UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Astrophysics and Cosmology

Final dissertation

Investigation of the ShapeFit analysis for future galaxy surveys

Thesis supervisor	**Candidate**
Alvise Raccanelli	**Samuele Magro**

Academic Year 2023/24

Contents

Introduction **3**

Premise: cosmological context **5**

1 Constructing the observed power spectrum **13**
- 1.1 True linear matter power spectrum . . . 14
  - 1.1.1 Perturbation theory . . . 14
  - 1.1.2 The power spectrum . . . 17
  - 1.1.3 Primordial power spectrum, transfer function and normalisation . . . 17
  - 1.1.4 Power spectrum shape . . . 20
- 1.2 Observed linear galaxy power spectrum . . . 24
  - 1.2.1 Alcock-Paczyński effect . . . 24
  - 1.2.2 Galaxy bias . . . 26
  - 1.2.3 Poisson shot noise . . . 27
  - 1.2.4 Redshift-space distortions . . . 28
  - 1.2.5 Redshift uncertainty . . . 31
- 1.3 Full linear model of the observed galaxy power spectrum . . . 32
- 1.4 Non-linear model of the observed galaxy power spectrum . . . 33
- 1.5 Multipole expansion . . . 35

2 Power spectrum as a forecasting tool **36**
- 2.1 Fisher matrix formalism: general description . . . 37
- 2.2 Fisher matrix formalism: application to galaxy surveys . . . 41

3 Cosmology with ShapeFit **49**
- 3.1 Classic approach: overview . . . 49
  - 3.1.1 Classic BAO analysis . . . 50
  - 3.1.2 Classic RSD analysis . . . 52

3.2 Full Modelling approach 54
3.3 Classic approach: pipeline in detail 55
    3.3.1 Model pipeline 55
    3.3.2 Data pipeline 56
    3.3.3 Compression step 59
    3.3.4 Cosmology inference step 59
3.4 Introducing ShapeFit 60
    3.4.1 Theoretical parameter dependencies 60
    3.4.2 Slope rescaling and slope parameters 65
    3.4.3 Variables complementarity 70
3.5 ShapeFit pipeline 71
    3.5.1 Computational advantage 73

4 Application to mocks **77**
4.1 ShapeFit vs Classic RSD 79
    4.1.1 N-body mocks analysis 81
4.2 ShapeFit vs FM 85
    4.2.1 Fixing $\boldsymbol{w_b}$ 85
    4.2.2 Varying $\boldsymbol{w_b}$ 86
    4.2.3 Varying $\boldsymbol{n_s}$ 87

5 Application to SDSS-III BOSS data **97**

Conclusions **101**

Appendix A **103**

Appendix B **105**

Appendix C **109**

Appendix D **111**

References **113**

Introduction

One of the most distinctive traits of our species is curiosity, and the will to keep asking questions. Since the dawn of human History, this characteristic of us has surely provided many advantages in the context of survival: asking questions leads to gathering knowledge about our surroundings, and knowledge allows to evaluate multiple courses of action and choose for the best and most convenient one. But apart from the context of a practical earthly life, outlined by the necessity to fulfill our most basic needs, our curiosity has often condensed into more profound, deep and yet thrilling questions. For thousands of years, we human beings have pursued the larger-than-life goal of understanding the nature of the Universe in which we happen to live: its origins, its structure and internal organization, and its eventual death. Over time, such visceral questions have become more refined, in concomitance with the advancements made in the field of physics and astronomy, and pushed by the emergence of the scientific method. All of this has eventually led to the development of cosmology: this is the science aimed at understanding the origin, overall structure and evolution of the Universe. As all sciences, cosmology has undergone a constant evolution through the years, propelled by the affirmation of Einstein's Theory of General Relativity, and then fed by the ever-increasing amount of data coming from galaxy surveys. These surveys have witnessed a constant improvement throughout the last century, and the next generation of surveys coming in the next years will delight us with strikingly precise performances. Among such surveys, we can cite the Euclid Telescope[1], which was launched in July 2023, the Dark Energy Spectroscopic Instrument[2] (DESI), the Large Synoptic Survey Telescope[3] (LSST), the Wide Field Infrared Survey Telescope[4] (WFIRST), the Spectro-Photometer for the History of the Universe, Epoch of Reionization and Ices Explorer[5] (SPHEREX), the Astrophysics Telescope for Large Area Spectroscopy Probe (ATLAS)[6], the

---

[1] http://www.euclid-ec.org/ (Amendola et al., 2018)

[2] http://desi.lbl.gov (Desi et al., 2016)

[3] https://www.lsst.org (Ivezić et al., 2019)

[4] https://wfirst.gsfc.nasa.gov (Spergel et al., 2015)

[5] https://www.jpl.nasa.gov/missions/spherex (Doré et al., 2014)

[6] (Wang et al., 2019).

Nancy Grace Roman Telescope[7]. We have finally entered the so-called age of ”precision cosmology”, characterised by new and more sophisticated instruments for our exploration of the Large Scale Structure of the Universe. Such increasing instrumental precision should be mirrored by equally powerful methods of statistical analyses, aiming at extracting as much information as possible.

One of the newest methods is ShapeFit, and it was introduced as a blending of pre-existing approaches, namely the Classical and Full Modelling ones. The idea is to take the best features from both predecessors (model-independence, tight constraints), while overcoming their respective limitations (e.g. computational cost).
The present work aims to introduce the motivations behind this new technique, outline the novelties of its pipeline, and finally demonstrate its advantages with respect to previous techniques, by direct application to simulated mocks and real BOSS data. In particular, when both ShapeFit and Full Modelling are applied to mock data, the former is able to better recover the expected parameter values, while featuring the same precision as the latter, and while also requiring a considerably lower computational effort. Furthermore, the application of both techniques to BOSS data shows that ShapeFit benefits from a greater robustness against systematic errors.

The work is structured as follows. In chapter 1, we will define and then model the galaxy observed power spectrum in a step-by-step fashion. Chapter 2 shows a way to employ the power spectrum as a predictive tool to perform cosmological inference. In chapter 3, we introduce and justify the ShapeFit methodology, highlighting its novelty in comparison with the older Classic and Full Modelling approaches. In chapter 4 we present the result of applying different techniques to some mock data: on the one hand, this will further highlight how such approaches perform differently; on the other hand, such analysis will show the emergence of some systematic shifts for ShapeFit. Finally, in chapter 5 we will briefly have a look at how ShapeFit fares when applied to actual BOSS data.

---

[7] https://roman.gsfc.nasa.gov (Wang et al., 2022)

Premise: cosmological context

In this early section of the work, we outline the basic cosmological framework which will be employed from now on in the present discussion. In what follows, all quantities are expressed in natural units, i.e. $c = 1$.

Following the Cosmological Principle assumption, stating that every point in space is equivalent to any other point, regardless of direction, the Universe on large scales will be treated as homogeneous and isotropic. This assumption is justified by observational evidence: when exploring the distribution of matter in the Universe, prominent fluctuations from the mean are present, but they are confined to scales smaller than a few tens of megaparsec; whereas, the average over such fluctuations reveals a Universe which actually looks isotropic and homogeneous(P. J. E. Peebles, 1993). Therefore, the distribution of matter in the Universe appears to follow that of a statistically homogeneous and isotropic random density field.

This corresponds to the Friedmann-Robertson-Walker (FRW) line element (Robertson, 1935, 1936):

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

where (t, r, θ, ϕ) are comoving coordinates, while $a(t)$ is the cosmic scale factor and the spatial curvature parameter K can be chosen to be $+1$ (positive spatial curvature), -1 (negative spatial curvature) or 0 (zero spatial curvature)[8]. The scale factor is the only dynamical variable upon which the FRW line element depends, attesting the high degree of symmetry of such metric (Kolb, 2018). The meaning of the scale factor can be further visualized in *Figure* 1: we can picture space as a grid which uniformly expands as time evolves. Points on the grid maintain

---

[8] Throughout the present work, the $K = 0$ case, corresponding to a flat Universe will be employed.

their coordinates, so the comoving distance between two points (i.e. simply the difference between the two coordinates) doesn't change. Meanwhile, the physical distance is proportional to the scale factor, so it evolves with time (Dodelson & Schmidt, 2020).

Formally, the comoving and physical distance between and object and an observer are defined as

$$r_{\text{comoving}} = \int_{t_{\text{e}}}^{t_{\text{o}}} \frac{dt}{a(t)} \tag{2}$$

$$r_{\text{physical}} = a(t) \int_{t_{\text{e}}}^{t_{\text{o}}} \frac{dt}{a(t)} \tag{3}$$

where t_{e} indicates the time at which a photon is emitted by the object, and t_{o} is the time at which the same photon is received by the observer.

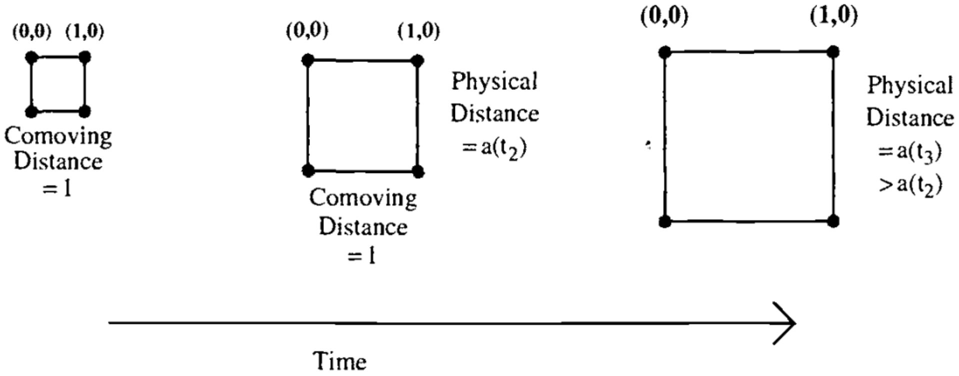

Figure 1: Expansion of the universe. The comoving distance between points on a hypothetical grid remains constant as the universe expands. The physical distance is proportional to the comoving distance times the scale factor, so it gets larger as time evolves. The picture is taken from chapter 1.1 of (Dodelson & Schmidt, 2020)

As we know, the Universe is not empty: the spacetime described by the FRW metric is inhabited by a variety of energy sources, both massive (like baryons or galaxies) and massless (like photons). These two actors on stage, spacetime and energy sources, play a game of mutual interaction, constantly shaping and modifying each other. On the one hand, the form of the metric determines how energy sources will move inside such metric; on the other hand, the mere presence of energy sources will bend, shrink and stretch the metric itself. The relation between metric curvature and energy sources is described by the Einstein Field Equations (Einstein, 1915; Einstein et al., 1916):

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{4}$$

The metric curvature is represented, on the left hand side, by the Einstein Tensor

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{R}{2} g_{\mu\nu} \tag{5}$$

where:

- $R_{\mu\nu}$ is the Ricci tensor, depending on the metric and its derivatives[9];

- $g_{\mu\nu}$ is the metric tensor;

- the contraction $R \equiv g^{\mu\nu} R_{\mu\nu}$ is the Ricci scalar.

On the right hand side on the Einstein Field Equations, sources of energy density are represented by the energy-momentum tensor $T_{\mu\nu}$, describing the energy and momentum of the spacetime matter content, while G represents Newton's gravitational constant. The content of the Universe is usually described in terms of perfect fluids, meaning that heat conduction, viscosity or other transport or dissipative processes are negligible. The perfect fluid is completely specified by two quantities: the rest-frame energy density ρ and an isotropic rest-frame pressure p (Hartle, 2003). Hence the energy-momentum tensor takes the following form in terms of the fluid four-velocity u^μ:

$$T_{\mu\nu} = (\rho + p) u_\mu u_\nu + p u_{\mu\nu} \tag{6}$$

---

[9] In terms of Christoffel symbols, $R_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\mu\alpha}$. Here, $\Gamma^{\alpha}_{\beta\gamma} \equiv \frac{1}{2} g^{\alpha\lambda}(g_{\lambda\beta,\gamma} + g_{\lambda\gamma,\beta} - g_{\beta\gamma,\lambda})$, and the ”,” stands for a partial derivative, $g_{\alpha\beta}, \gamma \equiv \partial_\gamma g_{\alpha\beta}$.

Applying Einstein's field equations of general relativity to the FRW line element, one obtains the Friedmann equations (Friedman, 1922), describing the time evolution of the scale factor $a(t)$ as a function of the curvature parameter K and the energy content of the Universe, made up by the total energy density ρ and pressure p:

$$H^2(t) \equiv \left[\frac{\dot{a}(t)}{a(t)}\right] = \frac{8\pi G}{3}\rho(t) - \frac{K}{a^2(t)} \tag{7}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left[\rho(t) + 3p(t)\right] \tag{8}$$

In these equations, G is the gravitational constant, and the derivation is made with respect to the cosmic time t. The Hubble parameter $H(t)$ has been defined as the expansion rate of the Universe at a given time.

The redshift can be linked to the scale factor as

$$z = \frac{a}{a_0} - 1 \tag{9}$$

where a_0 is the present-day scale factor, and it is normalized to $a_0 = 1$.
In terms of redshift, the Hubble expansion rate can be defined as

$$H(z) = H_0 E(z) \tag{10}$$

where $H_0 \equiv H(z = 0)$ is the Hubble parameter today: its value is still uncertain, so it is customary to absorb our ignorance about it into the dimensionless Hubble parameter h, by defining it as

$$H_0 = 100\,h\,\frac{\mathrm{km}}{\mathrm{s\ Mpc}} \tag{11}$$

The function $E(z)$ depends on the employed cosmological model.

As evident from (7), there exists a specific value of the density, known as the critical density ρ_c which flattens the geometry of the Universe, i.e. such that $K = 0$:

$$\rho_c(z) = \frac{3H^2(z)}{8\pi G} \tag{12}$$

From this definition, one can define, for each generic component i, the density parameter $\Omega_i(z) \equiv \rho(z)/\rho_c(z)$. Additionally, from (8) one can define the curvature density parameter $\Omega_K(z) = -K/\left[a^2(z)H^2(z)\right]$. Starting from these definitions, the following equation for the energy budget of the Universe must be fulfilled at all times:

$$\sum_{i=1}^{N} \Omega_i(z) + \Omega_K(z) = 1 \tag{13}$$

where N is the number of species involved in the chosen model.

Equations (7) and (8) can be combined into an energy conservation equation, linking the pressure to the energy density:

$$\dot{\rho} = -3H(\rho + p) \tag{14}$$

For each energy component of the Universe there exists a specific relation between ρ and p, namely an equation of state $\rho = \rho(p)$. It is customary to adopt a very simple form of such equation of state:

$$p = w\rho \tag{15}$$

where w is a constant equation of state parameter.

This choice allows one to describe a variety of contributions to the right-hand side of the Einstein equations — matter, radiation, vacuum energy, coherent scalar fields, cosmic strings, domain walls, and so on (Kolb, 2018).

This parametrization allows one to solve (14), thus finding, for the ith component:

$$\rho(a) \propto a^{-3(1+w_i)} \tag{16}$$

Non relativistic components, i.e. baryons, cold dark matter and massive neutrinos, are considered as a pressureless gas: $w_b = w_{cdm} = w_{\nu,massive} = 0$. Meanwhile, relativistic components such as photons and massless neutrinos behave as blackbody radiation: $w_\gamma = w_{\nu,massless} = 1/3$ (Hartle, 2003; Blanchard et al., 2020).

Flat ΛCDM

In the cosmological framework of general relativity, it is necessary to modify (4) in order to account for the observed accelerated expansion of the Universe (Hubble, 1929). The simplest way to do this is by adding a constant (hereafter, cosmological constant, Λ), which is able to mathematically solve the issue. The cosmological constant was originally placed by Einstein in the left hand side of (4), as an attempt to mathematically recover the situation of a static Universe. Only lately, Λ was moved to the right hand side, and new interpretations of it arose. As suggested (Zel'Dovich, 1968), this new form of energy, commonly known as dark energy, could be interpreted as vacuum energy:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} + 8\pi G T^{vac}_{\mu\nu} = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{17}$$

The addition of the cosmological constant is one of the ingredients of the flat ΛCDM model: this is the standard model for today's cosmology, i.e. that which describes at best our current observations.

Following from the previous addition, Friedmann equations are accordingly modified:

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{K}{a^2(t)} + \frac{\Lambda}{3} \tag{18}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left[\rho(t) + 3p(t)\right] + \frac{\Lambda}{3} \tag{19}$$

The cosmological constant contributes negatively to the pressure term, thus producing a repulsive effect, able to explain the accelerated expansion of the Universe. This negative pressure effect is also manifest from the equation of state parameter $w_\Lambda = -1$.

Moreover, within the context of the flat ΛCDM model, $\Omega_K = 0$, and the energy budget equation can be written as

$$\Omega_m = \Omega_b + \Omega_{cdm} + \Omega_{\nu,massive} \tag{20}$$

$$\Omega_{rad} = \Omega_\gamma + \Omega_{\nu,massless} \tag{21}$$

$$\Omega_m + \Omega_{rad} + \Omega_\Lambda = 1 \tag{22}$$

where:

- the matter energy density Ω_m is the sum of the energy densities of baryons (b), cold dark matter (cdm) and massive neutrinos ($\nu, massive$);
- the radiation energy density Ω_{rad} is the sum of the energy densities of photons (γ) and massless neutrinos ($\nu, massless$);

- Ω_Λ is the energy density of the dark energy component Λ, which accounts for the accelerated expansion of the Universe (see (P. Peebles & Ratra, 1988) (Weinberg, 1989) for more information about the cosmological constant).

Finally, within the flat ΛCDM model, the function $E(z)$ introduced in (10) takes the form:

$$E(z) = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0} + \Omega_{rad,0}(1+z)^4} \tag{23}$$

where the index 0 indicates the present-day value of the given quantity.

Chapter 1

Constructing the observed power spectrum

The present work revolves around different methods to extract information (and put constraints) on cosmological parameters, from the data collected by redshift-space galaxy surveys. In particular, such surveys allow us to measure the power spectrum of galaxies. Since the observed galaxy power spectrum is of utmost importance for the present discussion, the first chapter of this work is entirely dedicated to introducing and "building" such quantity. This is not a straightforward task, since the observed galaxy power spectrum differs from the "true" linear matter power spectrum in many ways. First of all, since we work under the terms of galaxy surveys, we can only deal with galaxies, while the "true" underlying matter distribution remains unreachable to us. Furthermore, our surveys map quantities in redshift space, thus distorting the "true" quantities. Moreover, other contributions such as the Alcock-Paczyński effect and instrumental shot noise must be added to the mix.

Therefore, the present section will start by outlining the main definitions and features of the "true" linear matter power spectrum, and only after we will add the contributions and alterations which stem from a practical observation of such quantity.

Finally, some words will be spent on the power spectrum multipole expansion.

1.1 True linear matter power spectrum

1.1.1 Perturbation theory

Today's large scale structure is the result of a process which started in the primordial Universe, when a random process generated minute density fluctuations. Such fluctuations subsequently evolved, ultimately giving rise to the Universe as it is known today.

In general, density fluctuations for a component i are characterised by the density contrast

$$\delta_i(\boldsymbol{x}, z) \equiv \frac{\rho_i(\boldsymbol{x}, z) - \bar{\rho}_i(z)}{\bar{\rho}_i(z)} \tag{1.1}$$

which quantifies the deviations of the density field $\rho_i(\boldsymbol{x}, z)$ around the mean spatial density $\bar{\rho}_i(z)$: $\boldsymbol{x}$ is a 3D comoving coordinate and z is the redshift.

To describe these fluctuations statistically, it is convenient to work in Fourier space by decomposing δ into plane waves:

$$\delta_i(\boldsymbol{x}, z) = \int \frac{d^3k}{(2\pi)^3} \tilde{\delta}_i(\boldsymbol{k}, z) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{1.2}$$

The advantage of this choice lies in the fact that fluctuations corresponding to different Fourier modes are uncorrelated, while each given mode has a nonzero variance. This crucial feature allows to easily separate different scales.

For non-relativistic matter and sub-horizon scales, the evolution of density fluctuations can be described by ideal fluid equations (P. J. E. Peebles, 2020). In particular, their growth obeys a second order differential equation. At early enough times, when those fluctuations are still small, the fluid equations can be linearised.

During matter domination, treating matter as a pressureless ideal fluid, the equation for the evolution of the density contrast becomes:

$$\ddot{\delta}_m(\boldsymbol{k}, z) + 2H\dot{\delta}_m(\boldsymbol{k}, z) - \frac{3H_0^2\Omega_{m,0}}{2a^3}\delta_m(\boldsymbol{k}, z) = 0 \tag{1.3}$$

where a is the scale factor and $\Omega_{m,0} = \Omega_{cdm,0} + \Omega_{b,0} + \Omega_{\nu,0}$[1] is the total matter contribution today, accounting for cold dark matter, baryons and massive neutrinos.

In a ΛCDM scenario with no massive neutrinos, and neglecting dark energy perturbations, the previous equation can be written as

$$\delta_m''(\boldsymbol{k}, z) + \left[\frac{H'(z)}{H(z)} - \frac{1}{1+z}\right]\delta_m'(\boldsymbol{k}, z) - \frac{3}{2}\frac{\Omega_m(z)}{(1+z)^2}\delta_m(\boldsymbol{k}, z) = 0 \tag{1.4}$$

The prime indicates derivatives with respect to the redshift, while $\Omega_m(z)$ is given by

$$\Omega_m(z) = \frac{\Omega_{m,0}(1+z)^3}{E^2(z)} \tag{1.5}$$

$$E^2(z) = \Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}(z) + \Omega_{K,0}(1+z)^2 \tag{1.6}$$

in which Ω_Λ and Ω_K stand for the density contributions of dark energy and spatial curvature respectively.

The solutions $\delta_m(\boldsymbol{k}, z)$ of equation (1.4) are scale-independent at late times: this suggests the introduction of a growth factor $D(z)$:

$$\delta_m(\boldsymbol{k}, z) = \delta_m(\boldsymbol{k}, z_i)\frac{D(z)}{D(z_i)} \tag{1.7}$$

where z_i is an arbitrary reference redshift in the matter dominated era.

In the present analysis, we will employ the following fitting form for $D(z)$ (Lahav, Lilje, Primack, & Rees, 1991; Carroll, Press, & Turner, 1992):

---

[1] As customary, we employ the definition $\Omega_i \equiv \rho_i/\rho_{crit}$, where ρ_{crit} is the critical density associated to a flat Universe.

$$D(z) = \frac{1}{1+z}\frac{5}{2}\Omega_m(z)\left\{\Omega_m^{4/7}(z) - \Omega_\Lambda(z) + \left[1 + \frac{\Omega_m(z)}{2}\right]\left[1 + \frac{\Omega_\Lambda(z)}{70}\right]\right\}^{-1} \tag{1.8}$$

where $\Omega_m(z)$ is given by(1.5), while

$$\Omega_\Lambda(z) = \frac{\Omega_{\Lambda,0}}{E^2(z)} \tag{1.9}$$

From the growth factor, we can define the growth rate parameter f as

$$f(a) = \frac{d\,\mathrm{ln}D(a)}{d\,\mathrm{ln}a} \tag{1.10}$$

or, in terms of redshift,

$$f(z) = -\frac{d\,\mathrm{ln}D(z)}{d\,\ln(1+z)} = -\frac{(1+z)}{D(z)}\frac{dD(z)}{dz} \tag{1.11}$$

This quantity is deeply related to redshift space distortions (hereafter, RSD), and it is usually involved in LSS analysis though galaxy redshift surveys. A further exploration of the meaning of parameter f will be carried out in the section 1.2.4.

For now, it may be sufficient to add that, within ΛCDM, the growth rate is well approximated by (Lahav et al., 1991)

$$f(z) = [\Omega_m(z)]^{0.6} + \frac{\Omega_\Lambda(z)}{70}\left[1 + \frac{\Omega_m(z)}{2}\right] \tag{1.12}$$

1.1.2 The power spectrum

The power spectrum P_i for the component i is defined as the variance of the distribution (Blanchard et al., 2020):

$$\langle\tilde{\delta}_i(\boldsymbol{k},z)\tilde{\delta}_i(\boldsymbol{k'},z)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k'})P_i(\boldsymbol{k},z) \tag{1.13}$$

where δ_D is a Dirac delta function, stating the independence of different modes, while the angular brackets indicate an average over the whole distribution. Accordingly to definition (1.13), the power spectrum will be small for a smoother distribution, while it will be large for a distribution which features many overdensities and underdensities.

Under the assumptions of homogeneity and isotropy, the power spectrum can only depend on $k = |\boldsymbol{k}|$ and z.

On top of these definitions, it is useful to define a dimensionless power spectrum as:

$$\Delta_i^2(k,z) \equiv \frac{k^3}{2\pi^2}P_i(k,z) \tag{1.14}$$

which accounts for the fact that the power spectrum has dimensions of k^{-3} or length^3.

1.1.3 Primordial power spectrum, transfer function and normalisation

According to the standard inflationary model, the primordial dimensionless power spectrum Δ_R^2 of curvature perturbations is given by a power law:

$$\Delta_R^2(k) = A_s\left(\frac{k}{k_0}\right)^{n_s-1} \tag{1.15}$$

where A_s is the amplitude of the primordial scalar perturbation, k_0 is a pivot scale, while the scalar spectral index n_s measures the deviation from scale invariance (corresponding to

$n_s = 1$).

Therefore, the primordial power spectrum is given by:

$$P_R(k) = \frac{2\pi^2}{k^3} A_s \left(\frac{k}{k_0}\right)^{n_s - 1} \tag{1.16}$$

The primordial power spectrum in (1.16) does not match the one we observe today. In fact, the presence of many cosmological components, along with their evolution and mutual interactions, alters the predictions of perturbation theory. This, in turn, affects the shape[2] of the original power spectrum, modifying its overall look. However, it must be noted that causality precludes such effect at arbitrarily large scales. All of these shape-changing effects are modelled by the transfer function $\mathcal{T}(k, z)$, which compares the amplitude of the resulting perturbations to the amplitude they would have had if causal physics had been neglected (Eisenstein & Hu, 1998). These shape-changing processes, which get encoded into the transfer function, will be better explored in section 1.1.4: now we outline its main effect, which is linked to the epoch of matter-radiation equality.

During radiation domination, curvature perturbations with comoving scales smaller than the horizon are suppressed, while super-horizon scales stay unaffected until they enter the horizon. Instead, during the matter dominated era, curvature perturbations remain constant on all scales. As a consequence, a characteristic scale corresponding to the epoch of matter-radiation equality gets imprinted on the transfer function, thus determining the main shape of the power spectrum.

Therefore, the power spectrum at any z and k is expressed as

$$P_i(k, z) = P_R(k)\mathcal{T}_i^2(k, z) = \frac{2\pi^2}{k^3} A_s \left(\frac{k}{k_0}\right)^{n_s - 1} \mathcal{T}_i^2(k, z) \tag{1.17}$$

Within the ΛCDM model (no massive neutrinos are included), the late-time growth of matter

---

[2] In subsequent sections, the term *shape* will indicate the no-wiggle component of the power spectrum. For now, we just employ the term *shape* with its usual meaning.

perturbations is scale-independent, so that the matter transfer function can be split into:

- a scale-dependent part, $T_m(k)$, normalised so that $T_m \to 1$ for $k \to 0$;
- the scale-independent growth factor $D(z)$.

With that in mind, we can come up with a convenient way to express the linear matter power spectrum (Blanchard et al., 2020):

$$P_{m,lin}(k,z) = \left(\frac{\sigma_8}{\sigma_N}\right)^2 \left[\frac{D(z)}{D(0)}\right]^2 T_m^2(k) k^{n_s} \tag{1.18}$$

where

$$\sigma_8^2 = \frac{1}{2\pi^2}\int dk k^2 P_m(k,0)|W_{TH}(kR_8)|^2 \tag{1.19}$$

$$\sigma_N^2 = \frac{1}{2\pi^2}\int dk k^2\, k^{n_s} T_m^2(k)|W_{TH}(kR_8)|^2 \tag{1.20}$$

$$W_{TH}(x) = 3\,\frac{\sin x - x\cos x}{x^3} \tag{1.21}$$

W_{TH} is the Fourier transform of the top-hat filter, and $R_8 = 8h^{-1}$Mpc; σ_8 is the rms of present-day linearly evolved density fluctuations in spheres of radius R_8, while σ_N is a normalization constant.

From these definitions, it is evident that the generic power spectrum (1.17) relates to the linear matter power spectrum (1.18) when the transfer function of the component i is

$$\mathcal{T}_m^2(k,z) = \left(\frac{\sigma_8}{\sigma_N}\right)^2 \left[\frac{D(z)}{D(0)}\right]^2 \frac{k^4 k_0^{n_s-1}}{2\pi^2 A_s} T_m^2(k) \tag{1.22}$$

1.1.4 Power spectrum shape

Physical situation

As previously stated, the shape of the ("true") power spectrum we would observe today is determined by the matter transfer function, which accounts for all the modifications that have occurred to the primordial power spectrum (1.16). In turn, the form of the transfer function is defined both by the evolution of the various cosmological components through different epochs and the early-time expansion history of the Universe. In this section, we will explore this process in a simple and schematic way, following the work of (Eisenstein & Hu, 1999).

Consider linear adiabatic perturbations around a Friedmann-Robertson-Walker metric, for a cosmology involving photons, baryons, cold dark matter, massive and massless neutrinos.

The crucial quantity in this whole discussion is for sure the Jeans scale, which we indicate here with λ_J. Perturbations on a scale $\lambda < \lambda_J$ are not subject to gravitational instability, due to pressure support or, in case of collisionless particles, sufficient rms velocity. Instead, perturbations on scales $\lambda > \lambda_J$ grow at the same rate, regardless of the scale. In general, the Jeans scale of a gravitating species marks a change in the species' behaviour, thus affecting the power spectrum shape evolution by getting imprinted in it.

In order to see how this process actually plays out, one can take it step by step, starting from the simplest scenario and then progressively increasing its complexity.

In a Universe only inhabited by cold dark matter and radiation, the λ_J of the whole system grows up to the size of the particle horizon at matter-radiation equality λ_{eq} (corresponding to a wavevector k_{eq}), then it shrinks to zero as the Universe becomes matter dominated. As a consequence, the matter transfer function (which in this case only accounts for cold dark matter) is bent in correspondence of λ_{eq}, and the original power spectrum (1.17) turns around at the same scale. Importantly, well after equality, the Jeans scale has dropped below all scales of interest (i.e. the regime of LSS measurements), thus perturbations grow at the same rate. This allows one to split the transfer function into a scale-independent growth factor $D(z)$ and a scale-dependent part $T_m(k)$ which we already introduced (in particular, the shape of $T_m(k)$ will depend only on k/k_{eq}).

Now we add baryons (i.e. nucleons and electrons) to the mix. Initially, baryons are dynamically coupled to photons through Compton (between photons and electrons) and Coulomb (between nucleons and electrons) interactions, thus originating a photon-baryon fluid with sound speed $c_s(z)$. As the Universe approaches recombination, the rate of Compton scattering progressively decreases, up to the point at which baryon-photon decoupling takes place. This is known as the end of Compton drag epoch, and it corresponds to a redshift $z_d \approx z_{rec}$. For $z > z_d$, the baryonic Jeans scale tracks the sound horizon, while it drops rapidly after recombination. Therefore, the sound horizon at the end of Compton drag epoch, indicated by r_d and defined as

$$r_d = \int_{z_d}^{\infty} dz' \frac{c_s(z')}{H(z')} \, , \tag{1.23}$$

is the scale that gets imprinted in the transfer function and power spectrum (its scale corresponds to the mode k_d). If we considered the shape of the galaxy 2-point correlation function, i.e. the real space counterpart of the power spectrum, then the presence of baryons would manifest as a little spike in correspondence of a scale $\lambda = r_d$. Since the power spectrum is non other than the Fourier transform of the 2-point correlation function, when visualizing its shape, the aforementioned spike promptly turns into a series of troughs and valleys. These wiggles correspond to the baryon acoustic oscillations, hereafter, BAO. A deeper discussion about the BAO is postponed to 1.1.4.

It must be added that, also in this case, the Jeans scale drops quickly after recombination: once again, we can split scale and redshift dependences inside the transfer functions, and now $T_m(k)$ will depend on k/k_d as well. We stress this in order to highlight the contrast with the last case we are going to examine.

Now, we add massive neutrinos to the mix. At sufficiently high temperatures, massive neutrinos behave as radiation and, accordingly, the neutrino Jeans scale (known as free-streaming scale) grows with the particle horizon. As the Universe expands and temperature drops, neutrinos become non-relativistic, and their free-streaming scale starts to shrink.

A neutrino perturbation with scale $\lambda < \lambda_J$ lies in the free-streaming regime, and it collisionlessly damps out (while cold dark matter and baryon perturbations grow more slowly due to the loss of a gravitational source). Instead, a neutrino perturbation with scale $\lambda > \lambda_J$ is out of the free-streaming regime: here free-streaming is halted, and neutrino perturbations track those of the other species, neutrinos falling into their potential wells.

This last scenario is much more complicated, since for eV mass neutrinos, the free-streaming scale today lies in the regime of LSS measurements (Bond & Szalay, 1983). This means that some scales today still lie in the free-streaming regime (despite the free-streaming scale decreasing with time). Consequently, at low-z, the growth function D will depend both on redshift and scale.

Eisenstein-Hu fit

All of these behaviours and interactions must be taken into account if one wants to reconstruct the shape of the power spectrum. Numerically, this corresponds to integrating the coupled Boltzmann equations for each mode as a function of time. But some analytical fits have been provided as well: in (Eisenstein & Hu, 1998), Eisenstein and Hu provide a fitting formula for the transfer function shape, in absence of massive neutrinos, and for the two cases in which BAO wiggles are and are not included.

