

---

# Optically selected galaxy clusters as a cosmological probe

Annalisa Mana

---



München, 2013



---

# Optically selected galaxy clusters as a cosmological probe

Annalisa Mana

---

Dissertation  
an der Fakultät für Physik  
der Ludwig-Maximilians-Universität  
München

vorgelegt von  
Annalisa Mana  
aus Fossano (CN), Italien

München, den 2. September 2013

Erstgutachter: Prof. Dr. Jochen Weller

Zweitgutachter: PD Dr. Klaus Dolag

Tag der mündlichen Prüfung: 9. Oktober 2013

*Ai miei genitori*



# Contents

<b>Zusammenfassung</b>	<b>x</b>
<b>Summary</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The homogeneous Universe . . . . .	1
1.1.1 Cosmological principle . . . . .	2
1.1.2 Friedmann-Lemaître-Robertson-Walker metric . . . . .	2
1.1.3 Einstein’s field equations . . . . .	3
1.1.4 Friedmann equations . . . . .	3
1.1.5 The critical density . . . . .	5
1.1.6 Energy density components . . . . .	5
1.1.7 Hubble’s law . . . . .	8
1.1.8 Cosmological distances . . . . .	10
1.2 The theory of structure formation . . . . .	11
1.2.1 Cosmic inflation . . . . .	11
1.2.2 Jeans gravitational instability . . . . .	12
1.2.3 Evolution of inhomogeneities . . . . .	12
1.2.4 Linearised perturbation equations . . . . .	14
1.2.5 Perturbation equations in an expanding Universe . . . . .	15
1.2.6 Power spectrum of density fluctuations . . . . .	17
1.3 Cosmological probes . . . . .	19
1.3.1 Supernovae Type Ia . . . . .	19
1.3.2 Baryon Acoustic Oscillations . . . . .	20
1.3.3 The Cosmic Microwave Background Radiation . . . . .	21
1.4 Galaxy Clusters . . . . .	25
1.4.1 History of galaxy clusters observations . . . . .	25
1.4.2 Main features, components and observables . . . . .	26

1.4.3	Cluster mass proxies . . . . .	28
1.4.4	Formation of galaxy clusters . . . . .	29
1.4.5	Clusters as cosmological probes . . . . .	29
1.5	$\Lambda$ CDM standard model . . . . .	31
1.5.1	Cosmological constraints from observations . . . . .	33
<b>2</b>	<b>Galaxy Clusters from theory side</b>	<b>37</b>
2.1	Cluster masses . . . . .	37
2.1.1	Definition . . . . .	38
2.1.2	Halo density distribution . . . . .	38
2.1.3	Weak Lensing signal . . . . .	39
2.2	Mass function . . . . .	43
2.2.1	Press-Schechter formalism . . . . .	44
2.2.2	N-body simulations calibration . . . . .	46
2.2.3	Cosmology dependence of the mass function . . . . .	48
2.3	Modelling cluster counts and total masses . . . . .	51
2.4	Clustering of clusters . . . . .	53
2.4.1	Halo bias . . . . .	53
2.4.2	Cluster power spectrum . . . . .	55
2.5	Primordial non-Gaussianity . . . . .	58
2.5.1	Definition of $f_{\text{NL}}$ parameter . . . . .	59
2.5.2	Modified mass function . . . . .	60
2.5.3	Modified bias . . . . .	62
<b>3</b>	<b>Observations, data and errors</b>	<b>65</b>
3.1	Multi-wavelength surveys of galaxy clusters . . . . .	65
3.1.1	X-ray surveys . . . . .	66
3.1.2	SZ surveys . . . . .	66
3.1.3	WL surveys . . . . .	67
3.1.4	Optical surveys . . . . .	67
3.1.5	Future surveys . . . . .	68
3.1.6	Cosmological constraints from cluster catalogues . . . . .	68
3.2	The Sloan Digital Sky Survey . . . . .	70
3.2.1	MaxBCG catalogue . . . . .	70
3.3	MaxBCG cluster number counts . . . . .	72
3.3.1	Cluster abundances . . . . .	72
3.3.2	Counts covariance matrix . . . . .	73
3.4	MaxBCG cluster masses . . . . .	76

3.4.1	Mean cluster masses from weak lensing observations . . . . .	76
3.4.2	Mass-richness scaling relation . . . . .	78
3.4.3	Cluster total masses . . . . .	80
3.5	MaxBCG cluster power spectrum . . . . .	81
3.5.1	Cluster power spectrum calculation . . . . .	81
3.5.2	Cluster power spectrum covariance matrix . . . . .	84
3.6	The counts-clustering off-diagonal covariance . . . . .	85
<b>4</b>	<b>Cosmological analysis</b>	<b>89</b>
4.1	Parameter estimation . . . . .	89
4.1.1	Bayes theorem . . . . .	90
4.1.2	Gaussian $\chi^2$ statistics . . . . .	90
4.1.3	$C$ -statistics . . . . .	92
4.1.4	Confidence regions and marginalisation . . . . .	92
4.2	Sampling methods . . . . .	93
4.2.1	Markov chains . . . . .	94
4.2.2	Monte Carlo methods . . . . .	94
4.2.3	MCMC methods . . . . .	95
4.3	The Cosmological Monte-Carlo . . . . .	98
4.4	Combined maxBCG analysis . . . . .	101
4.5	Results . . . . .	103
4.5.1	$\Omega_m - \sigma_8$ contours . . . . .	103
4.5.2	Scaling relation parameters contours . . . . .	106
4.5.3	$\log(10^{10} A_s) - \sigma_8$ contours . . . . .	110
4.5.4	$f_{\text{NL}} - \Omega_m$ and $f_{\text{NL}} - \sigma_8$ contours . . . . .	110
<b>5</b>	<b>Clusters-galaxies cross correlation</b>	<b>115</b>
5.1	Measurement by pixelization . . . . .	115
5.1.1	Pixelization method . . . . .	116
5.1.2	The mask and the catalogues . . . . .	117
5.2	Angular correlation function estimator . . . . .	118
5.3	Theoretical prediction . . . . .	118
5.4	Error estimates . . . . .	119
<b>6</b>	<b>Conclusions</b>	<b>121</b>
	<b>Acronyms</b>	<b>128</b>
	<b>Danksagung</b>	<b>146</b>



# Zusammenfassung

Aktuell werden großräumige Himmelsdurchmusterungen bei vielen verschiedenen Wellenlängen durchgeführt. Diese Beobachtungen dienen der Errichtung und Bestätigung eines kosmologischen Standardmodells für unser Universum. In den letzten Jahren wurden große Fortschritte in Theorie und Beobachtungen gemacht, um Galaxienhaufen als Testbett für die Kosmologie zu nutzen. Galaxienhaufen sind die größten gravitativ gebunden Strukturen und ihre Verteilung folgt der Entwicklung der großskaligen Struktur im Universum. Die Anzahldichte der Galaxienhaufen ist zudem sensitiv auf das zu Grunde gelegte kosmologische Modell. Durch die Beobachtung von Galaxienhaufen können die kosmologischen Parameter, zusätzlich zu anderen Messungen, eingeschränkt werden.

Diese Dissertation behandelt den wichtigen Beitrag von Galaxienhaufen zur Verifizierung des kosmologischen Standardmodells in einem von dunkler Materie und dunkler Energie dominierten Universum. Insbesondere untersuchen wir das Clustering von optisch selektierten Galaxienhaufen als zusätzlichen Parameter zu den üblichen kosmologischen Observablen. Das Clustering von Galaxienhaufen ergänzt die traditionellen Methoden der Zählung von Galaxienhaufen und der Vermessung von Masse-Observablen Relationen, weil die Analyse des Clusterings von Galaxien in den High-Peak, High-Bias Bereich vorangetrieben wird. Diese Methode ist ein mächtiges Werkzeug um bestehende Entartungen zu durchbrechen und genauere kosmologische Parameter zu gewinnen.

Als Erstes legen wir die wichtigsten theoretischen Grundlagen und Beobachtungen für das heutige Standardmodell der Kosmologie dar. Anschließend behandeln wir die grundlegenden Eigenschaften von Galaxienhaufen und insbesondere ihren Beitrag als Testbett für kosmologische Modelle.

Als nächstes entwickeln wir den theoretischen Rahmen für die Zählung von Galaxienhaufen und die Bestimmung des Leistungsspektrums. Wir überarbeiten die For-

mulierung und Kalibrierung der Halomassenfunktion, welche im Bereich hoher Massen von Galaxienhaufen bevölkert ist. Zusätzlich geben wir ein Rezept zur Modellierung des Leistungsspektrums von Galaxienhaufen mit dem Ort und der Rotverschiebung. Hierbei ist die Modellierung des schwach nicht-linearen Beitrags eingeschlossen und eine beliebige photometrische Glättung mit der Rotverschiebung ermöglicht. Zuletzt zeigen wir welchen Beitrag Galaxienhaufen bei der Beschränkung der Parameter für nicht Gauß-verteilte primordiale Anfangsbedingungen liefern können.

Anschließend widmen wir ein Kapitel der Präsentation unserer Basisdaten, dem Sloan Digital Sky Survey maxBCG Katalog. Wir beschreiben die Ableitung unserer Datensätze aus diesem Katalog von Galaxienhaufen und die entsprechenden dazugehörigen Fehlerabschätzungen. Speziell verwenden wir, jeweils mit den entsprechenden Kovarianzmatrizen, die Häufigkeit von Galaxienhaufen in verschiedenen Reichhaltigkeitsbereichen, Abschätzungen für die schwachen Linsenmassen und das Leistungsspektrum über Ort und Rotverschiebung. Zusätzlich, durch eine empirische Skalierungsrelation, setzen wir die Masse der Galaxienhaufen mit ihrer beobachteten Reichhaltigkeit in Verbindung und quantifizieren die Streuung der Daten.

Im nächsten Kapitel zeigen wir die Ergebnisse unserer Monte-Carlo-Markov-Ketten-Analyse und die daraus abgeleiteten Beschränkungen der kosmologischen Parameter. Mit dem maxBCG Datenset können wir sowohl die kosmologischen Parameter einschränken, als auch gleichzeitig die Masse-Observable-Relation vermessen. Wir finden, dass die Berücksichtigung des Leistungsspektrums eine  $\sim 50\%$  Verbesserung des Fehlers in der Fluktuationsamplitude  $\sigma_8$  und der Materiedichte  $\Omega_m$  ergibt. Für die anderen kosmologischen Parameter finden wir weniger signifikante Verbesserungen. Außerdem verwenden wir das mit WMAP7 gemessene Leistungsspektrum der kosmischen Hintergrundstrahlung, zusätzlich zu den Daten über Galaxienhaufen, und erhalten eine weitere Beschränkung der Vertrauensregionen. Zuletzt wenden wir unsere Methode auf Modelle des frühen Universums an, und bestimmen den Anteil der nicht Gauß-verteilten Fluktuationen des primordialen Dichtefelds (lokaler Typ). Unsere Ergebnisse sind konsistent mit den aktuellsten Beobachtungen.

Im letzten Kapitel präsentieren wir vorläufige Rechnungen zur Kreuzkorrelation zwischen Galaxienhaufen und Galaxien. Diese Rechnungen sind in der Lage die kosmologischen Modelle noch weiter einzuschränken.

Abschließend fassen wir unsere wichtigsten Ergebnisse zusammen und geben einen Ausblick auf mögliche weiterführende Forschungsprojekte.

# Summary

Multi-wavelength large-scale surveys are currently exploring the Universe and establishing the cosmological scenario with extraordinary accuracy. There has been recently a significant theoretical and observational progress in efforts to use clusters of galaxies as probes of cosmology and to test the physics of structure formation. Galaxy clusters are the most massive gravitationally bound systems in the Universe, which trace the evolution of the large-scale structure. Their number density and distribution are highly sensitive to the underlying cosmological model. The constraints on cosmological parameters which result from observations of galaxy clusters are complementary with those from other probes.

This dissertation examines the crucial role of clusters of galaxies in confirming the standard model of cosmology, with a Universe dominated by dark matter and dark energy. In particular, we examine the clustering of optically selected galaxy clusters as a useful addition to the common set of cosmological observables, because it extends galaxy clustering analysis to the high-peak, high-bias regime. The clustering of galaxy clusters complements the traditional cluster number counts and observable-mass relation analyses, significantly improving their constraining power by breaking existing calibration degeneracies.

We begin by introducing the fundamental principles at the base of the concordance cosmological model and the main observational evidence that support it. We then describe the main properties of galaxy clusters and their contribution as cosmological probes.

We then present the theoretical framework of galaxy clusters number counts and power spectrum. We revise the formulation and calibration of the halo mass function, whose high mass tail is populated by galaxy clusters. In addition to this, we give a prescription for modelling the cluster redshift space power spectrum, including an effective modelling of the weakly non-linear contribution and allowing for an

arbitrary photometric redshift smoothing. Some definitions concerning the study of non-Gaussian initial conditions are presented, because clusters can provide constraints on these models.

We dedicate a Chapter to the data we use in our analysis, namely the Sloan Digital Sky Survey maxBCG optical catalogue. We describe the data sets we derived from this large sample of clusters and the corresponding error estimates. Specifically, we employ the cluster abundances in richness bins, the weak-lensing mass estimates and the redshift-space power spectrum, with their respective covariance matrices. We also relate the cluster masses to the observable quantity (richness) by means of an empirical scaling relation and quantify its scatter.

In the next Chapter we present the results of our Monte Carlo Markov Chain analysis and the cosmological constraints obtained. With the maxBCG sample, we simultaneously constrain cosmological parameters and cross-calibrate the mass-observable relation. We find that the inclusion of the power spectrum typically brings a  $\sim 50\%$  improvement in the errors on the fluctuation amplitude  $\sigma_8$  and the matter density  $\Omega_m$ . Constraints on other parameters are also improved, even if less significantly. In addition to the cluster data, we also use the CMB power spectra from WMAP7, which further tighten the confidence regions. We also apply this method to constrain models of the early universe through the amount of primordial non-Gaussianity of the initial density perturbations (local type) obtaining consistent results with the latest constraints.

In the last Chapter, we introduce some preliminary calculations on the cross-correlation between clusters and galaxies, which can provide additional constraining power on cosmological models.

In conclusion, we summarise our main achievements and suggest possible future developments of research.

# Chapter 1

## Introduction

In this Chapter we introduce the theoretical and experimental research which has built the current concordance cosmological model. We first introduce the framework of a homogeneous Universe, based on Einstein equations for General relativity applied to the Universe as a whole. Secondly, we describe the basics of the evolution of primordial perturbations, which have led to the formation of the structures we see today. We then present the main cosmological probes which enable us to estimate cosmological parameters: the Supernovae Type Ia, the Baryon Acoustic Oscillations and the Cosmic Microwave Background. An entire Section is dedicated to the clusters of galaxies, their properties and their role in cosmology. Finally, we present the state-of-the-art of the constraints on  $\Lambda$ CDM parameters, obtained by combining galaxy clusters together with other cosmological probes.

### 1.1 The homogeneous Universe

In this Section, we introduce the mathematical background of modern cosmology based on Einstein's theory of gravity, in the assumption of a homogeneous and isotropic Universe. We describe how the Friedmann-Lemaître-Robertson-Walker metric, together with Einstein's field equations, leads to the Friedmann equations: the latter combine the description of the dynamics of the Universe, which depends on the energy density and pressure of the components, and the energy conservation of the components themselves.

### 1.1.1 Cosmological principle

On sufficiently large scales ( $> 100\text{Mpc}$ ), the Universe is *isotropic*, namely its properties are independent of the direction from which it is observed. This feature, combined with the cosmological principle which states that there is no preferred position in the Universe, implies that the Universe is also *homogeneous* on large scales. Among the four force interactions (electromagnetic, strong, weak, gravitational), only gravity plays a role on these scales.

### 1.1.2 Friedmann-Lemaître-Robertson-Walker metric

The effects of the gravitational force are described by the General Relativity (GR) framework (Einstein 1916). GR defines the space-time as a 4-dimensional manifold with a  $4 \times 4$  metric tensor  $g_{\mu\nu}$ , ten components of which are independent (time-time component  $g_{00}$ , three space-time components  $g_{0i}$  and six space-space components  $g_{ij}$ ). According to standard notation, Greek indices run from 0 to 3, where the 0-component is time, and refer to 4-d quantities (space-time), while Latin indices run from 1 to 3 and are used for 3-d (spatial) quantities. Considering the line element given by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu , \quad (1.1)$$

we can obtain the comoving spatial coordinates for fundamental observers by setting  $dx^i = 0$ , which implies  $g_{00} = c^2$ , where  $c$  is the speed of light. In addition to this, isotropy condition sets  $g_{0i} = 0$ . Thus, Eq. (1.1) can be simplified in terms of a time-dependent dimensionless scale factor  $a(t)$  and a 3-dimensional line element  $dl$  for an isotropic and homogeneous space, as

$$ds^2 = c^2 dt^2 - a^2(t) dl^2 . \quad (1.2)$$

Alternatively, the most common reformulation in comoving spatial polar coordinates  $(r, \theta, \phi)$  is

$$ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - Kr^2} + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \right] , \quad (1.3)$$

known as the *Friedmann-Lemaître-Robertson-Walker* metric (FLRW). Here  $r$  has a length dimension, while  $K$  has units of inverse squared length and represents the curvature scale of the Universe:  $K$  can assume values of  $0, +1, -1$  respectively in a flat (Euclidean), spherical (closed) or hyperbolic (open) model of Universe. Note that the curvature of space is equivalent to gravity: it is a measure of the energy

content in the Universe. The scale factor  $a(t)$  defines also the *deceleration parameter*

$$q = -\frac{\ddot{a} a}{\dot{a}^2}, \quad (1.4)$$

where  $\ddot{a} < 0$  ( $q > 0$ ) represents a decelerating Universe, while  $\ddot{a} > 0$  ( $q < 0$ ) an accelerating one.

### 1.1.3 Einstein's field equations

A step further, leads us to Einstein's field equations, which describe the dynamics of Eq. (1.3) by coupling the metric to the energy content of the Universe, as follows:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (1.5)$$

where  $G_{\mu\nu}$  is the Einstein tensor,  $G$  is the gravitational constant,  $R_{\mu\nu}$  the Ricci tensor and  $R$  the Ricci scalar. An additional term involving the so-called **cosmological constant**  $\Lambda$  was originally introduced by Einstein to achieve a static Universe, but then removed because of the evidence of an expanding Universe observed by Hubble (see 1.1.7).  $T_{\mu\nu}$  is the energy momentum tensor for the various component of the Universe, given by

$$T_{\mu\nu} = \left(\frac{P}{c^2} + \rho\right)u_\mu u_\nu - Pg_{\mu\nu}, \quad (1.6)$$

with the 4-velocity  $u^\mu = (c, 0, 0, 0)$ , where  $P$  is the pressure and  $\rho$  the mass density. From this definitions, it becomes clear how matter and space are related: matter tells space how to curve, while space tells matter how to move.

### 1.1.4 Friedmann equations

We assume hereafter that dots represent time derivatives, e.g.  $\dot{a} = da/dt$ . From Eq. (1.3), Christoffel symbols, Ricci tensor and Ricci scalar can be computed and inserted into Eq. (1.5). By solving then the time-time component  $G_{00}$  and the space-space components  $G_{ij}$  we obtain the so called **Friedmann equations** (FE), which describe the expansion of the Universe and its evolution in time:

$$\frac{\dot{a}^2}{a^2} + \frac{K c^2}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3}, \quad (1.7)$$

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

Here  $\Lambda$  has been reintroduced to explain the observed accelerated expansion of the Universe, being however still poorly motivated by particle physics (see 1.5). The pressure  $P$  is related to the mass density  $\rho$  by means of the perfect fluid equation of state  $P = w\rho c^2$ , where  $w$  is a constant dimensionless number and  $c$  is the speed of light, typically set to unity: so we do hereafter.

By differentiating Eq. (1.7) and inserting it in Eq. (1.8), the FE can be recast into a single equation, known as the *continuity equation*, which represents the mass-energy conservation:

$$\dot{\rho} + 3 \frac{\dot{a}}{a} (\rho + P) = 0 . \quad (1.9)$$

It is convenient to introduce the **Hubble parameter**, defined as

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} , \quad (1.10)$$

which represents the relative expansion rate of a homogeneous and isotropic FLRW Universe. For convention, the scale factor  $a(t)$  today ( $t = t_0$ ) is set to unity, i.e.  $a(t_0) = 1$ . With this definition, Eqs. (1.7) and (1.9) can be rearranged into the following:

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3} \left( \sum_i \rho_i + \rho_\Lambda \right) , \quad (1.11)$$

$$\sum_i \dot{\rho}_i + 3H \sum_i (\rho_i + P_i) = 0 . \quad (1.12)$$

We have introduced an energy density associated to the cosmological constant as

$$\rho_\Lambda \equiv \frac{\Lambda}{8\pi G} , \quad (1.13)$$

and we have replaced the density  $\rho$  with  $\sum_i \rho_i + \rho_\Lambda$ , where  $i$  refers to the various energy components we are considering. In particular,  $i = m$  for non-relativistic matter density (dust, or more precisely baryons and cold dark matter),  $i = r$  for radiation density (relativistic matter),  $i = \Lambda$  for the cosmological constant (or vacuum energy or dark energy, DE). Note that, even if the conservation of the total mass-energy holds because our Universe is an isolated system, there could be exchange/decay between different species.

### 1.1.5 The critical density

By demanding that the Universe is flat ( $K = 0$ ), Eq. (1.11) gives the definition of the critical density of the Universe:

$$\rho_c(t) = \frac{3H^2(t)}{8\pi G} , \quad (1.14)$$

and its value today is given by

$$\rho_{c,0} = \rho_c(t_0) = \frac{3H_0^2}{8\pi G} = 1.86 \times 10^{-29} h^2 \text{ g cm}^{-3} . \quad (1.15)$$

This also shows that the gravitational potential of a sphere of radius  $a(t)$  filled with matter at critical density is equivalent to its kinetic energy. The value of  $\rho_c$  today corresponds to approximately a galaxy mass per  $\text{Mpc}^3$ . The shape of the Universe and its finiteness depends on the balance between its expansion rate and the counter action of gravity, which is itself related to the matter density  $\rho_m$ :

- i) If  $\rho_m > \rho_c$ , the Universe is closed with positive curvature ( $K > 0$ ), like a sphere surface; it will eventually stop expanding and start collapsing in on itself (so-called Big Crunch).
- ii) If  $\rho_m < \rho_c$ , the Universe is open with negative curvature ( $K < 0$ ), like a saddle surface; it will expand forever.
- iii) If  $\rho_m = \rho_c$ , the Universe is flat with zero curvature ( $K = 0$ ), like a plane surface; it will expand forever, decreasing the rate of expansion. Recent measurements suggest that our Universe is most likely flat (see Section 1.3.3).

### 1.1.6 Energy density components

The energy density contents of the Universe are expressed by dimensionless parameters in units of the critical density  $\rho_c$ , i.e.

$$\Omega_i(t) \equiv \frac{\rho_i(t)}{\rho_c(t)} , \quad \Omega_{i,0} \equiv \frac{\rho_{i,0}}{\rho_{c,0}} , \quad (1.16)$$

where the label ‘0’ refers always to the present value. By combining Eqs. (1.13) and (1.14), the DE dimensionless parameter turns out to be:

$$\Omega_\Lambda(t) \equiv \frac{\rho_\Lambda(t)}{\rho_c(t)} = \frac{\Lambda}{3H^2(t)} , \quad \Omega_{\Lambda,0} = \frac{\Lambda}{3H_0^2} . \quad (1.17)$$

**Table 1.1:** Evolution of energy densities components of the Universe, classified by type, pressure, equation of state parameter and corresponding scale factor evolution.

