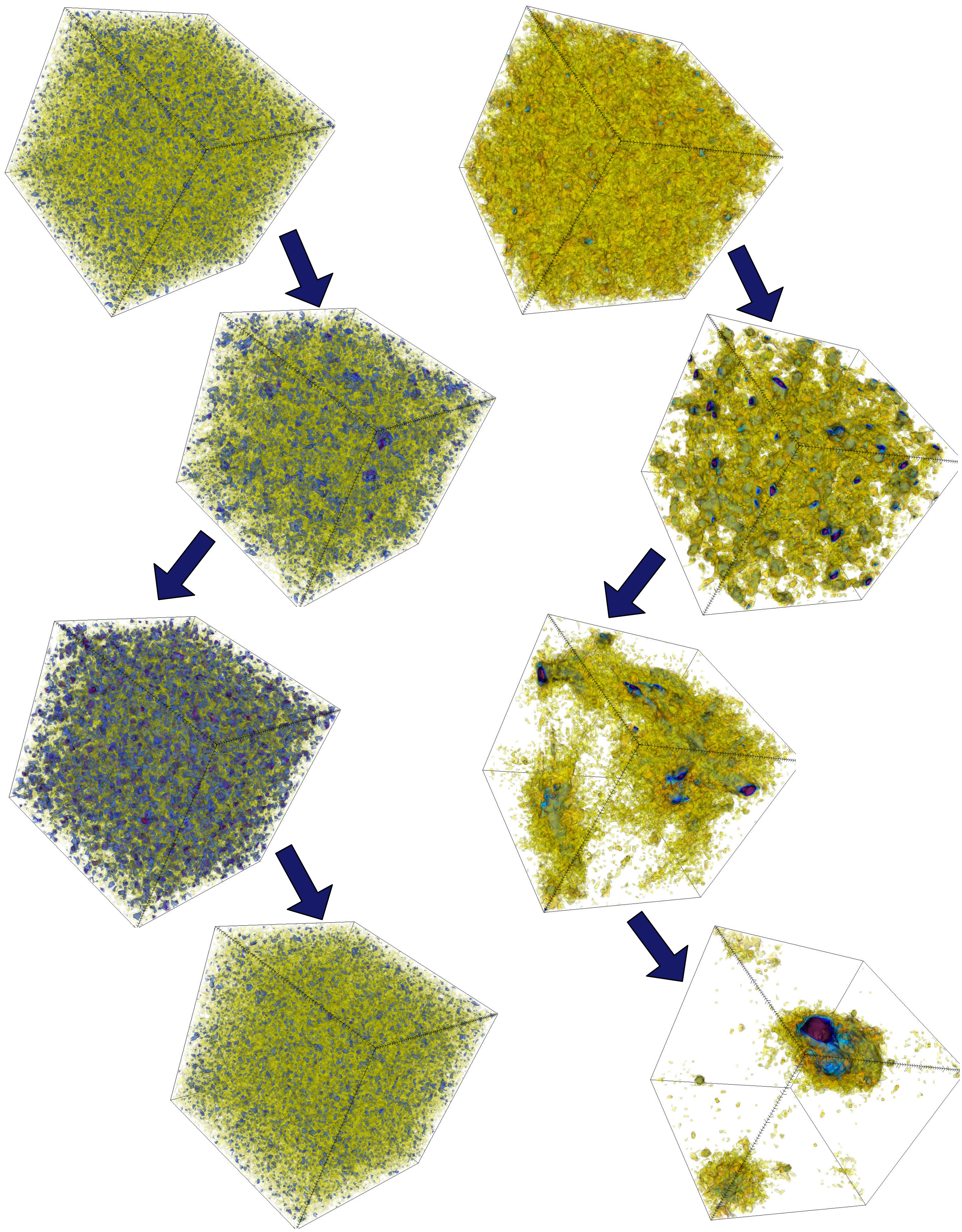


FIGURE 7.1: Snapshots of the number density contrast of neutrinos $\delta n_\nu(\vec{x}) \equiv n_\nu(\vec{x})/\bar{n}_\nu - 1$ at different times. **Left:** Model M2, at scale factors $a = 0.45$, 0.7, 0.75 and 0.95 from top to bottom. The overdensity oscillates between values close to 1 (represented as yellow tones) at early times, where there are no lumps, to values close to 10 (dark blue and purple tones), where several concentrated lumps form at intermediate times. At later times lumps dissolve and the overdensity decreases back to values close to unity. **Right:** Model M4, at scale factors: $a = 0.35$, 0.42, 0.53 and 0.64 from top to bottom. The neutrino lumps start growing at early times and merge progressively into larger and more concentrated structures. At the

In contrast, the behavior for large neutrino masses is qualitatively different. The concentration of the lumps continues to grow after their first formation. Lumps merge, and typically do not dissolve. The neutrino number density contrast reaches high values at late times. This is demonstrated on the right panel in fig.7.1 for an average value of the neutrino mass $m_{\nu,av} = 0.4$eV in the range $0.4 < a < 0.6$. This behavior resembles the one found for a constant cosmon-neutrino coupling in **ayaita_neutrino_2013**; **ayaita_structure_2012**; **baldi_oscillating_2011**

Due to the increasing value of the concentration and the increasing cosmon-neutrino coupling, the characteristic time scale becomes very short and gradients very large. This exceeds the present numerical capability of our simulations, typically at a value of the scale factor somewhat larger than $a = 0.6$. In fig.7.2 we show snapshots for two different values of neutrino masses shortly before the simulation breaks down.

The transition between the "heating regime" for small neutrino masses and the "concentration regime" for large neutrino masses occurs in the range $\langle m_\nu \rangle [0.4 : 0.6] \approx 0.07\text{eV} - 0.14\text{eV}$, where the time average is taken for $0.4 < a < 0.6$. The present value of the neutrino mass can be substantially larger due to oscillations and the continued increase of the mass and the temperature. For example, the phenomenologically viable model with $\langle m_\nu \rangle [0.4 : 0.6] = 0.07$eV corresponds to a present neutrino mass of around 0.08eV, but the time oscillations grow the neutrino mass to values of up to 0.5eV for very short intervals in the scale factor a. (compare with fig. 7.5 below).