We could model these two fitting formulae through the following notation, where dependencies on cosmological parameters are also shown:

$$T_{EH98}^{\text{wiggle}}(k, \Omega_b, \Omega_{cdm}, h) \tag{1.24}$$

$$T_{EH98}^{\text{no wiggle}}(k, \Omega_b, \Omega_{cdm}, h) \tag{1.25}$$

We mention them here since the ShapeFit analysis will make use of these fitting formulae, especially the one corresponding to a BAO-less spectrum: it leads to a fit of the linear matter power spectrum which does not feature any wiggles, while baryon damping is present, so that the broadband shape closely tracks that of the "true" power spectrum. However unrealistic this fit may appear compared to the one including BAO wiggles[3], it becomes quite useful in the context of ShapeFit, whose aim is to maximize the information that can be extracted from the data, while minimizing computational effort. As we will see in the next chapters, the most relevant part of the information lies in the broadband shape, while BAO wiggles do not help us that much more.

---

[3] According to (Eisenstein & Hu, 1998), the BAO-less fitting form agrees with numerical simulations to 1% through CMB and large-scale structure regimes (Hu, Scott, Sugiyama, & White, 1995). This fitting formula's performance is also highlighted in *Figure* 6 of (Eisenstein & Hu, 1998), where various zero-bayon transfer functions, including those presented in (P. Peebles, 1982), (Shandarin & Zeldovich, 1989), (Bardeen, Bond, Kaiser, & Szalay, 1986), are compared to the numerical simulations of CMBfast (Seljak & Zaldarriaga, 1996).

BAO as a standard ruler

The BAO provide a characteristic scale, r_d, that is "frozen" in the galaxy distribution. Such scale is determined by the matter and baryon densities, which can be precisely measured with CMB anisotropy data, effectively providing a standard ruler that can be exploited to get valuable information.

In particular, the BAO standard ruler effectively provides a measurement of the angular diameter distance $D_A(z)$ and of the Hubble parameter $H(z)$ as a function of redshift. The knowledge of these two quantities enables one to reconstruct the expansion history of the Universe, thus gaining new information about the nature of dark energy (Seo & Eisenstein, 2003).

The angle subtended by a structure (a galaxy, a cluster or - as in this case - a matter density oscillation) of intrinsic length $r_\perp$ in the direction perpendicular to the line of sight is given by

$$\Delta\theta(z) = \frac{r_\perp}{D_A(z)(1+z)} \tag{1.26}$$

where the angular diameter distance and the comoving distance are respectively defined as

$$D_A(z) = \frac{r(z)}{1+z} \tag{1.27}$$

$$r(z) = \int_0^z \frac{dz'}{H(z')} \tag{1.28}$$

If the BAO scale r_d is known, then measuring the angle $\Delta\theta(z)$ as a function of redshift allows one to reconstruct the redshift evolution of the angular diameter distance $D_A(z)$.

Analogously, the redshift interval subtended by a structure with intrinsic length $r_\parallel$ along the line of sight is

$$\Delta z = r_{\parallel}\, H(z) \tag{1.29}$$

Once again, if r_d is known, a measure of the subtended redshift Δz allows one to reconstruct the redshift evolution of the Hubble parameter $H(z)$.

It must be noted that the BAO standard ruler is still model-dependent, since r_d actually depends on matter and baryon abundances, and on the number of massive neutrino species.

1.2 Observed linear galaxy power spectrum

Up to now, we have just taken into account the ”true” form of the linear matter power spectrum, neglecting all the distortion effects that stem from our nature of earthling observers, from the fact that we are actually looking at galaxies, and from the practical limitations of our instrumentation. Now, we will explore such effects.

1.2.1 Alcock-Paczyński effect

The first distortion we take into account is that coming from the Alcock-Paczyński (hereafter, AP) effect (Alcock & Paczyński, 1979). This effect essentially originates from the way our galaxy surveys work.

The data collected by a galaxy survey (either photometric or spectroscopic) essentially correspond to measurements of the redshift and angular positions of the galaxies which inhabit the surveyed volume: what we don't have is a direct measurement of their distances from us, nor of their intrinsic scales. Given this limitation as earthling observers, we are then forced to define cosmic distances in terms of these two observable quantities.

The trick lies in the fact that such definitions of cosmic distances also depend on cosmological parameters, and we do not know the true cosmology of our Universe. We are then forced to adopt a so-called ”reference cosmology”: this will most likely differ from the true cosmology of the Universe[4], but at least it will enable us to obtain the distance values we crave for. Inevitably, the obtained distances will be distorted compared to the true ones, leading to an

---

[4] Even though one could adopt any cosmology as their reference, it is customary to choose the one resulting from latest and most precise observations.

overall geometric distortion of the LSS. To be more specific, the assumption of an incorrect cosmology would lead to an effective squashing of space along the line of sight, thus producing an anisotropy in the inferred density field which affects the galaxy clustering statistics or, in other words, the power spectrum.

In order to see how the AP distortion affects the wavevector $\boldsymbol{k}$, it might be wise to define the quantity $\mu = \hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{k}}$, where $\hat{\boldsymbol{z}}$ is the unit vector along the line of sight. Therefore, μ represents the cosine of the angle between the wavevector and the line of sight.

We now split $\boldsymbol{k}$ into the two components $k_{\parallel}$ and $k_{\perp}$, respectively parallel and perpendicular to the line of sight.

The AP distortion modifies these components via two scaling factors, which account for the difference between true and reference cosmology, along and across the line of sight. Such distortion produces a new wavevector $\boldsymbol{k'}$, whose components are (Ballinger, Peacock, & Heavens, 1996)

$$k'_{\parallel} = q_{\parallel} k_{\parallel} \qquad\qquad k'_{\perp} = q_{\perp} k_{\perp} \tag{1.30}$$

Following (Blanchard et al., 2020), we define the two scaling factors as the ratios between angular diameter distance and Hubble parameter in the true and reference cosmology:

$$q_{\parallel}(z) = \frac{H^{ref}(z)}{H(z)} \qquad\qquad q_{\perp}(z) = \frac{D_M(z)}{D_M^{ref}(z)} \tag{1.31}$$

So, using (1.30), the definition of μ and the fact that $k_{\parallel} = \mu k$, we find that

$$k(k', \mu') = \frac{k'}{q_{\perp}} \left[1 + \mu'^{\,2} \left(\frac{q_{\perp}^2}{q_{\parallel}^2} - 1 \right) \right]^{1/2} \tag{1.32}$$

$$\mu(\mu') = \mu' \frac{q_{\perp}}{q_{\parallel}} \left[1 + \mu'^{\,2} \left(\frac{q_{\perp}^2}{q_{\parallel}^2} - 1 \right) \right]^{-1/2} \tag{1.33}$$

As a result, the power spectrum is being measured at wrong wavevectors $\boldsymbol{k'}$. Since we are interested in modelling the observed power spectrum, we should write it as a function of such distorted quantities:

$$P_m(k', \mu', z) = \frac{1}{q_\perp^2 q_\parallel} \, P(k(k', \mu'), \mu(\mu'), z) \tag{1.34}$$

where the overall scaling $\frac{1}{q_\perp^2 q_\parallel}$ follows from a dilation in the volume element.

1.2.2 Galaxy bias

It is now time to address the fact that, through surveys, we only have access to galaxies' angular positions and redshifts, while the actual underlying distribution of matter remains out of our reach.

As a consequence, the only power spectrum we can obtain from our data is an observed galaxy power spectrum, not the "true" matter one we have been discussing so far. The link between the two can be expressed by a bias parameter, which quantifies how well the galaxy distribution is able to trace the underlying matter one.

The existence of a bias is certified by the fact that galaxies selected in different ways (e.g. optical galaxies versus IRAS galaxies) feature correlation functions with different amplitudes (while shapes are basically the same), meaning that they cannot all be unbiased tracers of the matter distribution (Peacock & Dodds, 1994; Oliver et al., 1996; Peacock, 1997).

We define δ_g and δ_m as the galaxy and matter overdensity fluctuations[5] respectively. The simplest bias model postulates that these two quantities are connected by a spatially-constant linear bias factor b,

$$\delta_g(\boldsymbol{r}) = b(z) \, \delta_m(\boldsymbol{r}) \tag{1.35}$$

---

[5] I.e. fluctuations with respect to mean spatial density, see (1.1).

while galaxy velocities are assumed to faithfully follow the velocity of matter at all redshifts:

$$\boldsymbol{v}_g = \boldsymbol{v}_m \tag{1.36}$$

Of course, $b = 1$ corresponds to unbiased tracers.

From (1.35), it is easy to infer the corresponding relation between power spectra:

$$P_{g,lin}(k, z) = b^2(z) P_{m,lin}(k, z) \tag{1.37}$$

1.2.3 Poisson shot noise

To further complicate the picture, one should consider that galaxies actually come as discrete units: the common assumption is that the observed galaxies are drawn randomly from a hypothetical continuous underlying distribution of galaxies. The observed galaxies then form a Poisson process of the underlying population[6].

In order to see how the discreteness of a survey affects the observed galaxy power spectrum, we need to introduce the selection function $\bar{n}(\boldsymbol{r})$. This quantity, characteristic of a survey, represents the mean number of galaxies found at position $\boldsymbol{r}$ by the survey, given its own selection criteria (e.g. flux limit). But it can also be interpreted as the probability of including a galaxy at position $\boldsymbol{r}$ into the survey.

Let $n(\boldsymbol{r})$ denote the observed number density of galaxies at position $\boldsymbol{r}$ in a survey. Due to the discrete nature of the galaxies, $n(\boldsymbol{r})$ will be a sum of delta functions.

The observed galaxy overdensity is then defined as

$$\delta_g^{obs}(\boldsymbol{r}) \equiv \frac{n(\boldsymbol{r}) - \bar{n}(\boldsymbol{r})}{\bar{n}(\boldsymbol{r})} \tag{1.38}$$

---

[6] Therefore, (1.37) should be understood in a probabilistic sense: galaxies are unbiased if the probability of finding a galaxy in a volume element dV at position $\boldsymbol{r}$ is proportional to the amount of matter $\rho_m(\boldsymbol{r})dV$ in that volume element (Hamilton, 1998).

In the Poisson process model, the observed galaxy overdensity provides a discretized but unbiased estimate of the true overdensity (corresponding to what we defined in (1.1)).

As a result, the expectation value of the covariance of observed overdensities is a sum of the true 2-point correlation function $\xi(r_{12})$ with a Poisson sampling noise, also known as shot noise, term:

$$\langle \delta_g^{obs}(\boldsymbol{r}_1)\delta_g^{obs}(\boldsymbol{r}_2)\rangle = \xi(r_{12}) + \frac{\delta_D(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{\bar{n}(\boldsymbol{r}_1)} \tag{1.39}$$

and an analogous relation applies to the power spectrum (Kaiser, 1987).

The shot noise term involves the Dirac delta δ_D, and reflects the fact that the probability of finding yourself as a neighbour at zero separation is unity. This suggests that, in principle, the shot noise contribution can be eliminated by excluding from the computation all the self-pairs of galaxies.

However, when defining the complete form of the power spectrum to be employed in the analysis, a shot noise term will be included, accounting for an imperfect removal of the effect due to this Poisson process (Blanchard et al., 2020).

1.2.4 Redshift-space distortions

Up to this point, we have discussed about the power spectrum as a quantity that lives in real-space. The truth is that we gain access to power spectrum measurements only through redshift-space galaxy surveys. But the observed redshift is not entirely cosmological: it also contains a contribution due to the line of sight component of galaxies' peculiar velocities. Following the wrong assumption that the observed redshifts are entirely cosmological, when estimating distances to each galaxy, leads to a distortion of the density field, in a way that imprints a specific pattern of anisotropies onto the observed power spectrum. This is known as the redshift space distortion.

The topic of redshift-space distortions (RSD) has been extensively explored in (Kaiser, 1987; Hamilton, 1998; Bertacca, Maartens, Raccanelli, & Clarkson, 2012). Here we outline the basic understanding of linear RSD, which are essential to construct the observed power spectrum employed in the present analysis.

To a given object, e.g. a galaxy, we can associate two "distances": the redshift distance $s \equiv cz$ and the true distance $r \equiv H_0 d$. We used quotation marks since these two definitions of distance are not expressed in usual units of distance (e.g. Mpc), but rather in velocity units.

These two distances are connected by the following relation:

$$s = r + \hat{r} \cdot \boldsymbol{v} \equiv r + v \tag{1.40}$$

where $\hat{r}$ is the unit vector along the line of sight, while $\boldsymbol{v}$ is the peculiar velocity of the galaxy.

Equation (1.40) thus indicates that the distortion which arises when remapping redshift-space into real-space originates from the peculiar velocities of galaxies along the line of sight. As an aside, it is easy to see from (1.40) that, in absence of such peculiar velocities, the classic Hubble's law $cz = H_0 d$ would be valid.

But positions are not the only victims of RSD. For galaxy density fluctuations in the linear regime, the remapping is embodied by the linear redshift distortion operator $\boldsymbol{S}$:

$$\delta_g^s(\boldsymbol{s}) = \boldsymbol{S}\delta_g(\boldsymbol{r}) \tag{1.41}$$

Here, $\delta_g^s(\boldsymbol{s})$ represents a galaxy density fluctuation in redshift-space, located at redshift position $\boldsymbol{s}$, while $\delta_g(\boldsymbol{r})$ represents a galaxy density fluctuation in real-space, located at true position $\boldsymbol{r}$[7].

The full form of $\boldsymbol{S}$ is[8]:

---

[7] In this case, we are considering an ideal case in which both fluctuations are perfectly known, as if they had been measured by a perfect and noiseless instrument. Therefore, they do not coincide with the δ_g^{obs} that we employed in 1.2.3. In a more realistic scenario, a shot noise contribution must be added

[8] Actually, this expression is valid only in the frame of a randomly located observer, at rest with respect to the CMB, and for a specific definition of δ_g. It is evident that this is not our case, but we keep this form, since other simplifications will be made.

$$\boldsymbol{S} = \mathbf{1} + \beta \left(\frac{\partial^2}{\partial r^2} + \frac{\alpha(\boldsymbol{r})}{r} \frac{\partial}{\partial r} \right) \nabla^{-2} \tag{1.42}$$

$$\alpha(\boldsymbol{r}) \equiv \frac{\partial \ln (r^2 \bar{n}(\boldsymbol{r}))}{\partial \ln r} \tag{1.43}$$

The β appearing in (1.42) is defined as the ratio between the linear growth rate and the bias:

$$\beta(z) \equiv \frac{f(z)}{b(z)} \tag{1.44}$$

while the $\bar{n}$ in (1.43) represents the selection function of the survey.

This expression of $\boldsymbol{S}$ is quite complicated, but things simplify a lot if one adopts the plane-parallel limit. In this limit, the observer is considered very distant from the surveyed volume, such that all peculiar velocities along the line of sight are plane parallel. In this limit, $\alpha \to 0$, so that $\boldsymbol{S}$ becomes

$$\boldsymbol{S}^P = \mathbf{1} + \beta \frac{\partial^2}{\partial z^2} \nabla^{-2} \tag{1.45}$$

where z marks the distance along the line of sight, and the superscript P indicates that we are assuming the plane-parallel approximation.

Switching to Fourier space, we have that

$$\frac{\partial^2}{\partial z^2} \nabla^{-2} \;\to\; \frac{k_z^2}{k^2} \equiv \mu^2 \tag{1.46}$$

with $\mu \equiv \hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{k}}$ being the cosine of the angle between the wavevector and the line of sight.

Therefore, if we now consider Fourier modes for density fluctuations (marked by a hat), we can rewrite (1.41) as

$$\hat{\delta}_g^s(k, \mu) = (1 + \beta\mu^2)\, \hat{\delta}_g(k) \tag{1.47}$$

Equation (1.47) tells us that, within the plane-parallel approximation, unredshifted density fluctuations appear amplified by a factor $1 + \beta\mu^2$ in redshift space. Therefore, for the corresponding power spectra, the relation becomes

$$P^s_{g,lin}(k, \mu, z) = (1 + \beta(z)\mu^2)^2\, P_{g,lin}(k, z) \tag{1.48}$$

meaning that we should expect to measure discrepancies in the clustering strength at different angles with respect to the line of sight.

This distortion is still a consequence of galaxies' peculiar velocities, as it appears clear if one considers the linearised continuity equation for galaxies (assuming linear bias): $\beta\delta_g + \boldsymbol{\nabla} \cdot \boldsymbol{v}_g = 0$. Peculiar velocity displacements are sourced by the true underlying density field, so the RSD pattern provides us with an additional source of cosmological information on the relation between density and velocity fields. The latter depends on the growth rate parameter f in the linear regime, and this explains the importance of such parameter in standard LSS analysis.

1.2.5 Redshift uncertainty

In practice, the power spectrum is measured in redshift bins: whenever we refer to observed quantities we should write, e.g., $P_g^{obs}(k, \mu, z_{eff})$ instead of $P_g^{obs}(k, \mu, z)$, where z_{eff} represents the effective redshift of the bin (which can be simply taken as the central redshift of the bin). Anyway, we decide to stick to our original notation for a matter of simplicity, but the meaning of z from now on is obvious.

Such uncertainty over the assumed redshift position results in a modification of the power spectrum, which can be modelled as (Wang, Chuang, & Hirata, 2013)

$$P_g^{obs}(k, \mu, z) \;\rightarrow\; F_z(k, \mu, z) P_g^{obs}(k, \mu, z) \tag{1.49}$$

with

$$F_z(k, \mu, z) \equiv \mathrm{e}^{-k^2\mu^2\sigma_r^2(z)} \tag{1.50}$$

Here, the factor F_z accounts for the smearing of the observed galaxy density field along the line of sight direction $k_\parallel = k\mu$ due to possible redshift errors[9]. Such errors propagate into a comoving distance error modelled as

$$\sigma_r(z) = \frac{\partial r}{\partial z}\sigma_z(z) = \frac{c}{H(z)}(1+z)\sigma_{z,0} \tag{1.51}$$

where $\sigma_z(z) = (1+z)\sigma_{z,0}$ is a linear scaling of the redshift error.

1.3 Full linear model of the observed galaxy power spectrum

We can finally write down the form of the model for the observed linear galaxy power spectrum that we have constructed so far:

$$P_{g,lin}^{obs}(k', \mu', z) = \frac{1}{q_\perp^2(z)q_\parallel(z)} b^2(z)(1+\beta(z)\mu^2)^2 P_{m,lin}(k, z) F_z(k, \mu, z) + P_{shot}(z) \tag{1.52}$$

where:

- $\dfrac{1}{q_\perp^2(z)q_\parallel(z)}$ is the volume rescaling due to the AP effect;

---

[9] These errors could come, e.g., from the uncertainty resulting from fitting the centroid of an emission line in a spectroscopic survey.

- $b(z)$ is the linear bias factor;

- $(1+\beta(z)\mu^2)^2$ is the power amplification due to RSD;

- $P_{m,lin}(k,z) = \left(\frac{\sigma_8}{\sigma_N}\right)^2 \left[\frac{D(z)}{D(0)}\right]^2 T_m^2(k) k^{n_s}$ is the true linear matter power spectrum;

- $F_z(k,\mu,z)$ is the smearing factor accounting for redshift uncertainties;

- P_{shot} is a shot noise term due to the Poisson process of the survey;

- $k = k(k', \mu')$ and $\mu = \mu(\mu')$ are the true variables expressed in terms of those distorted by the AP effect.

1.4 Non-linear model of the observed galaxy power spectrum

In this section, the previous results are generalized to the non-linear case: this is crucial if one aims at correctly modelling the power spectrum on small scales. In recent years, many different approaches to model the non-linear power spectrum of galaxy clustering, particularly in terms of non-linear RSD (Scoccimarro, 2004; Matsubara, 2008b, 2008a; Carlson, White, & Padmanabhan, 2009; Reid & White, 2011; Taruya, Nishimichi, & Saito, 2010). The present analysis follows the perturbation theory (PT) approach outlined in (Beutler et al., 2014).

In order to recover the non-linear counterpart of (1.52), we need to take a few steps back, and express the density (δ) and velocity (θ) real-space[10] auto and cross power spectra, $P_{m,\delta\delta}$, $P_{m,\theta\theta}$ and $P_{m,\delta\theta}$). If we are simply dealing with dark matter, then the density power spectrum is expected to increase in the non-linear regime, while velocities are randomised, thus damping the velocity power spectrum on small scales (Scoccimarro, 2004). Moreover, dark matter power spectra have to be related to those of galaxies, which are biased tracers. Here, following (Beutler et al., 2014), no velocity bias is assumed ($\theta_g = \theta_m$), while every possible galaxy bias term is included (McDonald & Roy, 2009). Therefore, at 1-loop standard perturbation theory, the galaxy auto and cross power spectra can be written as:

---

[10] In this case the "real space" corresponds to an unredshifted space.

$$\begin{aligned}P_{g,\delta\delta}(k) &= b_1^2 P_{m,\delta\delta}(k) + 2b_2 b_1 P_{m,b2\delta}(k) + 2b_{s2} b_1 P_{bs2,\delta}(k) + b_2^2 P_{m,b22}(k) + \\ &\quad + 2b_2 b_{s2}(k) P_{m,b2s2} + b_{s2}^2 P_{bs22}(k) + 2b_1 b_{3nl}\sigma_3^2(k) P_{m,lin}(k) \\ P_{g,\delta\theta}(k) &= b_1 P_{m,\delta\theta}(k) + b_2 P_{m,b2\theta}(k) + b_{s2} P_{m,bs2\theta}(k) + b_{3nl}\sigma_3^2(k) P_{m,lin}(k) \\ P_{g,\theta\theta}(k) &= P_{m,\theta\theta}\end{aligned} \tag{1.53}$$

Here, $P_{m,lin}(k)$ represents the linear matter power spectrum (1.18), while $P_{b2,x}$ and $P_{bs2,x}$ (with $x = \delta$ or θ) represent 1-loop corrections to the linear bias expansion[11]. Biasing is parametrized by four bias parameters, the first and second order biases b_1, b_2 (Fry & Gaztanaga, 1993), and the second and third order non-local[12] biases b_{s2} and b_{3nl}(McDonald & Roy, 2009)[13].

The real-space non-linear galaxy power spectra in (1.53) can be used to build the redshift-space counterpart of the same quantity (Scoccimarro, 2004; Taruya et al., 2010):

$$\begin{aligned}P_{g,RSD}(k,\mu) = \left[1 + \frac{(k\mu\sigma_P)^2}{2}\right]^{-2} &[P_{g,\delta\delta}(k) + 2f\mu^2 P_{g,\delta\theta}(k) + f^2\mu^4 P_{g,\theta\theta}(k) + \\ &+ b_1^3 A^{TNS}(k,\mu,f/b_1) + b_1^4 B^{TNS}(k,\mu,f/b_1)]\end{aligned} \tag{1.54}$$

Here, the prefactor is a Lorentzian damping term accounting for the non-linear RSD effect. The effect, also known as "Fingers of God" effect, is due to the velocity dispersion of satellite galaxies inside the host dark matter halos, which damps the power spectrum at small scales. The quantity σ_P represents a phenomenological incoherent velocity dispersion parameter. (1.54) is based on the TSN model, and the definitions of A^{TNS} and B^{TNS} can be found in (Taruya et al., 2010).

Therefore, the full non-linear model of the observed galaxy power spectrum is given by:

---

[11] The exact expression of these terms and σ_3 can be found in *Appendix A*.

[12] Here, the subscript nl stands for "non local", while the same subscripts means "non-linear" when it is associated to a power spectrum quantity.

[13] Further discussion about these biases can be found in (Baldauf, Seljak, Desjacques, & McDonald, 2012; Saito et al., 2014; Chan, Scoccimarro, & Sheth, 2012; Barreira, Lazeyras, & Schmidt, 2021)

$$P^{obs}_{g,nl}(k', \mu', z) = \frac{1}{q^2_\perp(z) q_\parallel(z)} P_{g,RSD}(k, \mu) F_z(k, \mu, z) + P_{shot}(z) \tag{1.55}$$

with $k = k(k', \mu')$ and $\mu = \mu(\mu')$.

On a final note, it would be wise to specify that while the true power spectrum defined in (1.13) has dimensions of $[\mathrm{Mpc}]^3$, in the context of LSS analysis the same quantity is usually expressed in $[\mathrm{Mpc/h}]^3$. This choice makes it explicit that distances are defined in units of a reference cosmology, and thus the observed power spectrum is a model-dependent quantity.

1.5 Multipole expansion

Once the forms (1.52) or (1.55) are reached, one could expand the μ-dependence in the Legendre-polynomials orthonormal base, thus allowing to describe the line of sight dependence through a series of multipoles (Brieden, Gil-Marín, & Verde, 2021b):

$$P^{(\ell)}(k') = (2\ell + 1) \int_{-1}^{1} P^{obs}_g(k, \mu) \mathcal{L}_\ell(\mu) d\mu \tag{1.56}$$

In principle, an infinite series of multipoles would be required, but in standard practice only the first few non-null ones are used. In fact, although the non-linear terms A^{TNS} and B^{TNS} include μ^6 and μ^8 contributions, the amount of information of these in the scales of interest, and thus the information contained in higher-order multipoles, is very small. This is why standard cosmological analyses up-to-date stop at the hexadecapole ($\ell = 4$), while odd multipoles contain no information under the flat-sky approximation and in absence of selection effects (Brieden et al., 2021b)[14].

Combining the monopole ($\ell = 0$) and quadrupole ($\ell = 2$) signals allows to break the large-scale degeneracy between linear bias and growth rate.

The hexadecapole ($\ell = 4$) helps breaking the degeneracy between AP effect and RSD.

---

[14] However, such approximation could be updated with the introduction of wide-angle effects, modelling indow functions at very large scales, see for example (Beutler, Castorina, & Zhang, 2019; Bertacca et al., 2012)

Chapter 2

Power spectrum as a forecasting tool

The quantity defined and modelled in the previous chapter, i.e. the power spectrum, can be employed to extract valuable cosmological information. Different kinds of analyses provide information of different nature and purpose.

For example, power spectrum values can be measured from galaxy surveys, and used to build a dataset. A theoretical model, depending on some parameters, is then fit to the collected data: thus, from this fitting procedure, constraints on the model parameters are extracted. In this case, the analysis can be thought as being "projected backwards": the data we collect today is the starting point, and from that we can extract information on the physical and cosmological variables which have led to that specific data realization. Some examples of this kind of analysis are the Classic RSD and FM approaches, which will be presented in subsequent chapters.

Alternatively, one can decide to employ the power spectrum as a predictive tool, thus opting for a "forward projected" analysis. In particular, the theoretical model of the power spectrum can be used to predict the entity of the constraints which could be obtained, when a specific experimental setup is employed. This is achieved thanks to a powerful statistical tool, the Fisher information matrix, whose application to cosmology was pioneered in mid '90s.

The purpose of this chapter is to better illustrate this kind of predictive analysis. Therefore, we will lay down the basics, aims and advantages of the Fisher forecast technique: in doing so, we will closely follow some works of Max Tegmark (Tegmark, Taylor, & Heavens, 1997; Tegmark, 1997).

We will start by describing its most general form and expression, and then specify how it adapts to the case of redshift surveys.

2.1 Fisher matrix formalism: general description

Once data are collected for an experiment of any sort, they can be stored in a vector $\boldsymbol{x}$[1].

Assuming that the result of each measurement is random, the data vector $\boldsymbol{x}$ constitutes a random variable: within the context of a theoretical model, its likelihood distribution will also depend on some model parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, ..., \theta_n\}$. Consequently, the likelihood can be written as $L(\boldsymbol{x};\boldsymbol{\theta})$[2].

In principle, for each of these parameters there exists a true value, but it remains unknown to the observer. One can only build an estimator for each of the model parameters, and try to make the most of the information contained in the collected data.

So, for example, we could denote the true parameters with $\boldsymbol{\theta_0} = \{\theta_{0,1}, \theta_{0,2}, ..., \theta_{0,n}\}$, while $\boldsymbol{\theta}$ stands for their estimate vector. Of course, we want this to be a good estimate of the true value. In the best case scenario, we will be able to define the Best Unbiased Estimator (BUE) for $\boldsymbol{\theta_0}$. In order to be the BUE, our estimator should:

- be unbiased, meaning that the ensemble average of such estimator gives the true value, i.e. $\langle\boldsymbol{\theta}\rangle = \boldsymbol{\theta_0}$;

- give the smallest possible error bars, i.e. minimize standard deviations, defined as $\sigma_{\theta_i} \equiv \sqrt{\langle\theta_i^2\rangle - \langle\theta_i\rangle^2}$.

The theoretical lower bound for the standard deviation of an estimator is known as the Cramér-Rao bound. We will later stress the importance of knowing this quantity in the context of setting up a future experiment. Additionally, we can define $\boldsymbol{\theta_{ML}}$ as the maximum likelihood estimator, i.e. that parameter vector which maximizes the likelihood function.

The relationship between our estimate $\boldsymbol{\theta}$ for the model parameters and the shape of the likelihood function L is encoded in the Fisher information matrix, which can be defined as:

$$F_{ij} \equiv \left\langle \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j} \right\rangle, \quad \text{where } \mathcal{L} \equiv -\ln L \tag{2.1}$$

and the angled brackets represent an ensemble average, and θ_i is the ith parameter belong-

---

[1] These data could be, for example, the temperature values of each of the pixels in a CMB sky map, or the counts-in-cells of a galaxy redshift survey.

[2] Notice that this is customarily called the "likelihood function": in fact, assuming that x is a fixed dataset, the likelihood will only depend on the parameters $\boldsymbol{\theta}$ (Tegmark et al., 1997)

ing to the parameter vector $\boldsymbol{\theta}$.

The quantities defined so far have been proven to satisfy the following theorems (Kenney & Keeping, 1951; Kendall & Stuart, 1969):

1) for any unbiased estimator, $\sigma_{\theta_i} \geq 1/\sqrt{F_{ii}}$;
2) if there exists a BUE $\boldsymbol{\theta}$, then it is $\boldsymbol{\theta_{ML}}$ or a function thereof;
3) $\boldsymbol{\theta_{ML}}$ is the BUE asymptotically (i.e., for a large set of data).

The first theorem is known as the Cramér-Rao inequality, and it holds only if all the other parameters (i.e. parameters $\theta_j \in \boldsymbol{\theta}$, with $j \neq i$) is known. Whereas, if all the other parameters are estimated from the data as well, then the minimum standard deviation rises to $\sigma_{\theta_i} \geq \sqrt{F_{ii}^{-1}}$. This latter case is the most realistic, since we rarely have an almost complete knowledge on the systems we try to investigate.

Nevertheless, the Cramér-Rao inequality expresses the power of the Fisher matrix as a statistical tool. In fact, knowing the Fisher information matrix allows us to compute the Cramér-Rao bound, corresponding to the minimum theoretical uncertainty we can associate to a parameter. The word "theoretical" means that this lower bound corresponds to an ideal scenario in which the experiment (whatever it is) is being conducted in perfect conditions, so that all kinds of instrumental noise can be neglected (more in general, every effect arising from practical applications which may worsen the precision of the measure can be neglected).

The relevance of this result lies in the fact that before actually setting up the real experiment, one can already get an idea of what precision to expect, given their forecast on how the experiment will play out: in this way, one becomes able to tune the design of their experiment in order to achieve the maximum precision.

The Fisher information matrix is tightly linked to the Covariance matrix, defined as (Tegmark et al., 1997):

$$C \equiv \langle \boldsymbol{\theta}\boldsymbol{\theta}^T \rangle - \langle \boldsymbol{\theta} \rangle \langle \boldsymbol{\theta} \rangle^T \tag{2.2}$$

Diagonal elements represent the variances of the parameters, while off-diagonal terms represent the cross-covariance of the parameters, and tell to which extent the parameters are correlated.

By Taylor-expanding $\mathcal{L}$ around $\boldsymbol{\theta_{ML}}$ (where the function L has a maximum by definition), one can express the Covariance matrix as the inverse of the Hessian matrix:

$$(C^{-1})_{ij} = \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j} \tag{2.3}$$

From such definition, it is rather easy to grasp the tight link between the Fisher matrix and the Covariance matrix. In fact, the Fisher matrix can be seen as the expectation value of such quantity at the point $\boldsymbol{\theta} = \boldsymbol{\theta_0}$ (which coincides with the $\boldsymbol{\theta_{ML}}$ on average if this estimator is unbiased), namely:

$$F_{ij} = \langle (C^{-1})_{ij} \rangle \tag{2.4}$$

If one is looking for a more intuitive definition, then the Fisher matrix can be also viewed as a measure of how fast, on average, the likelihood L falls off around its peak, i.e. the maximum-likelihood point. The larger the value of F_{ij}, the steeper the slope of the L around its maximum, the greater the information we draw from our analysis, and the lower the Cramér-Rao bound.

One more useful feature of the Fisher information matrix is the way it simplifies its own form in the case of a multivariate normal distribution (hereafter, MVN): this case is not so rare, since experimental noise can be often modelled as gaussian.

For example, consider a MVN likelihood. By definition, if a N-dimensional vector of random variables $\boldsymbol{x} = \{x_1, ..., x_N\}$ is MVN distributed, with average $\boldsymbol{\mu}$ and Covariance C, then its likelihood takes the form:

$$L(\boldsymbol{x}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}(\det C)^{1/2}}\ \mathrm{e}^{-\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \tag{2.5}$$

In the most general case, both mean and Covariance depend on the parameters: $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ and $C = C(\boldsymbol{\theta})$.