Type	Pressure	$w$	$\rho(t)$	$a(t)$
non-relativistic matter	0	0	$\propto a^{-3}(t)$	$\propto t^{2/3}$
radiation	$\rho/3$	1/3	$\propto a^{-4}(t)$	$\propto t^{1/2}$
curvature	$-\rho/3$	-1/3	$\propto a^{-2}(t)$	$\propto t$
vacuum energy	$-\rho$	-1	$\propto a^0(t)$	$\propto \exp(Ht)$

Since  $\Omega \equiv \Omega_{\text{tot}} = \sum_i \Omega_i = 1$ , the curvature parameter is defined as:

$$\Omega_k(t) = 1 - \Omega_m(t) - \Omega_r(t) - \Omega_\Lambda(t) = -\frac{K c^2}{H^2(t) a^2(t)}, \quad \Omega_{k,0} = -\frac{K c^2}{H_0^2}. \quad (1.18)$$

With this notation, we can calculate explicitly solutions to FE for each density component of the Universe. Namely, if each component is separately conserved, the continuity equation (1.12) can be integrated (assuming  $K = 0$ ) to give

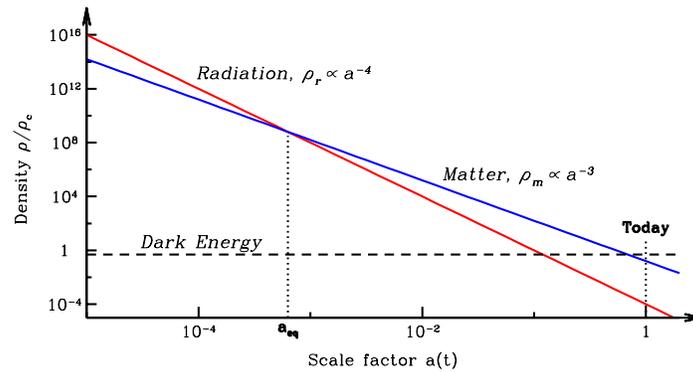
$$\rho_i \propto a^{-3(1+w_i)}, \quad a(t) \propto t^{\frac{2}{3(1+w_i)}}, \quad (1.19)$$

where the latter is obtained by combining with Eq. (1.11) and represents the evolution of the scale factor. Table 1.1 lists the behaviours of the various components of the Universe. Fig. 1.1 shows the evolution of  $\rho_m$ ,  $\rho_r$ ,  $\rho_\Lambda$  with respect to the cosmic size. Fig. 1.2 instead is representing the evolution of the scale factor in time for different models of the Universe: accelerating Universe, empty Universe, high/critical/low density Universe. We can finally reformulate in compact form Eq. (1.11) as

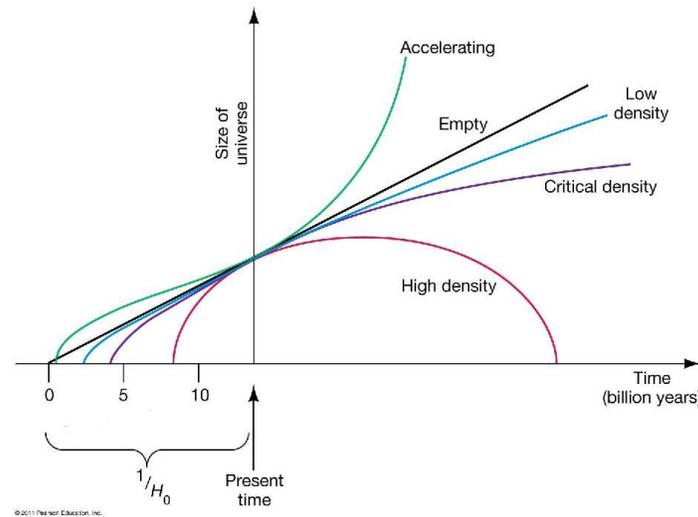
$$\begin{aligned} H^2(z) &= H_0^2 E^2(z), \\ E^2(z) &\equiv \Omega_m (1+z)^3 + \Omega_\Lambda (1+z)^{3(1+w)} + \Omega_k (1+z)^2 + \Omega_r (1+z)^4. \end{aligned} \quad (1.20)$$

The relevance of each energy component is evidently dependent on time: the Universe had a *radiation-dominated* epoch, up to the matter-radiation equality ( $\rho_r = \rho_m$ ) at  $a_{\text{eq}}$ , followed by a *matter-dominated* era. At late times ( $z \sim 0$ ), the DE component  $\rho_\Lambda$  starts to dominate, starting the *DE-dominated* epoch and driving the present day accelerated expansion of the Universe. Note that if we have a Universe and we populate it with ordinary particles, it will contract under the effect of gravity. If we instead populate this space with particles having a negative pressure (like DE), the space will expand, while GR would be still valid: the negative pressure is the cause

of the accelerated expansion of the Universe. As an example, if we throw an apple in a DE-dominated Universe, it will not fall, not because there is no gravity, but because while falling the space in between is expanding.



**Figure 1.1:** Log-log plot of energy density components of the Universe and their dependence on the scale factor  $a(t)$ : radiation energy density (red) scales as  $\propto a^{-4}$ , matter energy density (blue) as  $\propto a^{-3}$  and dark energy (black dashed) is constant with respect to  $a(t)$ . The scale factor is set to unity today ( $a_0 = 1$ ). The present value of the ratio  $\rho/\rho_c = \Omega$  is unity (i.e.  $\rho_0 = \rho_c$ ,  $\Omega_0 = 1$ ).



**Figure 1.2:** Scale factor as a function of time for different models of the Universe: accelerating Universe, empty Universe, high/critical/low density Universe. Credit: Pearson Education, Inc. 2011, <http://physics.uoregon.edu/>

### 1.1.7 Hubble's law

The discovery that the Universe was not static but expanding, by the astronomer Hubble (1929), can be considered as the dawn of observational cosmology. The phenomenon of galaxies appearing to recede from us at a rate proportional to their distance from Earth, can be quantified in terms of redshift of a galaxy spectrum. In fact, the intrinsic wavelength of light is stretched linearly, due to the expansion of the Universe, i.e.  $\lambda(t) \propto a(t)$ . More precisely, we can define the **cosmological redshift** (or simply redshift)  $z$  for relatively nearby objects as

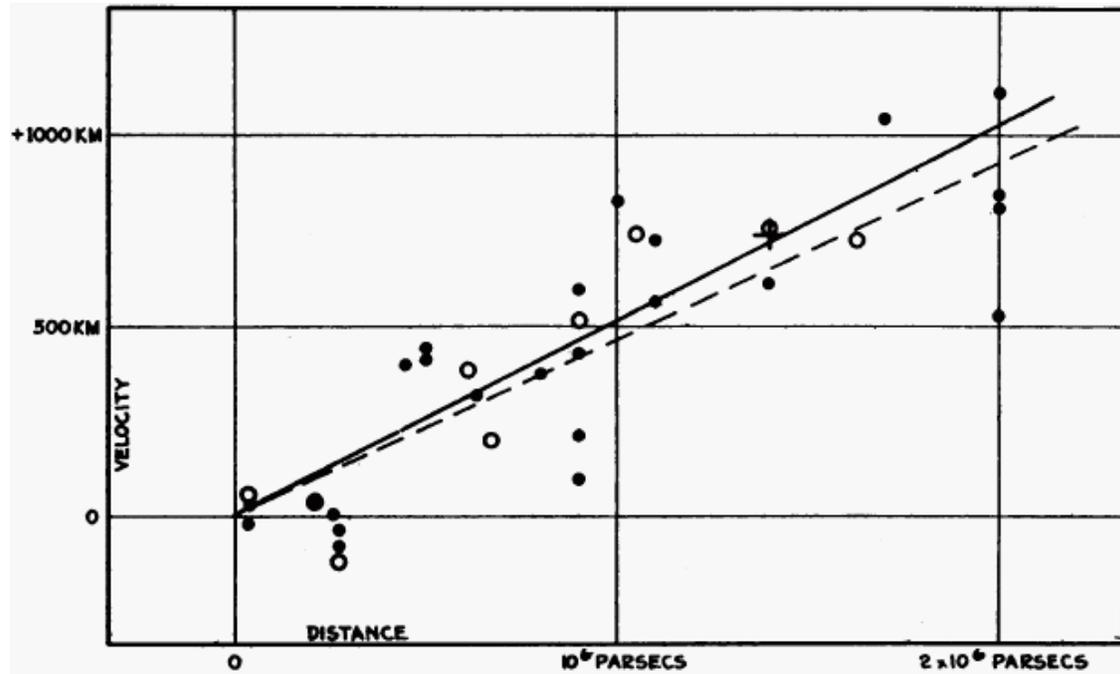
$$z \equiv \frac{\lambda_{\text{obs}}}{\lambda_{\text{em}}} - 1 = \frac{\nu_{\text{em}}}{\nu_{\text{obs}}} - 1 = \frac{a(t_{\text{obs}})}{a(t_{\text{em}})} - 1, \quad (1.21)$$

where  $\lambda_{\text{obs}}$  and  $\lambda_{\text{em}}$  are the observed and the emitted wavelengths, respectively, while  $\nu_{\text{obs}}$  and  $\nu_{\text{em}}$  are the observed and the emitted frequencies, respectively. If we locate the observer at today, as  $a_0 = 1$ , we obtain the relation  $a = 1/(1+z)$ .

Hubble's observations revealed that the light from galaxies which move away from Earth is shifted toward the red, while the light from galaxies which move toward Earth is shifted to the blue. This implies that the more distant a galaxy is, the longer (redder) is the observed wavelength of its emitted light, the greater its redshift is, and the faster it is moving away from Earth. The mathematical expression for Hubble's law is

$$v = H_0 D, \quad (1.22)$$

where  $v$  is the galaxy radial recession velocity in km/s,  $D$  is the distance between galaxy and Earth in Mpc and  $H_0 \equiv H(t_0)$  is the value of the **Hubble constant** at present time in  $\text{km s}^{-1} \text{Mpc}^{-1}$ . The Hubble constant is a scaling factor representing the today expansion rate of the Universe. It can be also written as  $H_0 = 100 h \text{ km s}^{-1} \text{Mpc}^{-1}$ , where  $h$  is a dimensionless number. In Fig. 1.3 we show the original Hubble diagram, displaying the velocities of distant galaxies (in km/s) with respect to the distance (in parsec). Filled points, whose best fit is the solid line, are corrected for the motion of the Sun, while open points, whose best fit is the dashed line, are not corrected for this effect. The slope in the diagram is the Hubble constant itself. After Hubble's discovery, it was thought that gravity acting on matter was slowing the expansion of the Universe. In 1998, however, a campaign of observations of distant Supernovae Ia, carried out with the Hubble Space Telescope (HST) revealed that the expansion of the Universe was instead accelerating, giving hints on an unknown component of DE (Garnavich et al. 1998; Schmidt et al. 1998; Riess et al. 1998b,a; Perlmutter et al. 1999).



**Figure 1.3:** The original Hubble diagram (Hubble 1929). Velocities of distant galaxies in km/s are plotted with respect to the distance in parsec. Solid line is the best fit to the filled points, which are corrected for the motion of the Sun. Dashed line is the best fit to the open points, which are not corrected for this effect. As velocity increases linearly with distance, there is an evident slope, i.e. the Hubble constant. Credit: Hubble (1929).

Finally, the inverse of the Hubble constant defines the **Hubble time**, i.e. an estimate of the age of the Universe, which assumes the following value from the latest Planck data (Planck Collaboration et al. 2013b):

$$t_H = \frac{1}{H_0} = 13.813 \pm 0.058 \times 10^9 \text{yr} \text{ (68\%c.l.)} . \quad (1.23)$$

The **Hubble radius** or *length* is instead the speed of light times the Hubble time:

$$r_H = \frac{c}{H_0} = 3.01 \times 10^3 h^{-1} \text{Mpc} = 9.30 \times 10^{25} h^{-1} \text{m} . \quad (1.24)$$

### 1.1.8 Cosmological distances

The expansion of space-time forces us to generalise the Euclidean concepts of distances. In a flat Universe, photons travelling to us satisfy  $c dt = a(t) dr$ . Thus, the comoving radial distance can be defined as

$$r = c \int_t^{t_0} \frac{dt'}{a(t')} = c \int_0^z \frac{dz'}{a_0 H(z')}, \quad (1.25)$$

where  $H(z)$  is given by Eq. (1.20).

The angular diameter distance  $D_A$  is given by the scale factor times the comoving radial distance

$$D_A(z) = a(z) r = \frac{c}{1+z} \int_0^z \frac{dz'}{H(z')}. \quad (1.26)$$

This distance will be used in the Alcock-Paczynski effect for the cluster power spectrum in our analysis (see Eq. 2.47).

The luminosity distance  $D_L$ , instead, links the bolometric observable flux  $F$ , namely the energy per unit time per unit area from the source to the observer, and bolometric intrinsic luminosity  $L$  of the source:

$$D_L = \sqrt{\frac{L}{4\pi F}}. \quad (1.27)$$

This means that farther objects appear dimmer. By observing the apparent luminosity of light sources, whose intrinsic luminosity is known (*standard candles*), we can infer the luminosity distance. Moreover, in a FLRW metric and assuming that light travels on null geodesics, the following relation holds

$$D_L(z) = a_0 (1+z) r = (1+z)^2 D_A(z). \quad (1.28)$$

This method has been applied to Supernovae Type Ia, which we will introduce in Section 1.3.1.

## 1.2 The theory of structure formation

This Section is entirely dedicated to the process of cosmic structure formation. We first introduce cosmic inflation and its importance in solving the horizon, flatness and magnetic monopoles problems. Then, we describe the Jeans gravitational instability theory, which is at the base of the structure formation scenario. We also present the evolution of density inhomogeneities of cold dark matter and baryons by means of linearised perturbation equations and their generalisation to an expanding Universe. Finally, we introduce the power spectrum of density fluctuations as a fundamental tool for the statistical description of the large-scale structures.

### 1.2.1 Cosmic inflation

Another key element of our current understanding of structure formation in the Universe is **cosmic inflation** (Guth 1981; Sato 1981). The decelerated expansion of the standard Big Bang scenario during the radiation-dominated and matter-dominated eras is not sufficient to solve few questions. One of these questions is known as the *horizon* problem: it asks why the Universe had almost the same temperature across the whole sky at  $t = 300,000$  yrs (as seen from the last scattering surface), when regions could not have been in causal contact due to the finite speed of light. Another problem is related to the *flatness* of the Universe: even if  $\Omega$  should shift away from unity in an expanding Universe, present observations suggest that  $\Omega \sim 1$  (i.e. the current density of the Universe is observed to be very close to this critical value) and thus was most likely very close to unity in the past too. This implies an accurate fine-tuning of initial conditions, otherwise the Universe would have already collapsed or expanded too fast to form structures. Finally, the *magnetic monopoles* problem refers to the observed absence of magnetic monopoles in the present Universe: this contradicts the Grand Unified Theories, unifying electromagnetic, strong and weak forces, which predict magnetic monopoles of about the same abundance as protons in the early Universe and thus expected to be present today. Therefore, a rapid epoch of accelerated, exponential expansion in the early Universe of a factor  $\sim 10^{26}$  in size ( $\sim 10^{78}$  in volume), from  $t = 10^{-33}$  to  $t = 10^{-30}$  s after the Big Bang, driven probably by a negative-pressure vacuum energy, is theorised to address these questions. This means that all the observable Universe is originated in a small causally-connected region. As a consequences of inflation, the Universe grows up so quickly that there is no time for the homogeneity to be broken, justifying the smooth temperature distribution of the last scattering surface. Furthermore, the quick enormous expansion can force  $\Omega_k$  down to zero, or around it, allowing for a

tiny growth up to the currently observed value. Finally, despite the huge number of magnetic monopoles in the early Universe, the chances of observing even one are infinitesimally small in such an extended Universe. After setting the initial conditions of the Universe, cosmic inflation amplifies also the tiny quantum fluctuations already present before inflation, generating the seeds of cosmic structures which then have been evolving in time till today.

### 1.2.2 Jeans gravitational instability

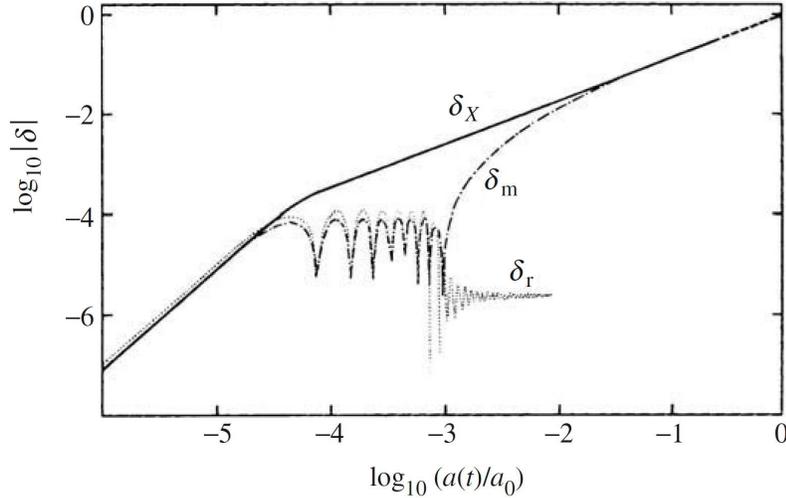
**Jeans gravitational instability** studies are the starting point of our standard cosmic structure formation scenario. Jeans (1902) investigated the gravitational instability in clouds of gas to explain how stars and planets form. It was proved that, in a static homogeneous and isotropic background fluid, small perturbations in density and velocity can occur and evolve in time. In particular, if pressure is negligible, an overdense region tends to become denser because it attracts material from the surroundings, and eventually collapse into a gravitational bounded system. The gravitational Jeans instability which causes the region to collapse can be quantified in terms of the **Jeans length** of a fluid

$$\lambda_J = c_s \left( \frac{\pi}{G\rho} \right)^{1/2}, \quad (1.29)$$

which represents the length scale to exceed (i.e.  $\lambda > \lambda_J$ ) for the fluctuations to grow, where  $G$  is the gravitational constant,  $c_s$  the speed of sound and  $\rho$  the background fluid mean density. In the case of  $\lambda < \lambda_J$ , instead, fluctuations oscillate as acoustic waves. This simple theory can be generalised to an expanding cosmological model, with the additional complications of a matter density which decreases with time ( $\rho \sim G^{-1}t^{-2}$ ) and a slower growing of perturbations, alternatively in accreting and decaying modes.

### 1.2.3 Evolution of inhomogeneities

Before the decoupling of photons from baryons, radiation pressure and gravitational collapse of matter competed with each other, producing oscillations of the baryon-photon plasma, known as Baryon Acoustic Oscillations (BAO, see Section 1.3.2). On the other hand, cold dark matter (CDM) inhomogeneities, by means of gravitational interaction only, could start to condensate and grow: the gravitational Jeans instability for dark matter (DM) particles allowed compact structures to form because it was not constrained by any force, such as radiation pressure. After recombination,



**Figure 1.4:** Evolution of density perturbations in cold dark matter  $\delta_X$ , baryonic matter  $\delta_m$  and radiation  $\delta_r$  components, at mass scale  $M \sim 10^{15} M_\odot$ , in a Universe with  $\Omega = 1$ ,  $h = 0.5$ . Credit: Coles & Lucchin (1995).

when baryons decoupled from radiation, the first local overdensities in the baryon density field could form and accrete in amplitude, because no radiative pressure could counteract the gravitational collapse anymore. The baryonic matter collapsed directly into the potential wells already created by the DM, forming structures much faster than it would have done without the presence of DM itself. Without DM, in fact, stars and galaxies formation would have occurred much later in the Universe than is observed. Even if, at this point, we can treat the evolution of perturbations in baryons and DM with the same physics description, the power spectrum of fluctuations in baryonic matter and DM are quite different. In particular, BAO dominate the baryon density power spectrum at early times, while their signature is almost negligible in the DM distribution. We will describe mathematically the evolution of perturbation in Sections 1.2.4 and 1.2.5. Fig. 1.4 exhibits the evolution of density perturbations in CDM  $\delta_X$ , baryonic matter  $\delta_m$  and radiation  $\delta_r$  components, at a mass scale of  $M \sim 10^{15} M_\odot$ , in a Universe with  $\Omega = 1$ ,  $h = 0.5$ . It is clearly shown how the perturbation in the baryon-photon fluid oscillates before decoupling, and how it grows rapidly to match the dominant dark matter perturbation, after decoupling.

### 1.2.4 Linearised perturbation equations

In order to describe quantitatively the evolution of the density perturbations, it is useful to introduce the dimensionless density contrast as

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (1.30)$$

where  $\rho(\vec{x}, t)$  is the matter density field as function of comoving coordinate  $\vec{x}$  and time  $t$ , while  $\bar{\rho}(t)$  is the average density of the Universe as a function of time  $t$ . The most common representation of this quantity is however in Fourier space:

$$\delta(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \hat{\delta}(\vec{k}, t) e^{-i\vec{k}\cdot\vec{x}}, \quad \hat{\delta}(\vec{k}, t) = \int \frac{d^3x}{(2\pi)^3} \delta(\vec{x}, t) e^{i\vec{k}\cdot\vec{x}}. \quad (1.31)$$

It is also useful to define the *power spectrum*  $P(k)$  and its dimensionless expression  $\Delta(k)$  as

$$P(k) \equiv \langle |\hat{\delta}(\vec{k}, t)|^2 \rangle, \quad \Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2}. \quad (1.32)$$

If  $\delta(\vec{x}, t)$  is a Gaussian random field, then  $P(k)$  completely describes the statistics of the perturbations field. We will examine the properties of this useful statistical tool in Section 1.2.6.

If we assume that matter (DM + baryonic) is accreting only via gravitational interactions, we can use the ideal fluid approximation. The evolution of primordial fluctuations can be described in the linear regime, if perturbations are small, i.e.  $|\delta(\vec{x})| \ll 1$ . The set of linearised fluid equations is the following:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{u}) &= 0 && \text{Continuity equation (conservation of mass)} \\ \frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \vec{\nabla}) \vec{u} + \frac{1}{\rho} \vec{\nabla} P + \vec{\nabla} \Phi &= 0 && \text{Euler's equation (conservation of momentum)} \\ \nabla^2 \Phi - 4\pi G \rho &= 0 && \text{Poisson's equation,} \end{aligned}$$

where  $\rho = \rho(\vec{x}, t)$  is the density,  $\vec{u}(\vec{x}, t)$  is the flow velocity,  $\vec{\nabla}$  is the gradient of a scalar field or the divergence of a vector field with respect to the spatial component,  $\nabla^2$  is the Laplace operator (i.e. the divergence of the gradient),  $\Phi$  is the gravitational potential. The static solution of this system of equations is  $\vec{u}_0 = 0$ ,  $\rho_0$  constant. The latter can be perturbed as  $\rho = \rho_0 + \delta\rho$ ,  $P = P_0 + \delta P$ ,  $\vec{u} = \vec{u}_0 + \delta\vec{u}$ ,  $\Phi = \Phi_0 + \delta\Phi$ : then the system can be recast into a single second order differential equation in  $\delta\rho$  and solved.