In fig.7.1 we show the distribution of the number (over)density contrast $\delta n_\nu(\vec{x}) \equiv n_\nu(\vec{x})/\bar{n}_\nu - 1$ at four different times and for two different models considered here, namely M2 (left panels) and M5 (right panels). For M2 as well as for models with smaller masses (not shown here), the neutrino lumps form and dissolve very quickly. The lumps are never stable and neutrinos accelerate to relativistic velocities when they fall into the gravitational potentials. The small lumps are also distributed homogeneously across the simulation box (see the third panel from above on the right of fig.7.1). The lumps reach maximal number density contrasts of about $\delta n_\nu \approx 10$. For M5 and for bigger masses, the neutrino lumps become stable, accreting more and more particles with the passing of time and increasing their concentrations. This leads to strong backreaction effects, changing the background cosmological evolution. After some time all neutrinos are concentrated in very big lumps, reaching very high values of $\delta_\nu \approx 50 - 100$, where $\delta_\nu \equiv \rho_\nu(\vec{x})/\bar{\rho}_\nu - 1$, see fig.7.2. After this point, the numerical framework for the growing neutrino quintessence evolution breaks down and we can no longer solve reliably the coupled system of equations.

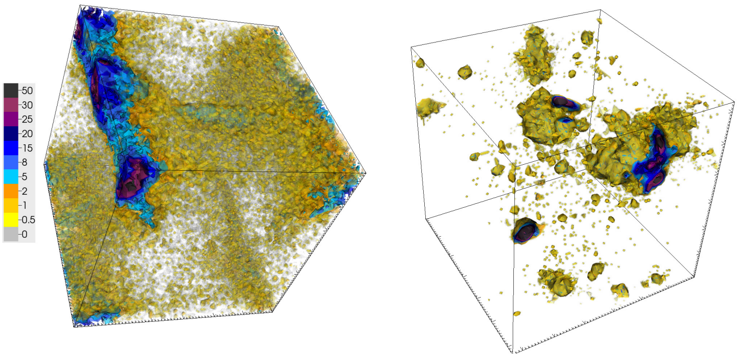


FIGURE 7.2: Snapshots of the neutrino overdensity field $\delta_\nu(\vec{x})$ for models M4 (top left) and M6 (bottom left) at scale factors of $a = 0.64$, and 0.62, respectively. In these models, neutrino lumps cluster into large stable structures with a high concentration, starting from a bottom-up approach, as was shown for model M4 in fig. 7.1. Neutrino structures occupy large parts of the simulation box, corresponding to scales of ~50 Mpc. At these scale factors, the forces introduced by the cosmon coupling are too strong to be resolved by our numerical approach and our simulation breaks down.

7.5 Cosmological evolution in the light neutrino regime

As we have seen in the previous section, there is a qualitative difference between the cosmological evolution of a model with a light or a heavy neutrino mass, the boundary being a present neutrino mass value of roughly ≈ 0.5eV (calculated in linear theory). In this section we explore more in detail the evolution of background quantities in the light mass model M2 whose parameters are shown in detail in table 7.1 for the linear calculation in nuCAMB (top panel) and for the N-body computation (bottom panel). We study in detail the differences appearing in the evolution of background quantities, when non-linear physics and backreaction are taken into account.

The standard definition for the homogeneous energy density fraction of the cosmon field ϕ is

$$\Omega_\phi = \frac{8\pi G}{3H^2}\bar{\rho}_\phi \ , \tag{7.27}$$

where $\bar{\rho}_\phi$ is the background energy density of a homogeneous scalar field $\bar{\rho}_\phi = K(\phi) + V(\phi)$ and $K(\phi)$ its kinetic energy. In linear theory, the homogeneous term would be the only term entering into Ω_ϕ; on the contrary, within the N-body simulation, the field is non-homogeneous and the combined energy density of the coupled neutrino-cosmon fluid receives also a contribution from the perturbations $\delta\rho_\phi$ of the non-homogeneous cosmon field, given by eq.7.18. The important quantity determining the evolution of a dynamical dark energy is not the energy density of the cosmon alone, but the energy density of the combined cosmon-neutrino fluid, given by

$$\Omega_{\phi+\nu} = \frac{8\pi G}{3H^2}\left(\bar{\rho}_\phi + \bar{\rho}_\nu\right) \ . \tag{7.28}$$

The average energy density of the neutrinos is not individually conserved and its evolution is given by the continuity equation with a coupling term on the r.h.s. **ayaita_structure_2012**; **baldi_hydrodynamical_2010**

$$\rho_\nu + 2\mathcal{H}\rho_\nu = -\beta(\phi)\phi'\rho_\nu \ , \tag{7.29}$$

where ϕ' is the time derivative of the field with respect to conformal time τ. The corresponding equation of state of the coupled fluid can then be defined as the sum of the pressure components divided by the sum of the density components

$$w_{\nu+\phi} = \frac{\bar{p}_\phi + \bar{p}_\nu}{\bar{\rho}_\phi + \bar{\rho}_\nu} \ . \tag{7.30}$$

In the literature **brookfield_cosmology_2007**; **das_super-acceleration_2006**; **perrotta_extended_1999**; **perrotta_dark_2002** there are several definitions of the effective equation of state or the observed equation of state in the case in which the scalar field is coupled to other particles. We argue that 7.30 is actually the equation of state one would observe from the evolution of the Hubble function (i.e. with Supernovae and standard candle methods of redshift distance measurements). In appendix 7.10 we comment further on this and show a comparison between the "observed" and theoretical equation of state of dark energy in fig.7.10.