The log-likelihood is

$$2\mathcal{L} = -2\ln L = N\ln(2\pi) + \log\,\det\,C + (\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \tag{2.6}$$

Since C is a symmetric square array, it satisfies the following property:

$$\ln(\det\,C) = \mathrm{Tr}\,[\ln\,C] \tag{2.7}$$

We now define the data matrix as $D_{ij} = D_{ij}(\boldsymbol{x},\boldsymbol{\theta}) = (x_i-\mu_i)(x_j-\mu_j)$. With that, we can rewrite:

$$(\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \sum_{i,j}(C^{-1})_{ij}(x_i-\mu_i)(x_j-\mu_j) = \mathrm{Tr}\left[C^{-1}D\right] \tag{2.8}$$

so that the log-likelihood becomes:

$$-2\ln L = \mathrm{const} + \mathrm{Tr}\left[\ln\,C + C^{-1}D\right] \tag{2.9}$$

In order to compute the Fisher matrix, we need to compute the derivatives of the log-likelihood with respect to parameter, and then take the ensemble average of both sides of (2.9). All steps for this derivation can be found in *Appendix B*. Here, we directly present the main result, i.e. the form of the Fisher matrix for a MVN likelihood:

$$F_{\alpha\beta} = \frac{1}{2}\mathrm{Tr}\left[C^{-1}C_{,\alpha}C^{-1}C_{,\beta} + C^{-1}(\boldsymbol{\mu}^T_{,\alpha}\boldsymbol{\mu}_{,\beta} + \boldsymbol{\mu}^T_{,\beta}\boldsymbol{\mu}_{,\alpha})\right] \tag{2.10}$$

This result is extremely powerful. If the data $\boldsymbol{x}$ are MVN distributed, and we know how the means $\boldsymbol{\mu}$ and the covariance matrix C depend on the parameters, we can calculate the Fisher matrix before carrying out the experiment: no data are needed.

2.2 Fisher matrix formalism: application to galaxy surveys

We now specify the expression of the Fisher information matrix when applied to analysing galaxy redshift surveys data (Tegmark, 1997; Seo & Eisenstein, 2003).

In this case, the elements in our vector $\boldsymbol{x}$ will correspond to values of the average power spectrum measured in thin shells in Fourier space . We decide to work in Fourier space since this choice brings great simplification: in fact, for a linear perturbative regime, the covariance of Fourier modes is diagonal, as a result of homogeneity and isotropy. The nth shell will have radius k_n, width dk_n and volume (Tegmark, 1997):

$$V_n = \frac{1}{(2\pi)^3} 4\pi k_n^2 dk_n \tag{2.11}$$

If we choose these shells to be thick enough so that each of them contains many uncorrelated modes, then the distribution of our data points $\boldsymbol{x}$ will approximate a multivariate gaussian, due to the central limit theorem (Tegmark, 1997). This allows one to exploit the simplification expressed by (2.10).

In particular, in our galaxy survey case, we can write (Tegmark, 1997; Feldman, Kaiser, & Peacock, 1993):

$$\boldsymbol{\mu}_n \approx P(k_n) \tag{2.12}$$

$$\boldsymbol{C}_{mn} \approx 2\frac{P(k_n)P(k_n)}{V_n V_{eff}(k_n)}\delta_{mn} \tag{2.13}$$

where

$$V_{eff}(k) \equiv \int \left[\frac{\bar{n}(\boldsymbol{r})P(k)}{1+\bar{n}(\boldsymbol{r})P(k)} \right]^2 d^3r \tag{2.14}$$

The symbol $\bar{n}(\boldsymbol{r})$ represents the survey selection function, which gives the theoretical expectation value for the number density of galaxies at a given position $\boldsymbol{r}$.

Therefore, V_{eff} can be interpreted as the effective volume, i.e. that portion of the total surveyed volume which is actually used to extract relevant information about the power spectrum. In fact, the effective volume only accounts for those regions in which the signal $P(k)$ exceeds the Poissonian shot noise $1/\bar{n}$.

We now introduce the flat-sky approximation, in which the survey box is considered to be far from the observer, so that the line of sight direction is the same through the whole survey. In this approximation, the selection function can be taken as independent on position, and we can introduce the quantity

$$\mu \equiv \frac{\boldsymbol{k} \cdot \hat{\boldsymbol{r}}}{k} = \frac{k_\parallel}{k} \tag{2.15}$$

where $\hat{\boldsymbol{r}}$ is the unit vector along the line of sight and $\boldsymbol{k}$ is the wavevector with norm $k = |\boldsymbol{k}|$. Therefore, μ corresponds to the cosine of the angle between the wavevector and the line of sight.

Consequently, in flat-sky we have

$$V_{\text{eff}}(k,\mu) = \left[\frac{\bar{n}P(k,\mu)}{1+\bar{n}P(k,\mu)} \right]^2 V_{\text{survey}} \tag{2.16}$$

where V_{survey} indicates the entire surveyed volume.

In the limit where $V_n V_{\text{eff}}(k,\mu) \gg 1$ (corresponding to assuming thick layers in k-space), the second term in (2.10) will be completely dominated by the first one (Tegmark, 1997).

Replacing (2.12), (2.13) into (2.10), we obtain the following expression for the Fisher information matrix in flat-sky, for a single redshift value (Seo & Eisenstein, 2003):

$$
\begin{aligned}
F_{\alpha\beta} &= \frac{1}{2(2\pi)^3}\int_{\boldsymbol{k}_{\min}}^{\boldsymbol{k}_{\max}} \frac{\partial \ln P(\boldsymbol{k})}{\partial \theta_\alpha}\frac{\partial \ln P(\boldsymbol{k})}{\partial \theta_\beta} V_{\text{eff}}(\boldsymbol{k}) d\boldsymbol{k} = \\
&= \frac{2\pi}{2(2\pi)^3}\int_{-1}^{1}\int_{k_{\min}}^{k_{\max}} \frac{\partial \ln P(k,\mu)}{\partial \theta_\alpha}\frac{\partial \ln P(k,\mu)}{\partial \theta_\beta} V_{\text{eff}}(k,\mu) k^2 dk d\mu
\end{aligned}
\tag{2.17}
$$

where we have exploited the azimuthal symmetry around the line of sight.

Since the survey size defines the maximum scale from which we can extract information, it is reasonable to take

$$
k_{\min} \approx 2\pi/(V_{\text{survey}})^{1/3} \tag{2.18}
$$

Conversely, the value of $k_{\max}$ stands for the minimum scale from which we can extract valuable information. Things get problematic when addressing increasingly smaller scales, since non-linear effects get more and more prominent. The choice of $k_{\max}$ is thus rather arbitrary, but this is one of the cases in which the Fisher forecast can help us choose the most suitable value, i.e. that which provides us with the most information.

While V_{eff} is linked to the technical features of our survey, the derivatives of the power spectrum with respect to the model parameters represent the portion of information strictly related to cosmology. Such derivatives must be computed at a chosen reference cosmology: the Fisher matrix thus provides the Cramér-Rao bound on an unbiased estimator of a given parameter, taking as the true value of the parameters those of the fiducial cosmology.

The $P(k,\mu)$ appearing in (2.17) stands for the observed linear power spectrum of galaxies, $P_{\text{g,obs}}(k_{\text{ref}},\mu_{\text{ref}})$. We already recovered the form of such quantity in chapter 1, after a thorough analysis.

In the remaining part of this chapter, we present a simple practical application of Fisher Forecasting, when applied to two different galaxy surveys: SPHEREx and ATLAS. The reference cosmology employed in the following analysis corresponds to that of (Ade et al., 2016), and reported in *Table* 2.1.

In this analysis, we aim at recovering constraints on the following set of redshift-dependent

Reference cosmology								
Cosmology	w_{cdm}	w_b	h	σ_8	n_s	M_ν $[eV]$	Ω_m	Ω_k
Planck	0.12	0.022	0.67	0.816	0.96	0.06	0.317	0.0

Table 2.1: Adopted reference values of cosmological parameters for the Fisher Forecast application: they are close to the Planck best fit cosmology (Ade et al., 2016).

parameters: $\{H(z), D_M(z), b(z), f(z), P_{sn}(z)\}$. They represent, respectively, the Hubble parameter (containing information about the expansion history of the Universe), the angular diameter distance, the linear galaxy bias, the growth rate of linear perturbations and the shot noise term (whose fiducial value is set to $P_{sn}(z) = 0$ at each redshift). Therefore, the number of parameters effectively entering in the Fisher matrix is given by $\{H(z), D_M(z), b(z), f(z), P_{sn}(z)\} \times N_z$, where N_z is the number of redshift bins considered by each survey.

Importantly, it must be noted that excluding a parameter from the Fisher matrix corresponds to assuming that one has a complete knowledge over such parameter. Therefore, those parameters excluded from the Fisher matrix will be effectively fixed to their fiducial values. In this analysis, a really strong assumption has been made: namely, that we have complete knowledge over all the (redshift-independent) cosmological parameters. Such assumption is rather unrealistic, but it is nonetheless a great simplification for this analysis, whose purpose is solely to highlight the differences in constraining power between the three galaxy surveys.

Specifications for the Fisher Forecasts performed with each of the two surveys can be found in *Appendix C*. There, redshift bins are indicated, as well as effective redshifts, computed at the center of each bin. We also show corresponding values of the galaxy number density, expressed in $(h/\mathrm{Mpc})^3$, and galaxy bias for each bin. All the three surveys follow the same prescription in the definition of the k and μ intervals over which the integrations in (2.17) are performed. In particular, k_{min} is computed from each bin's survey volume, as in (2.18), while $k_{\mathrm{max}} = 0.2$ for all the three surveys, and a total of 200 logarithmic spaced k-bins have been employed[3]; additionally, 50 values of μ ranging linearly from -1 to 1 have been employed.

For the SPHEREx survey analysis (Doré et al., 2014), N_z=11 redshift bins have been employed, ranging from $z_{\mathrm{min}} = 0.0$ to $z_{\mathrm{max}} = 4.6$, with a variable bin size: $\Delta z = 0.2$ for the first 5 bins, while $\Delta z = 0.2$ for the last 6 bins. The redshift uncertainty adopted for the SPHEREx analysis is $\sigma_z = 0.003$. For this survey, 75% of an all-sky coverage has been assumed, amounting to $\simeq 31000$ square degrees.

For the ATLAS survey analysis (Wang et al., 2019), N_z=18 redshift bins have been em-

---

[3] It has been tested that this number is sufficient to reach saturation in the Fisher Forecast code.

ployed, ranging from $z_{\min} = 0.5$ to $z_{\max} = 4.1$, with bin size $\Delta z = 0.2$. The redshift uncertainty adopted for the ATLAS analysis is $\sigma_z = 0.0001$, while the sky coverage amounts to 2000 square degrees. In this case, the bias is computed as $b(z) = 1.5 + 0.4(z - 1.5)$.

Therefore, a Fisher Forecast analysis is performed for each of the two considered galaxy surveys[4]. One of the results of these analyses is presented in *Figure* 2.1. Here, we show the relative uncertainty on the parameter f, obtained after marginalizing over all the other parameters, as a function of the redshift. This is displayed for the two cases: SPHEREx (blue) and ATLAS (orange). In other words, the plot is showing the (relative) Cramér-Rao bound for the parameter f, i.e. the lowest possible theoretical uncertainty one can achieve from a given galaxy survey, with the given specifications.

A few observations can be made about this plot.

First of all, we need to stress again that this analysis is based on the strong assumption of knowing the true values of cosmological parameters. If we had chosen to pursue a more realistic type of analysis, then we would have had to marginalize over those cosmology parameters as well, thus worsening the achievable precision.

Additionally, the two surveys, although being both quite deep in redshift, seem to produce visibly different results. SPHEREx is expected to constrain very well ($\approx$ few percent constraints) the growth rate of cosmic structure at low-z, since this is where such quantity is most rapidly changing. However, the constraining power decreases when moving to higher redshifts, since the galaxy number density visibly decreases, in accordance with SPHEREx specifications. Conversely, the ATLAS survey, albeit employing a considerably smaller sky coverage ($\approx 1/20$ of the full-sky) is expected to analyse a larger number of galaxies in total: this corresponds to larger values of $n(z)$, thus improving the effective volumes (see (2.14)) and, in turn, the constraining performance.

We close this chapter by presenting a further example of how the Fisher Forecast analysis can be exploited.

The Fisher matrix for a chosen set of parameters does not only contain information limited to those parameters. For example, after having constructed a Fisher matrix with a specific parameter set, one might decide to look for constraints on other parameters, belonging to a different set. In this case, one simply has to project the Fisher matrix already at hand into the new parameter space. Namely, given a starting set $p = (x, y, z)$ and a new set $p' = (a, b, c)$, the

---

[4] The derivatives appearing in (2.17) have been computed numerically via the five-point stencil method, and the derivative step has been opportunely adapted for each of the parameters.

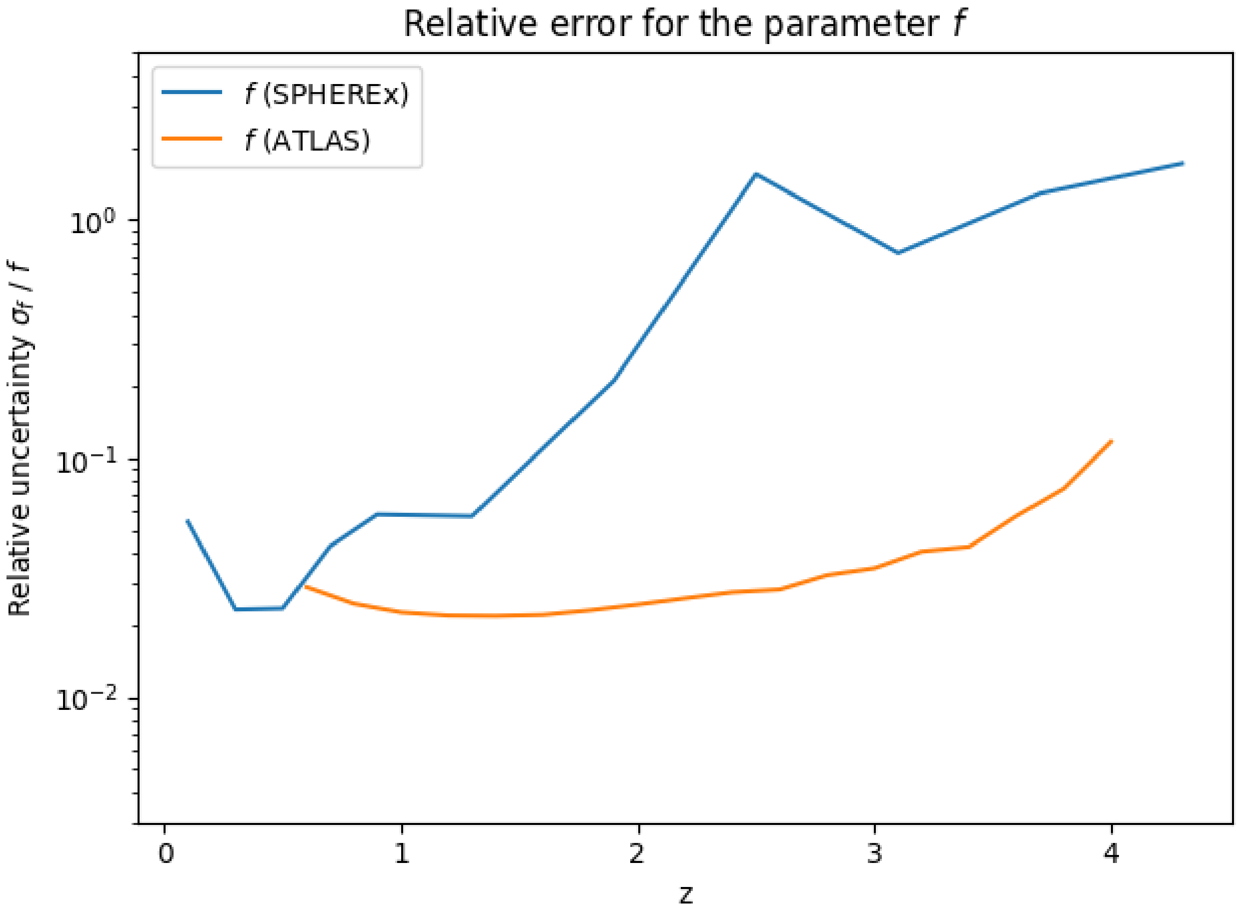

Figure 2.1: Relative uncertainty on the parameter f, obtained after marginalizing over all the other parameters, as a function of the redshift. This is displayed for the two cases: SPHEREx (blue) and ATLAS (orange).

new (projected) Fisher matrix can be obtained as follows (Coe, 2009):

$$F'_{mn} = \sum_{ij} \frac{\partial p_i}{\partial p'_m} \frac{\partial p_j}{\partial p'_n} F_{ij} \tag{2.19}$$

As a practical example, we employ this Fisher matrix projection to study a simple extension of the ΛCDM model. In particular, we are interested in a dynamical dark energy model, i.e. a model featuring an evolving equation of state for the dark energy component. A commonly used and well-tested redshift parametrization is the following (Blanchard et al., 2020):

$$w_{DE}(z) = w_0 + w_a \frac{z}{1+z} \tag{2.20}$$

where w_0 is the present-day value of the equation of state, while w_a is a measure of the time variation.

In this case, the evolution of the dark energy density obeys the law:

$$\rho_{DE}(z) = \rho_{DE,0}(1+z)^{3(1+w_0+w_a)} \exp\left[-3w_a \frac{z}{1+z}\right] \tag{2.21}$$

while the function $E(z)$ (defined in the Premise of this work for a flat ΛCDM) becomes

$$E(z) = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{DE,0}(1+z)^{3(1+w_0+w_a)} \exp\left[-3w_a \frac{z}{1+z}\right] + \Omega_{rad,0}(1+z)^4} \tag{2.22}$$

In order to put constraints on these new ΛCDM-extension parameters, i.e. w_0 and w_a (which respectively take value -1 and 0 within baseline ΛCDM), we start from the Fisher matrices computed beforehand for the ATLAS and SPHEREx galaxy survey, we marginalize with respect to the bias ($b(z)$) and shot noise ($P_{sn}(z)$) parameters, and finally project them to the new parameter space $\{w_0, w_a\}$. Once again, this analysis is based on the strong assumption that everything else about this ΛCDM-extended model is known (Ω_k fixed to 0, $\Omega_{DE,0}$ fixed to $1 - \Omega_{m,0}$). Results are shown in *Figure* 2.2, and they suggest that ATLAS will be able to put tighter constraints on the two dynamical dark energy parameters, with respect to SPHEREx.

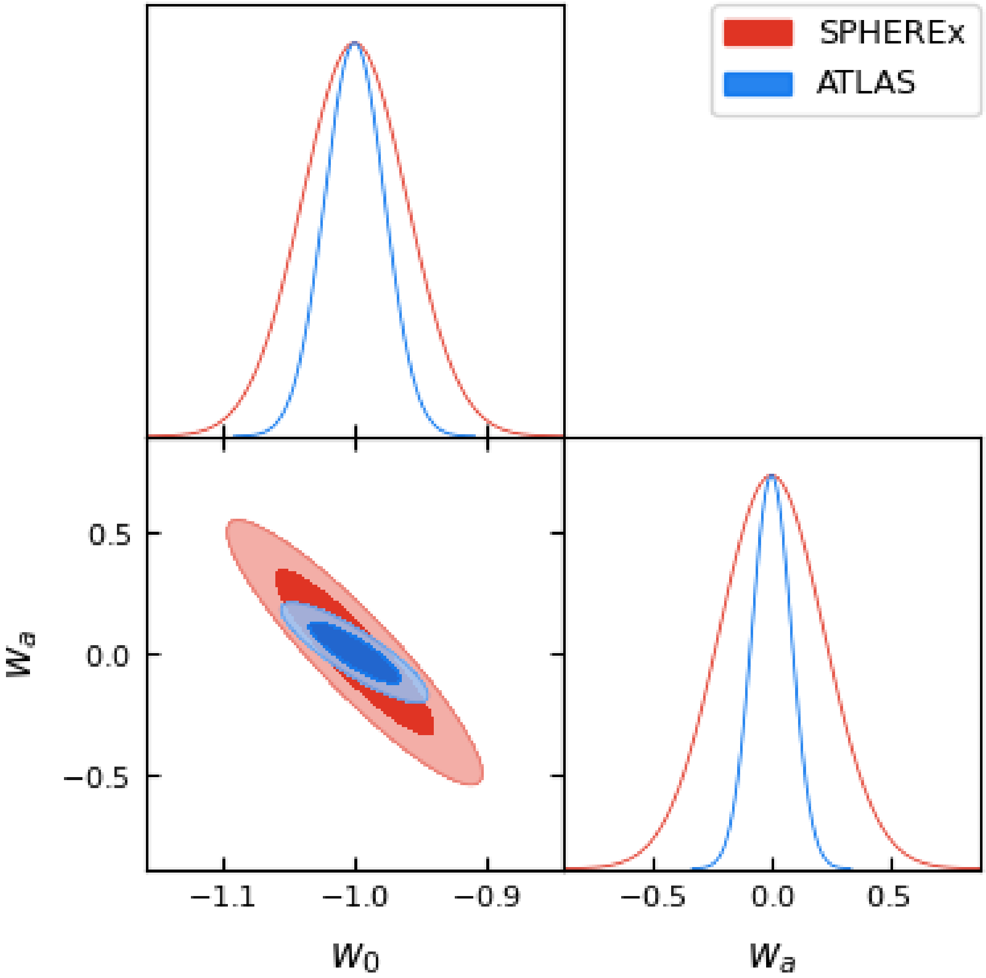

Figure 2.2: Posterior contours for the two dynamical dark energy model parameters, w_0 and w_a, as resulting from the projection of the Fisher matrices computed with the ATLAS (blue) and SPHEREx (red) survey specifications onto the new parameter space $\{w_0, w_a\}$, after marginalizing over bias and shot noise parameters.

Chapter 3

Cosmology with ShapeFit

Before introducing the ShapeFit analysis, it is time to recall the two main methods that have been employed to investigate the LSS and extract information about cosmological parameters. These are the Classic (either BAO or RSD) analysis, and the Full Model (FM) analysis.

This section closely follows (Brieden et al., 2021b), in which an extensive overview of the different techniques is presented.

3.1 Classic approach: overview

The Classic approach to the statistical analysis of LSS is based on exploiting the AP effect and the standard ruler provided by BAO, in order to obtain distance-redshift relations at the effective redshifts of surveys' samples.

Such approach gained popularity after the signal-to-noise ratio of BAO detection from redshift surveys became high enough. This new BAO signal provided new relevant information about the LSS of the Universe, so the community adopted this method in a widespread fashion, hence the name ”classic”.

To be more specific, the Classic approach can be split into Classic BAO and Classic RSD. Even though they feature a slightly different data pipeline, and different quantities are involved, the core technique is the same: they both manage to compress the power spectrum data into few physical observables, which are only sensitive to late-time physics. The great advantage of this kind of analysis is that, in order to compute such few variables, one needs to build a template power spectrum only once and from a reference cosmology, while no cosmological model is required to be chosen *a priori*. It is only after having computed the compressed variables that a cosmological model is chosen, and these model-independent variables are interpreted in light of it. The data compression employed by the Classic approach allows for a considerable computational advantage, as it will appear more clear from a comparison with the FM approach.

3.1.1 Classic BAO analysis

In the case of the Classic BAO, the analysis focuses on the information carried by the BAO peak position along and across the line of sight. The collected data is separated into two components, a wiggle and a broadband (non-wiggle) one: the former contains the BAO information, while the latter does not, and it is thus marginalized over. On the other side, a template power spectrum is constructed from the chosen reference cosmological model. The aim is to employ such template to perform a fit of the collected data.

Now, the power spectrum we observe (i.e. the data at our disposal) will inevitably be a function of the chosen reference cosmology, and the same can be said for the template. So, in principle, we should take these dependences into account by rescaling both quantities accordingly: to be more specific, we should use some late-time rescaling parameters for the data, and some early-time rescaling parameters for the template. But actually the standard habit consists in combining these together into new rescaling parameters (which now take into account both late and early time), and applying them to the template only. This choice, which is simply driven by practical reasons, means that the data remain untouched (i.e. they preserve their dependence from the reference cosmology), while the template is adapted to the data.

These new scaling parameters are defined as (Brieden et al., 2021b; Beutler et al., 2014; Alcock & Paczyński, 1979):

$$\alpha_\perp(z) = \frac{D_M(z) r_d^{ref}}{D_M^{ref}(z) r_d} \quad , \qquad \alpha_\parallel(z) = \frac{H^{ref}(z) r_d^{ref}}{H(z) r_d} \tag{3.1}$$

Here, D_M/r_d and $1/Hr_d$ respectively indicate the comoving angular diameter distance and the Hubble distance (in natural units), in units of the sound horizon at the end of drag epoch. The presence of the superscript ref indicates that the given quantity is being computed in the chosen reference cosmology, while the lack of it means that we are dealing with true quantities.

With this in mind, the scaling parameters (3.1) express the ratio between true and reference distances, in units of the sound horizon, across and along the line of sight. Expressing them in units of r_d is not a choice, but rather a necessity: the length of r_d is set by early-time physics and it cannot be recovered from late-time observations, such as LSS, and without early-time physics assumptions. Therefore, the two scaling parameters cannot be used to probe absolute distances, but only distances with respect to the standard ruler r_d.

It is obvious that if our chosen reference cosmology differs from the true one (which realistically is always the case), then $\alpha_\perp$ and $\alpha_\parallel$ will be different from 1, and all the observed geometrical quantities will be distorted compared to true ones.

Such distortion reflects the AP effect, and it is exploited to get information about the BAO peak position. In fact, if we assume that the Universe is homogeneous and isotropic, then the BAO should be a symmetric structure along all spatial directions. The AP effect would break this symmetry, leading to an apparent anisotropy, i.e. an excess in the relative BAO scales along and across the line of sight. Reverting back this line of thought, we can see that measuring such anisotropy will allow us to infer what the true cosmology of the Universe might be (Gil-Marín et al., 2020).

The AP effect also produces a distortion[1] in the power spectrum wavenumber k and in the cosine of the angle between the wavevector and the line of sight, μ

$$k \to \tilde{k} = \frac{k}{\alpha_\perp}\left[1 + \mu^2\left(\frac{\alpha_\perp^2}{\alpha_\parallel^2} - 1\right)\right]^{1/2} \tag{3.2}$$

$$\mu \to \tilde{\mu} = \mu\frac{\alpha_\perp}{\alpha_\parallel}\left[1 + \mu^2\left(\frac{\alpha_\perp^2}{\alpha_\parallel^2} - 1\right)\right]^{-1/2} \tag{3.3}$$

The transformations (3.2) and (3.3) are applied to the smooth component of the template. Therefore, the template power spectrum multipoles can be finally expressed as[2]:

$$P^{(\ell)}(k) = \frac{(2\ell+1)}{2\alpha_\perp^2\alpha_\parallel}\int_{-1}^{1} P_{model}(\tilde{k}(k,\mu), \tilde{\mu}(\mu))\mathcal{L}_\ell(\mu)d\mu \tag{3.4}$$

---

[1] Notice that these transformations greatly resemble those presented in section 1.2.1 , but on a closer look they appear slightly different. This is because here rescaling parameters are defined in a different way: we are assuming that $\tilde{k}_\parallel = k_\parallel/\alpha_\parallel$ and $\tilde{k}_\perp = k_\perp/\alpha_\perp$. Compare with (1.30) to appreciate the difference.

[2] This is an approximation, since a prefactor $(r_d^{ref}/r_d)^3$ should actually be included in the volume rescaling. In (Gil-Marín et al., 2020), the impact of such approximation is investigated.

By using this template to fit the measured multipoles, we are able to extract, in a given redshift bin, the value of $\alpha_\perp$ and $\alpha_\parallel$, i.e. two parameters containing information about the BAO position. Importantly, these two rescaling parameters describe the geometry and expansion history of the Universe in a model-independent way.

We now spend a few words to clarify the role of the reference cosmology in this approach[3].

In principle, these Classic techniques require the choice of two reference cosmologies: one to convert redshift and angular position measurements into distances, the other to build the power spectrum fitting template. Even though these cosmologies could in principle be different, the standard habit is to make them coincide.

Moreover, our final results will depend on our initial choice of the reference cosmology, which is in principle totally arbitrary . In this regard, mock catalogues become a valuable ally, since they allow us not only to check if our pipeline brings some systematic errors, but also to explore the impact of choosing different reference cosmologies. In (Gil-Marín et al., 2020), the Nseries mocks (Alam et al., 2017) have been employed in order to recover the values of $\alpha_\parallel$ and $\alpha_\perp$ when starting from different reference cosmologies. In particular, the highest deviations for these two parameters, with respect to $\alpha_\parallel$ and $\alpha_\perp$ computed from the true cosmology of Nseries, amount to 1% for $\alpha_\parallel$ and 0.8% for $\alpha_\perp$. These are obtained from an adopted reference cosmology which is very distinct from the true cosmology of Nseries, with shifts of $\Delta\Omega_m = 0.08$ and $\Delta\Omega_b = 0.019$, which are 10 and 50 sigma away, respectively, from the results reported by Planck (Aghanim et al., 2020). On the other hand, if a closer-to-standard ΛCDM reference cosmology is used, such as $\{\Delta\Omega_m = 0.025, \Delta\Omega_b = 0.011\}$, these shifts reduce to 0.5% on $\alpha_\parallel$ and 0.3% for $\alpha_\perp$. These results thus state the robustness of the technique.

3.1.2 Classic RSD analysis

This analysis follows almost the same steps as the Classic BAO one, but this time the focus is not centred only on the BAO position: BAO amplitude and anisotropy signal induced by RSD are now relevant as well.

Accordingly, in this case no decomposition is exerted on the power spectrum template, but the rescaling is instead applied to the full $P^{(\ell)}_{RSD}(k)$.

As we know, RSD correspond to an enhanced clustering signal along the line of sight due to the large-scale bulk velocity component of galaxies. Such bulk velocity is related to the growth

---

[3] The same considerations are valid for the Classic RSD analysis as well.

rate f. Thus, the Classic RSD approach provides information about the growth of structures: more specifically, $f\sigma_8$ is chosen as third compressed variable (thus joining $\alpha_\perp$ and $\alpha_\parallel$).

This choice is made to get rid of the degeneracy which haunts f and σ_8 in a 2-point statistics (in other words, any change in σ_8 is reabsorbed into f).

Actually, to be more precise, the third compression variable is not $f\sigma_8$, but $f\sigma_{8s}$, where[4]

$$\sigma_{8s} \equiv \sigma(R = s \cdot 8h^{-1}\mathrm{Mpc}, \mathbf{\Omega}) \tag{3.5}$$

and

$$s \equiv \frac{r_d}{r_d^{ref}} \tag{3.6}$$

Such redefinition of the matter fluctuation amplitude is promptly motivated. The original definition of σ_8 features the presence of $P(k, \mathbf{\Omega})$ in the integrand. But we know that this quantity is subject to a rescaling through $\alpha_\parallel$ and $\alpha_\perp$, which in turn depend on the chosen cosmology. Consequently, the value of σ_8 changes as the parameter space is explored during the RSD fitting process. The redefinition (3.5) is thus employed in order to keep σ_8 fixed to its reference value[5], as promptly shown:

$$\begin{aligned}
\sigma_{8s}^2(\mathbf{\Omega}) &= \int_0^\infty d(\ln k) k^3 P_{lin}(k, \mathbf{\Omega}) W_{TH}^2(ks \cdot 8h^{-1}\mathrm{Mpc}) && |\mathbf{\Omega} \to \mathbf{\Omega}^{ref} \\
&= \int_0^\infty d(\ln k) k^3 s^3 P_{lin}(sk, \mathbf{\Omega}^{ref}) W_{TH}^2(ks \cdot 8h^{-1}\mathrm{Mpc}) && |k' = ks \\
&= \int_0^\infty d(\ln k') k'^3 P_{lin}(k', \mathbf{\Omega}^{ref}) W_{TH}^2(k' \cdot 8h^{-1}\mathrm{Mpc}) \\
&= \sigma_8^2(\mathbf{\Omega}^{ref})
\end{aligned} \tag{3.7}$$

---

[4] In Eq. (1.19) we were defining a present-day matter fluctuation amplitude. A more general definition would be $\sigma^2(R, z, \mathbf{\Omega}) = \frac{1}{2\pi^2} \int dk k^2 P_m(k, z, \mathbf{\Omega}) |W_{TH}(kR)|^2$, making explicit the scale, redshift and cosmology dependences.

[5] If the data from a galaxy survey are combined with those coming e.g. from Planck, then this new set of data will fix σ_8 to its reference value. But this is not the case if only galaxy-survey data is employed.

In the above expression we employed the early-time rescaling of the power spectrum which, conceptually[6], can be written as

$$P^{lin}_{model}(k, \mathbf{\Omega}) = s^3 P^{lin}_{model}(sk, \mathbf{\Omega}^{ref}) \tag{3.8}$$

The rescaling (3.8) reflects the fact that, within the Classic approach, all early-time physics is assumed to be captured by the sound horizon r_d (i.e. r_d parametrizes all signatures of early-time physics). This assumption comes from fixing the power spectrum template after choosing a reference cosmology (as we will see, the FM approach follows a different route). This also explains why the smoothing scale of the redefined fluctuation amplitude σ_{8s} is defined relative to the sound horizon scale. As a final word, it may be useful to notice that all the three physical compressed variables measured by the RSD analysis, $\alpha_\parallel$, $\alpha_\perp$ and $f\sigma_{8s}$, are given in terms of the sound horizon ratio s.

3.2 Full Modelling approach

As previously stated, the great advantage of the Classic approach lies in the data compression step, which allows to obtain model-independent variables. But these variables only depend on late-time geometry and kinematics, while they remain independent of physical processes which take place in earlier epochs, and which determine the shape of the matter transfer function. This implies that the information we can extract from the Classic approach is somewhat limited: employing a fixed power spectrum template could be both a blessing and a curse.