### 1.2.5 Perturbation equations in an expanding Universe

If we want to extend this framework to an expanding Universe, then the above equations expressed in  $\delta$  (see Eq. 1.30) would be the following:

$$\begin{aligned}\frac{\partial\delta}{\partial t} + \frac{1}{a}\vec{\nabla}\cdot[(1+\delta)\vec{u}] &= 0 \\ \frac{\partial\vec{u}}{\partial t} + \frac{\dot{a}}{a}\vec{u} + \frac{1}{a}(\vec{u}\cdot\vec{\nabla})\vec{u} + \frac{1}{a}\vec{\nabla}\Phi &= 0 \\ \nabla^2\Phi - 4\pi G\rho a^2\delta &= 0.\end{aligned}$$

By assuming small perturbations and keeping only linear terms in  $\delta$ , we obtain the following linearised set of equations:

$$\begin{aligned}\frac{\partial\delta}{\partial t} + \frac{1}{a}\vec{\nabla}\cdot\vec{u} &= 0 \\ \frac{\partial\vec{u}}{\partial t} + \frac{\dot{a}}{a}\vec{u} + \frac{1}{a}\vec{\nabla}\Phi &= 0 \\ \nabla^2\Phi - 4\pi G\rho a^2\delta &= 0.\end{aligned}$$

The time evolution of linear matter density perturbations  $\delta$  in an expanding background fluid, neglecting radiation and DE contributions, can be finally reformulated in a single equation as

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\bar{\rho}\delta + \frac{c_s^2\nabla^2\delta}{a}. \quad (1.33)$$

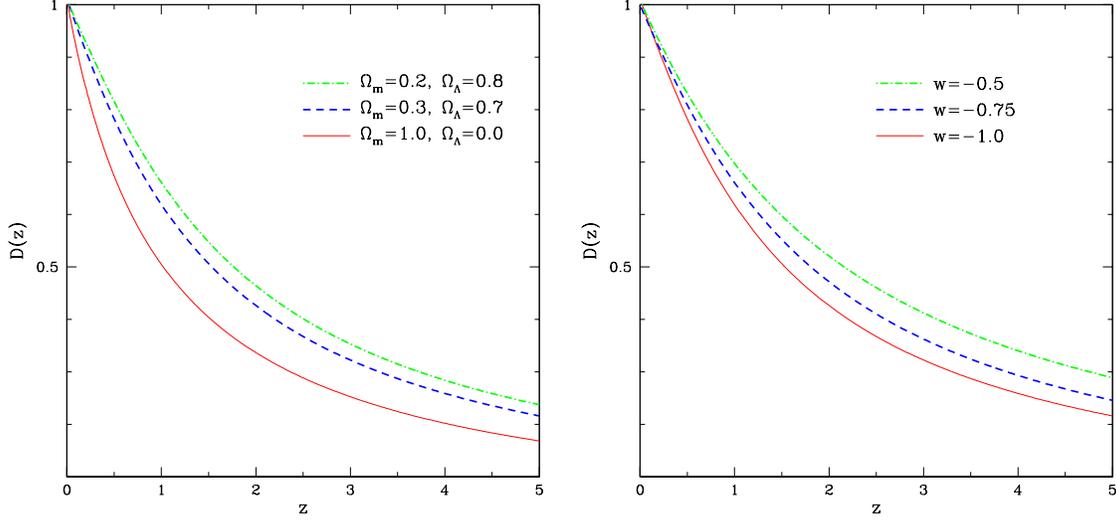
This represents a damped wave equation: on the left-hand side, the *drag* term including the Hubble parameter suppresses the growth of the perturbation; on the right-hand side, gravity and pressure act one against the other. Here  $c_s = \sqrt{\partial P/\partial\rho}$  is the adiabatic sound speed. Solution to Eq. (1.33) are given as

$$\delta(\vec{k}, t) = \delta_+(\vec{k}, t)D_+(t) + \delta_-(\vec{k}, t)D_-(t), \quad (1.34)$$

given that  $D_+$  and  $D_-$  correspond to the fluctuations growing and decaying modes, respectively. In the case of a collisionless fluid in a flat Universe with  $\Omega_m < 1$ , the growing mode is given by:

$$D_+(z) = \frac{5}{2}\Omega_m E(z) \int_z^\infty \frac{1+u}{E^3(u)} du, \quad (1.35)$$

where  $E(z)$  is defined as in Eq. (1.20). This redshift dependent quantity is very sensitive to cosmology and complementary to other probes such as luminosity and angular



**Figure 1.5:** The growth factor  $D(z)$  theory curve for different values of cosmological parameters, normalised at 1 for  $z = 0$ . Left panel: at fixed  $w = -1.0$ , we plot  $D(z)$  for  $\Omega_m = 0.2, \Omega_\Lambda = 0.8$  in green; for  $\Omega_m = 0.3, \Omega_\Lambda = 0.7$  in blue; for  $\Omega_m = 1.0, \Omega_\Lambda = 0.0$  in red. Right panel: at fixed  $\Omega_m = 0.3, \Omega_\Lambda = 0.7$ , we show  $D(z)$  for  $w = -0.5$  in green,  $w = -0.75$  in blue,  $w = -1.0$  in red.

distances. Fig. 1.5 shows the sensitivity of the growth factor  $D_+(z)$  to the variation in  $\Omega_m, \Omega_\Lambda$  and  $w$ . In the left panel we show the growth function at fixed  $w = -1.0$ , for varying  $\Omega_m = \{0.2, 0.3, 1.0\}$ ,  $\Omega_\Lambda = \{0.8, 0.7, 0.0\}$  in green, blue and red curves respectively: the more matter is present, the steepest the curve is, meaning that the formation of structure is more rapid but occurs later (at lower  $z$ ). A similar effect can be seen for a decreasing value of  $w$ , in the right panel: here we show, in fact, the growth function at fixed  $\Omega_m = 0.3, \Omega_\Lambda = 0.7$ , for varying  $w = \{-0.5, -0.75, -1.0\}$  in green, blue and red lines respectively.

We can model the growth of structures as a function of the cosmic time by parametrizing the linear growth rate of density perturbations on large scales,  $f(a)$ , as follows:

$$f(a) \equiv \frac{d \ln \delta}{d \ln a} \simeq \Omega_m^\gamma(a), \quad (1.36)$$

where  $\Omega_m(a) = \Omega_m a^{-3} E(a)^{-2}$ ,  $\delta \equiv \delta \rho_m / \rho_m$  is the ratio of the comoving matter density fluctuations with respect to the mean density of the Universe and  $\gamma$  is called

**growth index** (see Peebles 1980, 1993; Linder 2005). The growth index allows us to distinguish GR from modified gravity theories which can mimic the expansion history of the  $\Lambda$ CDM model. Several of these models predict a time and scale dependent growth index, i.e.  $\gamma(a, k)$ . It was obtained  $\gamma = 6/11 \simeq 0.55$  for  $\Lambda$ CDM (Wang & Steinhardt 1998), and, for example,  $\gamma = 11/16$  in the Dvali et al. (2000) (DGP) braneworld modified gravity model (Linder & Cahn 2007).

Non linear interactions between baryonic matter, dark matter and dark energy become important when perturbations are not small anymore, i.e.  $|\delta(\vec{x})| \sim 1$ . The complex evolution of structure formation in this regime can be studied only with numerical simulations (Kuhlen et al. 2012), such as the Millennium Simulation (Springel et al. 2005), and Millennium XXL (Angulo et al. 2012, 2013). Note also that for perturbations on large scales, the simple Newtonian approach we introduced is not valid anymore and we should perturb FLRW metric as  $g_{\mu\nu} = g_{\mu\nu}^0 + h_{\mu\nu}$ , where  $|h_{\mu\nu}| \ll g_{\mu\nu}$ .

### 1.2.6 Power spectrum of density fluctuations

The power spectrum of density fluctuations is an extremely useful tool for the statistical description of the large-scale structures in general. A correlated quantity is  $\sigma^2(M, z)$ , namely the variance in mass of the density fluctuation field, within identical volume elements corresponding to  $1/k$  length scale, in a linear evolution regime. To obtain an expression of  $\sigma^2(M, z)$ , we need to define the filtered density contrast by convolving it with a window function  $W_R$  as

$$\delta_R(\vec{x}, t) = \delta_M(\vec{x}, t) = \int d^3x' \delta(\vec{x}') W_R(|\vec{x} - \vec{x}'|), \quad (1.37)$$

where  $R = R(M) = (3M/4\pi\bar{\rho}_m)^{1/3}$  is the characteristic length scale below which we smooth out all the fluctuations, and  $W_R(x)$  is usually the spherical top-hat window function in real space

$$W_R(|\vec{x} - \vec{x}'|) = \begin{cases} 1, & \text{if } |\vec{x} - \vec{x}'| < R, \\ 0, & \text{otherwise.} \end{cases} \quad (1.38)$$

This leads to the definition of the variance of the density field:

$$\sigma^2(M, z) \equiv \sigma_M^2(z) \equiv \sigma_R^2(z) = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P(k, z) |\widehat{W}_R(k)|^2, \quad (1.39)$$

where  $\widehat{W}_R(k)$  is the Fourier Transformation (FT) of the top-hat filter function of  $R$ , given by

$$\widehat{W}_R(k) = \frac{3[\sin(kR) - kR \cos(kR)]}{(kR)^3}. \quad (1.40)$$

Here  $P(k, z)$  is the power spectrum of linear, independently evolving fluctuations, which can be expressed as

$$P(k, z) = P_{\text{in}}(k) T^2(k) D^2(z), \quad (1.41)$$

where  $P_{\text{in}}(k)$  is the primordial power spectrum,  $T(k)$  is known as the *transfer function* (Eisenstein & Hu 1998) and  $D(z)$  is the linear growing mode defined in Eq. (1.35). The power spectrum at primordial times is usually described by a power law as  $P_{\text{in}}(k) = A_s k^{n_s}$ . Here  $n_s$  is the primordial scalar spectral index, which is observed to be close to unity (Spergel et al. 2007), in agreement with inflationary models predictions (Harrison 1970; Zeldovich 1972), and  $A_s$  is the amplitude of the primordial power spectrum, which is by definition related to  $\sigma^2$ . The transfer function is carrying all scale-imprinting effects that modified the linear form of the primordial power spectrum during its evolution to the present day:

$$T(k) = \frac{\delta_k(z=0)}{\delta_k(z)D(z)}, \quad (1.42)$$

$z$  being here large enough for  $\delta_k(z)$  to mimic the original power spectrum. The scale  $k_{\text{eq}} = (2\Omega_m H_0^2 z_{\text{eq}})^{1/2}$  in the CDM model, which corresponds to the transition between the radiation-dominated phase and the matter-dominated epoch, breaks the transfer function shape: perturbations on small scales ( $k > k_{\text{eq}}$ ) are suppressed in amplitude (*Meszaros effect*), while they can grow on larger scales ( $k < k_{\text{eq}}$ ). Effectively,  $T(k) \propto k^{-2}$  for  $k \gg k_{\text{eq}}$  and  $T(k) \sim 1$  for  $k \ll k_{\text{eq}}$ . As a consequence, for higher  $\Omega_m$  perturbations are suppressed earlier and the peak of the matter power spectrum shifts to higher  $k$ . On small scales, dissipative processes from baryon-photon interactions leave their imprint (*Silk damping*, Silk 1968): the more baryons, the more damped the transfer function is. Finally, BAO appear in the transfer function as well: the position and amplitude of the wiggles depend on the amount of baryons and DM.

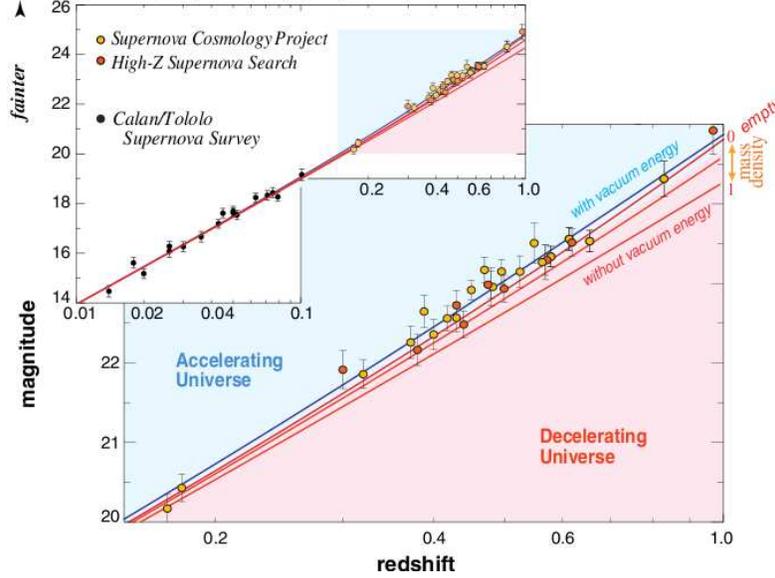
If the features of the power spectrum can be theoretically inferred, the normalisation has to be determined observationally. The latter is generally parametrised by the quantity  $\sigma_8$ , which is the variance defined in Eq. (1.39) having comoving radius  $R = 8 h^{-1} \text{Mpc}$ . This was motivated by Davis & Peebles (1983) results on early galaxy surveys, who found the variance of the galaxy number density on this scale to be about unity.

## 1.3 Cosmological probes

In this Section we summarise the main cosmological probes which enable us to measure cosmological parameters. Here we introduce the Supernovae Type Ia (SNIa), the Baryon Acoustic Oscillations (BAO) and the Cosmic Microwave Background (CMB), which respectively place constraints on  $\Omega_\Lambda$ ,  $\Omega_b$  and  $\Omega_k$ . The constraining power of a single cosmological probe is generally too weak to constrain simultaneously all cosmological parameters. However, by combining different probes, it is possible to place tight constraints on the cosmological parameters, to break degeneracies between them and reduce uncertainties. We will see the results obtained from the combination of these probes together with clusters of galaxies in Section 1.5.1.

### 1.3.1 Supernovae Type Ia

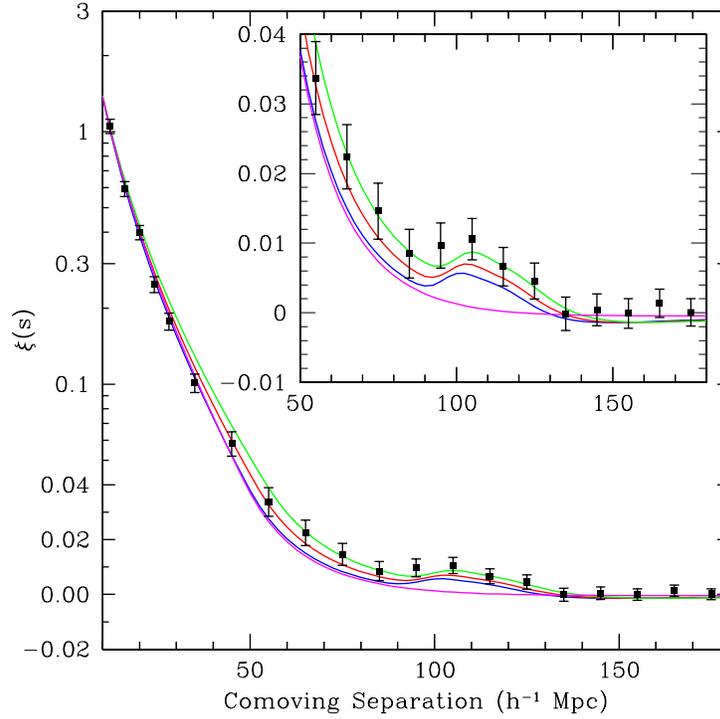
SNIa are thought to be the result of white dwarfs which accrete and explode upon reaching the Chandrasekhar mass limit. This process enable the Supernovae to have a characteristic intrinsic luminosity, which can be standardised empirically: thus, SNIa are potentially independent distance estimators, i.e. *standard candles*. Other types of Supernovae, instead, have more complex collapsing processes and different intrinsic luminosities, being thus less standardisable. In Fig. 1.6 we show the SNIa observations from the Supernova Cosmology Project and High-Z Supernova Search (high  $z$ ) and from Calan/Tololo Supernova Survey (Hamuy et al. 1993, 1995) (low  $z$ ), on a logarithmic redshift scale. The apparent magnitude of SNIa is proportional to the luminosity distance, which is associated to the redshift of the host galaxy. The measured luminosity distance can be compared to the theoretical prediction (see Eq. 1.27) to constrain  $\Omega_m$ ,  $\Omega_\Lambda$  and discriminate between different cosmological scenarios. In fact, here the SNIa observations are compared to few cosmological model: data are strongly inconsistent with  $\Lambda = 0$  models and favour models with  $\Lambda > 0$  (Perlmutter 2003). While high-redshift SNIa reveal that the Universe is now accelerating (Riess et al. 1998a; Perlmutter et al. 1999), nearby ones provide the most precise measurements of the present expansion rate,  $H_0$ . The most precise measurement of  $H_0$  comes from the luminosity calibration of nearby SNe Ia through Hubble Space Telescope observations of Cepheids in their host galaxies, carried on by the SH0ES program. With this method, Riess et al. (2011) obtained a value of the Hubble constant of  $H_0 = (73.8 \pm 2.4) \text{ km s}^{-1} \text{ Mpc}^{-1}$  (68% c.l.), including systematics. The combination of this result alone with the WMAP DR7 (see Section 1.3.3) constraints yields  $w = 1.08 \pm 0.10$  (68% c.l.).



**Figure 1.6:** Hubble diagram from SNIa, showing the apparent magnitude on a logarithmic redshift scale for nearby (Calan/Tololo Supernova Survey) and distant (Supernova Cosmology Project, High-Z Supernova Search) Type Ia Supernovae. At redshifts beyond  $z = 0.1$ , the cosmological predictions start to diverge, depending on the assumed cosmic densities. The red curves represent models with zero vacuum energy and mass densities from the critical density down to zero. The best fit (blue line) assumes a mass density of about  $\rho_c/3$  plus a vacuum energy density of about  $2\rho_c/3$ , implying an accelerating cosmic expansion. Credit: Perlmutter (2003).

### 1.3.2 Baryon Acoustic Oscillations

Prior to the decoupling phase of the Universe, when photons were scattered by electrons through Thomson scattering, radiation pressure opposed the gravitational collapse of matter, generating pressure waves, known as BAO. These oscillations left a signature in the distribution of matter on very large scales and in the features of CMB anisotropies (Hu & Dodelson 2002). This signature has been measured by galaxy surveys as an overdensity of galaxies at a characteristic comoving scale of  $100h^{-1}$  Mpc. For example, Fig. 1.7 shows the statistically significant bump on this comoving scale, revealed only by models which include baryons. This was obtained by Eisenstein et al. (2005), who measured the redshift-space correlation function  $\xi(s)$  of the Luminous Red Galaxies from the Sloan Digital Sky Survey (see Section 3.2), with median redshift  $z = 0.35$ , as a function of the comoving separation  $s$ .



**Figure 1.7:** The measured galaxy redshift-space correlation function of the SDSS LRG sample as a function of the comoving separation. The green, red and blue curves represent respectively models with  $\Omega_m h^2 = \{0.12, 0.13, 0.14\}$ , for fixed  $\Omega_b h^2 = 0.024$  and  $n_s = 0.98$ . The magenta line shows a pure CDM model ( $\Omega_m h^2 = 0.105$ ), with no evidence of an acoustic peak. Credit: Eisenstein et al. (2005).

### 1.3.3 The Cosmic Microwave Background Radiation

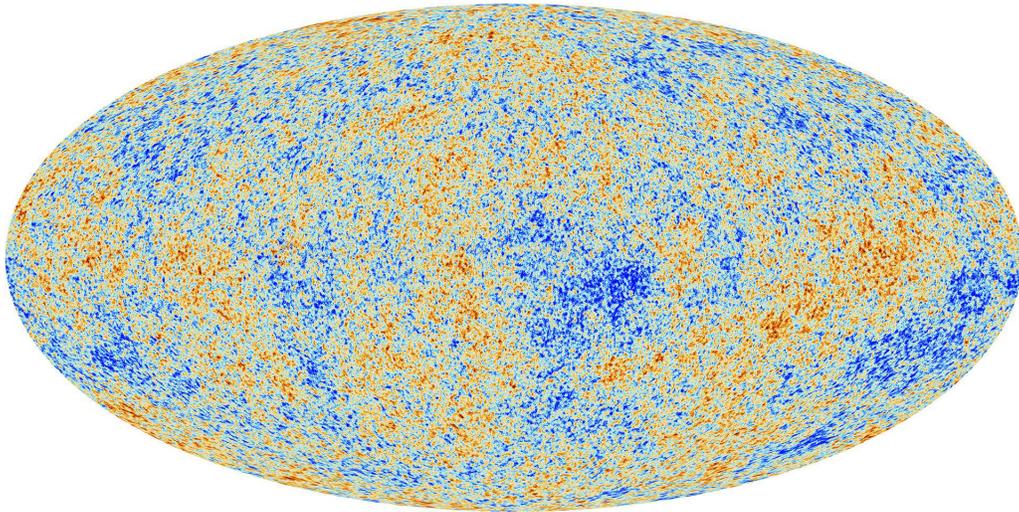
In 1964, the radio-astronomers Arno Penzias and Robert Wilson, accidentally discovered one of the most significant evidence of the Big Bang model: the existence of the CMB. Its presence probes that the Universe evolved from a hot ionised state where baryons and photons were strongly coupled. As the early Universe was expanding, the photon-baryon plasma cooled down to the temperature of  $T \sim 3000\text{K}$ , when baryons decoupled from photons to form neutral hydrogen. Photons could then free stream to us and be observed as the CMB. As photons from the last scattering surface have interacted weakly with ordinary matter, the CMB is a representative picture of the Universe when it was just 380,000 years old ( $z \sim 1100$ ).

The CMB appears to us as an isotropic radiation filling the whole Universe in all directions, with a characteristic black body spectrum at the temperature of approximately  $T_{\text{CMB}} = 2.73\text{K}$ . According to the cosmological principle, the Universe, and thus the CMB, is approximately isotropic and homogeneous on those large scales. More accurate investigations and more recent measurement, such as the ones by the COBE (Boggess et al. 1992; Smoot et al. 1992), WMAP (Bennett et al. 2013) and Planck (Planck Collaboration et al. 2013a) satellites show the presence of tiny temperature irregularities (see Fig. 1.8): these correspond to regions of slightly different densities, which represent the seeds of all structures we see today. More precisely, it has been observed that the distribution of the CMB is isotropic to the precision of  $10^{-3}$ : the background (monopole,  $l = 0$ ) appears completely uniform at a temperature of 2.73 K. Most of the residual anisotropy is due to the dipole anisotropy ( $l = 1$ ,  $\sim\text{mK}$ ), caused to the Doppler effect from the motion of the Sun with respect to the background radiation and the primordial anisotropy ( $l \geq 2$ ,  $\sim\mu\text{K}$ ), due to a scattering effect and a gravitational effect (*Sachs-Wolfe effect*, Sachs & Wolfe 1967). After subtracting all these contributions (including Milky Way emission visible in the central part of the map), we are left with density fluctuations of

$$\frac{\Delta T}{T} = \frac{\Delta\rho_m}{\rho_m} \approx 10^{-5} . \quad (1.43)$$

Note that the equality here between temperature and density fluctuations holds only if perturbations are adiabatic.

Statistical properties of the CMB are represented by the temperature power spectrum as a function of angular wavenumber  $l$  (small  $l$  correspond to large angular scales). The CMB power spectrum is a measure of the anisotropy power on different angular scales: the sky is divided up into polar coordinates and the observed temperature field is decomposed into spherical harmonics. The theoretical prediction of the CMB temperature power spectrum is related to the energy contents of the Universe and can be used to constrain cosmological parameters, by comparing with observed data. The CMB gives us information about the early Universe ( $z \sim 1100$ ), being less sensitive to the late Universe, as photons interact rarely with matter. The CMB anisotropies has been measured by COBE, WMAP, and more recently by Planck, South Pole Telescope and Atacama Cosmology Telescope up to  $l \sim 3000$ . For the cosmological analysis presented in this work, we include the CMB spectra measured by WMAP, for which we now provide some description.



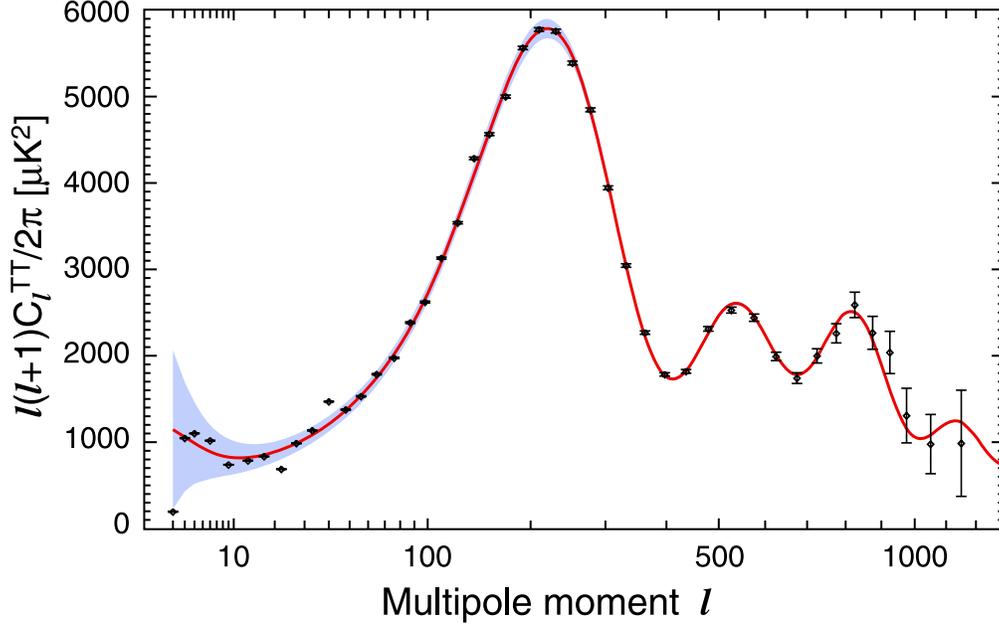
**Figure 1.8:** The anisotropies of the CMB as observed by Planck satellite. Cold spots are in blue, while hot are in red. Copyright ESA and the Planck Collaboration.