In fig.7.3 we plot for model M2 the background evolution of the neutrino energy density Ω_ν (orange lines) and the combined cosmon+neutrino fluid energy density $\Omega_{\nu+\phi}$ (blue lines) as defined in eq.7.28. The dashed lines correspond to the linear computation in nuCAMB, while the solid lines correspond to the results of the N-body simulation. One can see that the effect of non-linearities and backreaction is quite small and it is mostly just visible as a phase shift in the oscillations of Ω_ν, which is due to the dynamics of the oscillating lumps, that alter the field-dependent mass of the neutrinos as a function of time and space. The same trend is observed in model M1 (not shown here). This behavior tells us that for small neutrino masses, the effects of backreaction on the background evolution are practically negligible and a linear computation is enough to analyze those models further, with a considerable simplification with respect to a joint linear and non-linear analysis done in **pettorino_neutrino_2010**

We show in fig.7.4 the neutrino equation of state w_ν (orange lines) as well as the combined cosmon-neutrino fluid equation of state $w_{\nu+\phi}$ as defined in equation 7.30 (blue lines), both for the case of the linear computation with nuCAMB (dashed lines) and the non-linear computation (solid lines). In the linear analysis, the neutrinos are treated initially as relativistic particles: as the mass increases, they become more and more non-relativistic, reaching a w_ν of exactly zero at late times. On the contrary, the N-body simulation is able to follow the oscillations in the equation of state of neutrinos, which are caused by the fact that neutrinos get accelerated to relativistic velocities when they fall into deep gravitational and cosmon potentials. Once they are in these lumps, and they have acquired high speeds, their pressure increases and they tend to escape again from these lumps, causing the oscillating neutrino structures. When they are far away from the cosmon potentials, their velocities decrease and they become non-relativistic again. The fifth force acting among neutrinos attracts them again to the cosmon potential wells and the whole cycle repeats itself.

For the combined equation of state $w_{\nu+\phi}$, we find that the simulation predicts a slightly higher value than the linear one; this can be explained by studying how the neutrino fluid is heated due to the strong oscillations of the lumps. By falling repeatedly in the cosmon potential wells and increasing their kinetic energy, the neutrinos temperature increases and therefore neutrinos do not manage to become again completely non-relativistic. This can be seen in the dashed orange lines of

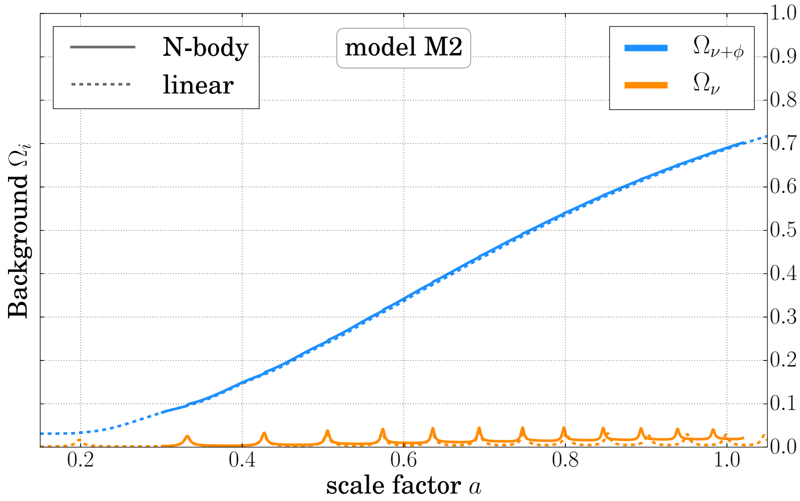


FIGURE 7.3: Evolution of $\Omega_{\phi+\nu}$ (blue lines) and Ω_ν (orange lines) for Model M2, compared between the linear output from nuCAMB (dashed lines) and the non-linear calculation of the N-body simulation (solid lines). The total cosmon-neutrino fluid has the same background evolution in the simulation as in the linear calculation. The neutrino energy density is somewhat larger in the simulation and shows a phase-shift in its oscillations, as discussed in the text.

plot 7.4, where the curve of w_ν does not touch the zero axis after $a \approx 0.5$. We will see in section 7.6 that neutrinos depart from their initial Fermi-Dirac distributions and reach temperatures which are high compared to the photon background.

In fig.7.5, we show the evolution of the spatial average of the neutrino mass in the N-body simulation as a function of the scale factor a. One can see that the value of $\bar{m}_v$ varies along an order of magnitude, from approx.10^{-2} to 10^{-1}, throughout a cosmological time interval. Due to a phase shift in the oscillation pattern, which sets in at around $a \approx 0.8$, the present day value of the average neutrino mass can be quite different to the one estimated with the linear analysis (and this change depends on the precise parameters of the model), so that the best estimate for the average cosmological neutrino mass today, is a time average of $\bar{m}_v(\phi)$ at late times, between $a = 0.8 - 1.0$. We can see that the big discrepancies between the masses of model M1 and M2 calculated in linear theory (e.g. 7.1) are washed away when non-linearities and backreaction effects are taken into account i.e. for small neutrino masses. Even if the present neutrino masses for model M1 and M2 differ by an order of magnitude in linear theory, we find a very similar time averaged value between the two models in the N-body simulation: respectively $\langle m_\nu \rangle [0.8 : 1.0] = 0.120$ and $\langle m_\nu \rangle [0.8 : 1.0] = 0.164$. The oscillation pattern of the neutrino mass for the more massive model (M2) contains higher peaks and has a smaller frequency than compared to the oscillations in the less massive model M1. For other models, this comparison can be seen in table 7.1.

7.6 heating of the neutrino fluid

The repeated acceleration of neutrinos to relativistic velocities during the periods of lump formation and dissolution lead to an effective heating of the neutrino fluid. While we do not expect a thermal equilibrium distribution of neutrino momenta and energies it is interesting to investigate how close the distribution is to the