The FM approach follows a different route: it avoids the compression step, and cosmological models are directly fitted to $P^{(\ell)}(k)$. During the fit, as cosmological parameter space is explored (usually via MCMC), the model prediction for the transfer function, along with non-linear corrections to the power spectrum, is computed for every choice of parameters. It appears evident that this requires a larger computational effort compared to the Classic analysis.

Moreover, while the Classic analysis doesn't require the choice of any cosmological model

---

[6] In reality, full angular dependence should be considered.

before the compression step, within FM one has to choose a model in order to start their analysis, meaning that the model-independence of the Classic analysis is lost in the context of FM.

But in exchange to these downsides, the FM approach grants better constraints on the inferred parameters. In fact, within FM, parameters will depend both on geometry and on the transfer function, thus leading to an interpretation which is not only geometrical (as it is the case in the Classic approach).

A tangible example from (Ivanov, Simonović, & Zaldarriaga, 2020; Philcox, Ivanov, Simonović, & Zaldarriaga, 2020): when imposing a Big Bang Nucleosynthesis (BBN) prior, FM is able to obtain a 1.6% constraint on the Hubble constant, while the same parameter is mildly constrained (10%) in the Classic approach (Alam et al., 2021).

Of course the increase in precision brought by FM is only evident when constraining cosmology with galaxy clustering data alone. Conversely, combining LSS data with other datasets, especially CMB data, would fix the sound horizon scale and the transfer function shape, leaving geometry as the main source of information to be exploited, and effectively taking us back to the philosophy of the Classic analysis.

3.3 Classic approach: pipeline in detail

Here, we take a deeper dive into the pipeline of the Classic approach, which essentially can be reduced in four main branches: constructing the model, acquiring the data, fitting one to other, and finally interpreting the result of the fit.

3.3.1 Model pipeline

After having chosen a reference cosmology, the template non-linear galaxy power spectrum in redshift space can be built. The first step is to compute the linear one $P_{lin}(k, \boldsymbol{\Omega^{ref}})$ (corresponding to (1.52)). This will be needed in order to compute the non-linear power spectrum $P_{nl}(k, \boldsymbol{\Omega^{ref}})$ (1.54). Lastly, the power specturm multipoles $P^{(\ell)}_{RSD}(k)$ are computed, via $\alpha_{\parallel}$, $\alpha_{\perp}$, f, using (3.4).

The final touch consists in adding the window function contribution: the observed power spectrum multipoles (which are extracted from the data) contain the effect of the survey selection function convolved with the actual galaxy power specturm signal. In order to perform an unbiased analysis, one needs to include the effect of the survey selection function in the theory model as well. This is done by multiplying the Hankel transform of the theory-predicted power spectra multipoles by some quantities $W^{\ell}(s)$, which will be better defined in the next paragraph.

3.3.2 Data pipeline

We will now illustrate the pipeline which starts from the data acquired by a survey (or generated by a mock catalogue) and leads to the quantity to which the template will be eventually fitted. Such data pipeline makes extensive use of mock catalogues, so it might be wise to introduce them.

Mock catalogues are designed to generate many realizations of objects distributions which reproduce the angular and radial selection function and small scale clustering of real data. In general, they are employed to estimate the covariance matrix, quantify the impact of systematic errors and to validate the pipeline and methods employed on the data.

Apart from mocks, this pipeline also makes use of so-called "random catalogues": they contain α_{ran}^{-1} times more objects than individual mocks, and feature the same selection function, but no intrinsic clustering (apart from that spuriously generated by the selection function).

Conversion into distances

The first step consists in converting the redshift and angular coordinates collected from the survey into distance measurements. This is obviously done on the basis of the chosen reference cosmology, and via the following equations[7] for the Hubble parameter H, the Hubble distance D_H and the comoving angular distance D_M:

$$H(z) = H_0\sqrt{(1+z)^4\Omega_r + (1+z)^3\Omega_m + \Omega_\Lambda} \tag{3.9}$$

$$D_H(z) = \frac{c}{H(z)} \tag{3.10}$$

$$D_M(z) = \int_0^z \frac{cdz'}{H(z')} \tag{3.11}$$

---

[7] Assuming we are working in the context of a flat ΛCDM model, the budget equation to be fulfilled is the following: $\Omega_\gamma + \Omega_{\nu,rel} + \Omega_{cdm} + \Omega_b + \Omega_{\nu,massive} + \Omega_\Lambda = 1$.

where c is the speed of light, and $\Omega_r = \Omega_\gamma + \Omega_{\nu,rel}$ is the radiation energy budget (i.e. the sum of the photon and massless neutrinos energy budgets).

Power spectrum estimator

The next step consists in measuring the power spectrum multipoles $P^{(\ell)}_{data}$, from the collected data. This requires the definition of a power spectrum estimator. The procedure to obtain such estimator is the following.

Once redshifts have been converted into distances, galaxies are placed into a cubic box of length L_{box} (expressed in [Gpc/h]). Galaxies and random objects are then assigned to a 512^3 cells grid, thus obtaining their respective densities per cell, $n(\boldsymbol{r})$ and $n_{ran}(\boldsymbol{r})$. These are used to define the weighted galaxy fluctuation field $F(\boldsymbol{r})$ (also known as Feldman-Kaiser-Peacock, or FKP, function) (Feldman et al., 1993):

$$F(\boldsymbol{r}) = \frac{w_{FKP}(\boldsymbol{r})}{\sqrt{I_2}}[w_{cp}(\boldsymbol{r})n_g(\boldsymbol{r}) - \alpha_{ran}n_{ran}(\boldsymbol{r})] \tag{3.12}$$

$$I_2 = \int d^3\boldsymbol{r} w_{FKP}(\boldsymbol{r})\langle w_{cp}(\boldsymbol{r})n_g(\boldsymbol{r})\rangle^2 \tag{3.13}$$

Here, $n_g(\boldsymbol{r}) = \sum_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)$, with $\boldsymbol{r}_i$ being the location of the i^{th} galaxy, and similarly for the random mock. The term w_{FKP} is the FKP weight, which can be expressed as a function of redshift as $w_{FKP}(z) = 1/(1 + \bar{n}(z)P_0)$, with $\bar{n}$ being the average number density. This weight is used to minimize the power spectrum variance at the typical BAO scale $P_0 \equiv P(k = 0.1\ \ h/\mathrm{Mpc})$.

Instead, the term w_{cp} is the close-pair weight, accounting for galaxy pair neighbours which are closer than the instrument angular resolution (limited by the fiber size)[8].

Therefore, the power spectrum multipoles are obtained via Fourier transformations following the Yamamoto approximation (Yamamoto, Nakamichi, Kamino, Bassett, & Nishioka, 2006;

---

[8] For example, in (Gil-Marín et al., 2020), this weight is defined as $w_{cp} = N_{targ}/N_{spec}$, i.e. the ratio between the number of targeted objects and the number of objects with actual spectroscopic observation. Anyway, both w_{FKP} and w_{cp} can be provided by mock catalogues.

Bianchi, Gil-Marín, Ruggeri, & Percival, 2015):

$$P^{(\ell)}_{\text{data}}(k) = \frac{2\ell+1}{2} \int \frac{d\Omega}{4\pi} \left[\int d\boldsymbol{r}_1 F(\boldsymbol{r}_1) e^{i\boldsymbol{k}\cdot\boldsymbol{r}_1} \int d\boldsymbol{r}_2 F(\boldsymbol{r}_2) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_2} \mathcal{L}_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}_2) \right] - P^{(\ell)}_{sn} \tag{3.14}$$

which assigns the varying line of sight towards one of the galaxies of each pair. The Poisson shot noise term $P^{(\ell)}_{sn}$ is subtracted from the monopole only, as it is null for $\ell > 0$.

Multipoles are computed in bins Δk.

Computing covariance matrix

In this step, mock catalogues are employed to compute the covariance matrix, which in turn provides error bars for the multipoles, and correlations between different k-bins.

In particular, if the mocks provide us with the power spectra multipoles $P_\ell(k)^n$, where $n \in \{1 \ldots N\}$ is an index which marks the nth realization, and N is the total number of realizations, then the covariance matrix can be computed as (Hartlap, Simon, & Schneider, 2007):

$$C_{ij} = \frac{1}{N} \sum_n^N \left[P_\ell(k_i)^n - \mu(k_i)\right] \left[P_\ell(k_j)^n - \mu(k_j)\right] \tag{3.15}$$

where the mean is computed as

$$\mu(k_i) = \frac{1}{N} \sum_n^N P_\ell(k_i)^n \tag{3.16}$$

Measuring the window function

The effect of the survey selection function is modelled by the quantities $W^\ell(s)$. They correspond to the "window function pair-counts" functions performed on the random catalogue

(Beutler et al., 2014) (Wilson, Peacock, Taylor, & de la Torre, 2017):

$$W^{\ell}(s) = \frac{2\ell+1}{I_2\alpha_{ran}^{-2}} \sum_{j>i}^{N_{ran}} \frac{w_{cp}(\boldsymbol{r}_i)w_{FKP}(\boldsymbol{r}_i)w_{cp}(\boldsymbol{r}_j+\boldsymbol{s})w_{FKP}(\boldsymbol{r}_j+\boldsymbol{s})}{2\pi s^2\Delta s}\mathcal{L}_\ell\left(\frac{\boldsymbol{x}_i\cdot\boldsymbol{s}}{x_i s}\right) \tag{3.17}$$

where s marks the position in redshift space.

The pair-count for each redshift bin is normalized by the associated volume given by $2\pi s^2\Delta s$, where Δs is the binning size of the s-count, and the $j > i$ condition prevents counting the same pairs twice. The window function is normalized by $I_2\alpha_{ran}^{-2}$ in order to account for the difference in number density between the random and data catalogue, and to ensure the same normalization of (3.14).

3.3.3 Compression step

At this point, one should have in their hands both the model power spectrum multipoles, including the $W^{(\ell)}(s)$ terms, and the multipoles obtained from the data, along with errors bars. It is now time to get constraints on the compressed parameters $\alpha_\parallel$, $\alpha_\perp$ and $f\sigma_{8s}$, and this is done via a MCMC fit. In practice, the model multipoles are fitted to the data ones, while the compressed parameters are being varied (along with nuisance parameters).

3.3.4 Cosmology inference step

Once the compressed variables have been found for every redshift bin, they are treated as new input data, and compared to any cosmological model of choice: the full covariance between the compressed variables is included in computing the likelihood, and the resulting parameters posterior is sampled via MCMC (in particular, the χ^2 is computed for the theoretical prediction for each quantity, obtained from an input cosmological model and parameters values).

It might be interesting to highlight that these last steps would appear different in a FM pipeline. In fact, within a FM context, they are actually replaced by one single step, during which nuisance parameters are varied along with cosmological ones, while the physical parameters ($\alpha_\parallel$, $\alpha_\perp$ and $f\sigma_{8s}$) are not varied, since they are derived from the chosen cosmological model.

The Classic approach pipeline is illustrated in *Figure 3.1.* We have put a lot of focus on illustrating this pipeline in order to better understand the novelty brought by ShapeFit. In fact, ShapeFit can be thought as an extension of the Classic RSD approach, so the ShapeFit pipeline will take the one illustrated in *Figure* 3.1 as a starting point, and simply add a few more steps and variables to it.

3.4 Introducing ShapeFit

In the previous sections we explored the advantages and disadvantages of the most widely used LSS analysis technique. It is now time to introduce the ShapeFit approach, a new technique initially presented in (Brieden et al., 2021b), which attempts to take the best from its predecessors, while overcoming their respective limitations.

The ShapeFit technique has much to share with the Classic RSD approach, such that it can easily be considered as an extension of it. ShapeFit follows the same philosophy of compressing information into few parameters, to be later interpreted: this means that this new technique preserves the model-independence of the Classic ones. But the revolution brought by ShapeFit consists in adding new compression parameters, which should be able to trace the signatures of early-time physics in the large-scale broadband shape of matter power spectrum. The presence of these additional parameters should allow ShapeFit to reach the same level of precision as the FM approach.

3.4.1 Theoretical parameter dependencies

In chapter 1, we constructed the full model of the observed galaxy power spectrum. As it can be appreciated from (1.52) and the following description, its overall shape and look depends on many parameters and variables. But the part which is most tightly connected to early-time physics is the true linear matter power spectrum $P_m(k, z)$. In the following discussion, we will numerically explore and visualize the way this quantity depends on different parameters.

If not stated otherwise, in what follows we will assume a flat ΛCDM model, and the reference cosmology shown in *Table* 3.1.

To begin with, in *Figure* 3.2 we are plotting the present-day linear matter power spectrum, obtained from the definitions (1.18) - (1.21) and the analytical transfer function provided by (Eisenstein & Hu, 1998). Some characteristic scales are shown as well, making it quite evident that the sound horizon at the end of drag epoch r_d plays a big role in determining the shape of

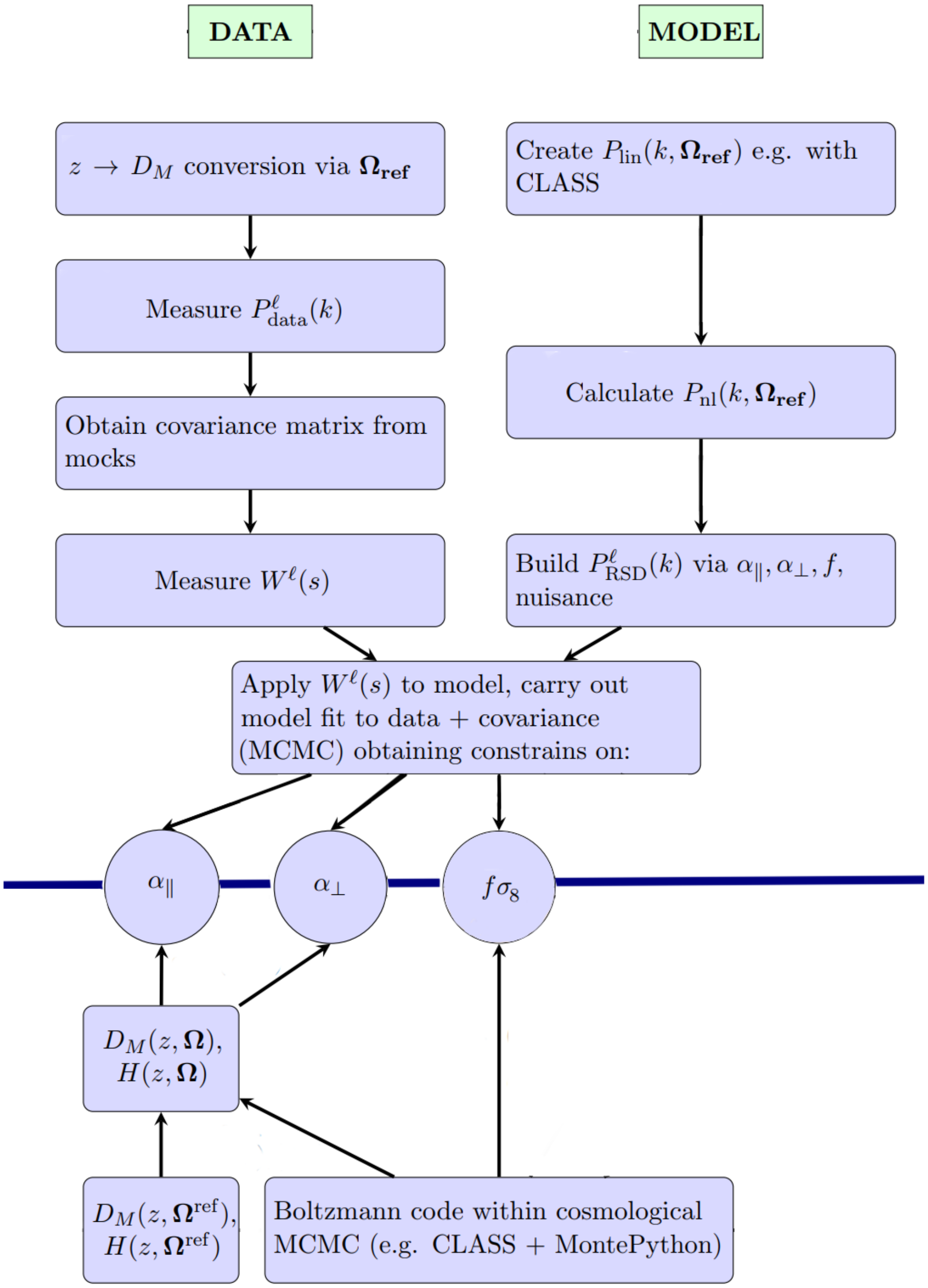

Figure 3.1: Classic approach pipeline. Steps above the blue line correspond to the data and model pipelines, while the cosmological interpretation step is located below the blue line. This scheme has been adapted from *Figure* 5 in (Brieden et al., 2021b).

Reference cosmology								
Cosmology	w_{cdm}	w_b	h	σ_8	n_s	M_ν $[eV]$	Ω_m	r_d $[Mpc]$
Planck	0.1190	0.022	0.676	0.8288	0.9611	0.06	0.31	147.78

Table 3.1: Adopted reference values of cosmological parameters: they are close to the Planck best fit cosmology (Aghanim et al., 2020).

the power spectrum. Hereafter, we list all the displayed scales.

- The Jeans scale for cold dark matter k_{eq}: it corresponds to the main turnaround of the shape.

- The Jeans scale for baryons $k_{damp} = 1/r_d$ (green dashed line): smaller scales get damped with respect to a baryon-less Universe, and wiggles (BAO) appear in the shape.

- The scales $k_{peak} = \dfrac{4n+1}{2}\dfrac{\pi}{r_d}$ (red dotted lines) and $k_{trough} = \dfrac{4n-1}{2}\dfrac{\pi}{r_d}$ (blue dotted lines), corresponding to the BAO peaks and troughs respectively (Eisenstein & Hu, 1998).

- The scale $k_{slope} = \pi/r_d$ (yellow dashed line), corresponds to the zero-crossing before the first BAO trough, and it marks the maximum slope of baryon suppression. This implies a further connection between r_d and the BAO wiggles position.

- Two scales, 0.008 [Mpc/h] and 0.5 [Mpc/h] (pink solid lines), delimiting the range of modes which are usually observed by spectroscopic galaxy surveys.

In *Figure* 3.3 we show the normalized derivative of the linear matter power spectrum with respect to some parameters. Each derivative is normalized by a factor Δ_i in order to match the effect of σ_8 at large scales. As one parameter is being varied, all others are kept fixed to the reference cosmology. As we can see, varying σ_8 only changes the power spectrum global amplitude, while varying n_s changes its global tilt.

Dependence on parameters $\{\omega_m, \omega_b, h\}$ is quite similar on large and small scales, but differences arise in the intermediate range of scales. In particular, BAO wiggles show changes both in their position and amplitude. In other words, when one of the parameters is varied, both BAO amplitude and BAO position are simultaneously modified. In order to break the ambiguity between these two kinds of modifications, it would be better to make the BAO wiggles' positions overlap, thus leaving the change in amplitude as the only effect.

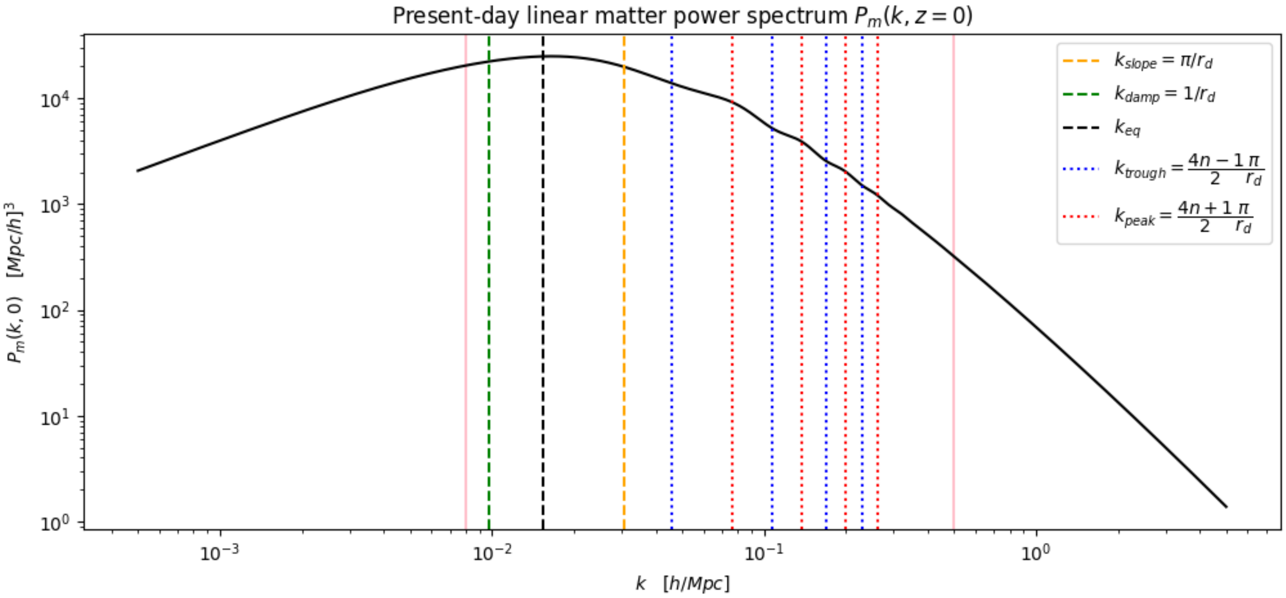

Figure 3.2: Linear matter power spectrum at redshift $z = 0$. The two pink vertical lines delimit the range of modes which are usually observed by spectroscopic galaxy surveys, $0.008 \leq k\,[\text{Mpc}/h] \leq 0.5$. k_{eq} denotes the Jeans scale for cold dark matter, and corresponds to the main turnaround of the shape. k_{damp} denotes the Jeans scale for baryons: smaller scales get damped with respect to a baryon-less Universe, and wiggles (BAO) appear in the shape. The wiggles troughs and peaks are denoted by k_{trough} and k_{peak} respectively. The scale denoted by k_{slope} corresponds to the maximum slope of baryon suppression.

It is rather easy to track down the cause behind the change in wiggles position. Essentially, varying a parameter among $\{\omega_m, \omega_b, h\}$ produces a shift in r_d, whose value determines the position of BAO peaks and troughs.

Therefore, in order to return wiggles back to their original positions, it will be sufficient to perform a rescaling of the power spectrum through the quantity s:

$$P_m(k,z) \;\rightarrow\; \hat{P}_m(k,z) = \frac{1}{s^3} P_m\left(\frac{k}{s}, z\right) \tag{3.18}$$

$$s = \frac{r_d(\Omega_i + \partial\Omega_i)}{r_d(\Omega_i)} \tag{3.19}$$

According to definition (1.19), σ_8 features a dependence on $P_m(k,z)$, implying that its value does not remain unchanged under the rescaling (3.19). To avoid such modification (i.e. to leave

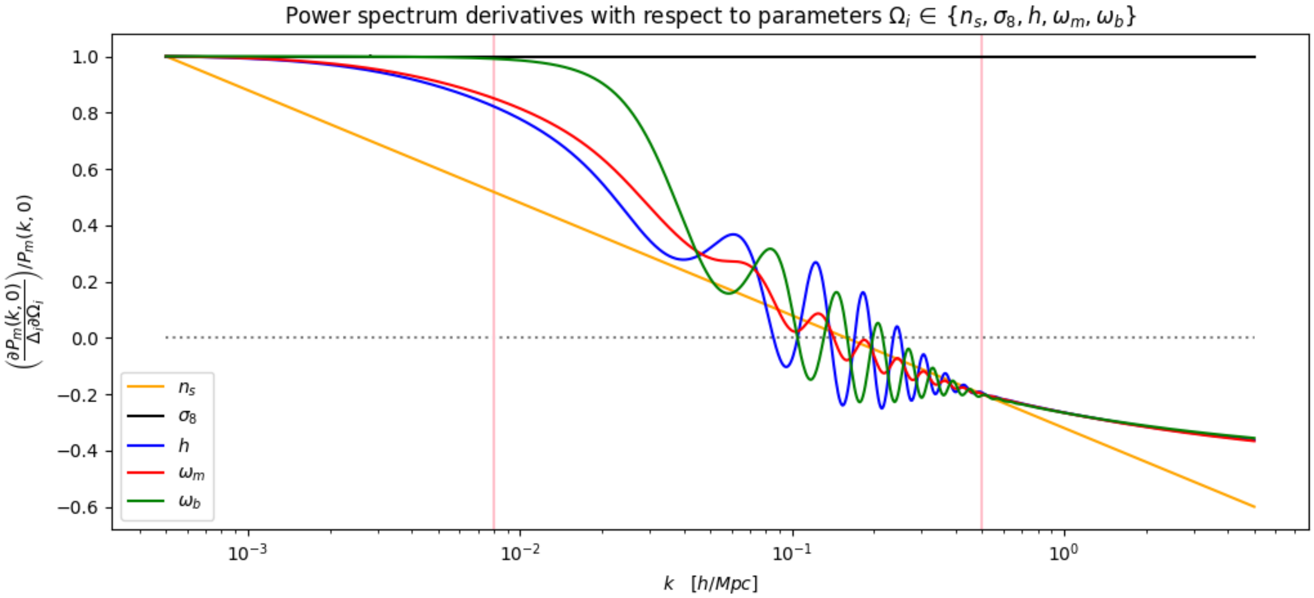

Figure 3.3: Present-day linear matter power spectrum derivatives with respect to parameters $\Omega_i \in \{n_s, \sigma_8, h, \omega_m, \omega_b\}$. Normalization factors Δ_i are chosen such that all parameters have the same impact on the power spectrum as σ_8 in the large scale limit. The two pink vertical lines delimit the range of modes which are usually observed by spectroscopic galaxy surveys, $0.008 \leq k\,[\mathrm{Mpc}/h] \leq 0.5$.

normalization unaffected by the rescaling), when computing the derivatives for $\hat{P}_m$ in *Figure* 3.4, we actually employ the redefinition ${\sigma_8}_s$[9].

From *Figure* 3.4 we can notice how the effect of h has been completely absorbed by the rescaling, since we are expressing r_d in [Mpc/h] units. On the other hand, ω_b and ω_m show a similar effect on the slope, with a small offset visible at k_{eq}. This can be explained by observing that, within standard ΛCDM, k_{eq} and r_d are closely related, so a shift in r_d will likely be accompanied by a shift in k_{eq} as well. Finally, it appears clear that the value of r_d is independent of σ_{8s} and n_s.

All in all, the picture presented by this simple theoretical analysis of the linear matter power spectrum shape is not so comforting: both *Figure* 3.4 and *Figure* 3.4 suggest some degeneracy,

---

[9] A technical note is needed. In the present work, we have employed the ratio $(\sigma_8/\sigma_N)^2$, defined in (1.18)-(1.21), as the matter power spectrum normalization. The way we numerically compute such normalization consists in leaving σ_8 fixed as a given value, while σ_N is computed via numerical integration. Now, the purpose of employing the redefinition σ_{8s} is to leave the normalization unchanged with respect to the chosen reference value: this is done by rescaling the top-hat filter scale R_8 by s, in order to account for the mode rescaling $k \to k/s$. Therefore, adapting this to our own case, we leave σ_8 as a fixed value, while σ_N is redefined as σ_{Ns} n the same fashion of Eq. (3.5).

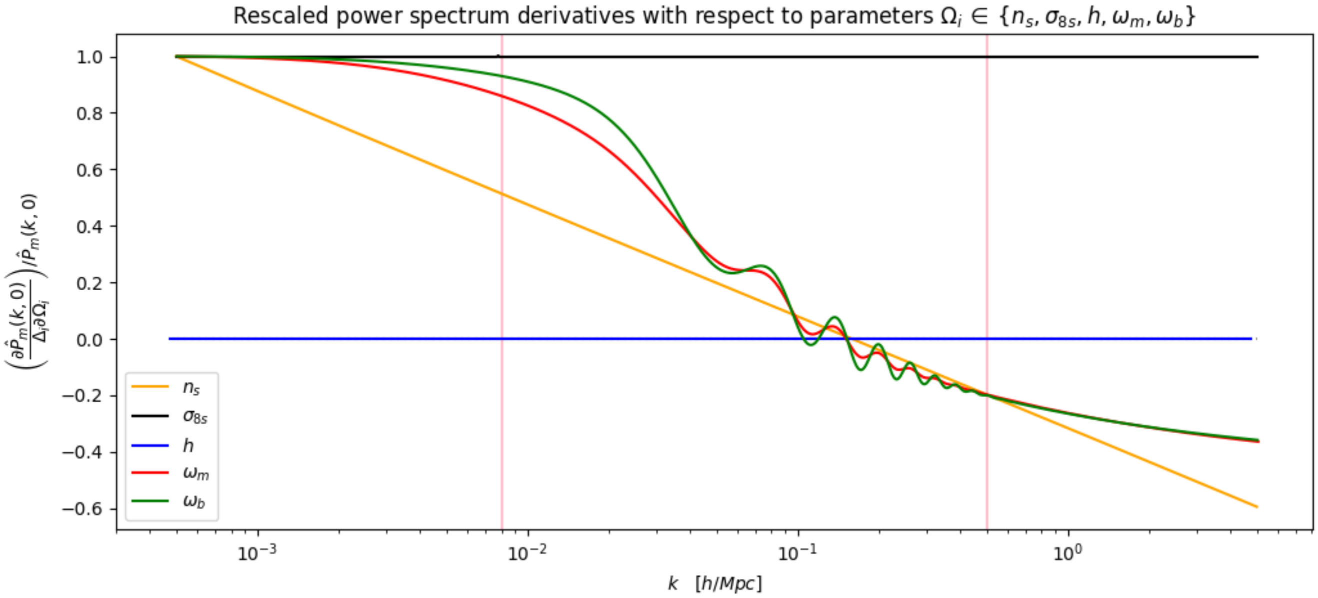

Figure 3.4: Present-day rescaled linear matter power spectrum derivatives with respect to parameters $\Omega_i \in \{n_s, \sigma_{8s}, h, \omega_m, \omega_b\}$. Notice that in this case we are employing σ_{8s} instead of σ_8. Normalization factors Δ_i are chosen such that all parameters have the same impact on the power spectrum as σ_8 in the large scale limit. The two pink vertical lines delimit the range of modes which are usually observed by spectroscopic galaxy surveys, $0.008 \leq k\,[\mathrm{Mpc}/h] \leq 0.5$. The dependence on h is lost since we are expressing r_d in [Mpc/h] units.

which makes it difficult to get reasonably good constraints for cosmological parameters. But, as the plots suggest, additional information coming from early-time physics, and affecting the power spectrum broadband shape via the transfer function, must be looked for in the large scale regime (and, to a smaller extent, in the amplitude of BAO wiggles). This additional, early-time, information is what ShapeFit exploits in order to improve parameter constraints with respect to the Classic approach (as we will see in chapter 4).

Another important point to take in is that, although the effect of n_s, ω_b and ω_m on the shape is expected to be degenerate, a crucial difference can be highlighted: ω_b and ω_m produce a slope change which is scale-dependent, while the one produced by n_s is scale-independent. This is a key aspect in the context of the ShapeFit methodology.

3.4.2 Slope rescaling and slope parameters

The Classic approach lives on the (unrealistic) assumption that all early-time physics can be captured by r_d, and the whole method consists in exploiting the standard ruler nature of such quantity. But unfortunately, as shown by the derivative plots, r_d does not depend on the parameters n_s and σ_8 (and thus on the primordial amplitude A_s, which is reabsorbed inside it). Moreover, as we anticipated in chapter 1 ((1.24) and (1.25)) and confirmed by derivative plots, the linear matter transfer function depends on many parameters, such as ω_b, ω_m, h. Crucially, such dependence lingers even after reabsorbing the effect of r_d (*Figure* 3.4).

In other words, the power spectrum slope features some early-time parameter dependence which is disjointed from the value of r_d, thus marking the limitations of the Classic approach.

The ShapeFit approach overcomes such limitations by compressing the bulk of this early-time information into new effective parameters. In this way, early-time and late-time information is still decoupled (as it was the case in the Classic approach), but it can be easily and consistently combined at the interpretation stage when constraining cosmological parameters.

This new method is able to preserve the same model-independent nature of the compressed variables of the Classic approach, while also reaching the same constraining power of the FM approach at the interpretation step.

Before introducing the new compressed parameters, one important aspect must be highlighted.

The ShapeFit approach is designed so that it extracts cosmological information from the BAO position, the overall power spectrum normalization and the power spectrum slope, but not from the BAO wiggles amplitude. This is done for two reasons. Firstly, the bulk of information is expected to come from the variation in the slope, not in the BAO amplitude[10]. Secondly, the BAO amplitude signal is not as robust as their position: in fact, despite being an early-time signal, it is heavily processed by late-time effects (e.g. non-linearities, bias, mode-coupling), which introduce some model-dependence).

We can thus proceed to model the slope of the linear power spectrum template, and its dependence on cosmological parameters. This is done by taking the reference template $P_r(k)$ and turning it into a new reference template $P'_r(k)$, defined such that

$$\ln\left(\frac{P'_r(k)}{P_r(k)}\right) = \frac{m}{a}\tanh\left[a\ln\left(\frac{k}{k_p}\right)\right] + n\ln\left(\frac{k}{k_p}\right) \tag{3.20}$$

---

[10] In order to get significant cosmological information from the BAO amplitude, one would require a survey with a volume way larger than 300 $[\mathrm{Gpc}/h]^3$, which is already unrealistically large for the next decade's standards.

where:

· n represents the overall scale-independent slope (completely degenerate with n_s);

· m represents a scale-dependent slope, which follows the transition of the linear power spectrum from the large scale to the small scale limit (in a ΛCDM model this is driven by the combined effect of ω_b and ω_m);

· the chosen pivot scale $k_p = \pi / r_d$ is the scale corresponding to the zero-crossing before the first BAO trough: interestingly, it corresponds to the scale of maximum baryon suppression, thus showing that the BAO wiggles position and the scale of suppression are linked by r_d;

· a is an amplitude which controls how fast the large and small scale limits are reached;

· the hyperbolic tangent is a generic sigmoid function reaching its maximum slope at the pivot scale k_p.