### Wilkinson Microwave Anisotropy Probe

The WMAP<sup>1</sup> is a NASA Explorer mission which collected a huge amount of data, now fully analysed to obtain important cosmological achievements. Charles Bennett and the WMAP team won the 2012 Gruber Cosmology Prize because of these published results. The WMAP instrument is composed of cooled microwave radiometers, with  $1.4 \times 1.6$  meter diameter primary reflectors, in five frequency bands (22-90 GHz) to allow the separation of the foreground signals from the CMB. WMAP measures the temperature difference between two points in the sky to an accuracy of  $10^{-6}$  degree: this means also that systematics have been carefully handled. The main achievement of this project has been the first fine-resolution (0.2 deg) full-sky map of the microwave sky. In addition to this, the inflationary model has been supported, as well as the Gaussian distribution of temperature fluctuations. Furthermore, the following constraints on cosmological parameters have been placed : the age of the Universe is 13.77 billion years old, within a 0.5%; the curvature of space is zero within 0.4%; the Universe contents are baryons (4.6%), dark matter (24.0%) and dark energy (71.4%). In our cosmological analysis, we include the CMB spectra from the WMAP Data Release 7, whose detailed cosmological results have been published by Komatsu et al. (2011). Fig. 1.9 shows the CMB temperature power

---

<sup>1</sup><http://map.gsfc.nasa.gov/>



**Figure 1.9:** The 7-year temperature (TT) power spectrum from WMAP. The curve is the  $\Lambda$ CDM model best fit to the 7-year WMAP data:  $\Omega_b h^2 = 0.02270$ ,  $\Omega_c h^2 = 0.1107$ ,  $\Omega_\Lambda = 0.738$ ,  $n_s = 0.969$ . The plotted errors include instrumental noise. The grey band represents cosmic variance. Credit: Larson et al. (2011).

spectrum  $l(l+1)C_l/2\pi$  as a function of multipole  $l$  ( $l = \pi/\theta$ ) as measured by WMAP DR7 (Larson et al. 2011). The locations and shapes of the first ( $l \sim 200$ ) and second peak ( $l \sim 500$ ) has been detected with high precision, while the third peak ( $l \sim 800$ ) is less constrained. The first peak location corresponds to the size of the sound horizon at the last scattering surface. As we can measure the distance to the last scattering surface, knowing the redshift of the CMB, we can locate a point in the Hubble diagram with very high accuracy, and probe the geometry of the Universe. This method measures the Universe to be spatially flat  $\Omega_k \sim 1$ . The other peaks instead represent combinations of  $\Omega_r$ ,  $\Omega_b$ ,  $\Omega_m$ . The cosmology results of WMAP DR9 have recently been published (Hinshaw et al. 2013).

## 1.4 Galaxy Clusters

Clusters of galaxies are a particularly rich source of information about the underlying cosmological model. They are the largest and most recent collapsed objects in the Universe. Studies of their evolution and properties can place strong constraints on the growth of structures and on the current cosmological paradigm. Here we briefly describe the history of galaxy clusters observations, their main constituents and observables, their formation process and their role as cosmological probes.

### 1.4.1 History of galaxy clusters observations

Galaxy clusters were discovered quite early in the history of modern astronomy by Messier (1784) and Herschel (1785), independently. The extragalactic nature of these objects was only later confirmed and galaxy clusters were considered as proper physical objects. Their nature was not recognised until the 1930's, when the dynamical analysis of Zwicky (1937) and Smith (1936) enable the first estimation of their mass. They showed the evidence for much more gravitational material than indicated by the stellar content of the galaxies in the cluster alone, giving the first hint of DM in the Universe. This was later confirmed by measurements of cluster masses using the velocity distribution of the galaxies by means of the Virial Theorem<sup>2</sup> (Rood 1974b,a). Then, the studies on galaxy clusters were extended to several aspects: origin and evolution, dynamical properties, distribution and characterization of the galaxies inside a cluster. Large catalogues of clusters (Abell 1958; Zwicky et al. 1968) based on eye estimates of the number of galaxies per unit solid angle were developed. The first all sky X-ray survey with the Uhuru satellite (Giacconi et al. 1972) confirmed that many clusters were spatially extended X-ray sources. More recently, the discovery of hot high-redshift clusters by Bahcall & Fan (1998) was the first suggestion of a DE component. Finally, last decades experienced the birth of numerous surveys in all wavelengths and an exponential increase of publications on galaxy clusters. More details about these latest scientific results will be covered in Chapter 3.

---

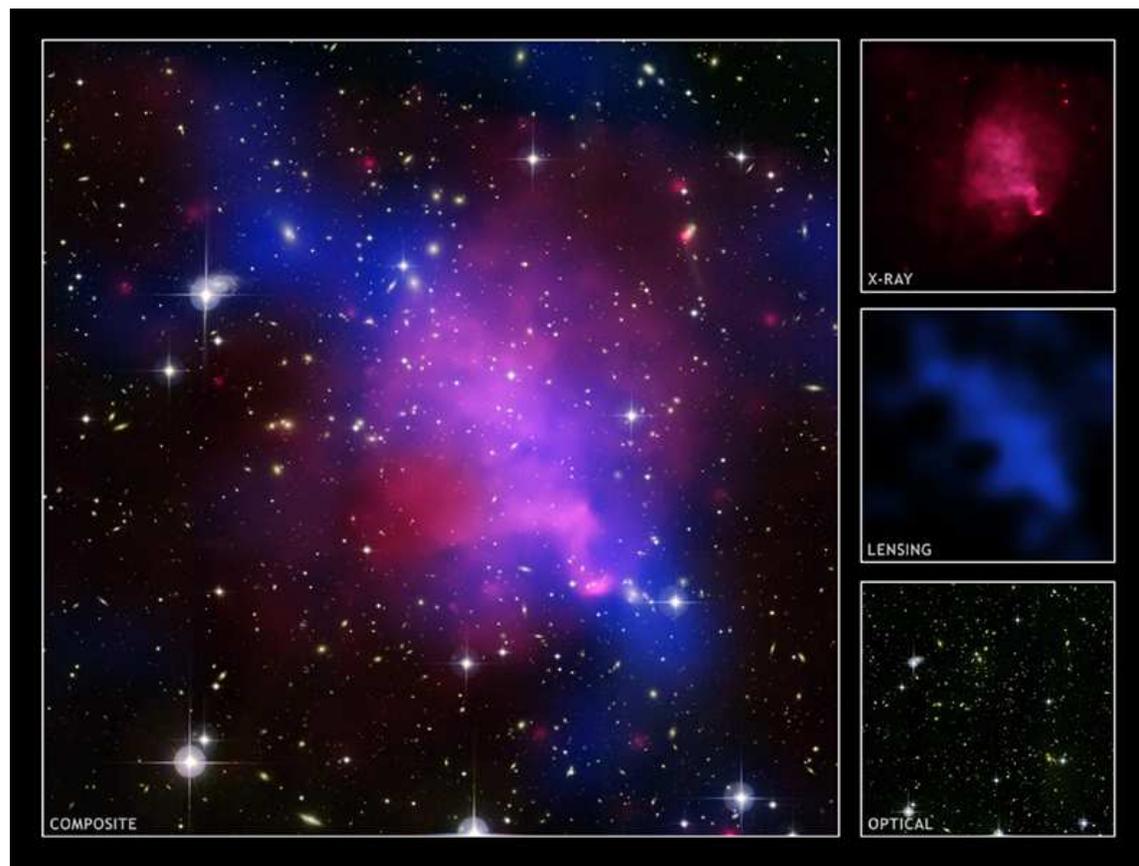
<sup>2</sup>The Virial Theorem states that, for a stable, self-gravitating, spherical distribution of objects of same mass, it holds  $E_k = -1/2 E_p$ , where  $E_k$  is the total kinetic energy of the objects and  $E_p$  is the total gravitational potential energy.

### 1.4.2 Main features, components and observables

Clusters typically have masses of  $10^{13}$ - $10^{15}M_{\odot}$ , sizes of the order of few Mpc, velocity dispersions of 800-1000 km/s and X-ray luminosities of  $10^{43}$ - $10^{45}$  erg/s. Clusters of galaxies are typically larger than groups and contain about 50 to 1000 members: this limits assign the denomination of *rich* and *poor* cluster, respectively. We can also distinguish clusters between *regular*, which are spherical with a central region of higher density, and *irregular* ones, which are instead not spherical and without a unique dense central region. Phenomenologically clusters are composed of:

- **Galaxies** (2-5%), which contain the condensed baryonic matter in the form of stars and cold gas. The typical population is composed of old and passive (red and dead) galaxies, which ended their star formation at  $z > 2$  and which sit on a *red-sequence* locus in a colour-magnitude diagram.
- **Intra-Cluster Medium** (ICM) (11-15%), which mainly consists of hydrogen and helium, represents most of the baryonic matter in a highly ionised form and low density ( $\sim 10^{-3}$ atoms/cm<sup>3</sup>). As a matter of fact, the ICM reaches temperature of approximately  $10^8$ K to balance the gravitational pull of the DM potential well, and emits in the X-ray band. The main X-ray emission processes from ICM are collisional: thermal *Bremsstrahlung* (free-free emission), recombination (free-bound emission), line radiation (bound-bound emission). The emissivity of the Bremsstrahlung mechanism is stronger in the densest innermost regions because is proportional to the squared number density of particles.
- **Dark Matter Halo** (80-87%): it follows a universal distribution known as the Navarro-Frenk-White (NFW) profile (Navarro et al. 1997), which depends on the central density and scale radius (see Section 2.1.2).
- **Intra-Cluster light**: it is the optical light from stars which are gravitationally bounded to the cluster itself.

As a consequence, they are accessible by multiple signals, across the whole electromagnetic spectrum. Fig. 1.10 shows the superposition of three views of the galaxy cluster Abell 520. The optical view represents the galaxy population; the hot ICM is captured in red in the Chandra X-ray Observatory image; finally, the gravitational lensing image is instead highlighting the collisionless core of DM component in blue. In general, the galaxy population and the intra-cluster light are visible in optical and near-infrared bands. The hot ICM is instead detected by the strong X-ray



**Figure 1.10:** Composite image of three views of the galaxy cluster Abell 520. The optical view shows the galaxies bound together by gravitational force. Diffuse, hot gas in between the galaxies emits X-rays: this is shown in red in the Chandra X-ray Observatory image. Gravitational lensing image is representing, instead, the collisionless core of dark matter component in blue. Credit: X-ray: NASA/CXC/UVic./A.Mahdavi et al.; Optical/Lensing: CFHT/UVic./A.Mahdavi et al..

emission, while at radio frequencies, synchrotron emission from relativistic electrons can be detected and provide information on the intra-cluster gas. Furthermore, at millimetre wavelengths, high-density regions within clusters cause distortions of the CMB spectrum by inverse Compton scattering, namely the *Sunyaev-Zel'dovich* (SZ) effect (Sunyaev & Zeldovich 1972): the low-energy CMB photons enhance their energy because of the collision with the high energy ICM electrons, causing a local frequency-dependent shift in the CMB spectrum observed through the cluster. This

effect is used to detect clusters with no redshift limitation and is quantified by the Compton  $y$ -parameter, i.e. the electron pressure integrated along the line of sight  $l$ :

$$y = \int \frac{k_B T_X(l)}{c^2 m_e} n_e(l) \sigma_T dl . \quad (1.44)$$

Here  $k_B$  is the Boltzmann constant,  $T_X$  is the X-ray temperature,  $m_e$  and  $n_e$  are the electron mass and number density respectively,  $\sigma_T$  is the Thomson cross-section. More practically, the quantity which is usually measured is the projection on the cluster area  $dA$ , namely the *integrated* Compton parameter  $Y_{SZ} \propto \int y dA$ . Finally, strong features are also detected in the gravitational lensing shear field, which gives information about the DM halo (see Section 2.1.3).

### 1.4.3 Cluster mass proxies

One of the key issues in the study of galaxy clusters is the determination of their true mass. Cluster total masses cannot be directly determined from observation, but instead they have to be deduced from some observational properties, called **mass proxies**, which correlate with the true mass via the so-called **scaling relations**. Various mass proxies in different wavelength and associated systematics have been used so far to determine the mass of clusters from observations, via the respective scaling relations and scatter around them. Here we only list the most common ones:

- i) the optical richness, i.e. number of red galaxies within  $R_{200}$ :  $N_{200}$  (Roza et al. 2010) - this is the observable throughout all our analysis;
- ii) the line-of-sight velocity dispersion:  $\sigma_v$ , which is related to the total mass as (Longair 2008)  $M \propto \sigma_v^2 R_{\text{vir}}$ , where  $R_{\text{vir}}$  is the virialization radius;
- iii) the X-ray temperature, bolometric luminosity, gas mass, gas total thermal energy;
- iv) the integrated SZ parameter at mm wavelength.

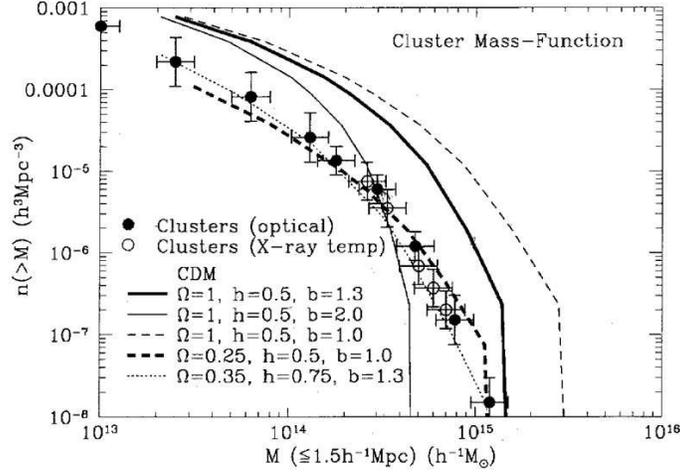
Note that another valid technique to measure cluster masses is gravitational lensing: it uses the distortions of background galaxies images caused by the space-time deformation, which is induced by the cluster halo mass.

### 1.4.4 Formation of galaxy clusters

Formation and evolution of clusters of galaxies trace directly the *hierarchical growth* of structures in the Universe. The first objects which start to collapse and virialize, deviating from the Hubble flow, have sub-galactic sizes. Then, these structures merge to originate the galaxies, which analogously can form galaxy clusters by merging. Fluctuations inside a region grow until they balance the local expansion: at this point, the expansion of the region is slowed down till it reaches a maximum radius. Having no more kinetic energy but only gravitational potential energy, the region collapses: baryons fall into the gravitational potential wells produced by the DM and potential energy is converted into kinetic one. This brings the gas to thermalisation, thus producing the hot plasma. When the Virial Theorem condition is satisfied, the dynamical equilibrium is reached. The kinetic energy of the galaxies moving randomly inside the cluster furnishes a pressure which counteracts the gravitational attraction: this gives stability to the cluster.

### 1.4.5 Clusters as cosmological probes

As in GR the geometry of the Universe is fully described by the total energy content (see Eq. 1.5), one can study the structure of the Universe by testing the geometry by means of probes such as SNIa, BAO and CMB. Alternatively, it is possible to test both the geometry and the structure with different probes and then compare the constraints. Clusters of galaxies are fundamental because they provide both an independent measure of cosmological parameters with different systematics to the CMB, SNIa and BAO, and a probe of the growth of structures. In particular, galaxy clusters are used to test cosmology by measuring their *mass function*, namely the number density of clusters as a function of their mass and redshift. The precise determination of the mass function and its evolution can place constraints on the energy components of the Universe. As an example, we show in Fig. 1.11 an early result for the cluster mass function obtained by Bahcall & Cen (1992). The optical data are based on richness, velocities and luminosity function of clusters, while the X-ray data refer to the temperature distribution of clusters. Here the observations of optical and X-ray galaxy clusters are compared with expectations from different cosmologies using CDM large-scale (box size of  $400h^{-1}\text{Mpc}$ ) simulations. The comparison shows that the cluster mass function is a powerful discriminant among models: the  $\Omega_m = 1$  model cannot reproduce the observations for any bias parameter. In fact, when normalised to the COBE CMB fluctuations on large scales (Smoot et al. 1992), this model predicts a much larger number of massive clusters than is observed. On the other hand, a low-density CDM model, with  $\Omega_m = \{0.25, 0.35\}$  and



**Figure 1.11:** Cluster mass function observations of optical and X-ray data, compared with CDM simulations. A model with  $\Omega_m = \{0.25, 0.35\}$  (with or without a cosmological constant), appears to match the observations. The  $\Omega_m = 1$  model, instead, fails in reproducing data. Credit: Bahcall & Cen (1992).

bias  $b = \{1.0, 1.3\}$ , with or without a cosmological constant, appears to fit well the observations. Precise observations of large numbers of clusters have later provided an important tool for understanding better their abundances. The full theoretical derivation, numerical calibration and discussion on the cosmology dependence of the mass function are provided in Section 2.2. In addition to a predicted mass function and a well-determined relation between the true cluster mass and the observable, a cluster experiment needs a large, clean, complete survey with a well-defined selection function. We list the main X-ray, millimetre, weak lensing and optical cluster surveys in Section 3.1. Complementary to the abundances, the *clustering* of galaxy clusters, i.e. their spatial distribution at  $z = 0$  and its evolution to higher redshifts, contains fundamental information on the underlying matter distribution as well. We give a detailed description of the cluster power spectrum and its cosmology dependence in Section 2.4.

Detailed theoretical modelling of clusters is a complicated astrophysics problem involving a variety of physical phenomena. Useful tools in this regards are numerical *simulations*. While the pure gravitational interactions of DM particles can be treated in a linear regime and their behaviour is well described, baryonic physics is far more complex, non-linear and involves hydrodynamical processes.

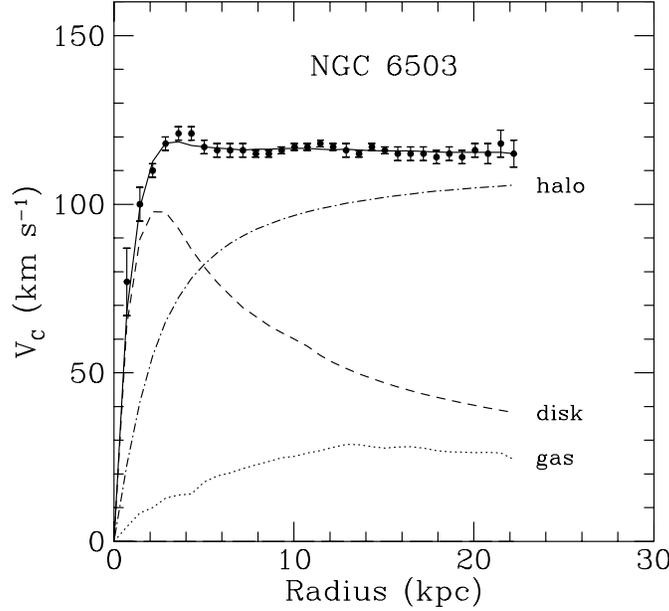
## 1.5 $\Lambda$ CDM standard model

Several observations over the past decades confirmed that the Universe is experiencing a phase of cosmic acceleration, driven by a dark form of energy with negative vacuum pressure. Perlmutter et al. (1999) with SNIa, Allen et al. (2004, 2008) with clusters of galaxies, Eisenstein et al. (2005) with Large-Scale Structure (LSS) and Komatsu et al. (2011) with the CMB, independently confirm the accelerated expansion epoch which is currently ongoing. Therefore the concordance **Lambda Cold Dark Matter** ( $\Lambda$ CDM) cosmological model has been formulated. It affirms that the Universe is composed of:

- $\sim 5\%$  of ordinary baryonic matter  $\Omega_b$ , mainly made up by hydrogen atoms ( $\sim 75\%$ ), Helium atoms ( $\sim 25\%$ ), while heavier elements are only a tiny fraction;
- $\sim 25\%$  of unknown (dark) form of matter  $\Omega_{\text{cdm}}$ , made up by species of sub-atomic particles that interact almost only gravitationally (and not electromagnetically) with ordinary matter, being thus totally collisionless;
- $\sim 70\%$  of unknown (dark) form of energy  $\Omega_\Lambda$ , responsible of the late time accelerating expansion;
- a radiation component  $\Omega_r$ , which is negligible today, as  $\Omega_r/\Omega_m \simeq 1/3250$ .

There are few probes of the existence of the DM component. One is related to the rotation curves of galaxies (see Fig. 1.12 and Freeman 1970) which do not reveal a Keplerian decline (namely the squared velocity is not proportional to the inverse radius), giving evidence of an undetected matter component. Furthermore, the gravitational lensing in galaxy clusters shows a mismatch between the amount of normal matter and the estimated total mass. In addition to this, the evidence of the collisionless nature of dark matter has been observed in few objects (e.g. the 'bullet cluster' in Markevitch et al. 2004; Clowe et al. 2004). A fundamental property of DM is that it is non-relativistic (i.e. cold): this is necessary to explain the structure formation model currently accepted. Possible candidates for a DM particle are provided by theoretical particle physics, e.g. Weakly Interacting Massive Particles (WIMPSs), which are massive particles interacting through the weak nuclear force and gravity.

The simplest way to define the DE dominant component is a positive value of the cosmological constant  $\Lambda$  introduced in Einstein field equations, with constant equation of state  $w = -1$ . However, few problems arise from this choice. First of all,



**Figure 1.12:** Rotation curve of galaxy NGC 6503: the data points with error bars are the observed velocities, the disk stars contribution is shown by the dashed line, while the contribution of the gas is represented by the dotted line. As Freeman (1970) first noticed that the expected Keplerian decline (i.e.  $v^2 \propto r^{-1}$ ) was not present in NGC 300 and M33 galaxies, also here there is clear evidence of an undetected dark matter halo component, with density  $\rho_{\text{DM}}(r) \propto r^{-2}$ . Credit: Kamionkowski (1998).

the *cosmological constant problem* appears if we associate  $\Lambda$  to the vacuum energy, i.e. the background energy in absence of matter: the observed cosmological constant is smaller by a factor of  $\sim 10^{120}$  than the value for the vacuum energy predicted by quantum field theories. In addition to this, the *coincidence problem* asks why we live at the special epoch where DE density is approximately equal to matter density. Numerous alternative theories try to explain the nature of this constituent (e.g. quintessence,...). For example, by assuming that the equation of state of DE evolves in time, we obtain  $w(z) = w_0 + w'z$  (Maor et al. 2001; Weller & Albrecht 2001, 2002), which diverges at high redshift, or  $w(z) = w_0 + w_1 z/(1+z)$  (Chevallier & Polarski 2001; Linder 2003). Alternatively, modification of gravity can be performed: they do not invoke a new form of energy, but instead introduce new physics which modifies Einstein's equations on large scales (e.g. Dvali et al. 2000).

### 1.5.1 Cosmological constraints from observations

For completeness, we now list the main cosmological parameters in the concordance  $\Lambda$ CDM model, which govern the global properties of the Universe and the spectrum of the initial density perturbations, together with their current constraints from the latest Planck mission (Planck Collaboration et al. 2013b) (see Table 1.2).

Symbol	Definition	Constraint
$\omega_b = \Omega_b h^2$	Baryon density	0.02214 $\pm$ 0.00024
$\omega_{\text{cdm}} = \Omega_c h^2$	Cold Dark Matter density	0.1187 $\pm$ 0.0017
$\Omega_k$	Spatial curvature	-0.0005 $^{+0.0065}_{-0.0066}$
$\Omega_\Lambda$	Dark Energy density	0.692 $\pm$ 0.010
$\ln(10^{10} A_s)$	Primordial pert. amplitude	3.091 $\pm$ 0.025
$\sigma_8$	RMS matter fluctuations	0.826 $\pm$ 0.012
$w$	Constant EoS of Dark Energy	-1.13 $^{+0.23}_{-0.25}$
$\tau$	Reionization optical depth	0.092 $\pm$ 0.013
$n_s$	Primordial scalar spectral index	0.9608 $\pm$ 0.0054
$\sum m_\nu$	Sum of the neutrino masses in eV	<0.230
$N_{\text{eff}}$	Effective number of neutrino-like species	3.30 $^{+0.54}_{-0.51}$
$H_0$	Hubble constant	67.80 $\pm$ 0.77
$t_0$	Age of the Universe (Gyr)	13.798 $\pm$ 0.037
$z_{\text{re}}$	Redshift of half-reionization	11.3 $\pm$ 1.1
$100\theta^*$	100 $\times$ angular size of sound horizon	1.04162 $\pm$ 0.00056

**Table 1.2:** List of the main cosmological parameters of  $\Lambda$ CDM model, together the constraints from Planck+WMAP+highL+BAO (Planck Collaboration et al. 2013b) for the following models: six parameter base  $\Lambda$ CDM model and derived parameters (blue, 68% limits) and extensions to the base  $\Lambda$ CDM model (green, 95% limits).