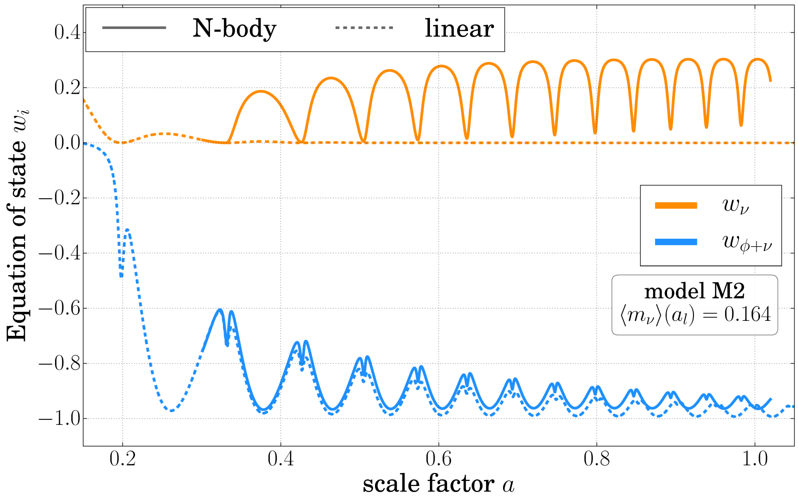


FIGURE 7.4: Equation of state of the combined neutrino-cosmon fluid $w_{\phi+\nu}$ (blue) and equation of state of neutrinos w_ν (orange). We compare the linear output (dashed lines) to the non-linear one obtained from the N-body simulation (solid lines) for model M2. This model has a time averaged RMS mass $\langle m_\nu \rangle(a_l) = 0.164$, where $a_l = 0.9$ in the label denotes the center of the time interval $a = [0.8 - 1.0]$ used to take the average. For w_ν, the linear output does not capture the oscillating equation of state of neutrinos due to the formation of structures, while for $w_{\phi+\nu}$, both codes agree relatively well. At late times the equation of state predicted by the simulation has a somewhat higher value and is phase-shifted due to the heating of the neutrino fluid.

Fermi-Dirac distribution of a free gas of massive neutrinos. This distribution depends on only two parameters, the neutrino mass and the temperature. At a given time we associate the neutrino mass to the space averaged neutrino mass. The temperature can be associated to the mean value of the momentum.

The energy of a relativistic particle is given by

$$E(p, m) = \sqrt{p^2 + m^2} \;, \tag{7.31}$$

while its kinetic energy $E_k = E(p, m) - m$. Equivalently, the kinetic energy is defined by

$$E_k = \int \vec{v} \cdot \mathrm{d}\vec{p} \;, \tag{7.32}$$

which yields $E_k = m(\gamma - 1)$ and reduces in the limit of very small velocities ($v \lll c$) to the usual $E_k = mv^2/2$. From there the Fermi-Dirac distribution as a function of momentum $p = |\vec{p}|$ can be obtained in the standard way. It depends on m and T. For $m \ll T$ it can be approximated by the relativistic distribution while for $m \gg T$ we recover the Maxwell-Boltzmann distribution. The distribution of particle momenta is then given by

$$\mathcal{P}(p)\mathrm{d}p = \frac{4\pi p^2}{\left(1 + e^{(E(p,m)-\mu)/T}\right)} \mathrm{d}p \tag{7.33}$$

where the factor 4π comes form the angular integration of the three-dimensional

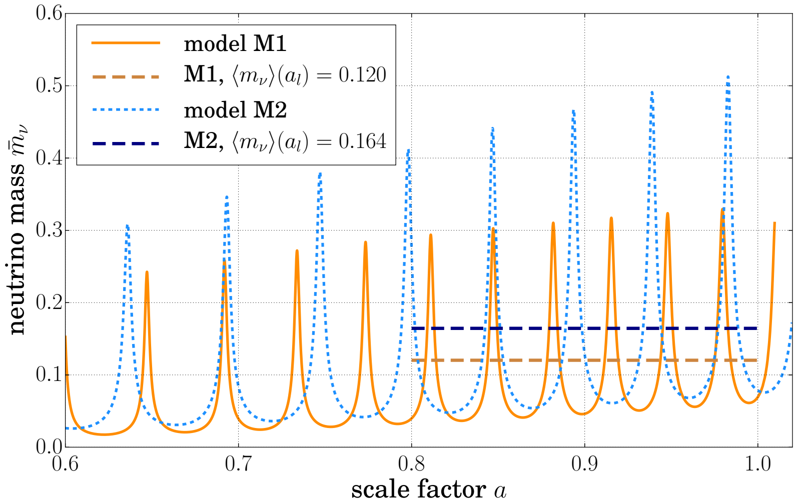


FIGURE 7.5: Neutrino mass $\overline{m}_\nu$ (average over simulation volume) in model M2 (firebrick red line) and model M1 (dodger blue line), as a function of the scale factor a, for the N-body simulation. The horizontal lines, show the time averaged RMS value at a late time $a_l = 0.9$, denoting the center of the time interval of $[0.8 : 1.0]$ considered for taking the average. For model M2, the time averaged neutrino mass is $\langle m_\nu \rangle(a_l) = 0.164$ (blue dashed lines), while for model M1 it is somewhat smaller $\langle m_\nu \rangle(a_l) = 0.120$ (red dashed lines). One can observe that the oscillation frequency is higher for the smaller $\langle m_\nu \rangle$ mass and the peaks are higher for the larger $\langle m_\nu \rangle$. The present neutrino masses of the two models calculated in linear theory differ by an order of magnitude, on the contrary the time averaged masses are very close to each other.

momentum. In the ultra-relativistic limit we can analytically integrate the momentum p over its distribution eq.7.33 and invert $\bar{p}(\overline{T})$ to yield the mean temperature as a function of the mean momentum

$$\bar{T} = \frac{180\zeta(3)}{7\pi^4}\bar{p} \tag{7.34}$$

We neglect the chemical potential in 7.33, because the exponential term in the denominator is 2 or 3 orders of magnitude larger than unity. Since in our case, the average momentum and mass of the neutrinos are of the same order, we cannot use either a non-relativistic or an ultra-relativistic limit. We need to consider both the mass and the momentum in the relativistic energy equation 7.31. Therefore for each model and each time, we numerically find $\overline{T}$ as a function of the mean momentum $\overline{p}$.