From the definition given in (3.20) it is evident that the two templates perfectly coincide for $k = k_p$ for all values of n and m, meaning that this slope rescaling is actually changing the power spectrum tilt by pivoting it around the scale k_p. In fact, the slope parameter m is added *a posteriori* to the linear power spectrum template in order to effectively parametrize a slope variation at the pivot scale k_p.

Additionally, the two spectra coincide for all values of k if $n = 0$ and $m = 0$: these two conditions mean that the primordial scale-independent slope n_s coincides with the reference one, and no further scale-dependent slope change is considered with respect to the original reference template.

The "goodness" of such slope rescaling has been computed numerically in *Figure* 3.5. These plots are displaying (solid lines) the ratio between two linear power spectra, both computed from the definitions (1.17) and (1.22) through the analytical no-wiggle EH98 transfer function (1.25), $P^{\text{no wiggle}}_{lin,EH98}(\mathbf{\Omega})/P^{\text{no wiggle}}_{lin,EH98}(\mathbf{\Omega}_{ref})$: here, the numerator is the response to a shift in ω_b, ω_m or n_s (as shown by the color bar in the plots), while the denominator is computed at the reference cosmology of *Table* 3.1. Importantly, the pivot scale adopted when computing the power spectra[11] is chosen to coincide with $k_{slope} = \pi/r_d$. Moreover, the power spectra involved in this ratio are opportunely rescaled by their corresponding value of r_d (i.e. the rescaling (3.18)-(3.19) is applied).

The plots also show the position of k_{eq}, after having accounted for the r_d rescaling (dashed-dotted vertical lines).

On top of the solid curves, the plots also show (dotted lines) the exponential of the right hand side of (3.2). In order to plot such quantity, the following definitions are employed:

$$n(\mathbf{\Omega}) = n_s(\mathbf{\Omega}) - n_s(\mathbf{\Omega}_{ref}) \tag{3.21}$$

$$m = \frac{d}{d\ln k}\left(\ln\left[\frac{P^{\text{no wiggle}}_{lin,EH98}\left(\frac{k}{s},\mathbf{\Omega}\right) \Big/ P_R\left(\frac{k}{s},\mathbf{\Omega}\right)}{P^{\text{no wiggle}}_{lin,EH98}\left(k,\mathbf{\Omega}_{ref}\right) \Big/ P_R\left(k,\mathbf{\Omega}_{ref}\right)}\right]\right)\Bigg|_{k=k_p} \tag{3.22}$$

where P_R is the primordial power spectrum. Meanwhile, k_p and a are left as free parameters, and tuned to get the best overlap between the two sets of curves.

By calibrating them for the cases of varying ω_m and ω_b, one finds a remarkable overlap for $a \approx 0.6$ and $k_p \approx 0.03\ h/\text{Mpc} \approx \pi/r_d^{ref} = k_{slope}^{ref}$ (the scale k_{slope} is also shown as dotted vertical lines). In other words, for these values of a and k_p, the proposed rescaling model, i.e. (3.20), is able to capture very well the power spectrum slope dependence on ω_m and ω_b. More specifically, the colored and grey curves agree up to $\sim 1\%$ (ω_m plot) and $\sim 3\%$ (ω_b plot) in the range $0.02 \leq k\ [h/\text{Mpc}] \leq 0.25$.

One can notice that the overlap is perfect in case of varying n_s. This is expected, since a shift in n_s only affects the primordial power spectrum, while leaving the transfer function unchanged; furthermore, as previously observed, r_d does not depend on n_s, resulting in $s = 1$. Therefore, the ratio between power spectra effectively results in a ratio between the primordial power spectra, which in turn reduces to the quantity $(k/k_p)^{(n_s - n_s^{ref})} = (k/k_p)^n$. At the same time, the natural logarithm in definition of m vanishes, since its argument essentially corresponds to a ratio between transfer functions[12] computed at the same scales. Therefore, m

---

[11] Namely, k_0 in (1.17).

[12] Here, we are not talking about the Eisenstein-Hu transfer function T_m, but the one denoted as $\mathcal{T}_m^2$ and appearing in (1.17).

vanishes in the right hand side of (3.20), thus leaving $\ln(k/k_p)^n$ as the only remaining term, and effectively matching the left hand side.

The slope rescaling (3.20) enters the ShapeFit pipeline right before the template rescaling by $\alpha_\parallel$ and $\alpha_\perp$. Effectively, ShapeFit consists in applying the Classic RSD fit to a reference template $P_r'(k)$, that is transformed at each step via (3.20) with free parameters m and n.

In principle, such transformation should be also applied to the reference power spectra appearing in the integrand of higher-order perturbation corrections. However, in order to avoid re-evaluating all perturbative terms at each step of the likelihood exploration, one can decide to apply this transformation as if it was scale-independent. In practice, all loop corrections are pre-computed using the linear power spectrum P_r, and then transformed in an approximated way during the likelihood evaluation. Here we show such approximation when evaluating the 1-loop correction $P_{\text{1-loop}} = P_{13} + P_{22}$ within SPT theory:

$$\begin{aligned}
P_{13}(k) &= P_r'(k)\int_0^\infty d^3q P_r'(q) F_3(k,q,-q) \approx \\
&\approx \left(\frac{P_r'(k)}{P_r(k)}\right)^2 P_r(k)\int_0^\infty d^3q P_r(q) F_3(k,q,-q) \\
P_{22}(k) &= \int_0^\infty d^3q P_r'(q) P_r'(|q-k|) F_2(k,q-k) \approx \\
&\approx \left(\frac{P_r'(k)}{P_r(k)}\right)^2 \int_0^\infty d^3q P_r(q) P_r(|q-k|) F_2(k,q-k)
\end{aligned} \tag{3.23}$$

where F_i are the respective kernels.

In (Brieden et al., 2021b) it is shown that such approximation, which aims to reduce the computational time of ShapeFit, does not hinder the analysis precision. In fact, the systematic errors made by factorizing m outside the loop integrals are found to be below 2% for $k \leq 0.15$ Mpc/h and below 3% for $k \leq 0.20$ Mpc/h.

3.4.3 Variables complementarity

Lastly, we try to visualize what the introduction of the slope parameter m means in the context of constraining different parameters. *Figure* 3.6 displays cosmological constraints for a standard flat ΛCDM model , obtained from two BOSS samples[13], and using three different sets of physical compressed variables. Grey curves correspond to the Classic RSD analysis based on the late-time variables $\{\alpha_{\parallel}, \alpha_{\perp}, f\sigma_8\}$; green curves come from a ShapeFit analysis, in which the compressed variables are $\{\alpha_{\parallel}, \alpha_{\perp}, f\sigma_8, m\}$; red curves arise from an analysis which only compresses the slope parameter m. This last kind of analysis has been chosen to better understand and visualize the complementarity between m and the classic RSD variables $\{\alpha_{\parallel}, \alpha_{\perp}, f\sigma_8\}$. Opaque and transparent contours correspond to performing the interpretation respectively with a Gaussian BBN prior $w_b = 0.02268 \pm 0.00038$ or with a broad uniform prior $0.005 < w_b < 0.04$.

The slope parameter m is an early-time variable, sensitive to early-time quantities which enter in the matter transfer function such as equality scale, sound horizon scale and primordial power spectrum. while the latter are late-time variables (only depending on geometry, expansion history and growth rate of the Universe). These variables are sensitive to physical processes affecting the Universe evolution in different epochs: we can therefore expect them to constrain different derived parameters. In particular, the BAO signal is able to constrain $h \cdot r_d$, while m is able to constrain $\Omega_m h$ (which determines the shape of the matter transfer function via matter-radiation equality epoch).

Therefore, the introduction of the new compressed parameter m to the classic set appears as a great improvement in terms of constraints, since it makes the analysis sensitive to early-time physics as well, while not hindering the ability to extract information from the classic late-time variables[14].

Here we suggest a small example of how one might exploit the complementarity between m and the classic BAO variables.

As previously observed, m is able to constrain the parameter $\Omega_m h$. Conversely, within a ΛCDM model, the purely late-time expansion history constrains the ratio $\alpha_{\parallel}/\alpha_{\perp}$, which in turn can be used to constrain Ω_m. Combining these two constraints, it would therefore be possible to

---

[13] Luminous Red Galaxies (LRG) samples of the SDSS-III BOSS survey (Alam et al., 2017) have been employed, and two uncorrelated redshift ranges have been considered: $0.2 < z < 0.5$ ($z_{eff} = 0.38$, $V_{eff} = 3.7\text{Gpc}^3$) and $0.5 < z < 0.75$ ($z_{eff} = 0.61$, $V_{eff} = 4.1\text{Gpc}^3$).

[14] In the left panel of *Figure* 4.2 it will appear clear that, being m (mostly) uncorrelated to the classic compressed parameters, introducing the former does not jeopardize the constraining precision of the latter.

obtain a constraint on H_0: this constraint would arise from a combination of early and late-time physical effects[15].

3.5 ShapeFit pipeline

As previously anticipated, the ShapeFit can be thought as an extension of Classic RSD. Consequently, its pipeline is almost identical to what we already presented in section 3.3. This shows from *Figure* 3.7, where the full ShapeFit pipeline is shown. Here, the new steps introduced by the ShapeFit methodology are displayed as orange boxes, while blue boxes are the Classic RSD steps.

The data pipeline is exactly the same as in Classic RSD, so we will not explore it further. Instead, the introduction of the new shape parameters m and n, which are the true protagonists of the ShapeFit technique, implies a few additions to the rest of the pipeline.

In particular, a new step is added to the model pipeline: now, right after having computed the non-linear template power spectrum $P_{nl}(k, \mathbf{\Omega}_{ref})$, the slope rescaling (3.20) is applied, in conjunction with the approximation (3.23), thus producing the new reference template $P'_{nl,r}(k, \mathbf{\Omega}_{ref})$, which includes the shape parameters m and n. This new quantity is then used to build the multipoles, as in the usual Classic approach.

Consequently, the fitting step is now used to get constraints on five compressed variables ($\alpha_\parallel$, $\alpha_\perp$, $f\sigma_{s8}$, m and n), and the cosmological interpretation step will be accordingly affected. In particular, the interpretation of the scaling parameters $\alpha_\parallel$ and $\alpha_\perp$ is identical to the Classic approach, while the other parameters deserve some more words.

Fluctuation amplitude and growth

The interpretation of $f\sigma_{8s}$ is nearly the same as in Classic RSD, with the only exception that now we are using σ_{8s} instead of σ_8.

Actually, from the definition (3.7) it is easy to see that modifying the power spectrum via the slope rescaling (3.20) with $(m, n) \neq (0, 0)$ would also modify σ_{8s}.

---

[15] In (Brieden, Gil-Marín, & Verde, 2021a), the $\Omega_m h$ measurement shown in *Figure* 3.6 ($\Omega_m h = 0.220^{+0.029}_{-0.019}$, without BBN prior on w_b) is combined with that obtained from the uncalibrated BAO of the full BOSS+eBOSS sample (*Table* 4 of (Alam et al., 2021)) $\Omega_m = 0.299 \pm 0.016$, thus obtaining $H_0 = 73.6^{+10.5}_{-7.5}$. Of course, the final value will depend on the chosen prior for w_b and on the chosen value of Ω_m.

Thus, it might be convenient to define the fluctuation amplitude at the pivot scale k_p as

$$A_p = P_{lin}^{\text{no wiggle}}(k_p, \mathbf{\Omega}) \tag{3.24}$$

which does not change with slope, since it is computed at the pivot point.

During the fitting procedure in the cosmological interpretation step, this amplitude parameter becomes

$$A_{sp} = \frac{1}{s^3} P_{lin}^{\text{no wiggle}}(k_p/s, \mathbf{\Omega}) \tag{3.25}$$

Then, the actual velocity fluctuation amplitude measurement is given as $f A_{sp}^{1/2}$, and thus

$$f\sigma_{8s} = \frac{(f\sigma_{8s})^{ref}}{(A_{sp}^{1/2})^{ref}} A_{sp}^{1/2} \tag{3.26}$$

which connects the more frequently used parameter ($f\sigma_{8s}$) with that more suitable to employ for parameter inference ($A_{sp}^{1/2}$).

Power spectrum slope

The scale independent slope n is interpreted as in (3.21), while the scale dependent slope m follows (3.22).

In case n is varied during the cosmological fit, (3.22) has to be applied to the power spectrum obtained with n_s fixed to n_s^{ref}: this ensures that a change in n_s does not lead to a different prediction for m, but only for n, via (3.21). In this way, n is obtained from the primordial power spectrum (whose slope is n_s), while m is obtained from the transfer function squared (which is the ratio between power spectrum and primordial power spectrum[16]). It is useful to

---

[16] So, the square brackets of (3.22) contain the ratio between the transfer functions associated to input cosmology and that computed at reference cosmology.

stress that the interpretation of m is not any late-time physics phenomenon, but a series of early-time processes modulating the power spectrum broadband shape.

3.5.1 Computational advantage

After having taken a deep dive into the ShapeFit pipeline, it will appear clear why this approach is computationally faster and more convenient than FM.

First of all, the ShapeFit procedure, whose goal is extracting the physical parameters $\{\alpha_\parallel, \alpha_\perp, f\sigma_{s8}, n, m\}$ (upper part of *Figure* 3.7), is model-independent, since it employs a reference template: therefore, it must be performed only once. In comparison, the FM approach requires to fit the power spectrum multipoles for each chosen model. Moreover, at the moment of cosmological interpretation (lower part of *Figure* 3.7), the MCMC deals with a much simpler likelihood surface and a much smaller number of variables (one compressed parameter for each chosen redshift). Conversely, within FM, the MCMC has to deal with a larger number of variables, i.e. one power spectrum measurement $P^{(\ell)}(k, z)$ for each considered multipole, k-bin, redshift and galactic hemisphere[17]. Consequently, the ShapeFit MCMC needs 5-10 times fewer sampled points to reach the same level of convergence. In conclusion, ShapeFit yields an overall speed-up factor of 40-80 (Brieden et al., 2021a).

This computational advantage with respect to FM, coupled with an improved statistical performance with respect to Classical analysis (as it will be shown in the chapter 4), mark the true benefits of adopting the ShapeFit methodology.

---

[17] See chapter 5 for a practical example.

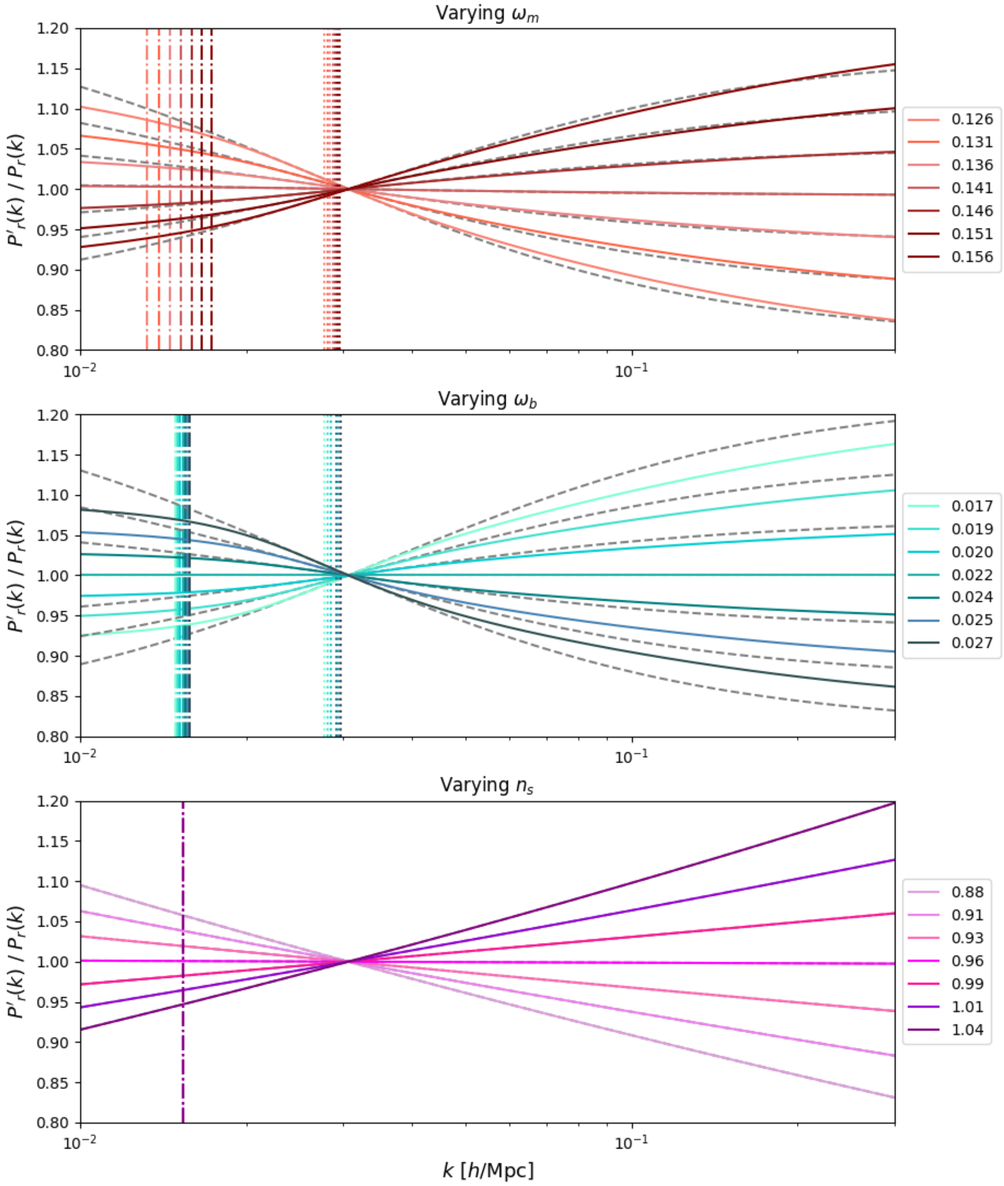

Figure 3.5: Rescaled EH98 prediction of the power spectrum shape (colored sold lines) comparison with the parametrization of (3.20) for $a = 0.6, k_p = 0.03\mathrm{Mpc}/h$. This choice fits the response to w_m (upper panel) very well and to w_b (middle panel) less well but still sufficient for our purposes. Dashed-dotted vertical lines show the rescaled location of k_{eq} for each model and dotted vertical lines highlight the positions where the scale-dependent slope reaches a maximum. This position is constant with varying cosmological parameters and very close to the expectation $k_p = \pi/r_d^{ref} \approx 0.03\mathrm{Mpc}/h$. The scale independent slope fits the prediction of varying n_s (bottom panel) perfectly by definition.

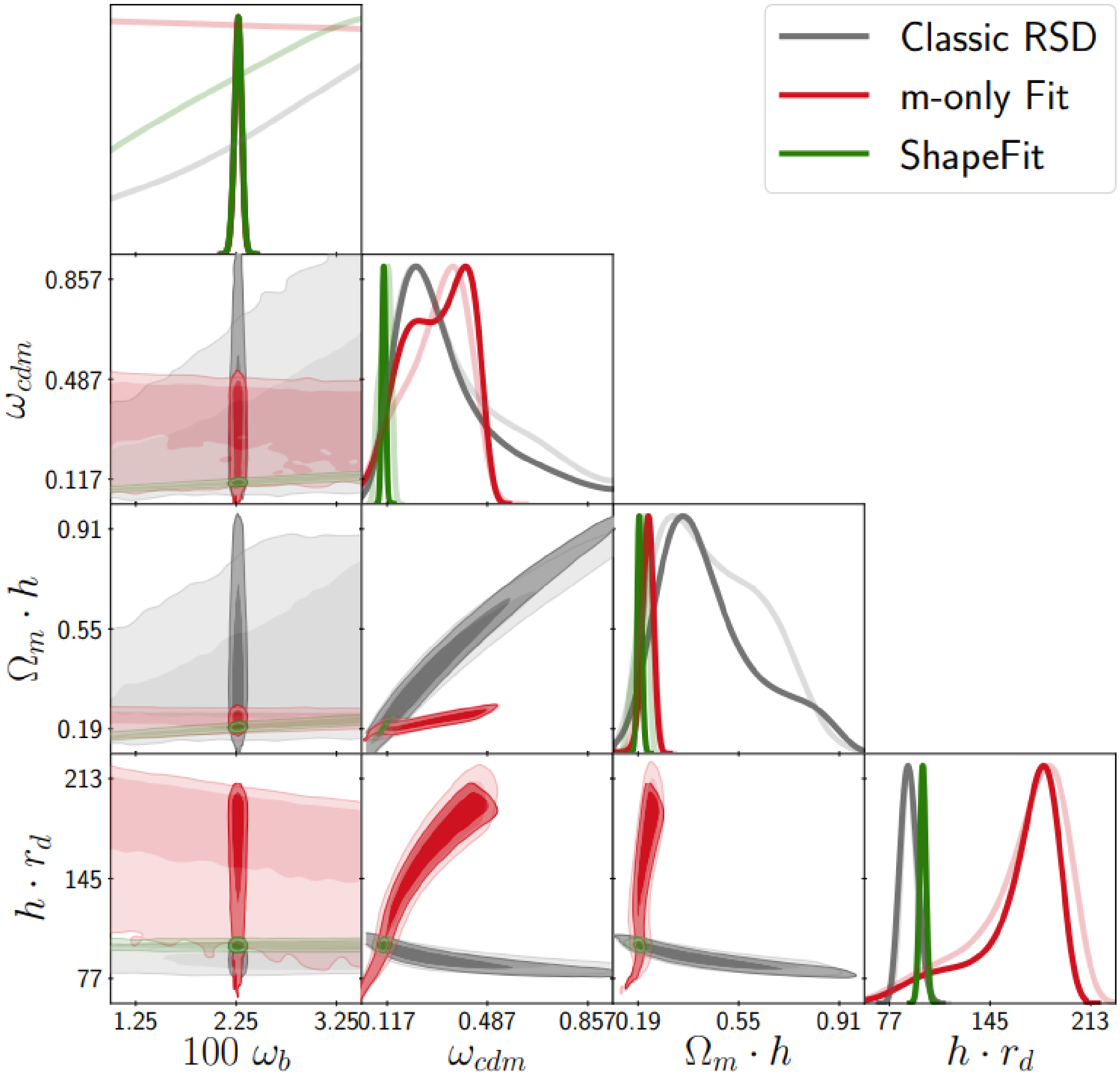

Figure 3.6: Interpretation within a flat-ΛCDM model with a Gaussian BBN prior on w_b (opaque contours) and without (transparent contours), of different physical variables constraints from the low- and high-redshift BOSS samples. Grey corresponds to classic RSD analysis based on late-time variables $\{\alpha_{\parallel}(z), \alpha_{\perp}(z), f\sigma_8(z)\}$, red corresponds to the early-time shape variable $m(z)$ only, and their combination is shown in green. Transparent contours are for a broad uniform prior $0.005 < w_b < 0.04$, the opaque contours for a Gaussian BBN prior $w_b = 0.02268 \pm 0.00038$. This image is actually borrowed from chapter 5 of the present text, but it better fits the discussion if put here.

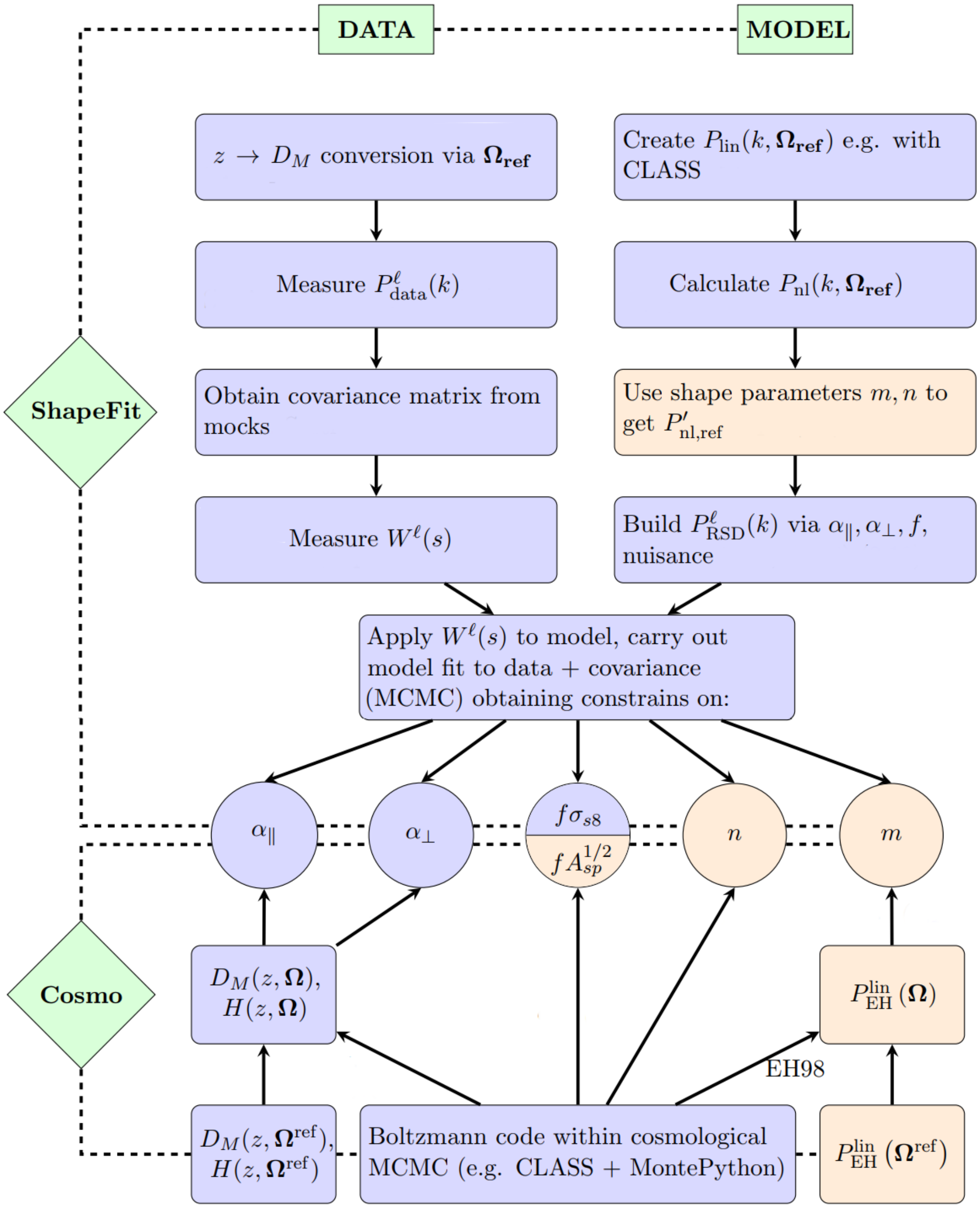

Figure 3.7: ShapeFit pipeline. Steps above the dashed lines correspond to the data and model pipelines, while the cosmological interpretation step is located below the dashed lines The purple fields represent steps in the Classic approach, while orange fields represent the ShapeFit additions. This image corresponds to *Figure* 5 in (Brieden et al., 2021b).

Chapter 4

Application to mocks

In this section, closely following chapter 5 of (Brieden et al., 2021b), we will see how this new ShapeFit methodology fares in comparison to the Classical and FM ones. In particular, we will compare the results of a fiducial analysis, conducted with the three methods, on the MultiDark-Patchy BOSS DR12 mocks[1], created by (Kitaura et al., 2016)(Rodríguez-Torres et al., 2016), and whose fiducial ΛCDM parameters are the following:

MultiDark simulation fiducial cosmology					
Cosmology	Ω_m	Ω_b	h	σ_8	n_s
MultiDark	0.307115	0.048206	0.6777	0.8288	0.9611

Table 4.1: Fiducial cosmology adopted by the PATCHY simulated mocks.

In particular, the present analysis has made use of the "ngc_z3" mocks, which correspond to a sample located at the north galactic cap, within a redshift range of $0.5 < z < 0.75$ ($z_{eff} = 0.61$). All 2048 realizations of the PATCHY mocks are employed in order to extract our dataset: multipole power spectra are measured for each individual mock catalogues, and the mode-weighted power spectra average of all 2048 realizations is then taken as data. To be more precise, multipoles are measured in bins $\Delta k = 0.005\ h\mathrm{Mpc}^{-1}$, and the analysis is conducted over a scale-range $0.01 \leq k\ \left[h\mathrm{Mpc}^{-1}\right]\ \leq 0.15$.

After computing the covariance matrix, this is rescaled by a factor 0.01, corresponding to the volume of 100 stacked mocks[2].

---

[1] https://fbeutler.github.io/hub/boss papers.html

[2] There is no need to "use" all the 2048 realizations: the effective volume of 100 stacked mocks is ≈ 300 Gpc^3, which is still significantly larger than the effective volume of the next generation of galaxy surveys.

Before showing the results, it is better to specify the prior ranges adopted for the cosmological, physical and nuisance parameters during the fitting process. These are collected in the table of *Figure* 4.1, but here we expand on them a little bit.

<table>
<tr><th colspan="2">Parameter</th><th colspan="4">Prior ranges</th></tr>
<tr><th>type</th><th>name</th><th>SF min</th><th>SF max</th><th>FM min</th><th>FM max</th></tr>
<tr><td>Cosmological</td><td>$\omega_{\rm cdm}$</td><td colspan="2">$[\rm None, None]$</td><td colspan="2">$[\rm None, None]$</td></tr>
<tr><td></td><td>h</td><td colspan="2">$[\rm None, None]$</td><td colspan="2">$[\rm None, None]$</td></tr>
<tr><td></td><td>$\ln\left(10^{10}A_s\right)$ / $A^{1/2}$</td><td colspan="2">$[\rm None, None]$</td><td colspan="2">$[0.2, 2.0]$</td></tr>
<tr><td>Physical</td><td>$\alpha_{\parallel}$</td><td colspan="2">$[0.5, 1.5]$</td><td colspan="2">–</td></tr>
<tr><td></td><td>$\alpha_{\perp}$</td><td colspan="2">$[0.5, 1.5]$</td><td colspan="2">–</td></tr>
<tr><td></td><td>f</td><td colspan="2">$[0, 3]$</td><td colspan="2">–</td></tr>
<tr><td></td><td>m</td><td colspan="2">$[-3, 3]$</td><td colspan="2">–</td></tr>
<tr><td>Nuisance</td><td>b_1</td><td colspan="2">$[0, 10]$</td><td colspan="2">$[0, 10]$</td></tr>
<tr><td></td><td>b_2</td><td colspan="2">[-10, 10]</td><td colspan="2">[-10, 10]</td></tr>
<tr><td></td><td>b_{s2}</td><td>lag.</td><td>[-10, 10]</td><td>lag.</td><td>[-10, 10]</td></tr>
<tr><td></td><td>$b_{3\rm nl}$</td><td>lag.</td><td>[-10, 10]</td><td>lag.</td><td>0</td></tr>
<tr><td></td><td>$c_0\,[h^{-2}{\rm Mpc}^2]$</td><td colspan="2">–</td><td>0</td><td>(0 ± 30)</td></tr>
<tr><td></td><td>$c_2\,[h^{-2}{\rm Mpc}^2]$</td><td colspan="2">–</td><td colspan="2">(0 ± 30)</td></tr>
<tr><td></td><td>$c_4\,[h^{-4}{\rm Mpc}^4]$</td><td colspan="2">–</td><td colspan="2">(500 ± 500)</td></tr>
<tr><td></td><td>$\sigma_P\,[h^{-1}\,{\rm Mpc}]$</td><td colspan="2">$[0, 10]$</td><td colspan="2">–</td></tr>
<tr><td></td><td>$A_{\rm noise}$ / $\Delta P_{\rm noise}\,[h^{-3}{\rm Mpc}^3]$</td><td colspan="2">[-5, 5]</td><td colspan="2">(0 ± 5000)</td></tr>
</table>

Figure 4.1: Prior ranges for parameters used for the ShapeFit (SF) and the FM fit. For each, we define a case with minimum ("min") and maximum ("max") freedom, where overlapping prior choices between the two choices are written in the center. Flat priors are given as [min; max], Gaussian priors are denoted as (mean ± std). "lag." stands for the local Lagrangian prediction prior. Parameters separated by " / " correspond to different conventions used between ShapeFit and FM fit for the same physical effect.

The model adopted for the ShapeFit is based on the 1-loop SPT + TSN model (1.54); whereas, for the FM fit, the EFT implementation of (Ivanov et al., 2020) has been chosen. These two different choices also imply different choices for some parameters within the two approaches (separated by " / " in the table of priors).

It may be useful to specify that the present implementation of ShapeFit is unable to correctly model large scale displacements and, consequently, the effects that their non-linear coupling has on the BAO amplitude. The same effect could be more easily modelled within the FM approach, by the introduction of the IR resummation (which brings a phenomenological damping of the BAO amplitude). But, for the sake of comparing the two approaches, the IR resummation has been neglected during the FM fit[3].

Two different cases are defined for nuisance parameters' prior ranges: minimum and maximum freedom. The minimum freedom scenario leans towards the fiducial setup of most Classic RSD analyses, where non-local bias parameters are fixed by the local Lagrangian prediction (Baldauf et al., 2012; Saito et al., 2014):

$$b_{s2} = -\frac{4}{7}(b_1 - 1), \qquad b_{3nl} = \frac{32}{315}(b_1 - 1) \tag{4.1}$$

while in the maximum freedom case b_1, b_{s2}, b_{3nl} are treated as independent parameters[4]. Some words about the counterterms c_0, c_2, c_4, which are characteristic of the EFT implementation of the FM approach[5]: in the minimum freedom scenario, c_0 is fixed to 0 to match the ShapeFit configuration, while the other counterterms are varied, since they correspond to ShapeFit's σ_P.