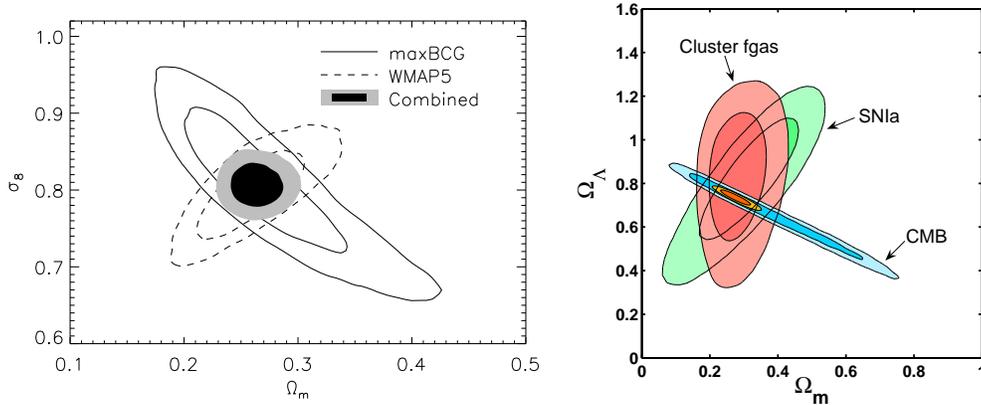
We conclude this Chapter by highlighting the constraining power on cosmological parameters of clusters of galaxies: in combination with other probes, such as SNIa, BAO and CMB, some parameters degeneracies can be broken and the errors tightened.

### $\Omega_m - \sigma_8$ constraints

Constraints on  $\Omega_m - \sigma_8$  plane were investigated by Mantz et al. (2010) comparing and combining three Rosita All Sky Surveys (RASS). Independent clusters studies of optical clusters (Rozo et al. 2010) (see left panel of Fig. 1.13), Sunyaev-Zeldovich clusters in combination with X-ray measurements (Benson et al. 2013) and X-ray clusters (Vikhlinin et al. 2009) showed consistent results. In the right panel of Fig. 1.13 we show Allen et al. (2008) constraints on the  $\Omega_m - \Omega_\Lambda$  plane, from the combination of Chandra measurements of the X-ray gas mass fraction  $f_{\text{gas}}$  of galaxy clusters, SNIa data and CMB measurements.

### Neutrinos

As any particle with a non-zero mass transits while cooling from a relativistic state to a non-relativistic state, the mass of neutrinos influences the background evolution and cosmic structure formation. The quantity typically used to describe neutrinos mass is  $\sum m_\nu$ , which is the species-summed mass. Constraints on this quantity come from clusters combined with CMB data (Burenin & Vikhlinin 2012). Few more considerations on this topic are included in Chapter 6.



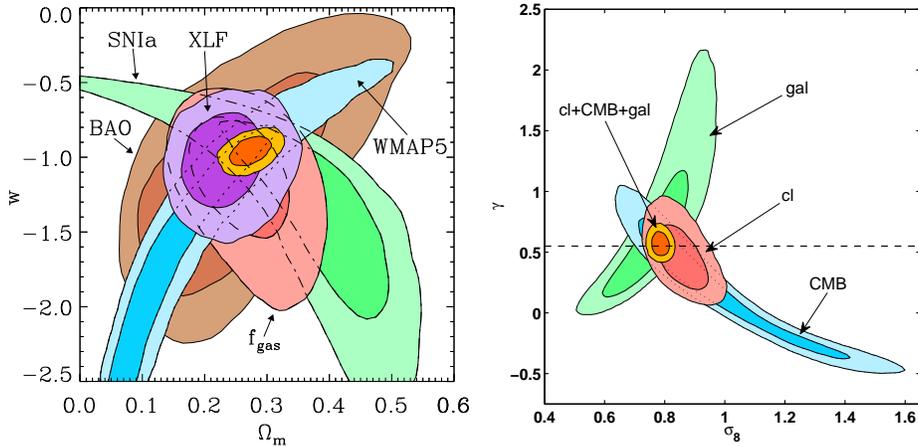
**Figure 1.13:** Left panel: Joint 68.3% and 95.4% confidence regions in the  $\Omega_m - \sigma_8$  plane from optical galaxy cluster of the maxBCG catalogue combined with WMAP5 (Dunkley et al. 2009). Right panel: contours for  $\Omega_m - \Omega_\Lambda$  from the combination of X-ray gas mass fraction (pink), CMB (blue) and SNIa (green). The orange contours show the constraint obtained from all three data sets combined. Credit: Rozo et al. (2010); Allen et al. (2008).

### DE equation of state

Allen et al. (2011) analysed the constraints on the DE equation of state together with  $\Omega_m$  (see left panel of Fig. 1.14) or  $\sigma_8$ . He combined the abundance and growth of RASS clusters (Mantz et al. 2010),  $f_{\text{gas}}$  measurements (Allen et al. 2008), WMAP5 results (Dunkley et al. 2009), Supernovae Ia data (Kowalski et al. 2008) and BAO measurements (Percival et al. 2010, 2011). Constraints on DE equation of state from data were also performed by Rapetti et al. (2005) with X-ray clusters+SN Ia+CMB, by Mantz et al. (2010); Benson et al. (2013) with X-ray clusters, while Vikhlinin et al. (2009) constrained  $w$  and  $\Omega_\Lambda$ .

### Cosmic growth $\gamma$

Rapetti et al. (2013) tested the cosmic growth predicted by GR ( $\gamma = 0.55$ ) with three independent measurements: galaxy clusters abundances and  $f_{\text{gas}}$  from RASS and Chandra, galaxy clustering from WiggleZ Dark Energy Survey, 6-degree Field Galaxy Survey and CMASS BOSS, and CMB from WMAP. The cosmic growth is modelled by the growth index  $\gamma$  defined in Eq. (1.36) and  $\sigma_8$ . We show in the right panel of Fig. 1.14 the constraints obtained on these parameters.



**Figure 1.14:** Left panel: Joint 68.3% and 95.4% confidence regions for  $w - \Omega_m$ , from the abundance and growth of RASS clusters (violet), X-ray gas mass fraction (pink), WMAP5 (blue), SNIa (green) and BAO (brown). Right panel: joint contours in the  $\sigma_8 - \gamma$  plane, from galaxy growth (green), CMB (blue) and cluster growth (red). The gold contours show the combination of the data sets. Credit: Allen et al. (2011); Rapetti et al. (2013).



# Chapter 2

## Galaxy Clusters from theory side

In this Chapter the theoretical framework of galaxy clusters one-point (number counts) and two-points statistic (power spectrum) is introduced, in order to understand why they are fundamental probes of the LSS of the Universe. For the study galaxy clusters, in fact, one needs to have first an estimate of their masses, which are not directly accessible. Here, we define the cluster masses and density profiles, with a particular emphasis to the weak lensing mass estimation, as this is the one we use in our analysis. Secondly, to understand how the cluster number counts change with the mass and with the cosmological model assumed, we revise the formulation and calibration of the cluster mass function and its sensitivity to cosmology. In addition to this, the spatial distribution of clusters can give additional information on cosmology. We thus introduce the concept of the model bias and a prescription for the clusters redshift space power spectrum. Finally, some definitions concerning the study of non-Gaussian initial conditions are provided, as clusters can be a good probe in this context.

### 2.1 Cluster masses

As the mass of galaxy clusters is not directly measurable, we describe here how to get an estimate of it. Cosmologists usually define the cluster mass with respect to the critical or the mean density of the Universe and assume a halo density profile. We then focus on the gravitational weak lensing technique to reconstruct the mass distribution, as it is part of the data sets we need for our combined cosmological analysis.

### 2.1.1 Definition

A common way to define a cluster mass is with respect to the critical density of the Universe,  $\rho_c(z)$  (see Eq. 1.14). One considers a spherical region of a virialized halo of radius  $R_\Delta$  and volume  $V_\Delta = 4\pi/3 R_\Delta^3$ , with mean density of  $\bar{\rho}(z) = \Delta \rho_c(z)$ . The total amount of matter contained in this region is defined to be the mass  $M_\Delta$  as

$$M_\Delta = V_\Delta \bar{\rho}(z) = \frac{4\pi}{3} R_\Delta^3 \Delta \rho_c(z) . \quad (2.1)$$

Typical values of  $\Delta$  are 200 (this work) or 500, because cluster properties are more evident to observe in high density contrast regions. Alternatively, the mass can be defined with respect to the background mean density of the Universe,  $\rho_m(z)$ : the mean density inside the virial radius is then  $\bar{\rho}(z) = \Delta \rho_m(z) = \Delta \Omega_m(z) \rho_c(z)$  (because of Eq. 1.16), and the value of  $\Omega_m(z)$  has to be assumed.

### 2.1.2 Halo density distribution

In order to estimate the mass of a cluster, we need to have a prescription for its density profile. Observation of galaxy clusters show that the velocity dispersion  $\sigma_v$  of galaxies inside a cluster remains nearly constant with distance from the cluster centre: this implies an underlying mass-density distribution,  $\rho(R) \propto R^{-2}$ . Binney & Tremaine (1987) found that the *Singular Isothermal Sphere* (SIS) was the simplest analytical model consistent with the observed mass distribution. Such a profile resembles the density structure of an isothermal self-gravitating system of particles, which encountered a violent relaxation process (Lynden-Bell 1967) from a chaotic initial state to a quasi-equilibrium. This density distribution is characterised by a constant and isotropic velocity dispersion  $\sigma_v$ , and has the following form

$$\rho_{\text{SIS}}(R) = \frac{1}{2\pi G} \frac{\sigma_v^2}{R^2} . \quad (2.2)$$

This approximation is, however, incomplete because the mass diverges with  $R$  linearly. Later, cluster formation numerical simulations obtained a more accurate form of density profile:

$$\rho(R) \propto R^{-p} (R + R_s)^{p-q} , \quad (2.3)$$

where  $p$  and  $q$  describe the slopes of inner and outer regions, and  $R_s$  is the scale radius where the profile steepens. Fits of Eq. (2.3) to simulated clusters give, for example, the Moore et al. (1999) profile ( $p = 1.5$ ,  $q = 3$ ) and Rasia et al. (2004)

profile ( $p = 1, q = 2.5$ ). Navarro et al. (1997) best fit values were instead  $p = 1, q = 3$ : more precisely they obtained the following expression (for  $\Delta = 200$ )

$$\rho_{\text{NFW}}(R) = \frac{\rho_c(z)\lambda_c}{(R/R_s)(1 + R/R_s)^2}, \quad (2.4)$$

where  $\lambda_c$  is the density contrast given by

$$\lambda_c = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}, \quad (2.5)$$

being  $c = R_{200}/R_s$  the concentration of the halo, namely the ratio of the virial radius  $R_{200}$  to the scale radius  $R_s$ . The concentration parameter is a function of halo mass and redshift and depends on cosmological parameters: typically, the earlier the halo forms, the higher is the value of  $c$ . Eq. (2.4) is known as the **Navarro-Frenk-White** (NFW) profile and it is the most widely used fitting formula for haloes density: we will assume this profile in this work.

Note that on large scales (clusters) CDM simulations have proven very successful, while on galaxy scales their predictions seem to be in conflict with some observations. Among these, the cusped central density profile given by the NFW profile, which is not observed in dwarf galaxies. Burkert (1995) empirically modified the analytic form of the NFW profile, trying to find the best-fitting density law to the observed rotation curves of dwarf galaxies. The Burkert's profile can be expressed as:

$$\rho_{\text{B}}(R) = \frac{\rho_0 R_s^3}{(R + R_s)(R^2 + R_s^2)}, \quad (2.6)$$

where the central density  $\rho_0$  and the scale radius are free parameters. This density law resembles an isothermal profile in the inner regions ( $R < R_s$ ) and predict a finite central density  $\rho_0$ , while in the outer regions it diverges logarithmically with radius, in agreement with the NFW profile.

### 2.1.3 Weak Lensing signal

According to GR, light rays propagate along null-geodesics of the space-time metrics. The presence of mass distribution perturbs the metrics, deflecting the light trajectories. This deflection distorts the shape of distant galaxies, allowing the study of the mass distribution of objects in the Universe. This distortion is called **gravitational lensing** effect. In 1804 Johann Soldner wrote the first paper on this topic,

predicting with Newtonian physics that a light ray passing close to the Sun is deflected by an angle of  $\sim 0.84$  arcsec. In 1919, Einstein obtained the same value  $\hat{\alpha} = 2GM_{\odot}/c^2R_{\odot} \sim 0.83$  arcsec, by applying Huygens principle to a wave front passing through a region in which the speed of the wave varies with position (the closer to the massive body, the smaller the speed is). After completing GR, he realised the correct equation for the deflection angle of a light ray passing at a distance  $R$  from an object with mass  $M$  was:

$$\hat{\alpha} = \frac{4GM}{c^2R} , \quad (2.7)$$

finding for the Sun the value of 1.74 arcsec. The additional factor of two is due to the spatial curvature which is missing in the Newtonian approach. More precisely, the full theory takes into account not just the variation of the time-time component of the metric, but also the variation of the spatial components.

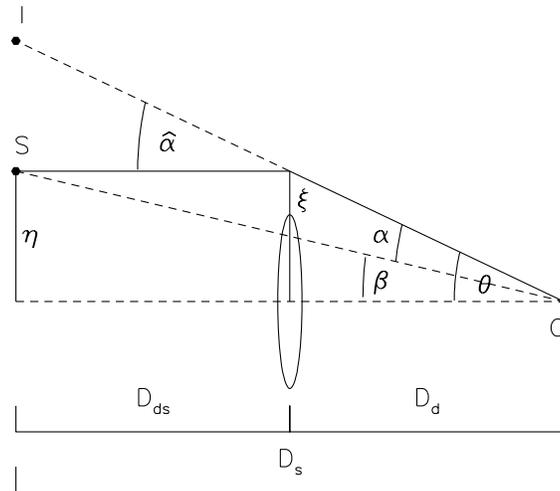
By measuring the gravitational lensing distortions of background galaxies caused by clusters of foreground galaxies, it is possible to probe their matter distribution. It plays a critical role, because it can measure the total cluster mass regardless of the baryonic content, star formation history, or dynamical state. Thus, lensing is crucial for calibrating the mass-observable relation, as we will detail in Chapter 3. In order to describe how gravitational lensing works, we need to introduce some basic notation following Bartelmann & Schneider (2001). If the lens is thin compared to the total length of the light path, the lens mass distribution can be projected on an orthogonal plane with respect to the line-of-sight, called *lens plane*. This mass sheet is characterised by its surface mass density

$$\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z) dz , \quad (2.8)$$

where  $\rho$  is the density and  $\vec{\xi}$  is a two-dimensional vector in the lens plane. The deflection angle at  $\vec{\xi}$  is the sum of all mass elements deflections in the lens plane:

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int \frac{(\vec{\xi} - \vec{\xi}')\Sigma(\vec{\xi}')}{|\vec{\xi} - \vec{\xi}'|^2} d^2\xi' . \quad (2.9)$$

The geometry of a typical gravitational lens system is shown in Fig. 2.1. A light ray from a source S, assumed to lie in the so-called *source plane*, is deflected by the angle  $\vec{\hat{\alpha}}$  at the lens plane before reaching observer O. The angle between the optical axis and the image I is  $\vec{\theta}$ , while the one between the optical axis and the true source is  $\vec{\beta}$ . The angular diameter distances observer-lens, lens-source, observer-source are



**Figure 2.1:** The gravitational lensing system representation. Angular diameter distances between observer (O) and source (S), observer and lens, lens and source are respectively  $D_s$ ,  $D_d$  and  $D_{ds}$ . A light ray is emitted from a source at transverse distance  $\vec{\eta}$  from the optical axis to the observer O, crossing the lens plane at transverse distance  $\vec{\xi}$ , deflected by an angle  $\vec{\alpha}$ . The angular separations of the observed image (I) and the source from the observer optical axis are  $\vec{\theta}$  and  $\vec{\beta}$ , respectively. The reduced deflection angle  $\vec{\alpha}$  and the real deflection angle  $\vec{\alpha}$  are related by Eq. (2.10). The picture is from Narayan & Bartelmann (1996).

respectively  $D_d$ ,  $D_{ds}$ , and  $D_s$ . Note that  $D_d + D_{ds} \neq D_s$ . With this notation, one can also define a reduced deflection angle as

$$\vec{\alpha} = \frac{D_{ds}}{D_s} \vec{\alpha} . \quad (2.10)$$

We now simply derive the *lens equation*, which relates the observed image and true source positions, from the geometry of Fig. 2.1:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) . \quad (2.11)$$

The solutions  $\vec{\theta}$  of this equation yield the angular positions of the images of the source at  $\vec{\beta}$ . Eq. (2.11) is generally non-linear and hence allows for multiple images corresponding to a single source position. It is also useful to introduce the *convergence*

parameter, which is the dimensionless surface mass density, as

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_d\vec{\theta})}{\Sigma_c} \quad \text{with} \quad \Sigma_c = \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}}, \quad (2.12)$$

where  $\Sigma_c$  is the critical surface mass density, and thus rewrite the reduced deflection angle as

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' \frac{(\vec{\theta} - \vec{\theta}')\kappa(\vec{\theta}')}{|\vec{\theta} - \vec{\theta}'|^2}. \quad (2.13)$$

The lens mapping can then be linearised locally: the Jacobian describing the distortion of images is given by

$$A = \frac{\partial \vec{\beta}}{\partial \vec{\theta}} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (2.14)$$

Here  $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\psi} = |\gamma|(\cos 2\psi + i \sin 2\psi)$ ,  $\psi$  being the angle between  $\vec{\alpha}$  and the x-axis, is the *tangential shear* which stretches the image around the lens, while the magnification of the image is described by the convergence  $\kappa$ . The *weak* lensing approximation is valid for  $\kappa \ll 1$  and  $|\gamma| \ll 1$ .

In order to estimate the mass of galaxy cluster, we have to employ a model for its density distribution, e.g. the NFW of Eq. (2.4). For any value of the NFW profile parameters (scale radius  $R_s$  and density contrast  $\lambda_c$ ), the convergence and the shear can be analytically inferred. By introducing a dimensionless radial distance  $x = r/R_s = \theta/\theta_s$  (where  $\theta_s = R_s/D_d$ ), we can in fact reformulate the convergence as

$$\kappa(x) = \begin{cases} (2R_s\lambda_c\rho_c)/[\Sigma_c(1-x^2)][1 - 2/\sqrt{1-x^2} \operatorname{atanh}\sqrt{(1-x)/(1+x)}] & x < 1, \\ (2R_s\lambda_c\rho_c)/3 & x = 1, \\ (2R_s\lambda_c\rho_c)/[\Sigma_c(x^2-1)][1 - 2/\sqrt{x^2-1} \operatorname{atan}\sqrt{(x-1)/(x+1)}] & x > 1. \end{cases}$$

The gravitational shear can instead be recast as

$$\gamma(x) = \begin{cases} (R_s\lambda_c\rho_c)/\Sigma_c g_<(x) & x < 1, \\ (R_s\lambda_c\rho_c)/\Sigma_c [10/3 + 4 \ln(1/2)] & x = 1, \\ (R_s\lambda_c\rho_c)/\Sigma_c g_>(x) & x > 1, \end{cases}$$

where the functions  $g_<(x)$  and  $g_>(x)$  are given by

$$g_<(x) = \frac{8 \operatorname{atanh}\sqrt{(1-x)/(1+x)}}{x^2\sqrt{1-x^2}} + \frac{4 \ln(x/2)}{x^2} - \frac{2}{(x^2-1)} + \frac{4 \operatorname{atanh}\sqrt{(1-x)/(1+x)}}{(x^2-1)(1-x^2)^{1/2}},$$

$$g_>(x) = \frac{8 \operatorname{atan}\sqrt{(x-1)/(1+x)}}{x^2\sqrt{x^2-1}} + \frac{4 \ln(x/2)}{x^2} - \frac{2}{(x^2-1)} + \frac{4 \operatorname{atan}\sqrt{(x-1)/(1+x)}}{(x^2-1)^{3/2}}.$$

In the case of galaxy clusters, the observable is the *reduced shear*  $\hat{\gamma}$ , which is obtained by averaging over observed galaxies shapes, and is defined as

$$\hat{\gamma}(\theta) = \frac{\gamma(\theta)}{1 - \kappa(\theta)}. \quad (2.15)$$

In order to estimate the mass of the cluster (lens), having assumed a density profile, we can compare the observed reduced shear with its theoretical prediction obtained combining the analytical expressions for  $\kappa$  and  $\gamma$ . We will use cluster mass estimates done with this technique in our cosmological combined analysis.

## 2.2 Mass function

Once the definition of the cluster mass is given, we can investigate the number density of clusters as a function of it, namely the *mass function*. This represents the expected number density of virialized dark matter halos at redshift  $z$ , with mass in the range  $[M, M + dM]$ , and can be expressed as

$$\frac{dn(M, z)}{d \ln M} = \bar{\rho}_m \left| \frac{d \ln \sigma^{-1}}{dM} \right| f(\nu). \quad (2.16)$$

Here,  $\bar{\rho}_m$  is the mean matter density of the Universe,  $\nu \equiv \delta_c / \sigma(M, z)$ ,  $\delta_c$  being the critical linear overdensity and  $\sigma^2(M, z)$  is the variance of the density fluctuation field filtered at scale  $M$  in the linear evolution regime (see Eq. 1.39).

To calculate the value for  $\delta_c$  analytically, we assume a **spherical top-hat collapsing model**. A closed spherical region of radius  $R$  and overdensity  $\delta$  (see Eq. 1.30), in a homogeneous expanding background field with mean density  $\bar{\rho}$ , evolves independently of the environment influenced by local properties only (Birkhoff's theorem). This region expands up to a maximum radius  $R_{\max}$  at the turn-around time  $t_{\text{turn}}$  and then decouples from Hubble expansion and recollapses to a bound system of radius  $R_{\text{vir}} = R_{\max}/2$  at the collapse time  $t_{\text{col}} = t_{\text{vir}} = 2t_{\text{turn}}$ , reaching the virial theorem equilibrium condition. For an Einstein-de Sitter Universe ( $\Omega_m = 1$ ,  $\Omega_k = 0$ ), the overdensity reaches the value of  $\Delta_{\text{vir}} \sim 178$  at  $t_{\text{vir}}$ . If one linearly extrapolates the density contrast at  $t_{\text{vir}}$ , according to Eq. (1.35), then obtains  $\delta_c \equiv \delta(t_{\text{vir}}) = 1.686$ . This value represents the threshold that a perturbation in the initial density field has to exceed to be able to collapse and virialize.

The analytic function  $f(\nu)$  was first derived theoretically by Press & Schechter (1974), then formalised by the excursion set approach (Bond et al. 1991) and generalised to non-spherical model by Sheth & Tormen (1999). More accurate estimations were obtained by means of a full fitting to  $N$ -body simulation (e.g. Jenkins et al. 2001; Tinker et al. 2008, 2010).