We extract for $a = 1$ the temperatures

$$\overline{T} = 0.077\text{eV (M1}, \bar{m}_\nu = 0.2404), \ \overline{T} = 0.065\text{eV (M2}, \bar{m}_\nu = 0.2327)\ . \tag{7.35}$$

They are higher by a factor 327 (M1) or 276 (M2) as compared to the CMB photon temperature 2.35×10^{-4}eV. This demonstrates the unconventional heating of the neutrino fluid due to the formation and dissolution of lumps. The high temperatures are connected with the almost relativistic equation of state of the neutrinos seen in fig. 7.4. Overall, the observed momentum distributions come rather close to the thermal equilibrium distribution. This also holds for the distribution of kinetic

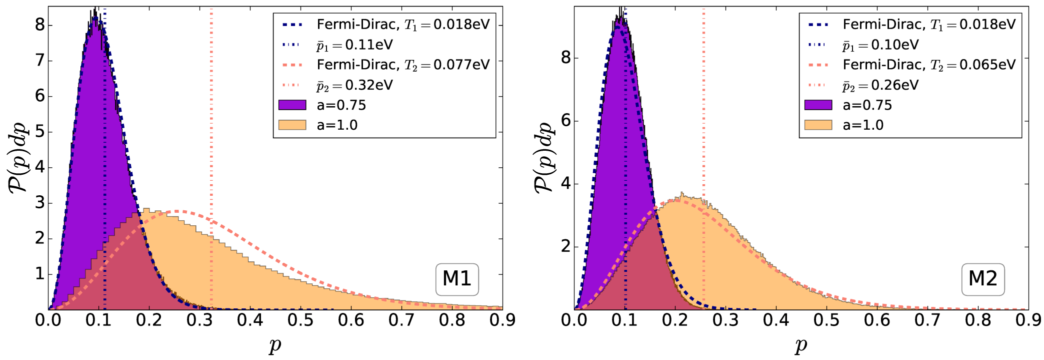


FIGURE 7.6: Distribution of the momenta of the neutrino particles in the simulation, for two different times, $a = 0.75$ (purple shade) and $a = 1.0$ (orange shade), compared against a Fermi-Dirac distribution with a temperature given by the mean of the distribution (dashed lines). **Left:** For model M1 the Fermi-Dirac fits very well for temperatures of $\bar{T} = 0.018$eV and $\bar{T} = 0.077$eV for each scale factor respectively. **Right:** For model M2 the fit is also good, the corresponding temperatures being $\bar{T} = 0.018$eV and $\bar{T} = 0.065$eV. The CMB photon temperature is 2.35×10^{-4}eV, this means that the non-linear cosmon-neutrino interactions heat the neutrino background by more than a factor 100.

energies. With the bulk quantities as momenta and kinetic energies roughly distributed thermally this is an example of prethermalization **berges_prethermalization_2004**

In fig. (7.6) we fit the distribution of momenta of the neutrino particles on the grid (shown with an histogram) with a Fermi Dirac distribution. The actual distribution of momenta fits the thermal equilibrium distribution very well. At later times (orange shade), the fit is slightly less good: neutrinos might be accelerating towards or away from lumps giving them an extra kick that shifts the peak of the distribution of momenta.

When comparing the equation of state of neutrinos obtained from the N-body simulation to a neutrino equation of state $w_\nu = p_\nu/\rho_\nu$, using our Fermi-Dirac fit to the particle distribution and eqns.7.15 and 7.14, we get a very good agreement, taking into account that for the Fermi-Dirac fit, we are neglecting the spatial variation of the neutrino mass $m_\nu(\phi)$. For model M2 at the scale factor $a = 0.75$ we obtain from the N-body simulation a neutrino equation of state of $w_\nu = 0.081$ while using the Fermi-Dirac fit to the distribution of particles with a mean temperature of $\bar{T} = 0.018$eV and an average neutrino mass $\bar{m}_\nu = 0.1835$, the proper calculation yields $w_\nu = 0.086$. For a later time, at $a = 1.0$ the N-body simulation gives us a value of $w_\nu = 0.207$ while the Fermi-Dirac fit with a mean temperature of $\bar{T} = 0.065$eV and an average neutrino mass $\bar{m}_\nu = 0.2327$ amounts to a neutrino equation of state of $w_\nu = 0.182$.

To visualize the evolution of w_ν, we can observe from fig.7.4 that neutrinos in the N-body simulation start as non-relativistic particles and oscillate between being almost relativistic and completely non-relativistic in the interval $a = [0.3, 0.6]$. However, at later times $a \gtrsim 0.7$ the neutrino equation of state still oscillates but never reaches a value of zero again. This is in agreement with our description of the heating of the neutrino fluid. Since the mean temperature of the neutrino fluid is increasing with time and therefore its mean kinetic energy and pressure, the minimum of the oscillations of the neutrino equation of state increases also in time and departs from zero, once neutrinos are heated to very high temperatures

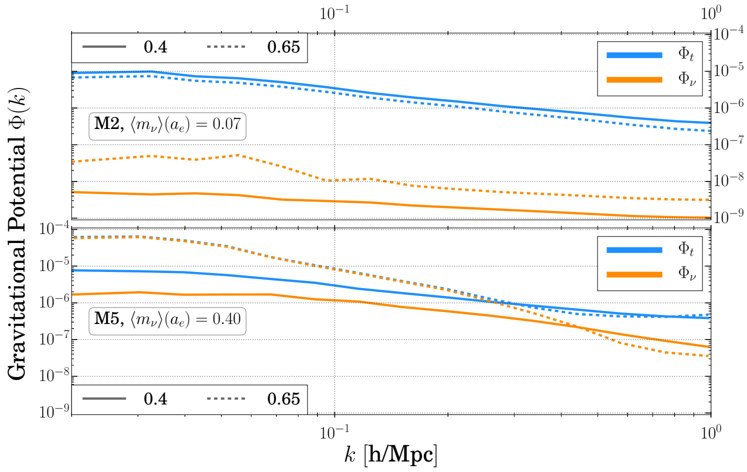


FIGURE 7.7: Power spectra of the total gravitational potential Φ_t (blue lines) and of the neutrino contribution Φ_ν (orange lines) for model M2 and model M5 at the scale factors $a = 0.40$ (solid lines) and $a = 0.65$ (dashed lines), as a function of scale. Model M2 has an RMS time averaged neutrino mass $\langle m_\nu\rangle(a_e) = 0.07$, where $a_e = 0.5$ stands for the central time in the interval $a = [0.4 - 0.6]$ used to take the average. Model M5 has in the same interval a higher RMS mass of $\langle m_\nu\rangle(a_e) = 0.40$. In the first model, Φ_t at large scales is of the order of 10^{-5}, while Φ_ν is 3 to 4 orders of magnitude smaller at both cosmological times. For model M5, in which neutrino lumps are stable and growing, one sees that at large scales, the total Φ_t starts with a value of 10^{-5} at $a = 0.4$, but reaches 10^{-4} at later times. At $a = 0.65$ the neutrino contribution is dominant and neutrino structures have migrated from small scales to large scales, as can be seen from the dip in Φ_ν at modes between $k = 0.2 - 1.0$ h/Mpc.

due to the collapsing and dissolving of the neutrino-cosmon lumps.