In general, the prior choices made for the maximum freedom case are dictated by the decision to follow as closely as possible the setup of (Ivanov et al., 2020).

Another difference revolves around the primordial amplitude: in the ShapeFit implementation, the employed parameter is $\ln(10^{10}A_s)$, while FM uses $A^{1/2} = (A_s/A_s^{Planck})^{1/2}$, with $A_s^{Planck} = 2.0989 \times 10^{-9}$.

Finally, while FM uses a Gaussian prior on the difference between shot and Poissonian noise, $\Delta P_{noise} = P_{noise} - PPoissonian$, ShapeFit employs a flat prior on the fractional difference A_{noise}.

4.1 ShapeFit vs Classic RSD

First, we compare the performance of ShapeFit and Classic RSD in the context of the compression step (so, for now, we leave out the interpretation part). It might be useful to take

---

[3] Actually, the inclusion of the IR resummation for ShapeFit would bear tangible effects only for a survey volume of at least $\approx 300\,[\mathrm{Gpc}/h]^3$

[4] Bias parameters describe the connection between dark matter and galaxy density fluctuations in real space. But (4.1) are motivated by theoretical considerations on the dark matter to halo connection, which doesn't necessarily translate equivalently into the dark matter to galaxy connection.

[5] They phenomenologically account for higher order non-linearities. c_0 corrects for dark matter behaving differently than a perfect fluid on small scales, while c_2 and c_4 account for non-linear RSD. Within the ShapeFit TSN model, instead of c_2 and c_4, there is one single parameter, σ_P (so, there is less freedom).

another look at the pipeline scheme presented in *Figure* 3.1 and *Figure* 3.7, in order to better appreciate the differences between the two methods.

In both approaches, the template power spectrum is built by adopting the MultiDark cosmology (see *Table* 4.1) as the reference one ($\boldsymbol{\Omega_{ref}}$).

The crucial difference between the two approaches is the absence of the new compressed parameters m and n in the Classic RSD. Therefore, the slope parameter m is varied only in the ShapeFit, while it is fixed to $m = 0$ for the Classic RSD.

Both "min" and "max" scenarios are evaluated, and the results are plotted in *Figure* 4.2, in which the dashed lines indicate the underlying parameter values of the simulation. Keep in mind that the error-bars displayed in these plots are relative to an effective volume of 300 [Gpc/h]^3: this is done to make explicit and quantify biases in recovered parameter estimates.

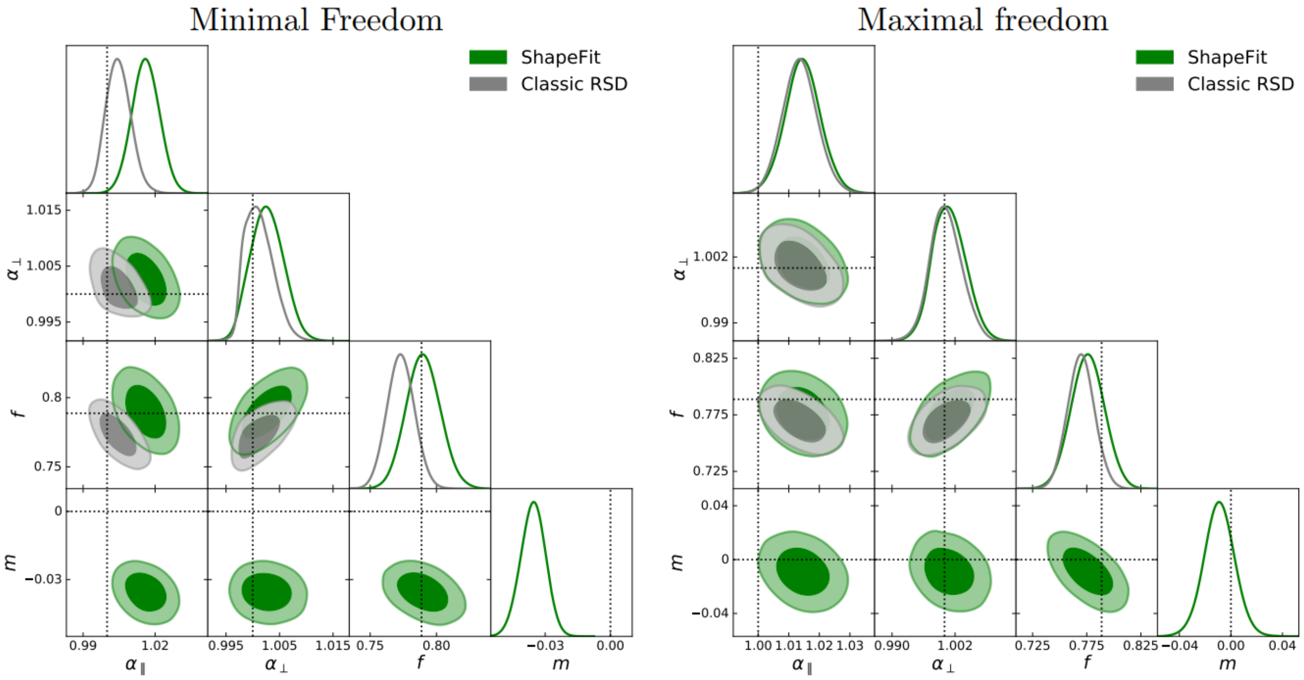

Figure 4.2: Results of the classic RSD fit (grey) and the ShapeFit (green) applied to the Patchy "ngc_z3" sample, showing the "min" (minimal freedom, non-local bias parameters to follow the Lagrangian prediction) case on the left; and the "max" (maximal freedom, fully free non-local bias parameters) case on the right panel. For the "min" case a systematic deviation of the slope parameter can be observed from the expectation $m = 0$, indicating that assumptions about the biasing scheme may introduce systematic shifts even at very large and linear scales. To make explicit and quantify biases in recovered parameter estimates, the error-bars are relative to an effective volume of 300 [Gpc/h]^3.

One can easily notice that, in the "maximal freedom" case, uncertainties on $\alpha_\parallel$ and $\alpha_\perp$ are almost indistinguishable for the two approaches; whereas, ShapeFit recovers a slightly larger error for the parameter f: this may be due to the correlation between f and m (while m is totally uncorrelated with $\alpha_{\parallel}$ and $\alpha_{\perp}$)[6]. The remaining biases on the found parameter values with respect to the expected ones can be due to both statistical uncertainty and incorrect modelling of non-linearities or RSD. But , in terms of uncertainties, ShapeFit is able to reproduce RSD's results almost perfectly.

Instead, looking at the "minimum freedom" case, the constraints recovered by Classic RSD appear closer to the actual theoretical predictions. This happens because, in the "min" scenario (i.e., non-local biases following the Lagrangian prediction), m undergoes a systematic shift towards $m = -0.036 \pm 0.006$, hence being in 6σ tension with the expected $m = 0$ value. This tension is not there for the RSD approach, since m is fixed to $m = 0$ from the start. Furthermore, this tension disappears when the non-local biases priors are relaxed, and the "maximal freedom" case is chosen: in fact, in the "max" scenario, the correct value of $m = 0$ is recovered by ShapeFit.

On the one hand, this finding highlights how priors (and, therefore, model) assumptions can lead to the emergence of systematic errors affecting parameter constraints: the authors of (Brieden et al., 2021b) then recommend to always allow maximal freedom for bias and nuisance parameters in forthcoming data analyses. This will slow down the MCMC chain convergence, but thanks to the design of ShapeFit, the fit only needs to be performed once.

On the other hand, the authors suggest that the characteristics of the employed mocks (ngc_z3) may have played a part in the emergence of these systematics: for example, such mocks may feature some inconsistencies with the local Lagrangian bias scheme (4.1). As a further check, the analysis has been repeated with a different type of mocks, namely, N-body mocks.

4.1.1 N-body mocks analysis

As just stated, some systematic errors in the previous analysis could stem from the characteristics of the employed mocks. In order to hunt for such systematics, further analyses have been conducted by the authors of the ShapeFit paper, which make use of N-body mocks instead of the PATCHY ones. First, a very simple and idealized case is analysed, in order to better highlight the systematic shift in parameter constraints. Then, a more complete and realistic

---

[6] The degradation of the constraints on f, due to this small degeneracy between f and m, is small (20% for a 300 $[\mathrm{Gpc}/h]^3$ volume), and it is expected to decrease for smaller, more realistic survey volumes (5% for a DESI-like volume of 30 $[\mathrm{Gpc}/h]^3$).

case is explored. Further details about the chosen N-body mocks and the followed procedure can be found in *Appendix D*. Here, we just outline the procedure in a few words, but the main goal is to present the main plots and results of such analysis.

Dark matter in real space

The analysis focuses on the very simple case of a set of dark matter particles, without survey geometry or selection function. No RSD are introduced either. Initial conditions are generated according to the second-order Lagrangian PT, and three redshift snapshots have been considered ($z = 0.5, 1.0, 1.5$). For simplicity, only the posteriors of $\{\alpha_\parallel, \alpha_\perp, f, m\}$ are explored, while n is fixed to $n = 0$ and the rest of nuisance parameters $\{b_1, b_2, \sigma_P, A_{noise}\}$ are marginalized.

The main results are illustrated by the dashed contours in *Figure* 4.3, showing the difference between the measured and the expected value for the 4 relevant physical parameters (accordingly scaled for visualization purposes).

The first lesson to take in is that, even if one chooses a very idealized situation, systematic shift are still visible. They are reported in the upper row of *Table* 4.2.

Geometric correction G	$\Delta^{sys}_{\alpha_\parallel}$	$\Delta^{sys}_{\alpha_\perp}$	Δ^{sys}_{f}	Δ^{sys}_{m}
without	$\simeq 0.01$ (1%)	$\simeq 0.005$ (0.5%)	$\simeq 0.003$ (0.3%)	$\simeq -0.04$ (4%)
with	$\simeq 0.005$ (0.5%)	$\simeq 0.0$ (0%)	$\simeq 0.003$ (0.3%)	$\simeq -0.01$ (1%)

Table 4.2: Shifts on parameter constraints, with and without the geometric correction defined in Eq. (3.2).

Furthermore, these shifts appear to be redshift-independent, hinting that they are not due to a theoretical limitation of the PT-model, or any biasing model assumption, but likely stem from a geometric effect. In what follows, we trace the cause of such shifts.

In the standard procedure, the power spectrum model is evaluated at the effective k-vector of each bin i, $P^{model}(k_{i,eff})$, where $k_{i,eff} \equiv \langle \boldsymbol{k} \rangle_i$ is the ensemble average over all possible directions of the k-vector within the i-bin. Conversely, the data-vector is measured by taking the average of $P(\boldsymbol{k}_i)$ across all $\boldsymbol{k}$-directions, i.e. $P^{data}(k_i) = \langle P^{data}(\boldsymbol{k}) \rangle_i$. During the analysis, these two quantities, $P^{data}(k_i)$ and $P^{model}(k_{i,eff})$, are subject to a fit, but they evidently do not represent the same thing, since $P(\langle k \rangle) \neq \langle P(k) \rangle$: this generates spurious signals, especially for the large-scale modes, where the number of modes per bin is small, and the discrepancy between the two quantities gets larger. So, the parameter shifts are ultimately due to mode discreteness, which is introduced by survey geometry (hence, it counts as a "geometric effect").

In principle, in order to overcome such issue, one should evaluate P^{model} for each $\boldsymbol{k}$ within the bin, and treat them in the same way as the data-vector, i.e. compute $\langle P^{model}(\boldsymbol{k}_i)\rangle$. But this would be too computationally expensive to be actually implemented with a MCMC.

The alternative solution proposed in (Brieden et al., 2021b) consists in the introduction of a geometric factor G, defined as

$$G(k) \equiv \frac{\langle P^{model}(\boldsymbol{k})\rangle}{P^{model}(\langle \boldsymbol{k}\rangle)} \tag{4.2}$$

which should accordingly modify the data-vector as

$$P^{data,G^{-1}}(k) = G^{-1} \times P^{data}(k) \tag{4.3}$$

and $\langle\,\rangle$ represents the ensemble average taken over all directions of the $\boldsymbol{k}$-vector[7].

The addition of such correction allowed to improve the parameter constraint shifts, as indicated in the bottom row of *Table* 4.2 and by the solid-line posteriors in *Figure* 4.3.

Here, posteriors have been derived by fitting the mean of 160 full N-body dark matter realization in real space, while the covariance has been rescaled to correspond to an effective volume of 2200 $[\mathrm{Gpc}/h]^3$. The model described in (1.53) and (1.54) and the local-Lagrangian bias scheme of (4.1) have been employed. Even though all nuisance parameters are varied during the fit, the plot only shows the physical parameters $f, \alpha_{\parallel}, \alpha_{\perp}, m$. The empty-dashed contours display the results of fitting the data-vector without any geometric correction (as usually done), whereas the filled contours account for the geometric correction through (4.2). For all parameters p, the expected value is $\Delta p = 0$.

The correction is particularly effective for m, with a shift of about 0.03 towards positive val-

---

[7] G slightly depends on the parameters at which the model is evaluated. Therefore, it is computed in an iterative way: the non-corrected ($G = 1$) data-vector is initially fit, and G is evaluated; then $P^{data,G^{-1}}$ is computed, and the process repeats until convergence in the derived posteriors is observed (one or two iterations should be enough).

Another sidenote: the entity of the G-correction also depends on the size of k-bins, and the size of the box in which the sample is embedded.

ues, for the specific set of geometric choices: $\Delta k = 0.01h/\mathrm{Mpc}$, size of the periodic box, $L = 2.4$ Gpc/h, and the k-range fitted, $0.02 \leq k\,[h/\mathrm{Mpc}] \leq 0.15$. Even after applying the correction, a residual systematic shift lingers. The authors of (Brieden et al., 2021b) estimate that, thanks to this geometric correction, such residual bias becomes comparable to the statistical error for an effective volume greater than 400 Gpc^3, while without the correction that would happen for a volume of only 25 Gpc^3.

Full N-body mocks

With the previous analysis, we explored a very simple scenario, in which many aspects of modelling the LSS, like galaxy bias and RSD, have been neglected. But an inaccurate modelling of such effects could possibly lead to systematic errors. It is thus wise to evaluate a more complete case scenario, in which all possible effects are taken into account. This is achieved by repeating the previous analysis, but this time with full N-body galaxy mocks (in particular, the Nseries mocks are used, see *Appendix D*). Results of this analysis are displayed in *Figure* 4.4. Here, posteriors are obtained from Nseries N-body mocks using $P^{(0,2,4)}$ at $0.02 \leq k[\mathrm{Mpc}/h] \leq 0.15$. Different colours highlight the comparison among the different analyses performed: RSD (purple) and ShapeFit (orange), both with local-Lagrangian bias parametrization of (4.1), and ShapeFit with free non-local biases (green). All cases correspond to the fit of the mean power spectra of the 84 independent Nseries realizations with an associated effective volume of Veff = 308 Gpc^3. Solid and dashed contours respectively correspond to posteriors with and without ($G = 1$) the G-correction. Nuisance parameters are marginalized over, while n is set to $n = 0$ for ShapeFit (for simplicity). Black-dashed lines indicate the expected values for the underlying cosmology of the Nseries mocks.

The geometric correction appears to have a smaller effect for these mocks, compared to what we saw in the previous case. One reason is that here a smaller (hence, more realistic) effective volume has been employed (308 Gpc^3 versus 6448 Gpc^3 of the previous case), thus increasing statistical errors. The other reason is that here a larger box (4 Gpc/h side length) has been chosen compared to the previous case (2.4 Gpc/h side length), yielding a finer sampling of k-modes, which reduces the geometric effect with respect to the dark matter case.

Realistic case

From *Figure* 4.4, it appears once again evident that the most significant shift is the one affecting the parameter m, and amounting to $\Delta_m^{sys} \approx 0.01 - 0.02$. At this point, it may be useful to investigate how such shift fares when contextualized in a realistic survey case.

In *Figure* 4.5, orange contours result from fitting the mean of the 84 Nseries mocks (as it is done in *Figure* 4.4): this corresponds to an associated effective volume of 308 Gpc^3. Conversely, blue contours result from averaging the individual 84 posteriors, each one with an associated effective volume of 3.67 Gpc^3, which is way closer to a real-life case. The main observation to make is that all the systematic errors reported in *Figure* 4.4 are quite subdominant with respect to statistical errorbars associated to a real-life volume, including Δ_m^{sys}. Therefore, such systematic shift should not provide any significant issue when fitting actual datasets.

4.2 ShapeFit vs FM

We now turn back to the PATCHY mocks analysis, and show a comparison between the ShapeFit and the FM approach (Classic RSD results are also displayed in the plots for completeness). In particular, the plots show resulting constraints for four parameters: $\{w_{cdm}, h, \Omega_m, \sigma_8\}$. Prior-wise, both maximum and minimum freedom scenarios are explored. During the fit, one can choose either to fix some of the parameters or to let them vary freely. Here, three different cases are presented: fixed w_b, varying w_b, varying n_s.

4.2.1 Fixing w_b

In this case, w_b is kept fixed to its mock expected value $w_b = 0.02214$: this choice is quite realistic, since constraints on w_b coming from LSS alone are usually not competitive with CMB or BBN ones.

Results are displayed both as posteriors in *Figure* 4.6 and in the table of *Figure* 4.7.

The table shows the error and the bias for each parameter, for each approach, for each prior freedom assumption, and also for two different survey volumes. In particular, constraints are shown for the parameters $\mathbf{\Omega} = \{w_{cdm}, h, \Omega_m, \sigma_8\}$, whose mean values are $\mathbf{\Omega}_{\boldsymbol{PATCHY}} = \{0.118911, 0.6777, 0.301175, 0.8288\}$. The uncertainty on the mean value Ω is indicated as σ_Ω, while the bias is given in units of uncertainty, as $(\bar{\Omega} - \Omega_{PATCHY})/\sigma_\Omega$. ShapeFit (SF), FM and

Classic RSD fits results are compared for the different bias model cases "min" and "max". Fits have been carried out using a covariance matrix corresponding to two diffrent volumes: a volume of 300 $[\mathrm{Gpc}/h]^3$ ($V \times 100$) and a more realistic DESI-like survey volume of 30 $[\mathrm{Gpc}/h]^3$ ($V \times 10$). The mean values cited here are obtained from the ($V \times 100$) runs, as these are more Gaussianly distributed. They do not necessarily coincide with the mean values of the ($V \times 10$) runs due to non-Gaussianity, but it has been checked that best-fits agree with each other.

The plot displays results of the ShapeFit and FM fits to the PATCHY "ngc z3" sample compared to the Classic RSD "max" results for reference.

From both the table and the plot it can be appreciated how well ShapeFit performs in recovering the expected parameter values (black dotted lines in the plot), while featuring the same precision as the FM approach: with a considerably lower computational effort, ShapeFit appears able to capture the same bulk of cosmological information captured by the FM. The fact that the size of the constraints is very similar in the two cases indicates that this bulk of information is concentrated on large linear scales and corresponds to the early-time physics imprint left on the matter transfer function[8]. Therefore, the cosmological information enclosed in the BAO amplitude appears subdominant with respect to the early-time one.

In particular, parameter values are very well recovered, as expected, in the maximum freedom case (within 0.5σ for the realistic volume case); whereas, the minimum freedom case appears clearly biased (by up to 2σ for the realistic volume case).

Such biases show up within the FM approach as well, and they are likely due to the choice of neglecting IR resummation in the power spectrum modelling (as anticipated in the introduction of chapter 4): introducing the IR resummation correction would modulate the BAO amplitude, but also broaden the constraints. ShapeFit does not use any BAO amplitude information, and therefore does not need to account for any IR resummation correction.

4.2.2 Varying w_b

Here we investigate the case in which no external prior constraints (e.g. from CMB or BBN) are included in our analysis: therefore, we repeat the same analysis as before, but now w_b is let free to vary with a uniform prior in the range [0.005,0,04].

Results are displayed in *Figure* 4.8, where only the maximum freedom case is tackled.

---

[8] Once again, we stress that the same bulk of information was out of reach for the Classic approach, since it followed a late-time geometrical interpretation of the fitted variables.

By looking at the ShapeFit results in the left panel of the figure, we can easily notice some degeneracies. These arise because the slope m is degenerate with w_b and w_{cdm}, so ShapeFit is unable to constrain them individually. Since h is also strongly dependent on the w_b prior, we can conclude that ShapeFit is not able to constrain the three parameters w_b, w_{cdm} and h, while it can put constraints on the derived parameters Ω_m,σ_8 and D_V/r_d[9] (displayed in the right panel).

Figure 4.8 makes it evident that the FM fit brings an additional constraining power with respect to ShapeFit: in fact, it is indeed able to constrain w_b, w_{cdm} and h. This additional constraining power comes from the BAO amplitude[10], which linearly depends on w_b/w_{cdm}, hence breaking the degeneracy between w_b and w_{cdm}. From the plots, one can observe the effect of IR-resummation for the FM fit. This correction is needed to account for all the late-time effects (e.g. mode coupling, bias, non-linearities) which affect the BAO amplitude (which, by itself, is an early-time physics imprint). The no-IR posteriors are evidently biased, and the introduction of IR-resummation has the effect of broadening them, while keeping their peak maxima: in the end, they are still biased, but tension with the expected values is reduced by the broadening effect.

The bottom line of this section is that, for both ShapeFit and FM, the case of varying w_b appears quite problematic in terms of recovering some of the parameters, and even BAO amplitude (which should in principle fix up things by breaking the degeneracies) doesn't look like a reliable probe for these LSS analyses.

4.2.3 Varying n_s

We now investigate the case in which the scalar tilt n_s is taken as an additional free parameter. In such case, the ShapeFit analysis must be run with both the scale-dependent and the scale-independent slope parameters, m and n. Since these two produce a similar effect on the power spectrum, they are strongly anti-correlated (due to the intrinsic degeneracy between w_b, w_{cdm} and n_s through the slope). See *Figure* 4.9 to visually appreciate such anti-correlation (the posteriors displayed here have been obtained from applying the ShapeFit analysis to Nseries mocks, in the minimum freedom case).

---

[9] $D_V \equiv [D_M^2(z)D_H(z)]^{1/3}$.

[10] As already stated in section 3.4.2, ShapeFit is designed so that it does not compress any BAO amplitude feature as a way to extract information: therefore, the degeneracy between w_b and w_{cdm} stays.

Posteriors for the parameters $\{w_{cdm}, h, n_s, \Omega_m, \sigma_8\}$ are shown in *Figure* 4.10 (keep in mind that now w_b is fixed). The shown posteriors have been computed for a flat ΛCDM model and from the PATCHY ”ngc_z3” sample. Green contours display ShapeFit results, while violet and blue contours show FM results respectively with and without IR resummation correction.

The first thing to observe is that both ShapeFit and FM + IR-resummation are able to recover the expected parameters quite well, while the FM without the IR-resummation is clearly biased.

The already cited degeneracy between n_s and w_{cdm} (and, consequently, n_s and Ω_m) through the slope leads to visible anti-correlations.

FM is able to break this degeneracy by introducing the IR corrections, which implies taking into account late-time physics in order to correctly model the damped BAO amplitude. Conversely, ShapeFit does not make use of the BAO amplitude, so there is no need to introduce late-time physics in the picture: constraints on n_s, w_{cdm} and Ω_m are purely driven by the early-time imprint on the power spectrum slope.

As a final note, it can be added that in the case where both n_s and w_b are let free to vary for the FM fit, the bias in w_b observed in *Figure* 4.8 propagates into a bias in n_s, regardless of IR corrections.

In the next chapter, we will highlight another advantage ShapeFit has with respect to FM, namely the greater robustness of the compressed physical variables against systematics.

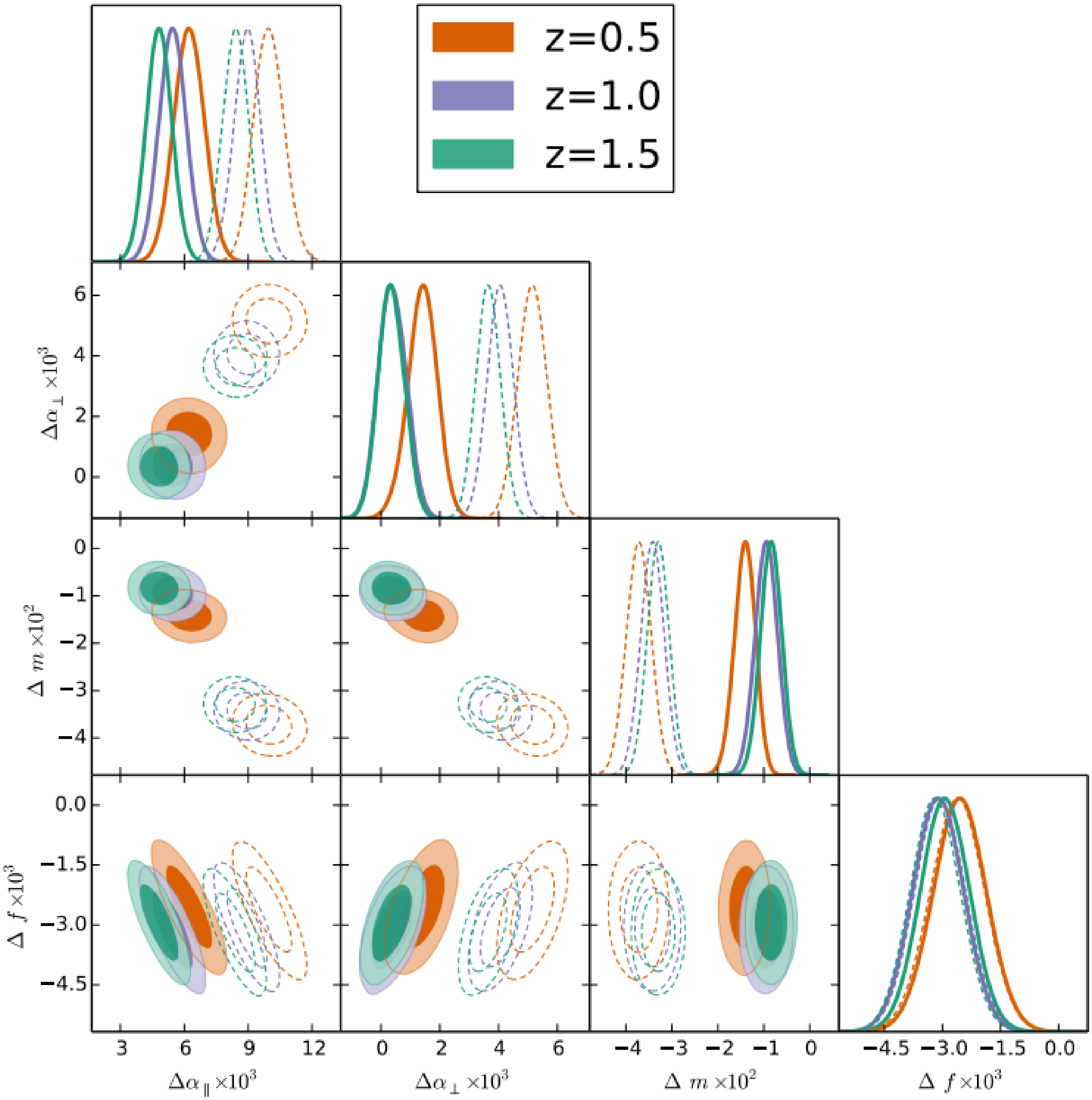

Figure 4.3: Posteriors derived when fitting the mean of 160 full N-body dark matter realization in real space. The covariance has been rescaled to correspond to an effective volume of 2200 $[\mathrm{Gpc}/h]^3$. We have used the model described in Eqq. (1.53) and (1.54) and the local-Lagrangian bias scheme of Eq. (3.1). In the plot we only show the physical parameters $f, \alpha_\parallel, \alpha_\perp, m$, although the remaining four nuisance parameters are also varied. The empty-dashed contours display the results of fitting the data-vector without any geometric correction (as usually done), whereas the filled contours account for the geometric correction through Eq. (3.2). For all parameters p, the expected value is $\Delta p = 0$. We observe that m is especially sensitive to the geometric correction with a shift of about 0.03 towards positive values, for the specific set of geometric choices: $\Delta k = 0.01 h/\mathrm{Mpc}$, size of the periodic box, $L = 2.4\ \mathrm{Gpc}/h$, and the k-range fitted, $0.02 \leq k\,[h/\mathrm{Mpc}] \leq 0.15$.

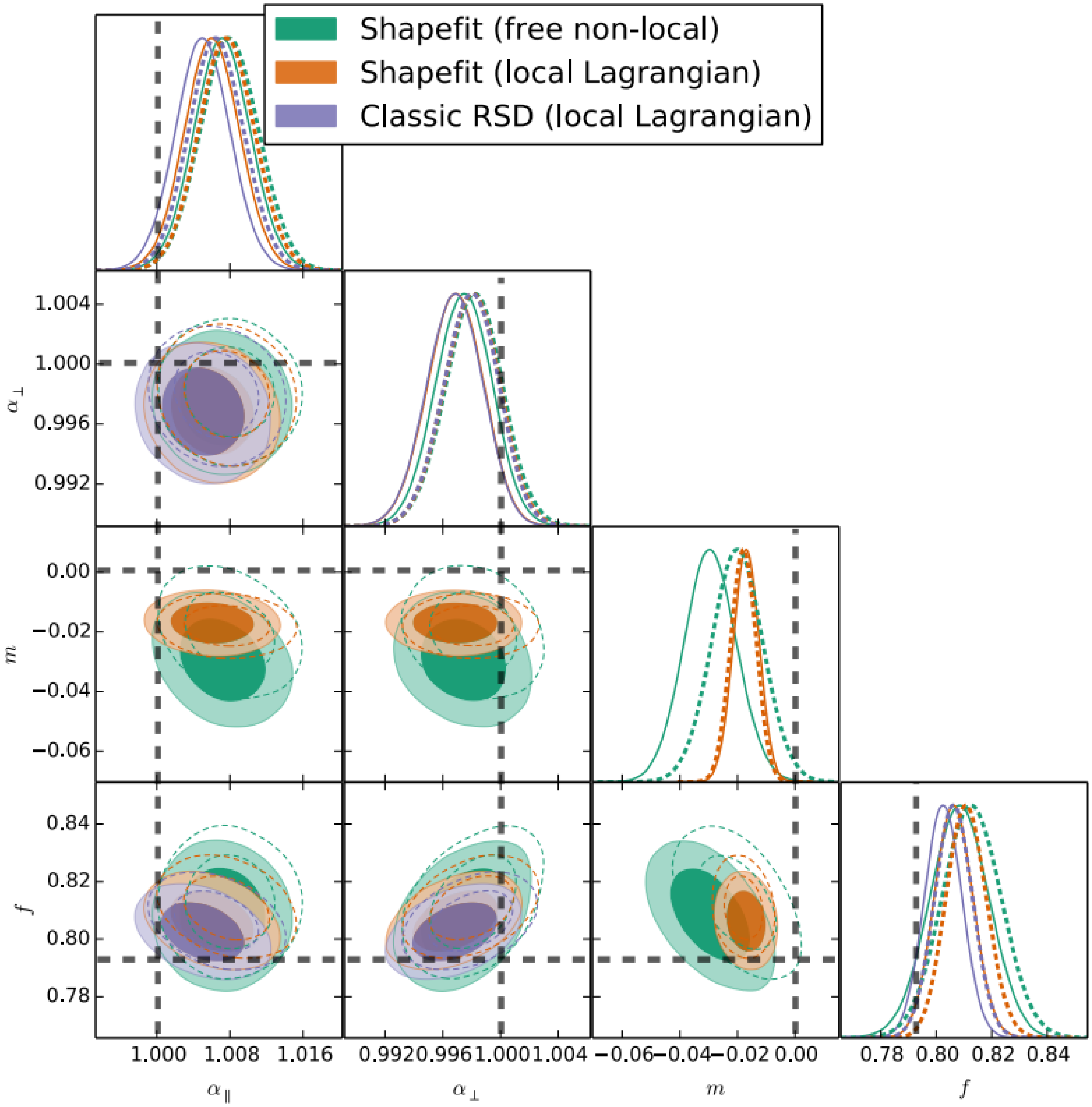

Figure 4.4: Posteriors from Nseries N-body mocks using $P^{(0,2,4)}$ at $0.02 \leq k[\mathrm{Mpc}/h] \leq 0.15$. We display the comparison of classic RSD (purple) and ShapeFit (orange), both with local-Lagrangian bias parametrization of (4.1), and ShapeFit with free non-local biases (green). All cases correspond to the fit of the mean power spectra of the 84 independent Nseries realizations with an associated effective volume of Veff = 308 Gpc^3. Horizontal black-dashed lines represent the true expected value for each compressed variable. The empty-dashed contours display the results without any geometric correction ($G = 1$) and the filled-solid contours with this correction included.