### 2.2.1 Press-Schechter formalism

Press and Schechter (Press & Schechter 1974, PS) combined the linear growth of density fluctuation field and the spherical top-hat collapse model to analytically derive a prescription of the mass function. They assumed that the density perturbations collapse and virialize when their linear density contrast exceeds a critical threshold  $\delta_c$  (as already seen,  $\delta_c = 1.686$  for spherical collapse). They also assigned a Gaussian probability distribution to the smoothed density field  $\delta_M$  (where position and time dependences are left understood),

$$p(\delta_M) = \frac{1}{\sqrt{2\pi}\sigma_M} \exp\left(-\frac{\delta_M^2}{2\sigma_M^2}\right), \quad (2.17)$$

where  $M = 4\pi R^3 \bar{\rho}_m/3$  and  $\sigma_M$  is the square root of Eq. (1.39). Then, the probability that a given point in space has an overdensity greater than the critical one, namely  $\delta_M > \delta_c$ , is simply

$$p_{\delta_M > \delta_c}(M) = \int_{\delta_c}^{\infty} p(\delta_M) d\delta_M = \frac{1}{2} \operatorname{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_M}\right). \quad (2.18)$$

Here  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 2/\pi \int_x^{\infty} \exp(-t^2) dt$  is the complementary error function. The Press and Schechter argument takes the latter probability to be proportional to the probability of a point being part of a collapsed object of mass greater than  $M$  (or equivalently scale greater than  $R$ ). This means that the fraction of space in which  $\delta_M > \delta_c$  corresponds to the fraction of cosmic volume filled with haloes of mass greater than  $M$ . Considering the whole mass range and in the limit of  $M \rightarrow 0$ , one should be able to account for the whole mass in the Universe and obtain  $\int_0^{\infty} p_{\delta_M > \delta_c}(M) = 1$ . In reality, this formalism gives only  $\int_0^{\infty} p_{\delta_M > \delta_c}(M) = 1/2$ , meaning that half of the mass is miscounted. The problem of this approach is the so-called *cloud-in-cloud* problem: a point with  $\delta_M < \delta_c$  at a certain mass scale  $M$ , has zero probability of reaching later the condition of  $\delta_{M'} > \delta_c$ , at some larger scale  $M' > M$ . In other words, a non-collapsed object at scale  $M$  has no chance of being

later part of another collapsed object on some scale  $M' > M$ . This means that at a given epoch we are accounting only for the objects which have just collapsed, i.e. for which  $\delta_M = \delta_c$ . Press and Schechter solved the problem by adding by hand a corrective factor of 2, but later Bond et al. (1991) justified the missing factor of 2, with the theory of excursion sets.

By differentiating now Eq. (2.18) with respect to the mass  $M$ , we obtain

$$\frac{dn(M, z)}{dM} = \frac{2}{V} \frac{\partial p_{\delta_M > \delta_c}(M, z)}{\partial M} = -\sqrt{\frac{2}{\pi}} \frac{\bar{\rho}_m \delta_c}{M \sigma_M^2} \frac{d\sigma_M}{dM} \exp\left(-\frac{\delta_c^2}{2\sigma_M^2}\right), \quad (2.19)$$

where we have divided by  $V = M/\bar{\rho}_m$  to obtain a quantity with units per volume. The previous formula gives the comoving number density of collapsed objects of mass  $M$  at redshift  $z$ , in terms of  $\sigma_M$  and  $\delta_c$ : this is indeed the **Press-Schechter mass function**. Following the notation of Eq. (2.16), we can write

$$f_{\text{PS}}(\nu) = \sqrt{\frac{2}{\pi}} \nu \exp\left(-\frac{\nu^2}{2}\right). \quad (2.20)$$

The shape of this mass function and its evolution is in reasonable agreement with what is measured in numerical simulations of hierarchical clustering from Gaussian initial conditions (e.g. Lacey & Cole 1994), although it underpredicts the abundances of low mass haloes and over-estimates the number of high mass ones. Sheth & Tormen (1999); Sheth et al. (2001); Sheth & Tormen (2002) (ST), following the method of Bond et al. (1991), incorporated into the PS mass function the effect of a non-spherical collapse approximation. In particular, they adopted an ellipsoidal model, where the collapse of a region depends not only on the initial overdensity, but also on the surrounding shear field. Since the mass of a region is related to its initial size, in this model there is a relation between the density threshold value required for collapse and the mass of the final object. However, they needed a fit to numerical simulations in order to estimate the mass function parameters, as we will describe in the next Section.

### 2.2.2 N-body simulations calibration

As we just mentioned, the ST semi-analytical mass function, which generalised the PS one, is a cosmology dependent fitting formula, which can be expressed as

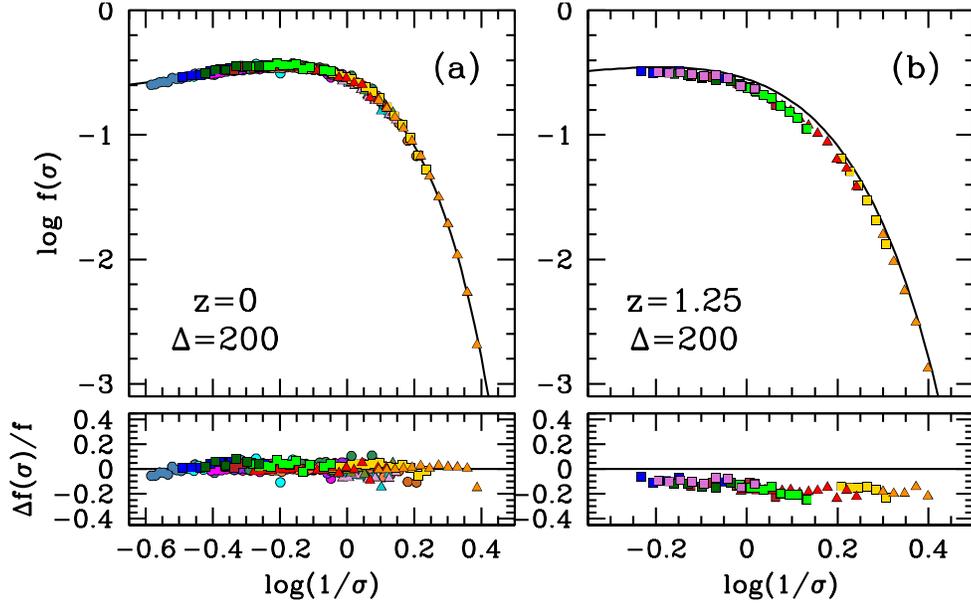
$$f_{\text{ST}}(\nu) = A_{\text{ST}} \sqrt{\frac{2\alpha_{\text{ST}}}{\pi}} \nu [1 + (\nu^2 \alpha_{\text{ST}})^{\beta_{\text{ST}}}] \exp\left(-\frac{\nu^2 \alpha_{\text{ST}}}{2}\right), \quad (2.21)$$

where the coefficients are obtained by a fit to GIF/Virgo Collaboration simulations of clustering (Kauffmann et al. 1999a). These simulations were performed with  $256^3$  particles, in two boxes of sizes  $L_1 = 85 \text{ Mpc}/h$  and  $L_2 = 141 \text{ Mpc}/h$ , for three cosmological models. They used a spherical overdensity (SO) group finder (Tormen 1998) to measure the mass function in the simulations. The best-fitting values they obtained are  $\alpha_{\text{ST}} = 0.707$  and  $\beta_{\text{ST}} = 0.3$ , while  $A_{\text{ST}} = 0.3222$  is derived assuming that all mass is collapsed into haloes (i.e. the integral of the mass function to infinity is equal to unity). The PS case follows easily from  $\alpha_{\text{ST}} = 1$ ,  $\beta_{\text{ST}} = 0$  and  $A_{\text{ST}} = 0.5$ . ST improve the analytical fit to N-body simulations results, but nevertheless remaining still poor in the agreement for rare high redshift haloes (Reed et al. 2007).

By means of larger and more elaborated N-body simulations, it has been found empirically that the mass function determined for a wide range of redshifts and cosmological models can be fitted accurately by a *universal* function (Jenkins et al. 2001; Evrard et al. 2002; Warren et al. 2006; Tinker et al. 2008). Some expressions, based on fits to simulation data, have been calculated and they agree at the 10–30%, with the largest discrepancy on the high mass tail. Jenkins et al. (2001) showed that the mass function of DM haloes from galaxies to clusters masses is quite well described by the ST function up to redshift  $z = 5$ , with some suppression at high masses. He analysed the results of the *Hubble Volume* simulation, a simulation of DM clustering in a cubic volume of size  $L = 3 \text{ Gpc}/h$ , with  $1024^3$  DM particles. This yields a DM particle mass of  $2.2 \times 10^{12} M_{\odot}$ , implying that a galaxy cluster halo typically contains 100-1000 particles. Despite the poor mass resolution, the very large volume allowed to explore the mass function on a broad range of masses, including the very high mass end, where clusters reside. They identified DM halos using the friends-of-friends algorithm (Davis et al. 1985). Jenkins proposed finally the following alternative analytic fitting formula to the simulations:

$$f_{\text{J}}(\sigma_{\text{M}}) = A_{\text{J}} \exp\left(-|\ln \sigma_{\text{M}}^{-1} + \alpha_{\text{J}}|^{\beta_{\text{J}}}\right), \quad (2.22)$$

where  $A_{\text{J}} = 0.301$ ,  $\alpha_{\text{J}} = 0.64$  and  $\beta_{\text{J}} = 3.82$ . Its accuracy is well tested by the *Millennium* simulation of Springel et al. (2005) and by the high precision mass function



**Figure 2.2:** Left panel: Tinker’s  $f(\sigma)$  at  $z = 0$  and for  $\Delta = 200$ , from all simulations. The solid line is the best fit function. The lower window shows the residuals with respect to the fitting function. Left panel:  $f(\sigma)$  at  $z = 1.25$  and for  $\Delta = 200$ . The lower window shows that the  $z = 1.25$  mass function is offset by  $\sim 20\%$  with respect to the  $z = 0$  one. Credit: Tinker et al. (2008)

at redshift zero measured by Warren et al. (2006). This last work was aimed to test the mass function over a wider range of mass scales than the one obtained from a single simulation. For this purpose, they simulated 16 boxes of different physical size but the same number of DM particles ( $1024^3$ ), nested in such a way that they defined a composite halo mass function covering five orders of magnitude in mass scale.

The current state-of-the-art halo mass function has been estimated by Tinker et al. (2008, 2010), who achieved a fit at the 5% precision level, for a  $\Lambda$ CDM cosmology. The simulations used to obtain this result were based on variants of the flat  $\Lambda$ CDM model, where the parameters referred to the first-year or three-year WMAP results (Spergel et al. 2003, 2007). They used fifty realizations of a simulation on a cubic box of  $L = 1280 \text{ Mpc}/h$  size, performed with the GADGET2 code (Springel 2005) and six simulations using the adaptive refinement technique of Kravtsov et al. (1997). They employed the standard SO algorithm Lacey & Cole (1994), but relocating the

centres of haloes at their density peaks, instead that on the centre of mass of the particles within the sphere. The results of all these simulations can be visualised in Fig. 2.2, showing the mass function best fit (solid line) at  $z = 0, 1.25$ , for  $\Delta = 200$ . The model they obtained, valid over wide redshift and mass ranges, has the following form:

$$f_T(\nu) = A_T \left[ 1 + (\beta_T \nu)^{-2\phi_T} \right] \nu^{2\eta_T+1} e^{-\gamma_T \nu^2/2}, \quad (2.23)$$

where  $A_T = 0.368$  and the other parameters evolve in redshift as

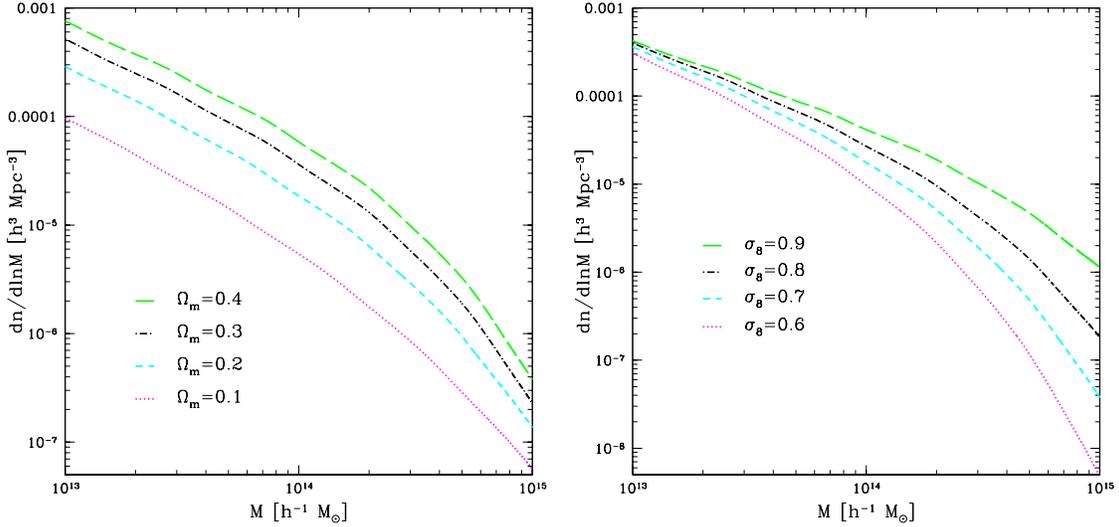
$$\begin{aligned} \beta_T &= 0.589 (1+z)^{0.20}, \\ \phi_T &= -0.729 (1+z)^{-0.08}, \\ \eta_T &= -0.243 (1+z)^{0.27}, \\ \gamma_T &= 0.864 (1+z)^{-0.01}. \end{aligned} \quad (2.24)$$

The above results are only valid for cluster mass at  $R_{200}$ , with an overdensity of  $\Delta = 200$  in units of the mean mass density of the Universe, which is the case we restrict to.

In none of the above calibrated mass functions, the effects of baryon physics is taken into account. An interesting analysis on the effect on the halo mass function caused by the inclusion of baryons has been performed by Cui et al. (2012). They employed two hydrodynamical simulations: one including radiative cooling, star formation and kinetic feedback from Supernovae, and a non-radiative simulation. These were based on the TreePM/SPH GADGET-3 code (Springel 2005), having a cubic volume of size  $L = 410 \text{ Mpc}/h$ , with  $2 \times 1024^3$  DM particles and using a SO halo finder algorithm. They obtained that the inclusion of baryons increases the mass of 1-2% at  $\Delta = 200$  and of 4-5% at  $\Delta = 500$ .

### 2.2.3 Cosmology dependence of the mass function

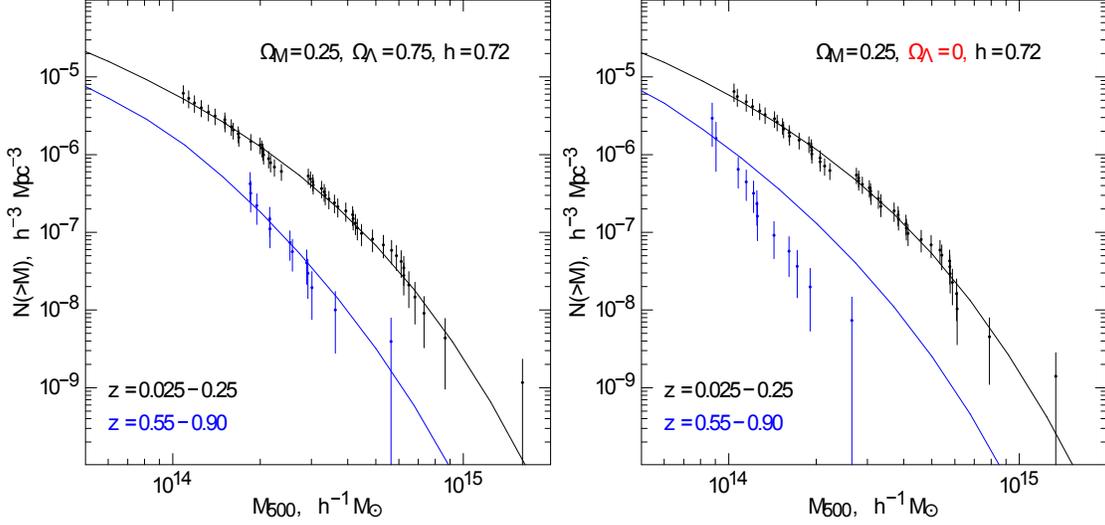
Galaxy clusters reside in the tail of the halo mass function and thus their numbers are exponentially sensitive to variations in cosmology, as evident from all the expressions for the cluster mass function that we have just shown (see Eqs. 2.20, 2.21, 2.22, 2.23). Thanks to this dependence, the mass function is a powerful probe of cosmological models: the cosmological parameters enter in the mass function through  $\sigma_M$ , which depends on the power spectrum and on the linear perturbation growth factor. In Fig. 2.3 we show the sensitivity of the Tinker mass function to variation in  $\Omega_m$  and  $\sigma_8$  parameters. For increasing values of  $\Omega_m$ , the number of clusters increases similarly on all mass scales. On the other hand, an increasing values of  $\sigma_8$  enhances



**Figure 2.3:** Left panel: mass function cosmological dependence on  $\Omega_m$ . Here we show the Tinker mass function at  $z = 0.2$  for  $\Omega_m = \{0.1, 0.2, 0.3, 0.4\}$  respectively in dotted magenta, dashed cyan, black dotted-dashed and green long-dashed curves and for fixed  $\sigma_8 = 0.8$ , assuming a flat Universe. Right panel: mass function cosmological dependence on  $\sigma_8$ . Here we show the Tinker mass function at  $z = 0.2$  for  $\sigma_8 = \{0.6, 0.7, 0.8, 0.9\}$  respectively in dotted magenta, dashed cyan, black dotted-dashed and green long-dashed curves and for fixed  $\Omega_m = 0.25$ , assuming a flat Universe.

mainly the number of massive clusters. If considering the mass function evolution in redshift, for a fixed mass, the higher the amplitude of mass fluctuations is, the more rapidly structures form at early times. On the other hand, the lower the amplitude of mass fluctuations is, the more slowly structures form at early times. Therefore, the normalisation of the power spectrum, has an evident influence on structure formation and evolution. Note that cluster counts and their redshift evolution are a powerful probe of cosmology, but nevertheless any systematics in the cluster mass measurements are enhanced by the steepness of the function itself.

Recently, the discovery of high redshift, massive clusters have led to possible tension with the  $\Lambda$ CDM model, because the cluster mass function is particularly sensitive to the cosmological parameters at high masses and redshifts. In Fig. 2.4 we show how high redshift clusters can help to discriminate between different cosmological models. Vikhlinin et al. (2009) presented the cosmological parameter constraints obtained from Chandra observations of high and low redshift clusters, detected in



**Figure 2.4:** Left panel: the measured mass function and predicted models for a nearly  $\Lambda$ CDM cosmology. Right panel: the data and the models are computed for a cosmology with  $\Omega_\Lambda = 0$ . The predicted number density of  $z > 0.55$  clusters is in strong disagreement with the data, and therefore the model with  $\Omega_m = 1$  and  $\Omega_\Lambda = 0$  has to be discarded. Credit: Vikhlinin et al. (2009)

X-ray by ROSAT survey. In particular, they used 37 clusters at  $\langle z \rangle = 0.55$  and 49 clusters at  $\langle z \rangle = 0.05$ . In the left panel, we show their measured mass function for a cosmology with  $\Omega_m = 0.25$  and  $\Omega_\Lambda = 0.75$ , which is reproduced for both the high- $z$  and low- $z$  clusters. In the right panel, the cosmology adopted is with  $\Omega_m = 0.25$  and  $\Omega_\Lambda = 0$  and the measured mass function is changed. The high- $z$  cluster number density is clearly in strong disagreement with the data, and therefore the model with  $\Omega_m = 0.25$ ,  $\Omega_\Lambda = 0.75$  has to be preferred to the  $\Omega_m = 0.25$ ,  $\Omega_\Lambda = 0$  one.

In addition to this, the high mass tail of the mass function, which corresponds to rare events, is sensitive to primordial non-Gaussianities, as we will investigate in Section 2.5.

## 2.3 Modelling cluster counts and total masses

We now are interested in predicting the number of observed galaxy clusters as a function of an observable  $M^{\text{obs}}$ , that we need to relate to the true mass  $M$  of the clusters via the scaling relation. In case of optical surveys, which is the case we restrict to, the observable is defined to be the richness  $N_{\text{gal}}^{\text{obs}}$ , i.e. the number of red galaxies within the radius  $R_{200}$  from the cluster centre. As already defined,  $R_{200}$  is the radius within which the average overdensity is  $\Delta = 200$  times the mean density of the Universe. We first consider the probability of observing  $N_{\text{gal}}^{\text{obs}}$  member galaxies at  $R_{200}$  for a given true mass  $M$  of the cluster. We can write this as:

$$p(N_{\text{gal}}^{\text{obs}}|M) = \int p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) p(N_{\text{gal}}|M) dN_{\text{gal}}, \quad (2.25)$$

where the distribution  $p(N_{\text{gal}}|M)$  is a delta function, because the relation between  $M$  and  $N_{\text{gal}}$  is the empirical mass-observable relation, which we will detail in Chapter 3 (see Eq. 3.11). We then assume  $p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}})$  to follow a log-normal distribution as suggested by Lima & Hu (2005)

$$p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) = \frac{1}{\sqrt{2\pi\sigma_{\ln N_{\text{gal}}^{\text{obs}}|M}^2}} \exp[-x^2(N_{\text{gal}}^{\text{obs}})], \quad (2.26)$$

where

$$x(N_{\text{gal}}^{\text{obs}}) = \frac{\ln N_{\text{gal}}^{\text{obs}} - \ln N_{\text{gal}}(M)}{\sqrt{2\sigma_{\ln N_{\text{gal}}^{\text{obs}}|M}^2}}, \quad (2.27)$$

and  $\sigma_{\ln N_{\text{gal}}^{\text{obs}}|M} \equiv \sigma_{\ln N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}}$  is the scatter around the mean  $N_{\text{gal}}(M)$  given by Eq. (3.11) (Lima & Hu 2005; Battye & Weller 2003): the equality between the two sigmas holds because  $p(N_{\text{gal}}|M)$  is a delta function. The cluster average number density within a richness bin  $[N_{\text{gal},i}^{\text{obs}}, N_{\text{gal},i+1}^{\text{obs}}]$  is then given by

$$\begin{aligned} n_i &= \int_{N_{\text{gal},i}^{\text{obs}}}^{N_{\text{gal},i+1}^{\text{obs}}} d \ln N_{\text{gal}}^{\text{obs}} \int d \ln N_{\text{gal}} \frac{dn}{d \ln N_{\text{gal}}} p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}) = \\ &= \int d \ln N_{\text{gal}} \frac{dn}{d \ln N_{\text{gal}}} \frac{1}{2} [\text{erfc}(x_i) - \text{erfc}(x_{i+1})], \end{aligned} \quad (2.28)$$

where  $p(N_{\text{gal}}^{\text{obs}}|N_{\text{gal}})$  is given by Eq. (2.26),  $x_i \equiv x(N_{\text{gal},i}^{\text{obs}})$  from Eq. (2.27). Note that

$$\frac{dn}{d \ln N_{\text{gal}}} = \frac{dn}{d \ln M} \frac{d \ln M}{d \ln N_{\text{gal}}}, \quad (2.29)$$

and for the mass function  $dn/d\ln M$ , defined in Eq. (2.16), we use the Tinker fit of Eq. (2.23). Hence, the total number of predicted galaxy clusters in the redshift range  $[z_{\min}, z_{\max}]$  and within the richness bin  $[N_{\text{gal},i}^{\text{obs}}, N_{\text{gal},i+1}^{\text{obs}}]$  can be calculated as

$$\Delta N_i = \Delta\Omega \int_{z_{\min}}^{z_{\max}} dz \frac{d^2V}{dz d\Omega} n_i, \quad (2.30)$$

where  $\Delta\Omega$  is the survey sky coverage and  $d^2V/dz/d\Omega$  is the comoving volume element. In the above formula, cosmology enters through both the mass function and the volume element. We will show the sensitivity to cosmology, in particular to  $\Omega_m$  and  $\sigma_8$ , of the number counts in the results of our analysis.

Similar to Eqs. (2.28) and (2.30), we can compute the total mass contained in the clusters. The average total mass contained in clusters within a bin can be obtained by weighting the integrand of Eq. (2.28) by  $m_i$ , i.e. the mean mass of clusters within the bin, estimated via the mass-observable relation:

$$(nm)_i = \int d\ln N_{\text{gal}} \frac{dn}{d\ln N_{\text{gal}}} \frac{1}{2} m_i [\text{erfc}(x_i) - \text{erfc}(x_{i+1})]. \quad (2.31)$$

The total mass of clusters in the redshift range  $[z_{\min}, z_{\max}]$  and within the richness bin  $[N_{\text{gal},i}^{\text{obs}}, N_{\text{gal},i+1}^{\text{obs}}]$  is thus

$$(\Delta N \bar{M})_i = \beta \Delta\Omega \int_{z_{\min}}^{z_{\max}} dz \frac{d^2V}{dz d\Omega} (nm)_i. \quad (2.32)$$

The parameter  $\beta$  is a nuisance parameter introduced to account for possible mismatch with the WL mass estimates, as discussed in Rozo et al. (2010): we will make this point more clear in Section 3.4.1.