7.7 Gravitational Potentials of Neutrino Lumps

The gravitational potential Φ is a good measure of the physics going on in structure formation. We know from observational constraints, that Φ is of the order of 10^{-5} on cosmological scales **pettorino_neutrino_2010**; **brouzakis_nonlinear_2011** In ΛCDM, the gravitational potential is sourced mainly by dark matter perturbations. In figures 7.7 and 7.8 we show that for models with small neutrino masses, the neutrino contribution to Φ remains several orders of magnitude smaller than the CDM contribution, at all scales and at all times. Moreover, one can observe an oscillation in time of the neutrino gravitational potential. For models with large neutrino masses, the neutrino contribution grows monotonically with time. At large scales $k \lesssim 0.3$h/Mpc and at late times the neutrino lump induced potential dominates over the cold dark matter gravitational potential. This renders the total potential Φ_{tot} too big to be compatible with present cosmological constraints.

We show the scale dependence of the total gravitational potential and the neutrino induced gravitational potential Φ_ν at two different cosmic time scales $a = 0.4$ and $a = 0.65$ in fig.7.7. While for $a = 0.4$ (solid lines) the neutrino contribution is still subdominant for both models M2 and M5, this changes at $a = 0.65$ for model M5 (bottom panel). For model M2 the total gravitational potential decreases in

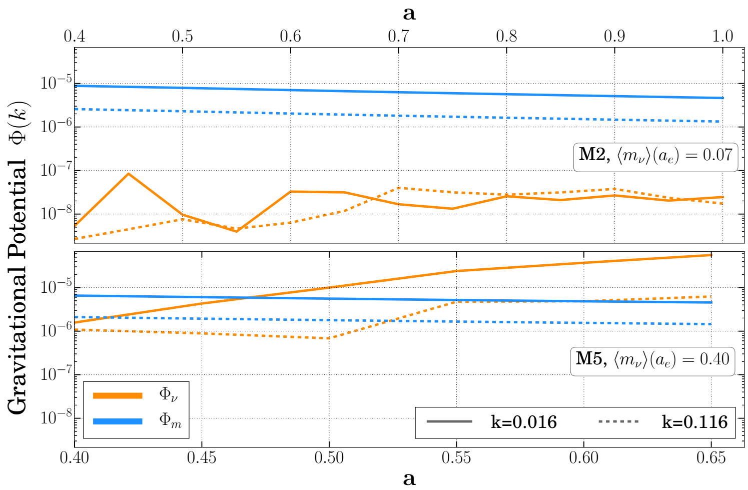


FIGURE 7.8: Power spectra of the matter gravitational potential Φ_m (blue lines) and of the neutrino contribution Φ_ν (orange lines) at two different scales, $k = 0.016$ (solid lines) and $k = 0.116$ (dashed lines) as a function of scale factor a with two different scales on the top and bottom axis. For model M2 the matter gravitational potential Φ_m is at most 10^{-5} at all times, while the neutrino contribution is 2-3 orders of magnitude smaller and displays time oscillations. In clear contrast, the neutrino contribution from model M5 for very large scales, reaches and dominates over the matter contribution for $a \gtrsim 0.5$ and pushes the total Φ to high values that would be ruled out by observations, see also fig. 7.7. The RMS neutrino mass has been taken in the same interval range as for fig.7.7, where $a_e = 0.5$.

time, as it is expected due to the effect of dark energy, while Φ_ν increases especially at large scales. For model M5, since the neutrino contribution dominates at $a = 0.65$, the total gravitational potential is raised to values of 10^{-4} at large scales, while at small scales ($k \gtrsim 0.4$) the neutrino contribution is still subdominant.

In fig.7.8 we show the power spectra of the total gravitational potential Φ (blue lines) and of the neutrino contribution to it (orange lines) at two different scales, $k = 0.016$ (solid lines) and $k = 0.116$ (dashed lines) as a function of the scale factor a. For model M2, corresponding to an average early time neutrino mass of 0.07 eV, the total gravitational potential $\Phi(k)$ is 10^{-5} at all times, while the neutrino contribution is 2-3 orders of magnitude smaller and shows time oscillations. In clear contrast, the neutrino contribution from model M5 (corresponding to an average early time neutrino mass of 0.40 eV)) for very large scales reaches and dominates over the matter contribution (at $a \gtrsim 0.5$) and pushes the total Φ to high values that would be ruled out by observations. This is due to the fact that neutrino lumps do not dissolve, but rather grow with continuously growing concentration and higher gravitational potential.

There is also anticorrelation between the neutrino structures and the neutrino induced gravitational potential, as expected from the fact that neutrinos will tend to fall into gravitational potential wells. In the left panel of fig.7.9, we plot the values of the neutrino number density contrast $\delta n_\nu = n_\nu(\vec{x})/\bar{n}_\nu - 1$ and the negative neutrino induced gravitational potential Φ_ν, along a diagonal line through the simulation box. The correlation of peaks and troughs (corresponding to an anticorrelation of δn_ν and Φ_ν) is very clear and it is valid for even small substructures of the order of a few Mpc. By plotting the neutrino density contrast δ_ν, we also show