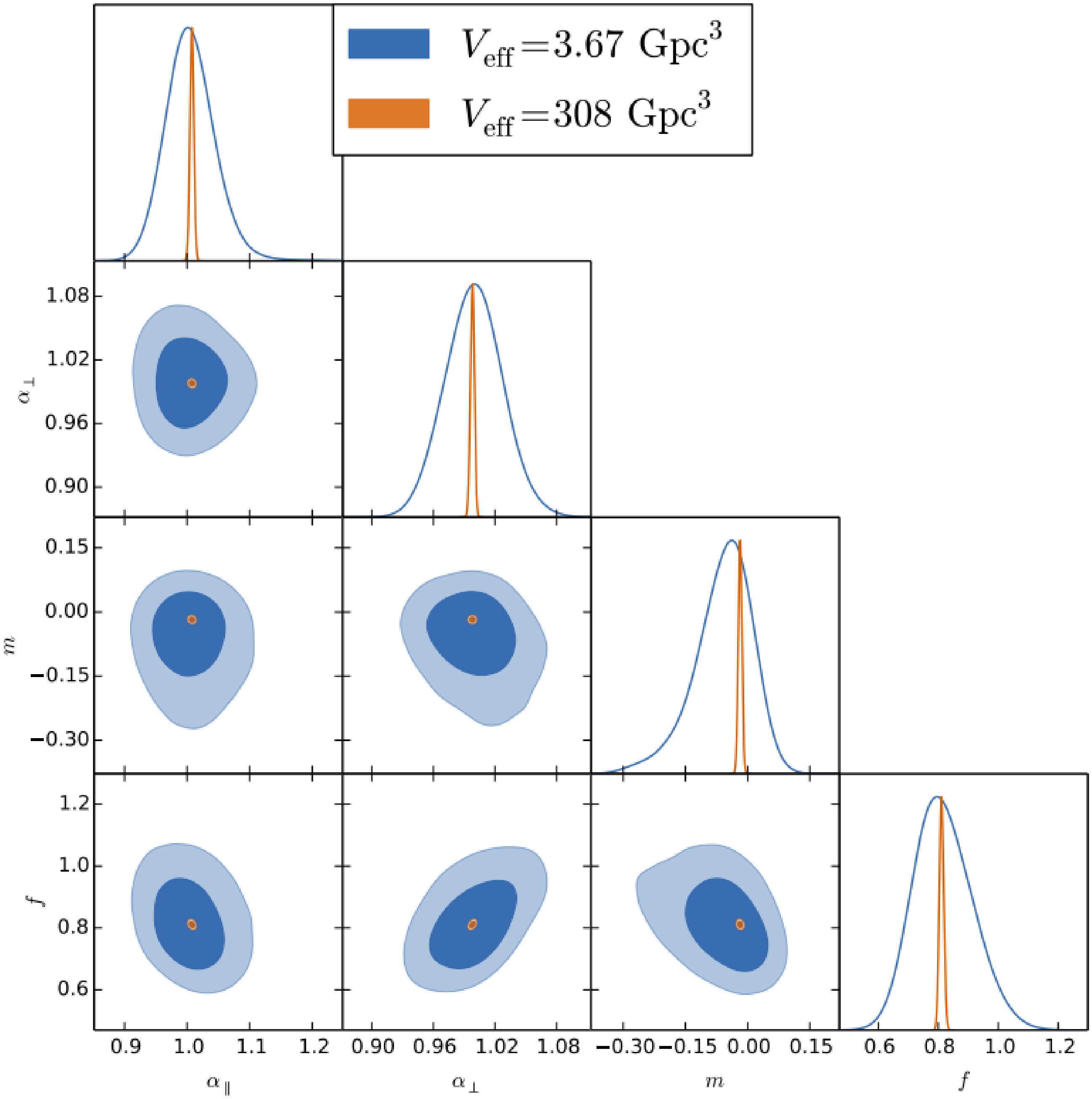

Figure 4.5: Case of ShapeFit with the local-Lagrangian bias assumption when fitting the mean of the 84 realizations (orange contours) and when fitting the averaged-posteriors of all 84 individual realizations (blue contours), which in this last case represents an effective volume similar to BOSS CMASS NGC, 3.67 Gpc^3.

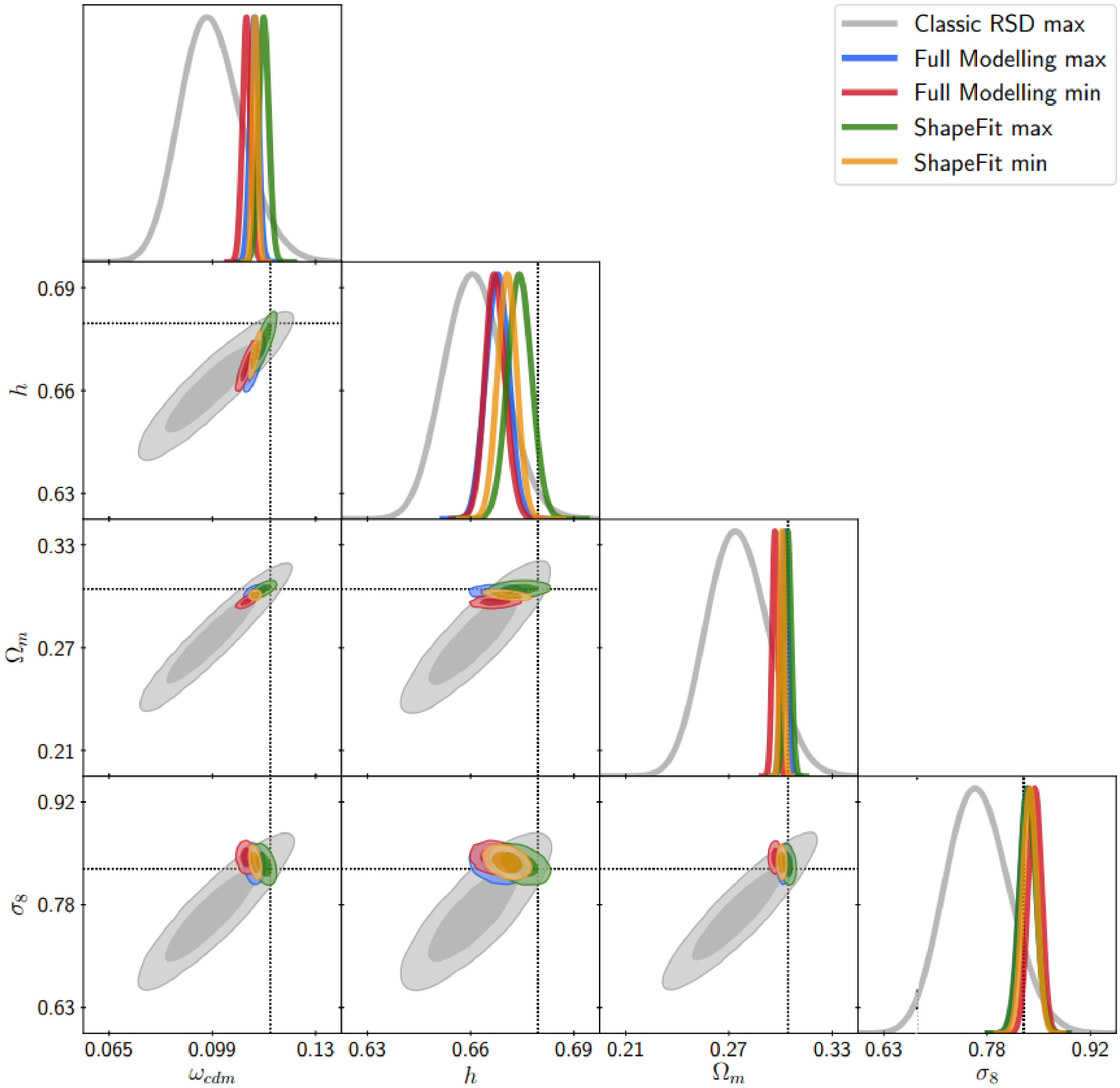

Figure 4.6: Results of the ShapeFit and FM fits to the PATCHY ”ngc_z3” sample compared to the Classic RSD ”max” results for reference. The size of the constraints are very similar in the two cases, indicating that ShapeFit captures the bulk of cosmological information captured by FM. The systematic shifts associated to the FM contours are partly caused by neglecting the IR resummation correction, which modulates the BAO amplitude, but also broadens the constraints. ShapeFit does not use any BAO amplitude information, and therefore does not need to crucially account for any IR-resummation correction. The fact that the size of the constraints is very similar in the two cases indicates that the cosmological information enclosed in the BAO amplitude is subdominant to the one enclosed in the large-scale shape of the power spectrum.

$\boldsymbol{\Omega}$	**Case**	**Fit**	**Mean**	**Error,** $V = 3(\mathrm{Gpc}/h)^3$		**Bias/**$\sigma_{\boldsymbol{\Omega}}$**,** $V = 3(\mathrm{Gpc}/h)^3$	
				$(V \times 100)$	$(V \times 10)$	$(V \times 100)$	$(V \times 10)$
ω_{cdm}	min	RSD	0.1162	0.0110	0.0400	-0.25	-0.10
		SF	0.1139	0.0009	0.0027	-5.57	-1.86
		FM	0.1110	0.0014	0.0040	-5.65	-1.97
	max	RSD	0.0997	0.0105	0.0372	-1.75	-0.69
		SF	0.1167	0.0017	0.0045	-1.24	-0.50
		FM	0.1136	0.0015	0.0046	-3.79	-1.21
h	min	RSD	0.6751	0.0079	0.0271	-0.35	-0.12
		SF	0.6695	0.0026	0.0079	-3.15	-1.02
		FM	0.6665	0.0027	0.0079	-4.00	-1.38
	max	RSD	0.6611	0.0079	0.0268	-2.21	-0.75
		SF	0.6729	0.0032	0.0102	-1.50	-0.49
		FM	0.6670	0.0031	0.0088	-3.45	-1.26
Ω_{m}	min	RSD	0.3031	0.0174	0.0620	-0.24	-0.08
		SF	0.3035	0.0012	0.0039	-2.78	-0.93
		FM	0.2996	0.0017	0.0049	-4.42	-1.57
	max	RSD	0.2783	0.0178	0.0623	-1.70	-0.54
		SF	0.3069	0.0021	0.0056	-0.11	-0.04
		FM	0.3052	0.0019	0.0057	-1.01	-0.33
σ_8	min	RSD	0.815	0.044	0.159	-0.33	-0.11
		SF	0.838	0.010	0.029	0.88	0.30
		FM	0.845	0.009	0.029	1.69	0.55
	max	RSD	0.765	0.044	0.153	-1.48	-0.53
		SF	0.835	0.012	0.033	0.50	0.18
		FM	0.836	0.012	0.032	0.57	0.21

Figure 4.7: Parameter constraints for $\boldsymbol{\Omega} = \{w_{cdm}, h, \Omega_m, \sigma_8\}$ given by the corresponding mean $\bar{\Omega}$, error σ_Ω and bias $(\bar{\Omega} - \Omega_{PATCHY})$ divided by σ_Ω with respect to the PATCHY cosmological parameters $\boldsymbol{\Omega_{PATCHY}} = \{0.118911, 0.6777, 0.301175, 0.8288\}$. ShapeFit (SF), FM and Classic RSD fits results are compared for the different bias model cases "min" and "max". Fits have been carried out using a covariance matrix corresponding to the volume of 100 stacked mocks $(V \times 100)$ and to 10 stacked mocks $(V \times 10)$, where $V = 3$ $(\mathrm{Gpc}/h)^3$. The mean values cited here are obtained from the $(V \times 100)$ runs, as these are more Gaussianly distributed. They do not necessarily coincide with the mean values of the $(V \times 10)$ runs due to non-Gaussianity, but it has been checked that best-fits agree with each other.

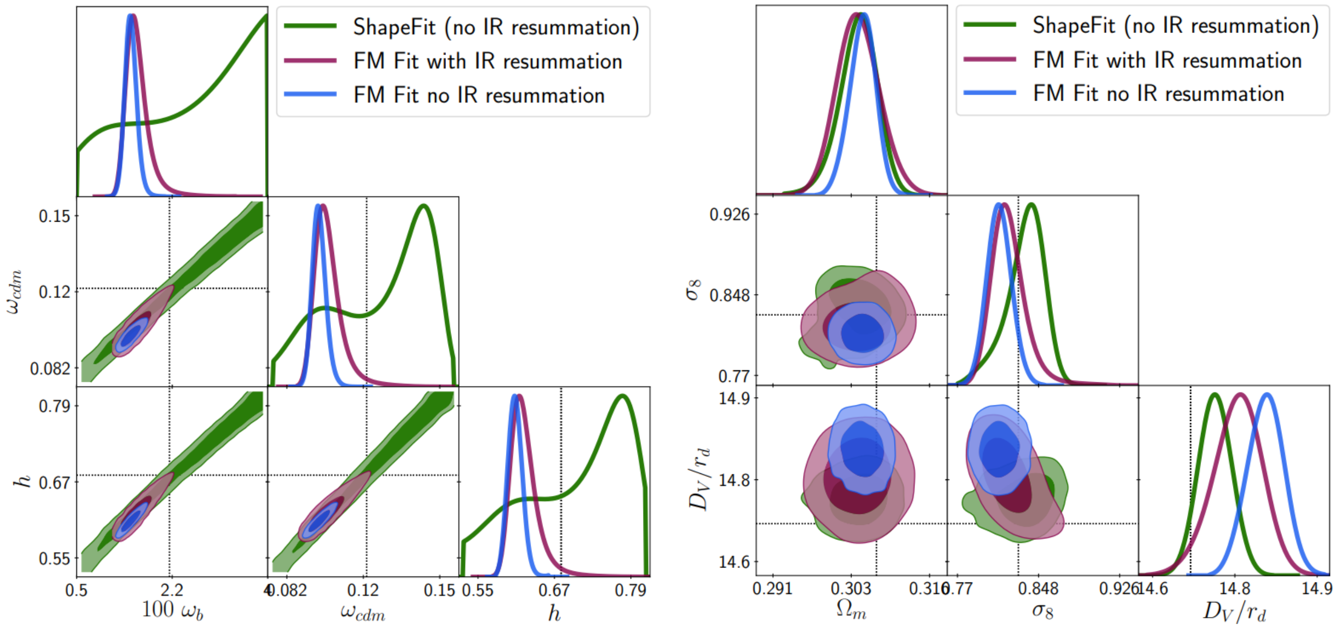

Figure 4.8: Results when analysing the PATCHY ”ngc_z3” sample with free w_b within a flat prior range of [0.005,0.04]. The green contours display the results from ShapeFit, whereas the blue and purple contours correspond to the FM fit case, with (purple) and without (blue) the IR resummation correction.

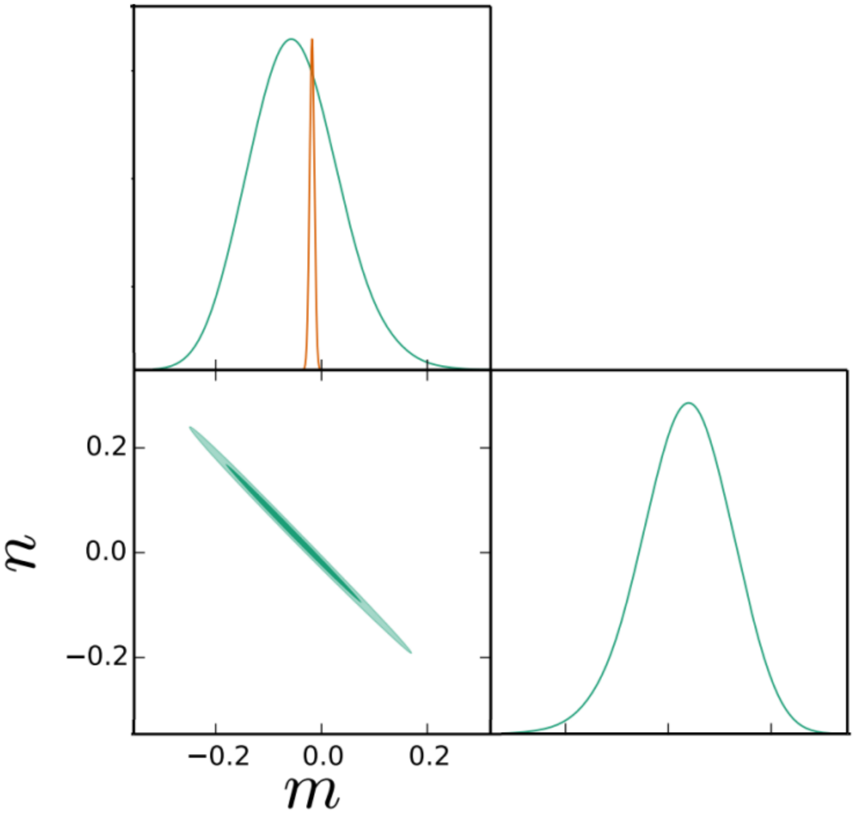

Figure 4.9: Posterior distribution for the mean of the Nseries mocks, corresponding to the minimum freedom case, for ShapeFit, with only m varying (in orange contours), and when both n and m are varied (green contours). The image is adapted from *Figure* 18 in *Appendix C* of (Brieden et al., 2021b).

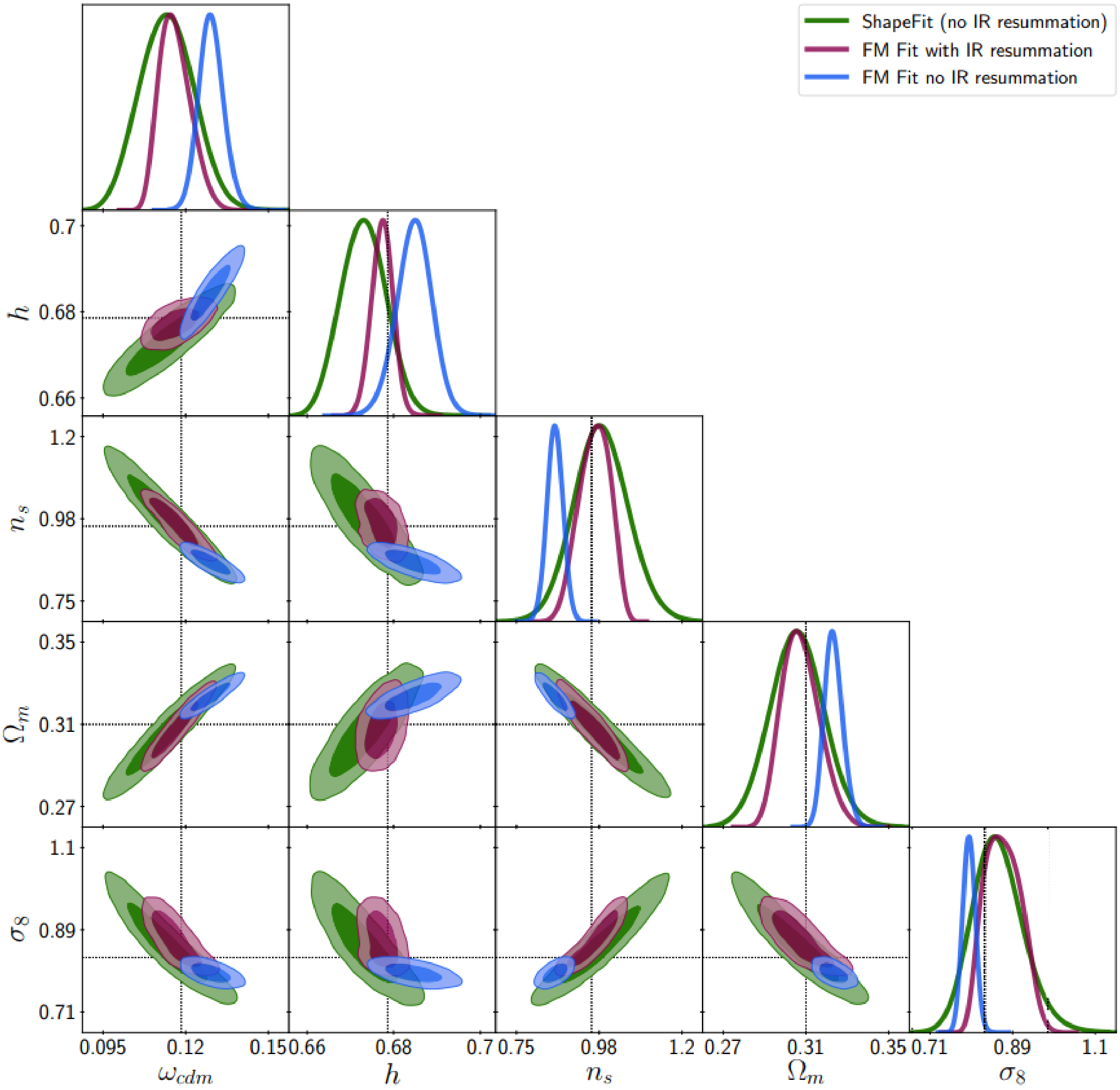

Figure 4.10: Derived posteriors for a flat ΛCDM model (for a fixed w_b) for ShapeFit (green contours) and the FM fit to the PATCHY ”ngc_z3” sample. The FM results are shown for the IR resummation correction turned on and off, in violet and blue contours, respectively. The differences between ShapeFit and FM posteriors are due to the extra constraining power from the BAO peak amplitude (not implemented in ShapeFit) which helps to break degeneracies between w_{cdm} and n_s. The BAO damping effect due to non-linear bulk flows, which IR resummation describes within FM, greatly reduces the BAO-amplitude-based constraining power, and hence broadens the n_s posteriors.

Chapter 5

Application to SDSS-III BOSS data

In chapter 4, the ShapeFit, Classic RSD and FM approaches have been applied to mock data. This allowed us to highlight some differences between the different methods that would otherwise remain unnoticed. For example, the larger effective volume provided by mocks makes it possible to appreciate some systematic shift that would otherwise be overcome by larger statistical errors.

The goal of this chapter, which will closely follow some results from (Brieden et al., 2021a), is to show how the different approaches fare when applied to a real dataset.

In particular, the analyses have been performed on the Luminous Red Galaxies (LRG) samples of the SDSS-III BOSS survey (Alam et al., 2017), covering two overlapping redshift ranges: $0.2 < z < 0.5$ ($z_{eff} = 0.38$, $V_{eff} = 3.7\text{Gpc}^3$) and $0.5 < z < 0.75$ ($z_{eff} = 0.61$, $V_{eff} = 4.1\text{Gpc}^3$). The two redshift ranges are considered uncorrelated.

The left panel of *Figure* 5.1 displays constraints on the compressed physical variables for Classic RSD ($\{\alpha_{\parallel}, \alpha_{\perp}, f\sigma_{s8}\}$), black dashed contours, and ShapeFit ($\{\alpha_{\parallel}, \alpha_{\perp}, f\sigma_{s8}, m\}$), solid green contours. In both cases the 1-loop SPT theory has been used to model the monopole and quadrupole signals for $0.01 \leq k[\text{Mpc}/h] \leq 0.15$. Only the high-$z$ bin of BOSS has been employed.

The lack of correlation between m and the other variables results in an almost perfect overlap between the two approaches' results. Actually, there exist a small correlation between m and $f\sigma_{s8}$, but it produces an increase in error of 5% only.

The physical variables constrained by ShapeFit in the left panel (along with another set of 4 parameters evaluated from the low-z bin of BOSS) have been taken as a starting point for the interpretation step. In particular, they are interpreted within a flat ΛCDM model with a Gaussian BBN prior $w_b = 0.02268 \pm 0.00038$. The resulting posteriors on the parameters $\{\Omega_m, w_{cdm}, h, \sigma_8\}$ are displayed in the right panel of *Figure* 5.1, along with those obtained from

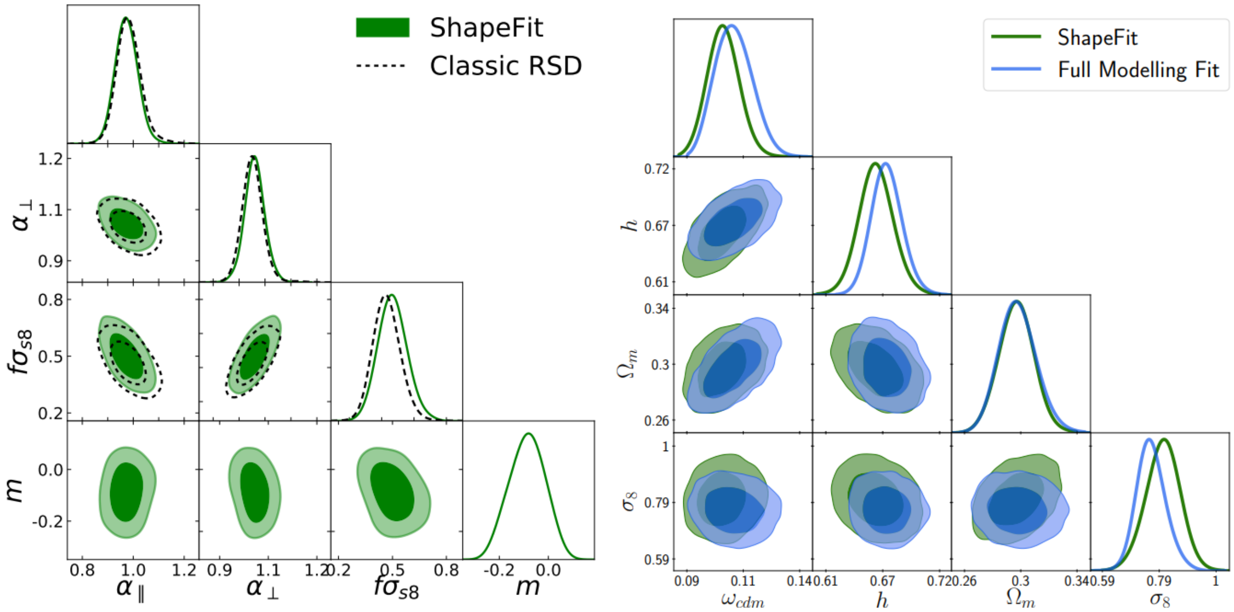

Figure 5.1: Left Panel: compressed physical parameters posteriors derived from power spectra measurements of the BOSS high-z sample, $z_{eff} = 0.61$ (constraints from the low-z sample show a very similar behaviour). Black dashed contours display the classic RSD results, while novel ShapeFit results are shown in green. In both cases the 1-loop SPT theory has been used to model the monopole and quadrupole signals for $0.01 \leq k[\mathrm{Mpc}/h] \leq 0.15$. Right Panel: posteriors derived from low and high-z samples of BOSS using the same scale-cuts as in the left panel. The blue contours correspond to the FM approach when a flat-ΛCDM model (+BBN Gaussian prior on w_b) is directly fitted to the 224 power spectra multipoles bins, $P^{(\ell)}(k,z)$, using EFT to model the power spectrum. Conversely, green contours are drawn from the 8 compressed physical variables of ShapeFit, interpreted under the same cosmological model as for the blue contours.

a FM fit (blue contours). The agreement between posteriors is evident from the plots, but it is necessary to highlight the difference in required computational effort. As already stressed in section 3.5.1, ShapeFit brings a significant computational advantage, since the fit must be taken only once, and the MCMC has to deal with less parameters. In this particular case, ShapeFit posteriors have been obtained from just 8 variables (4 physical variables for 2 redshift bins), while FM ones have been obtained from 224 $P^{(\ell)}(k,z)$ measurements (28 k-bins for two multipoles, two redshift bins and two galactic hemispheres).

Another advantage of ShapeFit with respect to FM lies in the greater robustness of the former approach against systematic effects, which could arise when dealing with realistic surveys. Such systematic effects appear to be "absorbed" by the early-time parameter m (which, as a result, gets shifted): this happens because late-time physics constraints from clustering measurements are significantly more robust than early-time ones.

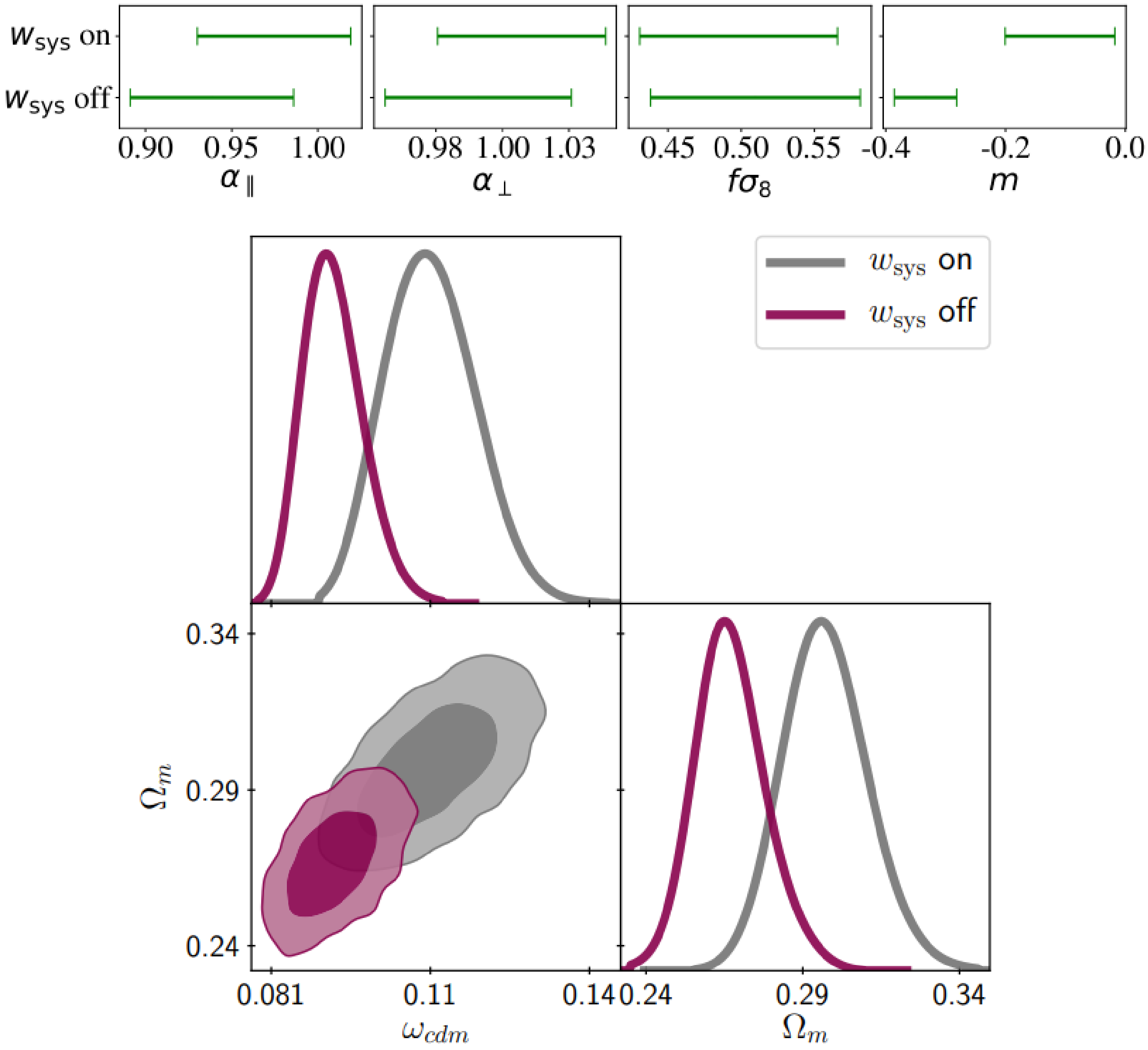

Figure 5.2: Effect of turning on and off the imaging systematic weights of BOSS data for ShapeFit in its compressed set of physical variables (upper panels); and for the FM fit in the $\{\Omega_m, w_{cdm}\}$ plane (lower panel). For ShapeFit $f\sigma_8$ and $\alpha_\perp$ are barely affected by this correction, whereas m absorbs most of the effect. For FM fit, w_{cdm} and Ω_m are significantly biased. The image corresponds to *FIG.3* of (Brieden et al., 2021a).

Such advantage is not present in the FM approach, as evident from *Figure* 5.2, showing results obtained from both the low-z and high-z BOSS samples. It displays constraints for the ShapeFit compressed variables (upper panel) and for the parameters Ω_m and w_{cdm} computed via FM. Results are shown for two scenarios: "w_{sys} off", corresponding to the case in which we account for the systematics (opportunely weighting them); "w_{sys} on", corresponding to the case in which systematic errors are present.

Switching the systematics on and off has two different effects for ShapeFit and FM. In the upper plots, we see that such action manifests as a shift in m, while other compressed variables stay mostly unchanged. In the lower panel, the same action produces a bias in the resulting constraints.

We can conclude that the compression step is able to isolate the part of cosmological signal whose information content is least affected by systematics.

Conclusions

In the present work, we have taken a deep dive into the ShapeFit approach, as a new methodology to extract information from LSS observations. This new technique presents itself as a step forwards in the field of statistical analysis of the LSS: it preserves the same model independence and computational advantage of Classic RSD, while reaching the same precision of FM.

This is achieved by the introduction of additional compressed parameters (i.e. the slope parameters m and n), with respect to Classic RSD, which provide a new source of information by modelling the power spectrum slope dependence on early-time physics. In particular, n models the scale-independent slope, while m corresponds to the slope of the matter power spectrum at a specific pivot scale. The complementarity between m and the classic compression variables has been tested by direct application of the ShapeFit technique to both mocks and real datasets (see *Figure* 3.6).

Conversely, ShapeFit is not designed to extract information the BAO amplitude, while this is achieved within the FM approach through the IR corrections, and contributes to the accuracy of a FM fit. This design choice lies in the fact that BAO amplitude information becomes relevant only when no CMB or BBN w_b prior is employed (see *Figure* 4.8 and *Figure* 13 in *Appendix A* of (Brieden et al., 2021b)); additionally it comes from an early-time physics imprint which is heavily processed by late-time effects, and therefore not robust enough. Implementing a IR resummation correction for ShapeFit would produce a tangible statistical improvement only for a survey volume of at least 300 $[\mathrm{Gpc}/h]^3$, to be compared with a more realistic volume $\lesssim 100$ $[\mathrm{Gpc}/h]^3$

The performed tests have confirmed that ShapeFit is indeed able to recover the same precision of FM (see *Figure* 4.6), while saving up on computational effort and appearing more robust against systematics (see *Figure* 5.2): more specifically, the computational time required by ShapeFit to perform the cosmological inference step is effectively indistinguishable from that of the Classic approach, and ~ 30 times faster than FM.