To conclude, for our analysis we will need both Eqs. (2.30) and (2.32) and the respective binned data sets, in order to fit the richness-mass relation and constrain cosmological parameters simultaneously.

## 2.4 Clustering of clusters

Galaxy clusters are tracers of the LSS of the Universe (Mo et al. 1996), as they select the highest-density regions of the dark matter overdensity field, namely the **dark matter haloes**. The overdensity of the latter is related to the underlying dark matter distribution by the **halo bias** parameter, whose derivation from the mass function is provided here. We then give a prescription for the cluster two-points statistics (power spectrum) calculation. As the effect of baryons is negligible for the clustering properties of the clusters, in the following we use the naming 'cluster' and 'halo' interchangeably. This simplification is an advantage with respect to the clustering of galaxies, where additional assumption on the halo occupation distribution must be made to model correctly the galaxy clustering.

### 2.4.1 Halo bias

Fluctuations in the cluster number density on large scales are more pronounced than the fluctuations of the underlying matter density field (Bahcall & Soneira 1983; Klypin & Kopylov 1983): the ratio between the two is known as the **bias** parameter. Equivalently, the squared bias can be defined as the ratio between the halo power spectrum and the linear matter power spectrum. The halo bias can be derived from a theory of the mass function by using the peak-background split formalism (Cole & Kaiser 1989; Mo & White 1996), in which the density perturbations are split between long-wavelength  $\delta_l$  and short-wavelength  $\delta_s$  modes. This method gives a prediction for the halo bias in Lagrangian space  $b^L(\vec{q}, M)$ , that at linear order looks as

$$\delta_h^L(\vec{q}, M) = b_0^L + b^L(\vec{q}, M) \delta(\vec{q}) . \quad (2.33)$$

The Lagrangian coordinate  $\vec{q}$  represents the position in space of the initial conditions, which will move to the position  $\vec{x}$  after evolution. Assuming linear Gaussian perturbations, spherical collapse model and neglecting a large-scale velocity bias, the Lagrangian bias can be related to the Eulerian one by (Mo & White 1996)

$$b = b^E = 1 + b^L , \quad (2.34)$$

leaving out the known dependences. The bias parameter can be assumed to be scale-independent (except for the modifications in the presence of primordial non-Gaussianity), as long as only linear scales are considered.

Cluster bias can be seen also as a variation of  $\delta_c$  parameter by long-wavelength density modes (Kaiser 1984; White et al. 1987). According to the peak-background

split model, the long modes modulate the overall amplitude, so that the halo overdensity is given by (Cole & Kaiser 1989; Mo & White 1996)

$$\delta_h^L(\vec{q}, M) = \frac{n[M, \delta_1(\mathbf{q})]}{\bar{n}} - 1 = \frac{f\left[\frac{\delta_c - \delta_1(\vec{q})}{\sigma(M)}\right]}{f\left[\frac{\delta_c}{\sigma(M)}\right]} - 1, \quad (2.35)$$

where the second equality follows from expressing the number of halos  $n$  and their mean  $\bar{n}$  using the mass function  $f(\delta_c/\sigma) \equiv f(\nu)$ . The formula for the Lagrangian bias at any order then follows by expanding this equation in a Taylor series of  $\delta_1$ . For example, at linear order, the Lagrangian bias as a function of mass  $M$  and redshift  $z$  is simply

$$b^L(M, z) = -\frac{f'(\nu)}{f(\nu)\sigma}, \quad (2.36)$$

where the derivative of the mass function is taken with respect to  $\nu$ , and the mass and redshift dependences of  $\nu$  and  $\sigma$  are implicit. The linear Eulerian bias is then obtained by simply adding unity. Note that small values of  $\sigma(M)$  (large values of  $M$ ) correspond to large values of  $b(M)$ : this means that rare, high-mass objects are more likely to be observed in regions where the underlying matter density is much higher than the average one.

Considering the PS mass function of Eq. (2.20), the Eulerian bias function takes the form

$$b_{\text{PS}}(\nu) = 1 + \frac{\nu^2 - 1}{\delta_c} = 1 - \frac{1}{\delta_c} + \frac{\delta_c}{\sigma^2}. \quad (2.37)$$

However, the PS bias does not match properly the simulations (Sheth & Tormen 1999; Jing 1999), due to the already discussed discrepancy between the PS mass function and simulated data. When using the ST mass function of Eq. (2.21), instead, we recover the following improved expression (Sheth et al. 2001):

$$b_{\text{ST}}(\nu) = 1 + \frac{1}{\sqrt{a}\delta_c} \left[ \sqrt{a}(a\nu^2) + \sqrt{ab}(a\nu^2)^{1-c} - \frac{(a\nu^2)^c}{(a\nu^2)^c + b(1-c)(1-c/2)} \right], \quad (2.38)$$

where  $a = 0.707$ ,  $b = 0.5$  and  $c = 0.6$ . Finally, by plugging in the Tinker's mass function of Eq. (2.23) and keeping the leading order terms, the Eulerian bias is (Tinker et al. 2010):

$$b_{\text{T}}(M, z) \simeq 1 + \frac{\hat{\gamma}\nu^2 - (1 + 2\hat{\eta})}{\delta_c} + \frac{2\hat{\phi}/\delta_c}{1 + [\hat{\beta}\nu]^{2\hat{\phi}}}, \quad (2.39)$$

where all parameters are defined in Eq. (2.24).

We can finally calculate the average cluster bias  $\bar{b}$  over a certain mass range by weighting the bias function with the mass function (Lima & Hu 2005):

$$\bar{b}(z) = \frac{1}{\Delta N} \int_{M_{\min}}^{M_{\max}} d \ln M \frac{dn(M, z)}{d \ln M} b(M, z) , \quad (2.40)$$

where the normalisation factor  $\Delta N$  is the full integral of the mass function in the observed range and where we restrict to the Tinker's case, namely  $b(M, z) = b_{\text{T}}$  through all our analysis.

### 2.4.2 Cluster power spectrum

Clustering is analysed statistically by the two-point correlation function, which is defined as the variance of the density field in two points separated by a distance  $\vec{r}$ :

$$\xi(\vec{x}, \vec{r}) \equiv \langle \delta(\vec{x} + \vec{r}) \delta(\vec{x}) \rangle . \quad (2.41)$$

By assuming homogeneity and isotropy conditions,  $\xi$  depends only on the modulus of the distance between the two points, i.e.  $\xi(|\vec{r}|) \equiv \xi(r)$ . The corresponding function in Fourier space is the power spectrum. We also know that for a DM halo it holds

$$\tilde{\delta}_h(\vec{k}, M) = \tilde{b}(\vec{k}, M) \tilde{\delta}(\vec{k}) , \quad (2.42)$$

thus in linear approximation the halo-halo power spectrum  $P_{hh}$  can be related to the linear matter power spectrum  $P_{\text{lin}}$  as

$$P_{hh}(k, M, z) = b^2(M, z) P_{\text{lin}}(k, z) = b^2(M, z) D^2(z) P_{\text{lin}}(k, 0) , \quad (2.43)$$

where  $D(z)$  is the linear growth function and  $P_{\text{lin}}(k, 0) = P_{\text{in}}(k) T^2(k)$ ,  $P_{\text{in}}(k)$  being the primordial power spectrum and  $T(k)$  the transfer function (see Eq. 1.41). In order to integrate the mass dependence, we need to weight the bias as described in Eq. (2.40) and compute all quantities at a fixed redshift (which in our case is the mean redshift of the cluster sample,  $\bar{z} \simeq 0.2$ ). This is further justified by observing that the growth of  $\bar{b}(z)$  is compensated by a similar drop in  $D(z)$ ; we have checked that for our fiducial cosmology, in the observed range  $0.1 \leq z \leq 0.3$  the variation of  $\bar{b}(z) D(z)$  is at the percent level.

Furthermore, to be able to correctly fit models to the data, the following five effects have to be taken into account (Hütsi 2010):

- the convolution with the survey window, which suppresses the power on scales comparable (and larger) to the survey size;
- the photo-z errors, which are responsible for a damping of the spectrum on small scales;
- the weak non-linearities, which add power on small scales;
- the redshift-space distortions (RSD) contribution;
- the Alcock-Paczynski effect.

We first define the full power spectrum  $P_{\text{NL}}$  to be

$$P_{\text{NL}}(k) = (b^{\text{obs}})^2 (1 + q_{\text{NL}} k^{3/2}) s(k) P_{\text{lin}}(k) \left[ 1 + \frac{2}{3} \beta_z + \frac{1}{5} \beta_z^2 \right]. \quad (2.44)$$

Here the bias is rescaled as  $b^{\text{obs}} = \bar{b} B$ , where we include a nuisance parameter  $B$  to represent the uncertainty on the bias derived from the mass function. The second term in Eq. (2.44) models the effect of residual weak non-linearities with a simple effective fitting function with one free parameter  $q_{\text{NL}}$ , considering the fact that in our analysis we only use data up to  $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$ . This form is similar to the one of Cole et al. (2005), but the index 3/2 is found by Hütsi (2006c) to be a better fit. We then model the photo-z smoothing with the corrective factor

$$s(k) = \left( \frac{\sqrt{\pi}}{2 \sigma_z k} \right) \text{erf}(\sigma_z k), \quad (2.45)$$

assuming that photo-z errors follow a Gaussian distribution with dispersion  $\delta z$  and corresponding spatial smoothing scale  $\sigma_z = \delta z c / H_0$ . The last term of Eq. (2.44) is the linear correction due to redshift space distortions, for which we assume (Kaiser 1987)

$$\beta_z(\bar{z}) \simeq \Omega_m^{0.55}(\bar{z}) / b^{\text{obs}}(\bar{z}). \quad (2.46)$$

These distortions arise because of the cluster galaxies peculiar velocities: the redshift space power spectrum is modified with respect to the power spectrum in the real space. We have checked that the RSD correction changes at most at the percent level if we calculate it at the limits of our redshift range.

We finally take into account the Alcock-Paczynski effect (Alcock & Paczynski 1979): we rescale the full theoretical power spectrum with respect to the cosmology

used to convert redshifts to distances in the measurements (denoted by the superscript ‘fid’), assuming a single isotropic dilation applies (Eisenstein et al. 2005; Hütsi 2006c), i.e.

$$\tilde{P}^{\text{obs}}(k) = \frac{1}{c_{\text{isotr}}^3} P^{\text{obs}}\left(\frac{k}{c_{\text{isotr}}}\right). \quad (2.47)$$

Here  $c_{\text{isotr}} = (c_{\parallel} c_{\perp}^2)^{1/3}$ ,  $c_{\parallel} = H^{\text{fid}}/H$ ,  $c_{\perp} = D_A/D_A^{\text{fid}}$ ,  $D_A$  is the angular diameter distance of Eq. (1.26). All quantities are calculated at the mean redshift  $\bar{z}$ .

After including the convolution with the survey window, the total observed power spectrum  $P^{\text{obs}}$  is modelled as

$$P^{\text{obs}}(k) = \int d \ln \kappa \kappa^3 P_{\text{NL}}(\kappa) K(\kappa, k). \quad (2.48)$$

Here the kernel  $K(\kappa, k)$ , accounting for the effect of the finite survey area, can be described by the analytic fit of Hütsi (2010, 2006a):

$$K(\kappa, k) = K(k, \kappa) = \frac{c}{k\kappa} [g(k + \kappa) - g(k - \kappa)], \quad (2.49)$$

where

$$g(y) = \text{atan}\left(\frac{e^4 + 2d^2y^2}{e^2\sqrt{4d^4 - e^4}}\right). \quad (2.50)$$

By imposing the condition  $\int K(\kappa, k)\kappa^2 d\kappa = 1$  one obtains the normalisation constant

$$c = \frac{1}{\pi e \sqrt{2 - (e/d)^2}}. \quad (2.51)$$

In the case of the maxBCG catalogue survey geometry, the best fitting parameters are  $d = 0.0044$  and  $e = 0.0041$ .

## 2.5 Primordial non-Gaussianity

The simplest single-field, slow-roll model for inflation gives rise to nearly Gaussian distributed curvature perturbations (Maldacena 2003; Acquaviva et al. 2003). However, many generalizations, such as multi-field models, predict evidence for large **primordial non-Gaussianities** (PNG, Bartolo et al. 2004). In general, non-Gaussianity arises from any source of non-linearities. Free fields having no interactions have a precisely Gaussian wave function. In the case of a single scalar field  $\Phi$  (inflaton field), the equation of motion is given by

$$\ddot{\Phi} + 3H\dot{\Phi} + V'(\Phi) = 0 , \quad (2.52)$$

where  $V'(\phi) = dV(\Phi)/d\Phi$  is linear and the potential  $V(\Phi)$  can be expressed as

$$V(\Phi) = V_0 + \frac{1}{2}m^2\Phi^2 , \quad (2.53)$$

with a constant  $V_0$  and mass  $m$ . This leads to a linear system, with no trace of non-linearities, implying that if  $\Phi$  is Gaussian, it cannot gain deviations from Gaussianity. Self-interactions of the field  $\Phi$  are defined to be the higher orders terms in the potential, i.e.

$$V(\Phi) = V_0 + \frac{1}{2}m^2\Phi^2 + \frac{\mu}{3}\Phi^3 + \frac{\lambda}{4}\Phi^4 + \dots , \quad (2.54)$$

which cause  $V'$  to be non-linear anymore and thus giving rise to non-linearities. However these self-interactions cannot be large for the inflation to happen: they need to satisfy the slow-roll condition

$$\epsilon = -\frac{H}{H^2} \simeq \frac{1}{2} \left( \frac{V'}{V} \right)^2 \ll 1 , \quad (2.55)$$

which implies a severe limit on the flatness of the potential and for  $H(t)$  to vary slowly. The curvature perturbations are then related to the scalar field perturbations  $\delta\Phi$ : even if  $\delta\Phi$  is Gaussian, non-Gaussianities could be detected, because the metric perturbations are not linear in  $\delta\Phi$ , even if they are always of the order of  $\epsilon$ . To summarise, Gaussian quantum fluctuations by means of self-interactions generate non-Gaussian fluctuations in the inflaton field. The latter produce non-Gaussian curvature perturbations, which then generate non-Gaussian signature in the large-scale structures.

PNG would have multiple observable consequences, of which we will here consider only two: the halo mass function changes as a function of the non-zero (positive) skewness (LoVerde et al. 2008; Pillepich et al. 2010), and the halo bias becomes strongly scale-dependent due to the coupling of long- and short-wavelength modes (Dalal et al. 2008; Matarrese & Verde 2008; Slosar et al. 2008; Afshordi & Tolley 2008; Desjacques et al. 2009; Giannantonio & Porciani 2010; Desjacques & Seljak 2010; Schmidt & Kamionkowski 2010; Desjacques et al. 2011). As a consequence on the structure growth, we expect more massive collapsed structures at high redshift, because of the change in the timing of structure formation, and shape and evolution of the mass function, and a scale dependence on the linear large-scale bias. Therefore, the effects of PNG are mainly visible at high redshifts, high masses objects and on the large scale power spectrum.

Constraints on PNG can be investigated by measurements of abundances and clustering of galaxy clusters. This is the main motivation that leads us to extend our model to constrain primordial non-Gaussianity from both bias and abundances of the maxBCG clusters sample.

### 2.5.1 Definition of $f_{\text{NL}}$ parameter

We first introduce some useful notation, following Giannantonio & Porciani (2010). The  $f_{\text{NL}}$  parameter quantifies the amount of PNG in the simplest local, scale-independent case as

$$\Phi(\vec{x}, z_*) = \varphi(\vec{x}, z_*) + f_{\text{NL}} * [\varphi^2(\vec{x}, z_*) - \langle \varphi^2 \rangle(z_*)] , \quad (2.56)$$

where  $\Phi$  is the Bardeen's potential at redshift  $z_*$  and position  $\vec{x}$ ,  $\varphi$  is an auxiliary Gaussian potential with same dependences, and  $*$  represents convolution (which reduces to multiplication in case of constant  $f_{\text{NL}}$ ). Throughout this work, we define  $f_{\text{NL}}$  at early times ( $z_* \approx 1100$ ), according to the CMB convention (opposite to the LSS one, which refers instead to  $z = 0$ ). We then assume the local shape for the bispectrum

$$B_\varphi(\vec{k}_1, \vec{k}_2, \vec{k}_3) = 2 f_{\text{NL}} B_s^2 (k_1^{n'} k_2^{n'} + k_1^{n'} k_3^{n'} + k_2^{n'} k_3^{n'}) , \quad (2.57)$$

where  $B_s$  is the amplitude of  $P_\varphi(\vec{k})$  and  $n' = n_s - 4$ ,  $n_s$  being the index of the primordial density fluctuation spectrum  $P(k) = A_s k^{n_s}$ . The relation between the spectral amplitudes  $A_s$  and  $B_s$  is given by

$$\frac{B_s}{A_s} = \frac{9}{4} H_0^4 \Omega_m^2 .$$

At leading order in  $f_{\text{NL}}$  and neglecting trispectrum corrections, the potential power spectrum can be approximated by its Gaussian part:

$$P_{\Phi}(k) \simeq P_{\varphi}(k) . \quad (2.58)$$

The matter perturbations are related to the primordial potential by the Poisson equation:

$$\tilde{\delta}(\vec{k}, z) = \alpha(\vec{k}, z) \tilde{\Phi}(\vec{k}, z_*) , \quad (2.59)$$

where

$$\alpha(k, z) = \frac{2}{3} \frac{c^2 k^2 T(k) D(z)}{\Omega_{\text{m}} H_0^2} \frac{g(0)}{g(z_*)} . \quad (2.60)$$

Here  $T(k)$  is the transfer function,  $D(z)$  the linear growth function,  $c = 300000$  [km/s] is the speed of light and  $g(z) \propto (1+z)D(z)$ . We can then write for the matter power spectrum  $P$  as follows

$$P(k, z) = \alpha^2(k, z) P_{\Phi}(k, z_*) \simeq \alpha^2(k, z) P_{\varphi}(k, z_*) \quad (2.61)$$

and, as we consider linear theory only, we assume throughout  $P = P_{\text{lin}}$ .

## 2.5.2 Modified mass function

The halo mass function is modified in the presence of PNG as it gains a dependence on the skewness. We use the LoVerde et al. (2008) mass function (LV), which is given by

$$f_{\text{LV}}(\nu) = \sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}} \left[ \nu + S_3 \frac{\sigma}{6} (\nu^4 - 2\nu^2 - 1) + \frac{dS_3}{d \ln \sigma} \frac{\sigma}{6} (\nu^2 - 1) \right] , \quad (2.62)$$

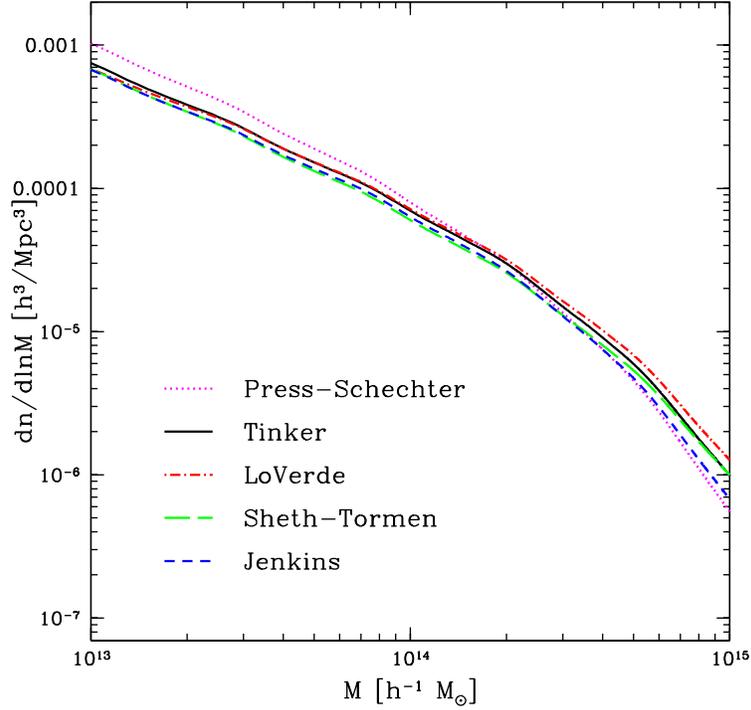
where  $\nu = \delta_c / \sigma$  and  $\sigma$  is defined in Eq. (1.39).  $S_3$  is the normalised skewness of the matter density field (the mass dependence is implicit), defined as

$$S_3 = \frac{\mu_3}{\sigma^4} = \frac{\langle \delta_{\text{M}}^3 \rangle}{\langle \delta_{\text{M}}^2 \rangle^2} , \quad (2.63)$$

in terms of the third-order momentum  $\mu_3$  of the smoothed density field.

$$\mu_3 \equiv \langle \delta_{\text{M}}^3 \rangle = \int \frac{d^3 k_1}{(2\pi)^3} \frac{d^3 k_2}{(2\pi)^3} \frac{d^3 k_3}{(2\pi)^3} \langle \delta_{\text{M}}(\vec{k}_1) \delta_{\text{M}}(\vec{k}_2) \delta_{\text{M}}(\vec{k}_3) \rangle , \quad (2.64)$$

and  $dS_3/d \ln \sigma$  is its first derivative with respect of  $\ln \sigma$ , defined as in Desjacques et al. (2009). To improve the agreement with  $N$ -body simulations, and for consistency with



**Figure 2.5:** We show a comparison between various mass functions, computed for our fiducial cosmology at  $z = 0$ . Press-Schechter (magenta dotted), Sheth-Tormen (green long-dashed), Jenkins (blue dashed), Tinker (black solid) and LoVerde (red dotted-dashed) for  $f_{\text{NL}} = 200$ .

the rest of our analysis, we replace the Gaussian limit of the mass function from the Press-Schechter form to the Tinker one, so that we will use the rescaled form defined as:

$$\tilde{f}_{\text{LV}}(\nu) \equiv \frac{f_{\text{T}}(\nu)}{f_{\text{PS}}(\nu)} f_{\text{LV}}(\nu) , \quad (2.65)$$

where  $f_{\text{PS}}$  and  $f_{\text{T}}$  are respectively given by Eqs. (2.20) and (2.23). In Fig. 2.5 we show a comparison between all the mass functions introduced so far, computed at  $z = 0$  for the fiducial cosmology of our analysis: PS (magenta dotted), ST (green long-dashed), Jenkins (blue dashed), Tinker (black solid) and LV (red dotted-dashed) mass function for  $f_{\text{NL}} = 200$ .

### 2.5.3 Modified bias

By applying the peak-background split formalism, we analytically derive the Lagrangian linear halo bias associated to the LV mass function of Eq. (2.62), using Eq. (2.36):

$$b_{\text{LV}}^{\text{L}}(\nu) = \frac{\delta_c}{\sigma^2} - \frac{1}{\sigma} \frac{6 + S_3 \sigma (4\nu^3 - 4\nu) + 2 \frac{dS_3}{d \ln \sigma} \sigma \nu}{6\nu + S_3 \sigma (\nu^4 - 2\nu^2 - 1) + \frac{dS_3}{d \ln \sigma} \sigma (\nu^2 - 1)}. \quad (2.66)$$

It follows that the Lagrangian bias associated to the rescaled mass function of Eq. (2.65) is hence

$$\tilde{b}_{\text{LV}}^{\text{L}}(\nu) = -\frac{1}{\sigma} \frac{\partial \tilde{f}_{\text{LV}}(\nu)}{\partial \nu} = b_{\text{LV}}^{\text{L}}(\nu) + b_{\text{T}}^{\text{L}}(\nu) - b_{\text{PS}}^{\text{L}}(\nu), \quad (2.67)$$

where  $b_{\text{LV}}^{\text{L}}$  is given in Eq. (2.66), while  $b_{\text{T}}^{\text{L}}$  and  $b_{\text{PS}}^{\text{L}}$  are derived from the Eulerian biases of Eqs. (2.39) and (2.37) by subtracting unity.