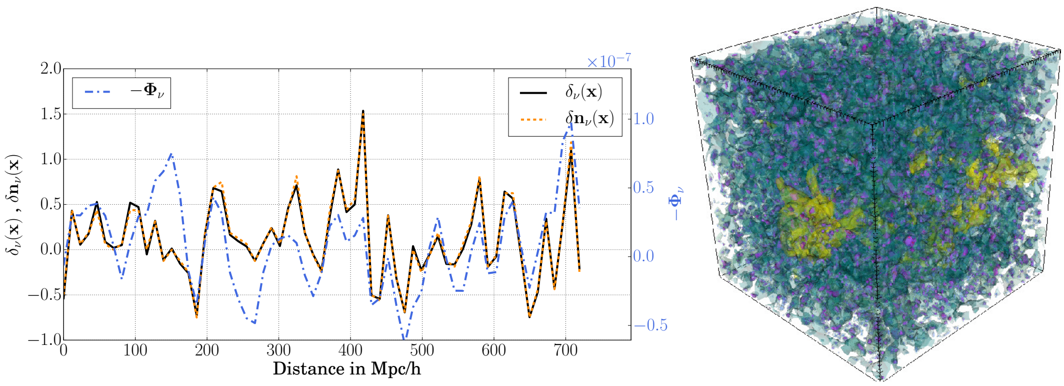


FIGURE 7.9: **Left:** Line plot through the main diagonal of the simulation box, for model M2 at $a = 0.75$. The negative of the neutrino contribution to the gravitational potential Φ_ν (blue dot-dashed lines) oscillates between $\pm 1.0 \times 10^{-7}$. This is correlated to the neutrino density contrast (black lines) and the number density contrast $\delta n_\nu = n_\nu(\vec{x})/\bar{n}_\nu - 1$ (orange dashed lines) reaching values of up to 1.5. **Right:** Snapshot of the same simulation, showing a equipotential contour of the gravitational potential, for $\Phi_\nu = +1.0 \times 10^{-7}$ in yellow, and the small but dense neutrino lumps in blue, purple and red, corresponding to density contrasts δn_ν of 1.5, 2.0 and 4.0 respectively.

that at this time $a = 0.75$, the neutrino number density and the energy density are proportional, meaning that neither local mass variations or relativistic speeds are having any effect in the neutrino total energy. In the right panel of fig.7.9 we visualize the neutrino induced gravitational potential as a yellow region marking the equipotential surface $\Phi_\nu = +1.0 \times 10^{-7}$ and the neutrino number overdensity structures colored blue, purple and red, corresponding to density contrasts δn_ν of 1.5, 2.0 and 4.0 respectively. For this model (M2) and at this specific time, the neutrino structures are spread almost homogeneously throughout the simulated volume.

7.8 Conclusions of Dynamics of Neutrino Lumps

We have investigated the dynamics of neutrino lumps in growing neutrino quintessence and how it depends on the mass of neutrinos. As a main result of this paper we found a characteristic divide in the qualitative behavior between small and large neutrino mass.

For light neutrino masses the combined effects of oscillations in the neutrino masses and the cosmon-neutrino coupling lead to rapid formation and dissociation of the neutrino lumps. The concentration in the neutrino structures never grows to very large overdensities. As a consequence, backreaction effects remain small. The effects of lump formation and dissociation lead to an effective heating of the neutrino fluid to temperatures much higher than the photon temperature. Due to this heating, the neutrino equation of state becomes again close to the one for relativistic particles. For a small present average neutrino mass $m_\nu = 0.06$eV it has been found earlier **pettorino_neutrino_2010** that the cosmology of growing neutrino quintessence resembles very closely a cosmological constant, making differences to the ΛCDM model difficult to detect. We extend this qualitative feature to a whole range of light neutrino masses.

For large neutrino masses, one finds a qualitatively different behavior. Big neutrino lumps form, due to the strong cosmon-mediated fifth-force between neutrinos. These lumps are stable and keep growing in concentration and density. The strong clumping of the cosmic neutrino background induces large backreaction effects on the overall cosmic evolution. As a result, the combined cosmon-neutrino fluid does not act effectively as a cosmological constant anymore and compatibility with observations is difficult to achieve. This situation is similar to the case of a constant cosmon-neutrino coupling **fuhrer_backreaction_2015**

The divide in the characteristic behavior reflects the competition between heating of the neutrino fluid and lump concentration. We have not yet established a quantitatively accurate value of the parameter $\hat{m}_\nu$ where the divide is located, since the numerics are rather time consuming. In principle, this divide will lead to an upper bound on the present neutrino mass, as seen in terrestrial experiments. For models in the vicinity of model M2, which seem compatible with observations so far, spatial average neutrino masses as large as 0.5eV can occur at the peak of oscillations, c.f. fig.7.5. We note that if we live inside a neutrino lump the neutrino mass will be reduced as compared to the cosmological value.

We have further computed the strength of the neutrino-induced gravitational potential. For light masses, this potential is found to be rather small, rendering a detection of the neutrino lumps difficult. As neutrino masses increase towards an intermediate mass region, before reaching the heavy mass range incompatible with observation, the neutrino-induced gravitational potentials will get stronger. By continuity we expect that in the intermediate mass region the clumped neutrino background becomes observable.

7.9 Initial parameters for generating the models in nuCAMB and in the N-Body simulation

CAMB values	M1	M2	M3	M4	M5	M6
input: $\Omega_\nu h^2$	0.048	0.048	0.075	0.018	0.038	0.098
$\tilde{m}_\nu$ amplitude factor	8.35×10^{-5}	2.0×10^{-4}	6.0×10^{-4}	9.9×10^{-3}	8.8×10^{-3}	8.8×10^{-3}
input $\hat{r}_{\nu eV}$ factor	1.5045	1.5045	2.3508	0.5642	1.1911	3.0718
V_i	0.99×10^{-7}	0.99×10^{-7}	0.99×10^{-7}	0.99×10^{-7}	0.99×10^{-7}	0.99×10^{-7}

TABLE 7.2: Table of initial parameters for each model computed with nuCAMB. $\tilde{m}_\nu$ is the neutrino mass amplitude used in the simulations, $\hat{r}_{\nu eV}$ is the neutrino mass units conversion factor between the simulations and nuCAMB. V_i is the initial value of the cosmon potential.

N-body parameters	Values	Meaning
L	428	Box side length in Mpc/h
N_g	64	Grid size per dimension
N_{pdm}	262144	Number of dark matter particles in the simulation
$N_{p\nu}$	524288	Number of neutrino particles in the simulation
ngs_{acc}	1.0×10^{-5}	Numerical accuracy for the NGS solver

TABLE 7.3: Table of numerical parameters for each N-body simulation run. Several tests were performed varying these parameters, but the overall behavior for the purposes of this paper was the same. We also tested model M2 with grid sizes of 128 and 8 times the number of particles, not noticing any qualitative difference in the dynamics. Only deviations of the order of 10% in perturbation quantities were observed when varying the grid size or the number of particles.