For the tests on mocks, two different prescriptions have been adopted in the choice of nuisance parameters' prior ranges: they have been denoted as "minimum freedom" and "maximum freedom". The former fixes non-local bias parameters by the local Lagrangian prediction, while the latter treats them as independent parameters. Since the broadband power spectrum shape, and hence the slope m, is sensitive to bias assumption, the authors of (Brieden et al., 2021b) advocate the adoption of "maximal freedom" in ShapeFit implementations, although this would slow down the MCMC chains convergence.

Further tests on mocks have highlighted the presence of some systematic shifts haunting the ShapeFit compressed variables. In particular, the employment of a geometric correction has been advocated, in order to counter the largest shift observed, the one in the parameter m for the "minimal freedom" case. However, such systematic shifts should have a very small impact on the next decade's surveys, since realistic effective volumes will still be relatively small, thus allowing statistical errors to dominate.

From these various applications and tests, ShapeFit appears as a very functional and convenient tool: therefore, its application to future surveys may provide new and interesting results.

Appendix A

Here the bias terms in (1.53) are displayed. These can be also found in (Beutler et al., 2014).

$$P_{m,b2\delta}(k) = \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) F_S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) \tag{A.1}$$

$$P_{m,b2\theta}(k) = \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) G_S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) \tag{A.2}$$

$$P_{m,bs2\delta}(k) = \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) F_S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) \tag{A.3}$$

$$P_{m,bs2\theta}(k) = \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) G_S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) \tag{A.4}$$

$$P_{m,b22\theta}(k) = \frac{1}{2} \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) \left[P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) - P_m^{\rm lin}(q) \right] \tag{A.5}$$

$$P_{m,b2s2\theta}(k) = -\frac{1}{2} \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) \left[\frac{2}{3} P_m^{\rm lin}(q) - P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) \right] \tag{A.6}$$

$$P_{m,bs22\theta}(k) = -\frac{1}{2} \int \frac{d^3q}{(2\pi)^3} P_m^{\rm lin}(q) \left[\frac{4}{9} P_m^{\rm lin}(q) - P_m^{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) S^{(2)}(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q})^2 \right] \tag{A.7}$$

The symmetrized second order PT kernels are given by

$$F_S^{(2)}(\boldsymbol{q_1}, \boldsymbol{q_2}) = \frac{5}{7} - \frac{\boldsymbol{q_1} \cdot \boldsymbol{q_2}}{2 q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \left(\frac{\boldsymbol{q_1} \cdot \boldsymbol{q_2}}{q_1 q_2} \right)^2 \tag{A.8}$$

$$G_S^{(2)}(\boldsymbol{q_1}, \boldsymbol{q_2}) = \frac{3}{7} - \frac{\boldsymbol{q_1} \cdot \boldsymbol{q_2}}{2q_1q_2}\left(\frac{q_1}{q_2} + \frac{q_2}{q_1}\right) + \frac{4}{7}\left(\frac{\boldsymbol{q_1} \cdot \boldsymbol{q_2}}{q_1q_2}\right)^2 \tag{A.9}$$

$$S^{(2)}(\boldsymbol{q_1}, \boldsymbol{q_2}) = \left(\frac{\boldsymbol{q_1} \cdot \boldsymbol{q_2}}{q_1q_2}\right)^2 - \frac{1}{3} \tag{A.10}$$

Additionally, we can define

$$D^{(2)}(\boldsymbol{q_1}, \boldsymbol{q_2}) = \frac{2}{7}\left[S^{(2)}(q_1, q_2) - \frac{2}{3}\right] \tag{A.11}$$

$$\sigma_3^2(k) = \frac{105}{16}\int \frac{d^3q}{(2\pi)^3} P_m^{\mathrm{lin}}(q)\left[D^{(2)}(-\boldsymbol{q}, \boldsymbol{k})S^{(2)}(\boldsymbol{q}, \boldsymbol{k} - \boldsymbol{q}) + \frac{8}{63}\right] \tag{A.12}$$

Appendix B

Here, we present the steps needed to recover the form of the Fisher Matrix for a MVN likelihood. Throughout the derivation, the following notation will be employed: $\partial f/\partial\theta_\alpha = f_{,\alpha}$.

Starting from (2.9), we calculate the derivatives of the right hand side terms ln C and $C^{-1}D$:

$$(\ln\, C)_{,\alpha} = C^{-1}C_{,\alpha} \tag{B.1}$$

$$(C^{-1}D)_{,\alpha} = C^{-1}_{,\alpha}D + C^{-1}D_{,\alpha} \tag{B.2}$$

By exploiting the identity $(C^{-1})_{,\alpha} = -C^{-1}C_{,\alpha}C^{-1}$, and using (B.1) and (B.2), we get:

$$(-2\mathrm{ln}L)_{,\alpha} = \mathrm{Tr}\left[C^{-1}C_{,\alpha} - C^{-1}C_{,\alpha}C^{-1}D + C^{-1}D_{,\alpha}\right] \tag{B.3}$$

Now, in order to obtain $(-2\mathrm{ln}L)_{,\alpha\beta}$, we need to further differentiate, this time with respect to β. As before, we first compute the derivatives of the right hand side terms $C^{-1}C_{,\alpha}$, $(C^{-1}C_{,\alpha})(C^{-1}D)$ and $C^{-1}D_{,\alpha}$:

$$\begin{aligned}(C^{-1}C_{,\alpha})_{,\beta} &= C^{-1}_{,\beta}C_{,\alpha} + C^{-1}C_{,\alpha\beta} = \\ &= -C^{-1}C_{,\beta}C^{-1}C_{,\alpha} + C^{-1}C_{,\alpha\beta}\end{aligned} \tag{B.4}$$

$$\begin{aligned}\left[(C^{-1}C_{,\alpha})(C^{-1}D)\right]_{,\beta} &= (C^{-1}C_{,\alpha})_{,\beta}(C^{-1}D) + (C^{-1}C_{,\alpha})(C^{-1}D)_{,\beta} = \\ &= -C^{-1}C_{,\beta}C^{-1}C_{,\alpha}C^{-1}D + C^{-1}C_{,\alpha\beta}C^{-1}D + \\ &\quad + (C^{-1}C_{,\alpha})(C^{-1}_{,\beta}D) + (C^{-1}C_{,\alpha})(C^{-1}D_{,\beta}) = \\ &= -C^{-1}C_{,\beta}C^{-1}C_{,\alpha}C^{-1}D + C^{-1}C_{,\alpha\beta}C^{-1}D + \\ &\quad - C^{-1}C_{,\alpha} - C^{-1}C_{,\beta}C^{-1}D + C^{-1}C_{,\alpha}C^{-1}D_{,\beta}\end{aligned} \tag{B.5}$$

$$(C^{-1}D_{,\alpha})_{,\beta} = -C^{-1}C_{,\beta}C^{-1}D_{,\alpha} + C^{-1}D_{,\alpha\beta} \tag{B.6}$$

Therefore, putting everything together we get

$$\begin{aligned}(-2\ln L)_{,\alpha\beta} =& \mathrm{Tr}\Big[- C^{-1}C_{,\beta}C^{-1}C_{,\alpha}) + C^{-1}C_{,\alpha\beta} + \\ & + C^{-1}(C_{,\beta}C^{-1}C_{,\alpha} - C_{,\alpha\beta} + C_{,\alpha}C^{-1}C_{,\beta})C^{-1}D + \\ & + C^{-1}C_{,\alpha}C^{-1}D_{,\beta} - C^{-1}C_{,\beta}C^{-1}D_{,\alpha} + C^{-1}D_{,\alpha\beta}\Big]\end{aligned} \tag{B.7}$$

Therefore, putting everything together we get

$$\begin{aligned}(-2\ln L)_{,\alpha\beta} =& \mathrm{Tr}\Big[- C^{-1}C_{,\beta}C^{-1}C_{,\alpha}) + C^{-1}C_{,\alpha\beta} + \\ & + C^{-1}(C_{,\beta}C^{-1}C_{,\alpha} - C_{,\alpha\beta} + C_{,\alpha}C^{-1}C_{,\beta})C^{-1}D + \\ & + C^{-1}C_{,\alpha}C^{-1}D_{,\beta} - C^{-1}C_{,\beta}C^{-1}D_{,\alpha} + C^{-1}D_{,\alpha\beta}\Big]\end{aligned} \tag{B.8}$$

When taking the ensemble average, we need to consider that all data dependence is inside the data matrix D. So, by exploiting the linearity of the ensemble average operator, we find:

$$\begin{aligned}\Big\langle -2\frac{\partial \mathrm{ln} L}{\partial\theta_\alpha\partial\theta_\beta}\Big\rangle = \mathrm{Tr}\Big[& -C^{-1}C_{,\beta}C^{-1}C_{,\alpha}+C^{-1}C_{,\alpha\beta}+ \\ & +C^{-1}(C_{,\beta}C^{-1}C_{,\alpha}-C_{,\alpha\beta}+C_{,\alpha}C^{-1}C_{,\beta})C^{-1}\langle D\rangle + \\ & +C^{-1}C_{,\alpha}C^{-1}\langle D_{,\beta}\rangle - C^{-1}C_{,\beta}C^{-1}\langle D_{,\alpha}\rangle + C^{-1}\langle D_{,\alpha\beta}\rangle\Big]\end{aligned} \tag{B.9}$$

Therefore, one needs to compute the ensemble average of the data matrix and its derivatives. According to the definition of Covariance, we can start by writing

$$\langle D_{ij}\rangle = \langle (x_i-\mu_i)(x_j-\mu_j)\rangle = C_{ij} \tag{B.10}$$

Since the data do not depend on parameters (which are denoted here by greek letters), and since, by definition, $\langle x_i\rangle = \mu_i$, we find:

$$\begin{aligned}\langle D_{ij,\alpha}\rangle &= \langle (x_{i,\alpha}-\mu_{i,\alpha})(x_j-\mu_j)+(x_i-\mu_i)(x_{j,\alpha}-\mu_{j,\alpha})\rangle = \\ &= \langle -\mu_{i,\alpha}(x_j-\mu_j)-\mu_{j,\alpha}(x_i-\mu_i)\rangle = \\ &= -\langle \mu_{i,\alpha}x_j\rangle + \langle \mu_{i,\alpha}\mu_j\rangle - \langle \mu_{j,\alpha}x_i\rangle + \langle \mu_{j,\alpha}\mu_i\rangle = \\ &= -\langle \mu_{i,\alpha}\mu_j\rangle + \langle \mu_{i,\alpha}\mu_j\rangle - \langle \mu_{j,\alpha}\mu_i\rangle + \langle \mu_{j,\alpha}\mu_i\rangle = \\ &= \langle -\mu_{i,\alpha}\mu_j + \mu_{i,\alpha}\mu_j - \mu_{j,\alpha}\mu_i + \mu_{j,\alpha}\mu_i\rangle = 0\end{aligned} \tag{B.11}$$

$$\begin{aligned}\langle D_{ij,\alpha\beta}\rangle &= \langle [-\mu_{i,\alpha}(x_j-\mu_j)-\mu_{j,\alpha}(x_i-\mu_i)]_{,\beta}\rangle = \\ &= \langle [-\mu_{i,\alpha}x_j+\mu_{i,\alpha}\mu_j-\mu_{j,\alpha}x_i+\mu_{j,\alpha}\mu_i]_{,\beta}\rangle = \\ &= \langle -\mu_{i,\alpha\beta}x_j-\mu_{i,\alpha}x_{j,\beta}+\mu_{i,\alpha\beta}\mu_j+\mu_{i,\alpha}\mu_{j,\beta}+ \\ &\quad -\mu_{j,\alpha\beta}x_i-\mu_{j,\alpha}x_{i,\beta}+\mu_{j,\alpha\beta}\mu_i+\mu_{j,\alpha}\mu_{i,\beta}\rangle = \\ &= \langle -\mu_{i,\alpha\beta}x_j+\mu_{i,\alpha\beta}\mu_j+\mu_{i,\alpha}\mu_{j,\beta}-\mu_{j,\alpha\beta}x_i+\mu_{j,\alpha\beta}\mu_i+\mu_{j,\alpha}\mu_{i,\beta}\rangle = \\ &= -\mu_{i,\alpha\beta}\mu_j+\mu_{i,\alpha\beta}\mu_j+\mu_{i,\alpha}\mu_{j,\beta}-\mu_{j,\alpha\beta}\mu_i+\mu_{j,\alpha\beta}\mu_i+\mu_{j,\alpha}\mu_{i,\beta} = \\ &= \mu_{i,\alpha}\mu_{j,\beta}+\mu_{j,\alpha}\mu_{i,\beta}\end{aligned} \tag{B.12}$$

Therefore, in vector matrix notation, we can write such results as

$$\langle D \rangle = C \tag{B.13}$$

$$\langle D_{,\alpha} \rangle = 0 \tag{B.14}$$

$$\langle D_{,\alpha\beta} \rangle = \boldsymbol{\mu}_{,\alpha}^{T} \boldsymbol{\mu}_{,\beta} + \boldsymbol{\mu}_{,\beta}^{T} \boldsymbol{\mu}_{,\alpha} \equiv M_{\alpha\beta} \tag{B.15}$$

which we can substitute in (B.9), thus finally obtaining:

$$\begin{aligned}
\Big\langle -2\frac{\partial \mathrm{ln} L}{\partial\theta_\alpha \partial\theta_\beta} \Big\rangle = \mathrm{Tr}\Big[& -C^{-1}C_{,\beta}C^{-1}C_{,\alpha} + C^{-1}C_{,\alpha\beta} + \\
& + C^{-1}(C_{,\beta}C^{-1}C_{,\alpha} - C_{,\alpha\beta} + C_{,\alpha}C^{-1}C_{,\beta})C^{-1}C + \\
& + 0 - 0 + C^{-1}M_{\alpha\beta}\Big] = \\
& = \mathrm{Tr}\left[C^{-1}C_{,\alpha}C^{-1}C_{,\beta} + C^{-1}M_{\alpha\beta}\right]
\end{aligned} \tag{B.16}$$

Appendix C

SPHEREx specifications				
$z_{\rm min}$	$z_{\rm max}$	$z_{\rm eff}$	$n(z)$ $[h/{\rm Mpc}]^3$	$b(z)$
0.0	0.2	0.1	9.97×10^{-3}	1.3
0.2	0.4	0.3	4.11×10^{-3}	1.5
0.4	0.6	0.5	5.01×10^{-4}	1.8
0.6	0.8	0.7	7.05×10^{-5}	2.3
0.8	1.0	0.9	3.16×10^{-5}	2.1
1.0	1.6	1.1	1.64×10^{-3}	2.7
1.6	2.2	1.3	3.59×10^{-6}	3.6
2.2	2.8	1.5	8.07×10^{-7}	2.3
2.8	3.4	1.7	1.84×10^{-6}	3.2
3.4	4.0	1.9	1.5×10^{-6}	2.7
4.0	4.6	1.7	1.13×10^{-6}	3.8

Table 5.1: Specifications adopted for the Fisher Forecast performed for the SPHEREx galaxy survey. The adopted redshift uncertainty is $\sigma_z = 0.003$. For this survey, 75% of an all-sky coverage has been assumed, amounting to $\simeq 31\,000$ square degrees.

ATLAS specifications				
z_{min}	z_{max}	z_{eff}	$n(z)$ $[h/\text{Mpc}]^3$	$b(z)$
0.5	0.7	0.6	2.88×10^{-2}	1.14
0.7	0.9	0.8	2.39×10^{-2}	1.22
0.9	1.1	1.0	2.81×10^{-2}	1.3
1.1	1.3	1.2	2.01×10^{-2}	1.38
1.3	1.5	1.4	1.95×10^{-2}	1.46
1.5	1.7	1.6	2.53×10^{-2}	1.54
1.7	1.9	1.8	1.29×10^{-2}	1.62
1.9	2.1	2.0	8.76×10^{-3}	1.7
2.1	2.3	2.2	6.76×10^{-3}	1.78
2.3	2.5	2.4	5.98×10^{-3}	1.86
2.5	2.7	2.6	8.63×10^{-3}	1.94
2.7	2.9	2.8	3.36×10^{-3}	2.02
2.9	3.1	3.0	3.06×10^{-3}	2.1
3.1	3.3	3.2	1.70×10^{-3}	2.18
3.3	3.5	3.4	1.77×10^{-3}	2.26
3.5	3.7	3.6	7.47×10^{-4}	2.34
3.7	3.9	3.8	4.42×10^{-4}	2.42
3.9	4.1	4.0	2.04×10^{-4}	2.5

Table 5.2: Specifications adopted for the Fisher Forecast performed for the ATLAS (WIDE) galaxy survey. The adopted redshift uncertainty is $\sigma_z = 0.0001$, with a sky coverage of 2000 square degrees. The bias is computed as $b(z) = 1.5 + 0.4(z - 1.5)$ (Wang et al., 2019)

Appendix D

Here we provide some more details about the N-body mocks employed for the analysis presented in section 4.1.1 .

Simple dark matter scenario

For this analysis, a N-body suite of 160 simulations has been employed. The simulations are based on a flat ΛCDM cosmology, consistent with WMAP (Hinshaw et al., 2013) bestfit cosmology. Each simulation corresponds to a box with size $L = 2.4$ Gpc/h, and $N_p = 768^3$ particles. The initial conditions have been generated at $z = 49$ by displacing the particles according to second-order Lagrangian PT from their initial grid points. In this analysis, the outputs at three redshifts have been considered: $z = 0.5, 1.0, 1.5$. No RSD are applied, out of simplicity.

The data-vector of each of the 160 realizations is obtained at each redshift bin, consisting of its monopole, quadrupole and hexadecapole signals between $0.02 \leq k[\mathrm{Mpc}/h] \leq 0.15$, sampled in bins of $\Delta k = 0.01$ Mpc/h size, and with a total of 13×3 elements. The average of the 160 data-vector realizations is used as dataset. The total associated volume of such data-vector amounts to 6448 Gpc^3. Covariance is estimated from the 160 realizations, and then rescaled to be that corresponding to the full effective volume available.

More information about these simulations can be found in section 3.1 of (Gil-Marín, Wagner, Norena, Verde, & Percival, 2014).

Nseries mocks

These mocks have been generated out of 7 independent periodic boxes of 2.6 Gpc/h side. From those, a total of 84 pseudo-independent realizations have been extracted, with sky geometry similar to the northern galactic cap of CMASS DR12 data, for $0.43 < z < 0.7$ ($z_{eff} = 0.56$). 2048^3 dark matter particles per box are considered. The identified halos are populated with galaxies following a halo occupation distribution model tuned to match the clustering of LGRs observed by BOSS. Covariance is estimated from the 2048 realizations of NGC CMASS DR12 PATCHY mocks catalogues, and then rescaled by a 10% factor, since the PATCHY mocks have 10% fewer particles than the Nseries mocks. The underlying cosmology of the Nseries mocks is close to that of WMAP. The effective volume sampled is $V_{eff} = 84 \times 6.67 = 308[\mathrm{Gpc}]^3$.

Further details on these mocks can be found in section 2.2.2 of (Gil-Marín et al., 2020).

References

Ade, P. A., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., … others (2016). Planck 2015 results-xiv. dark energy and modified gravity. *Astronomy & Astrophysics*, *594*, A14.

Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., … others (2020). Planck 2018 results-vi. cosmological parameters. *Astronomy & Astrophysics*, *641*, A6.

Alam, S., Ata, M., Bailey, S., Beutler, F., Bizyaev, D., Blazek, J. A., … others (2017). The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, *470*(3), 2617–2652.

Alam, S., Aubert, M., Avila, S., Balland, C., Bautista, J. E., Bershady, M. A., … others (2021). Completed sdss-iv extended baryon oscillation spectroscopic survey: Cosmological implications from two decades of spectroscopic surveys at the apache point observatory. *Physical Review D*, *103*(8), 083533.

Alcock, C., & Paczyński, B. (1979). An evolution free test for non-zero cosmological constant. *Nature*, *281*(5730).

Amendola, L., Appleby, S., Avgoustidis, A., Bacon, D., Baker, T., Baldi, M., … others (2018). Cosmology and fundamental physics with the euclid satellite. *Living reviews in relativity*, *21*, 1–345.

Baldauf, T., Seljak, U., Desjacques, V., & McDonald, P. (2012). Evidence for quadratic tidal tensor bias from the halo bispectrum. *Physical Review D*, *86*(8), 083540.

Ballinger, W., Peacock, J., & Heavens, A. (1996). Measuring the cosmological constant with redshift surveys. *Monthly Notices of the Royal Astronomical Society*, *282*(3), 877–888.

Bardeen, J. M., Bond, J., Kaiser, N., & Szalay, A. (1986). The statistics of peaks of gaussian random fields. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 304, May 1, 1986, p. 15-61. SERC-supported research.*, *304*, 15–61.

Barreira, A., Lazeyras, T., & Schmidt, F. (2021). Galaxy bias from forward models: linear and second-order bias of illustristng galaxies. *Journal of Cosmology and Astroparticle Physics*, *2021*(08), 029.

Bertacca, D., Maartens, R., Raccanelli, A., & Clarkson, C. (2012). Beyond the plane-parallel and newtonian approach: Wide-angle redshift distortions and convergence in general relativity. *Journal of Cosmology and Astroparticle Physics*, *2012*(10), 025.

Beutler, F., Castorina, E., & Zhang, P. (2019). Interpreting measurements of the anisotropic galaxy power spectrum. *Journal of Cosmology and Astroparticle Physics*, *2019*(03), 040.

Beutler, F., Saito, S., Seo, H.-J., Brinkmann, J., Dawson, K. S., Eisenstein, D. J., … others (2014). The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: testing gravity with redshift space distortions using the power spectrum multipoles. *Monthly Notices of the Royal Astronomical Society*, *443*(2), 1065–1089.

Bianchi, D., Gil-Marín, H., Ruggeri, R., & Percival, W. J. (2015). Measuring line-of-sight-dependent fourier-space clustering using ffts. *Monthly Notices of the Royal Astronomical Society: Letters*, *453*(1), L11–L15.

Blanchard, A., Camera, S., Carbone, C., Cardone, V., Casas, S., Clesse, S., … others (2020). Euclid preparation-vii. forecast validation for euclid cosmological probes. *Astronomy & Astrophysics*, *642*, A191.

Bond, J. R., & Szalay, A. S. (1983). The collisionless damping of density fluctuations in an expanding universe. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 274, Nov. 15, 1983, p. 443-468.*, *274*, 443–468.

Brieden, S., Gil-Marín, H., & Verde, L. (2021a). Model-independent versus model-dependent interpretation of the sdss-iii boss power spectrum: Bridging the divide. *Physical Review D*, *104*(12), L121301.

Brieden, S., Gil-Marín, H., & Verde, L. (2021b). Shapefit: extracting the power spectrum shape information in galaxy surveys beyond bao and rsd. *Journal of Cosmology and Astroparticle Physics*, *2021*(12), 054.

Carlson, J., White, M., & Padmanabhan, N. (2009). Critical look at cosmological perturbation theory techniques. *Physical Review D*, *80*(4), 043531.

Carroll, S. M., Press, W. H., & Turner, E. L. (1992). The cosmological constant. *Annual review of astronomy and astrophysics*, *30*(1), 499–542.

Chan, K. C., Scoccimarro, R., & Sheth, R. K. (2012). Gravity and large-scale nonlocal bias. *Physical Review D*, *85*(8), 083509.

Coe, D. (2009). Fisher matrices and confidence ellipses: a quick-start guide and software. *arXiv preprint arXiv:0906.4123*.

Desi, C., Aghamousa, A., Aguilar, J., Ahlen, S., Alam, S., Allen, L., … others (2016). The desi experiment part i: science, targeting, and survey design.

Dodelson, S., & Schmidt, F. (2020). *Modern cosmology.* Academic press.

Doré, O., Bock, J., Ashby, M., Capak, P., Cooray, A., de Putter, R., … others (2014). Cosmology with the spherex all-sky spectral survey. *arXiv preprint arXiv:1412.4872*.

Einstein, A. (1915). Die feldgleichungen der gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften*, 844–847.

Einstein, A., et al. (1916). The foundation of the general theory of relativity. *Annalen Phys*, *49*(7), 769–822.

Eisenstein, D. J., & Hu, W. (1998). Baryonic features in the matter transfer function. *The Astrophysical Journal*, *496*(2), 605.

Eisenstein, D. J., & Hu, W. (1999). Power spectra for cold dark matter and its variants. *The Astrophysical Journal*, *511*(1), 5.

Feldman, H. A., Kaiser, N., & Peacock, J. A. (1993). Power spectrum analysis of three-dimensional redshift surveys. *arXiv preprint astro-ph/9304022*.

Friedman, A. (1922). Über die krümmung des raumes. *Zeitschrift für Physik*, *10*(1), 377–386.

Fry, J. N., & Gaztanaga, E. (1993). Biasing and hierarchical statistics in large-scale structure. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 413, no. 2, p. 447-452.*, *413*, 447–452.

Gil-Marín, H., Bautista, J. E., Paviot, R., Vargas-Magaña, M., de La Torre, S., Fromenteau, S., … others (2020). The completed sdss-iv extended baryon oscillation spectroscopic survey:

measurement of the bao and growth rate of structure of the luminous red galaxy sample from the anisotropic power spectrum between redshifts 0.6 and 1.0. *Monthly Notices of the Royal Astronomical Society*, *498*(2), 2492–2531.

Gil-Marín, H., Wagner, C., Norena, J., Verde, L., & Percival, W. (2014). Dark matter and halo bispectrum in redshift space: theory and applications. *Journal of Cosmology and Astroparticle Physics*, *2014*(12), 029.

Hamilton, A. (1998). Linear redshift distortions: A review. *The Evolving Universe: Selected Topics on Large-Scale Structure and on the Properties of Galaxies*, 185–275.

Hartlap, J., Simon, P., & Schneider, P. (2007). Why your model parameter confidences might be too optimistic. unbiased estimation of the inverse covariance matrix. *Astronomy & Astrophysics*, *464*(1), 399–404.

Hartle, J. B. (2003). *Gravity: an introduction to einstein's general relativity.* American Association of Physics Teachers.

Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C., Dunkley, J., … others (2013). Nine-year wilkinson microwave anisotropy probe (wmap) observations: cosmological parameter results. *The Astrophysical Journal Supplement Series*, *208*(2), 19.

Hu, W., Scott, D., Sugiyama, N., & White, M. (1995). Effect of physical assumptions on the calculation of microwave background anisotropies. *Physical Review D*, *52*(10), 5498.

Hubble, E. (1929). A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the national academy of sciences*, *15*(3), 168–173.

Ivanov, M. M., Simonović, M., & Zaldarriaga, M. (2020). Cosmological parameters from the boss galaxy power spectrum. *Journal of Cosmology and Astroparticle Physics*, *2020*(05), 042.

Ivezić, Ž., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., … others (2019). Lsst: from science drivers to reference design and anticipated data products. *The Astrophysical Journal*, *873*(2), 111.

Kaiser, N. (1987). Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, *227*(1), 1–21.

Kendall, M. G., & Stuart, A. (1969). The advanced theory of statistics. vol. 3. *Biometrics*, *25*(2), 435.

Kenney, J. F., & Keeping, E. (1951). *Mathematics of statistics, vol. ii.* New York: D. Van Nostrand Co. Inc.

Kitaura, F.-S., Rodriguez-Torres, S., Chuang, C.-H., Zhao, C., Prada, F., Gil-Marin, H., … others (2016). The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: mock galaxy catalogues for the boss final data release. *Monthly Notices of the Royal Astronomical Society*, *456*(4), 4156–4173.

Kolb, E. (2018). *The early universe.* CRC press.

Lahav, O., Lilje, P. B., Primack, J. R., & Rees, M. J. (1991). Dynamical effects of the cosmological constant. *Monthly Notices of the Royal Astronomical Society*, *251*(1), 128–136.

Matsubara, T. (2008a). Nonlinear perturbation theory with halo bias and redshift-space distortions via the lagrangian picture. *Physical Review D*, *78*(8), 083519.

Matsubara, T. (2008b). Resumming cosmological perturbations via the lagrangian picture: One-loop results in real space and in redshift space. *Physical Review D*, *77*(6), 063530.

McDonald, P., & Roy, A. (2009). Clustering of dark matter tracers: generalizing bias for the coming era of precision lss. *Journal of Cosmology and Astroparticle Physics*, *2009*(08), 020.

Oliver, S., Rowan-Robinson, M., Broadhurst, T., McMahon, R., Saunders, W., Taylor, A., … Conrow, T. (1996). Large-scale structure in a new deep iras galaxy redshift survey. *Monthly Notices of the Royal Astronomical Society*, *280*(3), 673–688.

Peacock, J. (1997). The evolution of galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, *284*(4), 885–898.

Peacock, J., & Dodds, S. (1994). Reconstructing the linear power spectrum of cosmological mass fluctuations. *Monthly Notices of the Royal Astronomical Society*, *267*(4), 1020–1034.

Peebles, P. (1982). Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations.

Peebles, P., & Ratra, B. (1988). Cosmology with a time-variable cosmological'constant'. *Astrophysical Journal, Part 2-Letters to the Editor (ISSN 0004-637X), vol. 325, Feb. 15, 1988, p. L17-L20. NSF-supported research.*, *325*, L17–L20.

Peebles, P. J. E. (1993). *Principles of physical cosmology* (Vol. 27). Princeton university press.

Peebles, P. J. E. (2020). *The large-scale structure of the universe* (Vol. 98). Princeton university press.

Philcox, O. H., Ivanov, M. M., Simonović, M., & Zaldarriaga, M. (2020). Combining full-shape and bao analyses of galaxy power spectra: a 1.6% cmb-independent constraint on h0. *Journal of Cosmology and Astroparticle Physics*, *2020*(05), 032.

Reid, B. A., & White, M. (2011). Towards an accurate model of the redshift-space clustering of haloes in the quasi-linear regime. *Monthly Notices of the Royal Astronomical Society*, *417*(3), 1913–1927.

Robertson, H. P. (1935). Kinematics and world-structure. *Astrophysical Journal, vol. 82, p. 284*, *82*, 284.

Robertson, H. P. (1936). Kinematics and world-structure ii. *Astrophysical Journal, vol. 83, p. 187*, *83*, 187.

Rodríguez-Torres, S. A., Chuang, C.-H., Prada, F., Guo, H., Klypin, A., Behroozi, P., … others (2016). The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: modelling the clustering and halo occupation distribution of boss cmass galaxies in the final data release. *Monthly Notices of the Royal Astronomical Society*, *460*(2), 1173–1187.

Saito, S., Baldauf, T., Vlah, Z., Seljak, U., Okumura, T., & McDonald, P. (2014). Understanding higher-order nonlocal halo bias at large scales by combining the power spectrum with the bispectrum. *Physical Review D*, *90*(12), 123522.

Scoccimarro, R. (2004). Redshift-space distortions, pairwise velocities, and nonlinearities. *Physical Review D*, *70*(8), 083007.

Seljak, U., & Zaldarriaga, M. (1996). A line of sight approach to cosmic microwave background anisotropies. *arXiv preprint astro-ph/9603033*.

Seo, H.-J., & Eisenstein, D. J. (2003). Probing dark energy with baryonic acoustic oscillations from future large galaxy redshift surveys. *The Astrophysical Journal*, *598*(2), 720.

Shandarin, S. F., & Zeldovich, Y. B. (1989). The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium. *Reviews of Modern Physics*, *61*(2), 185.

Spergel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., … others (2015). Wide-field infrarred survey telescope-astrophysics focused telescope assets wfirst-afta 2015 report. *arXiv preprint arXiv:1503.03757*.

Taruya, A., Nishimichi, T., & Saito, S. (2010). Baryon acoustic oscillations in 2d: Modeling redshift-space power spectrum from perturbation theory. *Physical Review D*, *82*(6), 063522.

Tegmark, M. (1997). Measuring cosmological parameters with galaxy surveys. *Physical Review Letters*, *79*(20), 3806.

Tegmark, M., Taylor, A. N., & Heavens, A. F. (1997). Karhunen-loeve eigenvalue problems in cosmology: How should we tackle large data sets? *The Astrophysical Journal*, *480*(1), 22.

Wang, Y., Chuang, C.-H., & Hirata, C. M. (2013). Towards more realistic forecasting of dark energy constraints from galaxy redshift surveys. *Monthly Notices of the Royal Astronomical Society*, *430*(3), 2446–2453.

Wang, Y., Robberto, M., Dickinson, M., Hillenbrand, L. A., Fraser, W., Behroozi, P., … others (2019). Atlas probe: Breakthrough science of galaxy evolution, cosmology, milky way, and the solar system. *Publications of the Astronomical Society of Australia*, *36*, e015.

Wang, Y., Zhai, Z., Alavi, A., Massara, E., Pisani, A., Benson, A., … others (2022). The high latitude spectroscopic survey on the nancy grace roman space telescope. *The Astrophysical Journal*, *928*(1), 1.

Weinberg, S. (1989). The cosmological constant problem. *Reviews of modern physics*, *61*(1), 1.

Wilson, M. J., Peacock, J. A., Taylor, A. N., & de la Torre, S. (2017). Rapid modelling of the redshift-space power spectrum multipoles for a masked density field. *Monthly Notices of the Royal Astronomical Society*, *464*(3), 3121–3130.

Yamamoto, K., Nakamichi, M., Kamino, A., Bassett, B., & Nishioka, H. (2006). A measurement of the quadrupole power spectrum in the clustering of the 2df qso survey, pasj 58 (feb., 2006) 93–102. *arXiv preprint astro-ph/0505115*.

Zel'Dovich, Y. B. (1968). The cosmological constant and the theory of elementary particles. *Soviet Physics Uspekhi*, *11*(3), 381.