7.10 The effective and observed equation of state of dark energy

The equation of state w of a species is defined in terms of its pressure and density as

$$w = \frac{p}{\rho} \ . \tag{7.36}$$

For the case of coupled species, there have been several definitions in the literature (c.f. **brookfield_cosmology_2007**; **das_super-acceleration_2006**; **perrotta_extended_1999**; **perrotta_dark_2002**). This is due to the fact that when there is an exchange of energy and momentum between a particle and a scalar field, the time evolution of matter does not correspond anymore to the volume dilution rule $\rho(a) = \rho_0 a^{-3}$. Therefore the equation of state of the dark energy field will not be as simple as the equation of state of a homogeneous scalar field, namely

$$w_\phi = \frac{p_\phi}{\rho_\phi} = \frac{(1/2a^2)\dot{\phi}^2 + V(\phi)}{(1/2a^2)\dot{\phi}^2 - V(\phi)} \ . \tag{7.37}$$

In **das_super-acceleration_2006**; **brookfield_cosmology_2007** two definitions of the equation of state (e.o.s) of dark energy were investigated. The first one, called the effective e.o.s. w_{eff}, given by

$$w_{eff} \equiv w_\phi + \frac{\beta\dot{\phi}}{3H}\frac{\rho_\nu}{\rho_\phi} \tag{7.38}$$

and the second one named "apparent" e.o.s. w_{ap}, defined by

$$w_{ap} \equiv \frac{w_\phi}{1+x} \ , \tag{7.39}$$

with

$$x = -\frac{\rho_{\nu,0}}{a^3\rho_\phi}\left[\frac{m_\nu(\phi)}{m_\nu(\phi_0)} - 1\right] \ , \tag{7.40}$$

where a 0 subscript denotes quantities at $z = 0$. Notice that by construction, w_{ap} at the present epoch is identical to w_ϕ and that w_{ap} can be smaller than -1. We find that none of these two definitions of e.o.s. describe the dynamical dark energy field present in our model.

Since in our model the neutrino and cosmon field behave together as a tightly coupled fluid, we are interested in the conserved equation of state for the combined fluid. This is given by the sum of both contributions from the pressure, divided by the sum of both densities. Therefore, we define $w_{\nu+\phi}$ as

$$w_{\nu+\phi} \equiv \frac{\bar{p}_\phi + \bar{p}_\nu}{\bar{\rho}_\phi + \bar{\rho}_\nu} \ . \tag{7.41}$$

This definition should agree with what we can extract directly from observations of the Friedmann equation. The equation of state $w_{DE}(z)$ of a general dark energy component that fulfills the continuity equation (if it is composed by two coupled fluids, the continuity equation is fulfilled for the sum of densities and pressures), appears in the Friedmann equation as

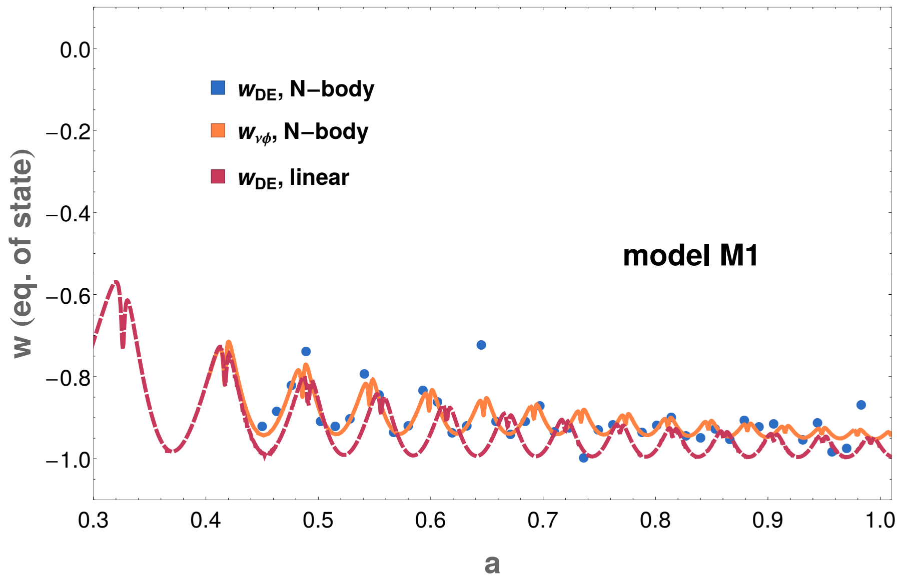


FIGURE 7.10: Different $w(z)$ curves in the case of model M1 computed with the N-body simulation and compared to linear theory. The red dashed lines represent w_{DE} computed from the Hubble function which is an output of the linear code nuCAMB. The solid orange line, representing $w_{\nu+\phi}$ is computed using the standard pressure and density outputs from the simulation. The blue dots stand for some set of computed values of $w_{DE}(z)$ obtained from the Hubble function calculated entirely within the N-body simulation. Due to numerical noise in the oscillating derivatives of $E(z)$, there is some scatter in the e.o.s. obtained in this case.

$$E^2(z) \equiv \frac{H^2(z)}{H_0^2} = \left[\Omega_{r,0}(1+z)^4 + \Omega_{m,0}(1+z)^3 + \Omega_{DE,0}\exp\left\{\int_0^z \frac{3(1+w_{DE}(\tilde{z}))}{1+\tilde{z}}\mathrm{d}\tilde{z}\right\} + \Omega_{k,0}(1+z)^2\right] \tag{7.42}$$

Moreover, in a flat Universe and with a negligible contribution from radiation, we can solve for $w_{DE}(z)$ and obtain **amendola_dark_2010**

$$w_{DE}(z) = \frac{(1+z)(E^2(z))' - 3E^2(z)}{3\left[E^2(z) - \Omega_{m,0}(1+z)^3\right]} \ . \tag{7.43}$$

Thus, from background expansion observations we can obtain a constrain on $w_{DE}(z)$, provided we know from large-scale structure or CMB observations the present value of Ω_m. In the fig. 7.10 we compare $w_{DE}(z)$ obtained from eq.7.43 and $w_{\nu\phi}(z)$ computed both consistently within the N-body simulation. In the former case, numerical noise in the derivatives of $E(z)$ create certain scatter in $w_{DE}(z)$ at late times.

Chapter 8

Conclusions

To do...

☐ 1 (p. 24): Comp. Add derivation of this equation

Appendix A

Appendix Title Here

Write your Appendix content here.