In the presence of PNG, the halo density perturbations depend not only on the dark matter perturbations  $\delta$ , but also on the potential  $\varphi$ . The latter can then be related back to the density in Fourier space by using the Poisson equation, so that the effective Eulerian bias can be written at a fixed redshift  $\bar{z}$  as

$$b_{\text{eff}}^{\text{obs}}(M, k, f_{\text{NL}}) = b(M, f_{\text{NL}}) + \Delta b(M, k, f_{\text{NL}}), \quad (2.68)$$

where the bias contains implicitly the following two corrections:

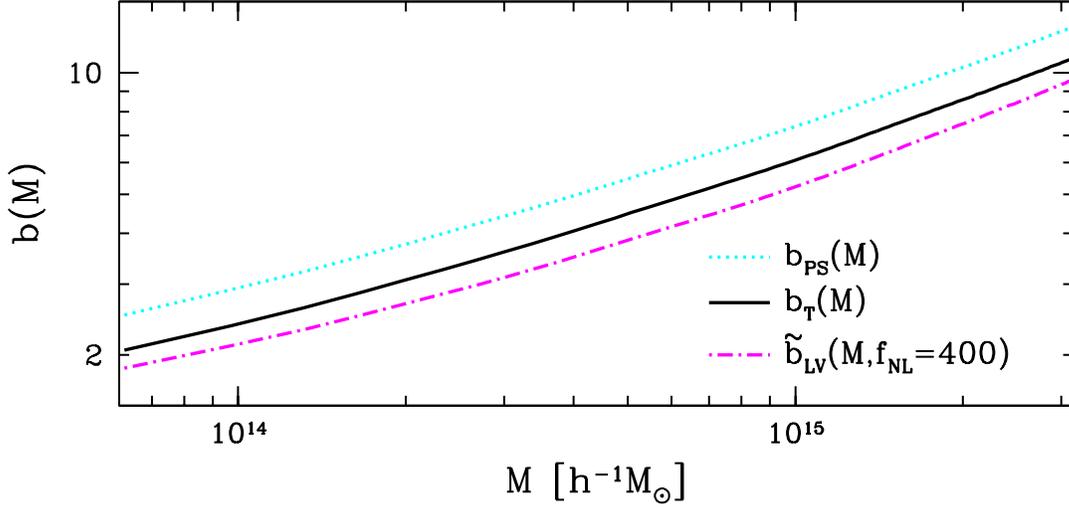
- (i) a scale-independent correction with respect to the Gaussian case, following from the difference in the mass function, given by

$$\delta b(f_{\text{NL}}) \equiv b(M, f_{\text{NL}}) - b(M, 0); \quad (2.69)$$

- (ii) a scale-dependent correction, given by

$$\Delta b(M, k, f_{\text{NL}}) \equiv \frac{2 f_{\text{NL}} \delta_c b^{\text{L}}(M, f_{\text{NL}})}{\alpha(k, \bar{z})}. \quad (2.70)$$

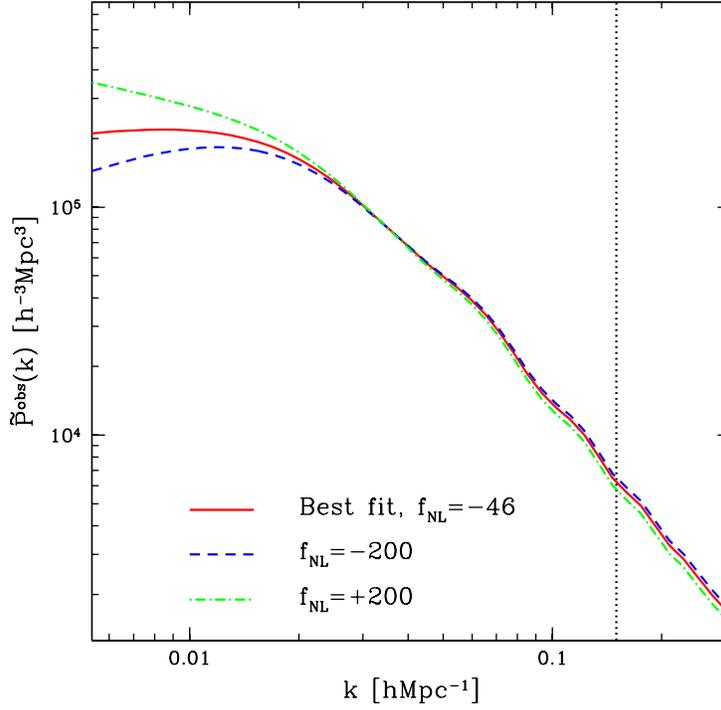
Fig. 2.6 shows the linear Tinker halo bias  $b_{\text{T}}(M)$  as a function of halo mass  $M$  at  $\bar{z} = 0.2$ , compared with the PS case, and with the scale-independent part of the bias in the presence of primordial non-Gaussianity described in Eq. (2.68): the cosmology is fixed to the best-fit model for our analysis. As in the Gaussian case, we will then



**Figure 2.6:** Mass dependence of the linear halo bias at  $\bar{z} = 0.2$  for three mass functions: Press-Schechter (cyan dotted), Tinker (black solid) and modified LoVerde mass function in the presence of primordial non-Gaussianity (magenta dot-dashed), with  $f_{\text{NL}} = 400$ . The cosmological model is fixed at the fiducial model in our analysis.

average the bias over the masses in our catalogue, following Eq. (2.40). To include the uncertainty on the assumption of a mass function, we introduce a nuisance parameter  $B$ , which rescales the bias as  $b^{\text{obs}} = \bar{b} B$ .

The scale-independent correction  $\delta b(f_{\text{NL}})$  is small, easily confused with other normalisation effects, and relies on the assumed form of the mass function and the peak-background split method. For these reasons, it is worth ensuring that the results do not depend on this contribution. We make sure this happens in our case because any constant rescaling of the bias can be equally explained by either a change in the nuisance parameter  $B$  or a change in  $f_{\text{NL}}$ . But since a model with  $f_{\text{NL}} \neq 0$  also predicts the scale-dependent bias, it will be favoured only in case such a feature is indeed observed in the data, otherwise the  $B \neq 1$  model will be assigned a better likelihood. In practice, we impose a Gaussian prior centred on  $B = 1$  (details in Chapter 4), but we have checked that the results on  $f_{\text{NL}}$  do not depend significantly on this choice.



**Figure 2.7:** The effect of PNG on the cluster power spectrum. We compare here the predictions for the best-fit model we obtain with  $f_{\text{NL}} = -46$  (red solid) and for two cases with  $f_{\text{NL}} = -200$  (blue dashed) and  $f_{\text{NL}} = 200$  (green dot-dashed). The dotted line at  $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$  represents the smallest scale we use in the analysis.

Finally, we show in Fig. 2.7 the full power spectrum  $\tilde{P}^{\text{obs}}(k)$  in the presence of PNG for a choice of  $f_{\text{NL}}$  values. High and positive values of  $f_{\text{NL}}$  increase evidently the power spectrum on large scales, while weakly suppressing it on small scales. The scale-dependent bias induced by PNG impacts on large scales, while the smaller scale-independent contribution affects the small scales. The survey window convolution partially suppresses the effect of PNG on the largest scales, which become comparable with the survey volume.

# Chapter 3

## Observations, data and errors

This Chapter is dedicated to the description of the data sets and error estimates we use in our combined cosmological analysis. First, we describe briefly galaxy cluster observations through the whole electromagnetic spectrum. We list the most important X-ray, millimetre (SZ), weak lensing and optical surveys of past, present and future times, as well as the cosmological constraints obtained from previous works with cluster catalogue. A particular focus is given to the Sloan Digital Sky Survey maxBCG optical catalogue, which is used in this work, and to the data sets we derived from it. We first describe the cluster counts and their covariance matrix. Secondly, we concentrate on the weak lensing mass estimates, the richness-mass scaling relation and the cluster total masses. Finally, the description of the redshift space cluster power spectrum is provided, together with its covariance error matrix. We end with some considerations on the cross-covariance matrix between counts and power spectrum.

### 3.1 Multi-wavelength surveys of galaxy clusters

Clusters have been detected across multiple wavelengths with varying degrees of success. To date, a few hundreds have been observed in the millimetre, few thousands in the X-ray, many tens of thousands in the optical. The efficiency of a survey is due to the combination of the technical properties of the detection instrument (e.g. flux sensitivity and angular resolution) and the physical features of the observed objects (e.g. intrinsic luminosity and redshift). Surveys are thus characterised by completeness (namely amount of objects that should have been detected), and purity (i.e. contamination due to spurious detected objects). This section includes a description

of the the various detection methods as well as the main galaxy cluster surveys and the derived catalogues.

### 3.1.1 X-ray surveys

The peculiar X-ray extended emission from the hot intra-cluster medium in galaxy clusters is relatively simple to identify. However, X-ray surveys are limited in surface brightness: too diffuse or too faint sources are missed. The deep knowledge of the X-ray emission from galaxy clusters enables to predict exactly the completeness and the purity of the survey, by means of extensive numerical simulations. The first X-ray cluster catalogues (Edge et al. 1990; Gioia et al. 1990) were derived from Ariel V<sup>1</sup> and HEAO-1<sup>2</sup> all-sky surveys and made use of the Einstein Observatory and EXOSAT<sup>3</sup> (Lahav et al. 1989). Later, the ROSAT satellite All-Sky Survey (RASS<sup>4</sup>, Voges et al. 1999) provided data used to create few catalogues: e.g. the ROSAT-ESO Flux Limited X-ray (REFLEX) galaxy clusters survey (Böhringer et al. 2004); the HIFLUGCS (Reiprich & Böhringer 2002) and the Massive Cluster Survey (MACS, Ebeling et al. 2010). Another catalogue of serendipitously observed clusters at high  $z$  is the ROSAT Deep Cluster Survey (RDCS, Rosati et al. 1998). Later NASA's Chandra X-ray Observatory<sup>5</sup> and ESA X-ray Multi-Mirror Mission Newton (XMM-Newton<sup>6</sup>) represented a breakthrough in X-ray observations, thanks to the improved angular resolution ( $\theta \sim 0.5''$ ), and the high sensitivity due to the large collecting area, respectively.

### 3.1.2 SZ surveys

Galaxy clusters can distort the spectrum of the CMB by the SZ effect, which is observed at millimetre wavelengths. SZ surveys are able to detect high-redshift, massive clusters, due to the redshift-independent SZ signature on the CMB: they can potentially observe the most distant clusters in the Universe, being only limited in the intrinsic SZ signal (related to the mass) and not in flux. However, the current mm telescopes have a poor spatial resolution ( $\theta \sim 60''$ ), being unable to resolve substructures or possible contaminating sources. The current operating SZ surveys are the Atacama

---

<sup>1</sup><http://heasarc.gsfc.nasa.gov/docs/ariel5/ariel5.html>

<sup>2</sup><http://heasarc.gsfc.nasa.gov/docs/heao1/heao1.html>

<sup>3</sup><https://heasarc.gsfc.nasa.gov/docs/exosat/exosat.html>

<sup>4</sup><http://www.xray.mpe.mpg.de/cgi-bin/rosat/rosat-survey>

<sup>5</sup><http://chandra.harvard.edu/>

<sup>6</sup><http://xmm.esac.esa.int/>

Cosmology Telescope (ACT<sup>7</sup>, Kosowsky 2006; Marriage et al. 2011; Menanteau et al. 2013), the South Pole Telescope (SPT<sup>8</sup>, Vanderlinde et al. 2010; Carlstrom et al. 2011; Reichardt et al. 2013) and Planck<sup>9</sup> satellite (Planck Collaboration et al. 2011).

### 3.1.3 WL surveys

Clusters have also been detected as peaks in the shear field of deep weak gravitational lensing surveys (Wittman et al. 2001, 2003; Dahle et al. 2003). This method directly detects the projected surface mass density of the clusters, but introducing relevant noise in the mass estimation for individual systems (White & Kochanek 2002; de Putter & White 2005). In order to detect a cluster via its gravitational lensing signal, numerous images of many background faint galaxies need to be taken, thus this is a quite expensive technique. Despite the fact that WL is probably not an optimal method for cluster detection, it is a valid tool in cluster mass calibration (Sheldon et al. 2001; Johnston et al. 2007a) for clusters detected with other methods: more details will come in Section 3.4.1.

### 3.1.4 Optical surveys

Optical selection is based on overdensity of galaxies on the sky. This method was used first in 1784 by Charles Messier in the discovery of the Virgo Cluster. It yielded to the first statistically complete sample of galaxy clusters built by George Abell (Abell 1958) using the Palomar Observatory Sky Survey, then extended to include more than 4000 objects (ACO, Abell et al. 1989), and to the catalogue by Zwicky et al. (1968). These catalogues are biased towards identification of massive objects, with centrally peaked galaxy distribution, and nearby systems. Another identification technique is based on the overdensity of red galaxies which typically populate the well-evolved galaxy clusters and results into a tight red-sequence in a colour-magnitude diagram. This method has been used by the Red-Sequence Cluster Survey (RCS, Gladders & Yee 2005) using the Canada-France-Hawai'i Telescope and Cerro Tololo Inter-American Observatory, leading to the first cosmological constraints in optical surveys (Gladders et al. 2007). The large amount of data from the Sloan Digital Sky Survey (SDSS<sup>10</sup>) with its multi-band imaging and hundreds of thousands spectroscopic targets (Ahn et al. 2012) has been used to select several

---

<sup>7</sup><http://www.princeton.edu/act/>

<sup>8</sup><http://pole.uchicago.edu/>

<sup>9</sup>[http://www.esa.int/Our\\_Activities/Space\\_Science/Planck](http://www.esa.int/Our_Activities/Space_Science/Planck)

<sup>10</sup><http://www.sdss.org/>

samples, such as the maxBCG cluster catalogue (Koester et al. 2007b) discussed in Section 3.2.1. This has later been extended to the GMBCG catalogue by Hao et al. (2010). Other larger SDSS cluster samples include the work of Wen et al. (2009) and Szabo et al. (2011). Although the large number of detected clusters, the main issue with optical clusters is the difficult choice of a low-scatter mass proxy, as these are affected by selection effects. Another problem is that, while in the local Universe most galaxies members of a clusters are known to be red, this may not be valid anymore at high redshift, where galaxies seem to have star formation processes (Fassbender et al. 2011), causing then the lack of red-sequence feature.

### 3.1.5 Future surveys

The optical and near-infrared current and upcoming generation of galaxy cluster surveys will include the Panoramic Survey Telescope and Rapid Response System (PanSTARRS<sup>11</sup>), the Dark Energy Survey (DES<sup>12</sup>), the Kilo-Degree Survey (KIDS<sup>13</sup>) and VISTA Kilo-degree INfrared Galaxy Survey (VIKING<sup>14</sup>), the Large Synoptic Survey Telescope (LSST<sup>15</sup>), and the Euclid mission<sup>16</sup>. X-ray observations will be improved by the eROSITA all-sky survey, while for the SZ signal the future data releases of Planck, ACT, SPT, and the Cerro Chajnantor Atacama Telescope (CCAT<sup>17</sup>) will play a key role. Great improvements in the mm-band observations are expected to come from the Atacama Large Millimetre Array (ALMA<sup>18</sup> Testi et al. 2010), reaching  $\theta \sim 0.01''$ .

### 3.1.6 Cosmological constraints from cluster catalogues

The cluster number counts from some of the catalogues listed above have been used to constrain cosmology via the mass function. Efforts on this front have been led by Rozo et al. (2010); Zu et al. (2012) who derived cosmological constraints from the maxBCG cluster sample. The tightest scaling relation between observable and cluster mass comes from  $f_{\text{gas}}$  measurements of X-ray data ( $< 10\%$  scatter, Allen et al.

---

<sup>11</sup><http://pan-starrs.ifa.hawaii.edu/public/>

<sup>12</sup><http://www.darkenergysurvey.org/>

<sup>13</sup><http://kids.strw.leidenuniv.nl/>

<sup>14</sup><http://www.eso.org/public/teles-instr/surveytelescopes/vista/surveys.html>

<sup>15</sup><http://www.lsst.org/lsst/>

<sup>16</sup><http://sci.esa.int/euclid>

<sup>17</sup><http://www.ccatobservatory.org/>

<sup>18</sup><http://www.almaobservatory.org/>

2008). Constraints on dark energy with  $\sim 20\%$  uncertainty were obtained from X-ray cluster samples studied by Mantz et al. (2008, 2010) and Vikhlinin et al. (2009). Preliminary results for SZ surveys show agreement with the concordance WMAP CMB results (Komatsu et al. 2011).

Data on the cluster masses obtained from weak-lensing analyses of background galaxies have been combined with the number counts to improve the constraining power (Johnston et al. 2007b; Sheldon et al. 2009; Mahdavi et al. 2007).

The statistics of rare events in the high-peak, high-mass limit has been used to constrain cosmology by Hotchkiss (2011); Hoyle et al. (2012).

The uncertainty in the scaling relation is one of the biggest obstacles in using galaxy clusters as cosmological probes (e.g. Haiman et al. 2001; Battye & Weller 2003). Majumdar & Mohr (2003) suggested to use the clustering of clusters as a complementary probe. So far, only limited efforts have been dedicated to the measurement of the clustering properties of galaxy clusters: Hütsi (2010) measured the power spectrum of maxBCG clusters resulting in weak detection of baryon acoustic oscillations, Estrada et al. (2009) measured the correlation function for the same catalogue, and Hong et al. (2012) measured the correlation function of the cluster catalogue by Wen et al. (2009). Collins et al. (2000) measured the spatial correlation function of the REFLEX X-ray cluster catalogue, while Balaguera-Antolínez et al. (2011) measured its power spectrum. From the same survey, Schuecker et al. (2003) derived cosmological constraints from cluster abundances and large-scale clustering.

Finally, the constraining power of these data on measurements of the amount of primordial non-Gaussianity in the initial density perturbations field, which can potentially rule out entire classes of inflationary models (Bartolo et al. 2004), has been investigated by several authors. Bounds from the LSS using multiple galaxy catalogues have been studied by Afshordi & Tolley (2008); Slosar et al. (2008); Xia et al. (2010a,b, 2011); Sefusatti et al. (2012); Giannantonio et al. (2013); Ross et al. (2013). In particular Slosar et al. (2008) found  $-29 < f_{\text{NL}} < 70$  at 95% c.l., while Xia et al. (2011) reported hints of detection at  $8 < f_{\text{NL}} < 88$  (95% c.l.) for the local model. Other constraints come from the CMB bispectrum analyses: from WMAP data it was obtained  $-3 < f_{\text{NL}} < 77$  at 95% c.l. (Hinshaw et al. 2013; Bennett et al. 2013), Shandera et al. (2013) found  $f_{\text{NL}} = -3_{-91}^{+78}$  at 68% c.l. from X-ray cluster data combined with the WMAP7 data, while latest constraints on  $f_{\text{NL}}$  local parameter from Planck are  $2.7 < f_{\text{NL}} < 5.8$  at 68% c.l. (Planck Collaboration et al.

2013c). Future galaxy surveys such as Euclid are expected to reach an accuracy of  $\Delta f_{\text{NL}} \sim 3$  (Giannantonio et al. 2012). Oguri (2009) suggested that measuring the variance of cluster counts can yield significant constraints on PNG, while Sartoris et al. (2010) showed in principle how such constraints can be improved by using the cluster power spectrum. Forecasts for cosmology and PNG have been also investigated by Pillepich et al. (2012), with the future eROSITA X-ray cluster survey.

## 3.2 The Sloan Digital Sky Survey

The SDSS is an imaging (5 bands (u, g, r, i, z), spanning from ultraviolet to infrared) and spectroscopic survey in the North Galactic Cap and a small region in the South, encompassing 10,000 deg<sup>2</sup>. It contains more than 930,000 galaxies and 120,000 quasars. SDSS-I operated between 2000 and 2005, while SDSS-II between 2005 and 2008. The final public data release from SDSS-II in 2008 was the Data Release 7 (DR7, Abazajian et al. 2009). Meanwhile, SDSS-III will operate till 2014 and already produced Data Release 8 (DR8, Aihara et al. 2011) in 2011 and Data Release 9 (DR9, Ahn et al. 2012), which contains the first results of the Baryon Oscillation Spectroscopic Survey (BOSS<sup>19</sup>). The survey employs a 2.5-meter telescope at Apache Point Observatory in New Mexico, equipped with a 120-mega-pixel camera and a pair of spectrographs. In 2009 the Nobel Prize in Physics was awarded for the optical fibres and the digital imaging detectors (CCDs) technologies used in SDSS. The great quality and quantity of these data were crucial to the development of significant research in astrophysics and cosmology, including structure and stellar populations of the Milky Way, properties of galaxies, LSS, DM and DE.

### 3.2.1 MaxBCG catalogue

The maxBCG catalogue <sup>20</sup> (Koester et al. 2007a) is a sample of 13,823 galaxy clusters compiled from SDSS DR5 photometric data. The clusters are chosen in an approximately volume-limited way from a 500 Mpc<sup>3</sup> region, covering  $\sim 7,500$  deg<sup>2</sup> of sky with a photometric redshift (photo-z) range of  $0.1 \leq z \leq 0.3$ . The photo-z errors are small and of the order of  $\Delta z = 0.01$ . Each of these clusters contains 10 to 190 E/S0 ridgeline galaxies, which are brighter than  $0.4 L_*$ , within the scale radius  $R_{200}$ . Two catalogues derived from SDSS data are used to construct the maxBCG sample:

<sup>19</sup><http://www.sdss3.org/surveys/boss.php>

<sup>20</sup>[http://umsdss.physics.lsa.umich.edu/catalogs/maxbcg\\_public\\_catalog.dat](http://umsdss.physics.lsa.umich.edu/catalogs/maxbcg_public_catalog.dat)

a photometric catalogue, used for the cluster detection, and a spectroscopic one, useful to test the validity of the photo- $z$  estimates. The maxBCG sample is assembled by selecting the brightest cluster galaxies and applying a *red-sequence method* to identify cluster members in its neighbourhood. This relies on two key features of galaxy clusters: there is an almost ubiquitous population of galaxies which occupies a tight-scattered region in a colour-magnitude diagram (the E/S0 ridgeline), and the brightest cluster galaxies (BCG), typically located in the centre of the clusters, have similar colours and luminosities. The maxBCG algorithm considers two spatial dimensions (RA,DEC), two colour dimensions (g-r, r-i), one brightness dimension (i) and the redshift, and calculates the membership likelihood of each galaxy.

The pipeline calculates two likelihoods for each galaxy: the likelihood to be part of the E/S0 ridgeline and the likelihood to check the similarity to a BCG at redshift  $z$ . These likelihoods are computed over a grid of redshifts in the range  $0.1 \leq z \leq 0.3$ . The redshift corresponding to the maximum likelihood gives a first estimate of the redshift of the object. Subsequently, the algorithm assigns the status of BCG to the galaxy which maximises the product of these likelihoods. Then, it calculates the number of galaxies  $N_g$  within  $1h^{-1}\text{Mpc}$  from the BCG, with luminosity greater than  $0.4L_*$  (where  $L_*$  is the SDSS typical galaxy luminosity) and smaller than the BCG luminosity. This quantity is converted by Hansen et al. (2005) empirical relation to  $R_{200}$ , which is the radius at which the overdensity of galaxies with an absolute r band magnitude of  $-24 \leq M_r \leq -16$  is 200 times that of the average galaxy density. The galaxies within  $R_{200}$ , with similar colours ( $\pm 2\sigma$ ) and magnitude of the BCG and within a  $\Delta z = 0.02$  are then promoted as members of a cluster. This selection process starts from the richest cluster, and removes the BCG and cluster members from the galaxy catalogue before identifying the next highest likelihood, and thus the next BCG and associated cluster. In this way, previously flagged BCG and corresponding cluster members can be absorbed by richer clusters. The procedure is iterated until all galaxies are either BCGs or cluster members. Finally a cut on the number of cluster members is imposed, i.e.  $N_g > 10$ .

An analysis of N-body mock galaxy catalogues shows that the maxBCG algorithm results in more than 90% purity and more than 85% completeness, for clusters with masses  $M \geq 10^{14}M_\odot$  (Wechsler et al. 2006; Koester et al. 2007a). These simulations ensure that each cluster has a BCG by locating the brightest galaxy within the halo at the bottom of the dark matter potential well. Then, two quantities are computed for each mock cluster detected by the maxBCG algorithm: the fraction  $f_c$  of cluster galaxies within the richest DM halo, and the fraction  $f_h$  of galaxy halos

within the cluster halo. The first quantity represents the purity of the cluster finding algorithm: for each cluster, the halo with the maximum fraction of cluster members is identified. The second measurement, instead, estimates the completeness: for each halo, the cluster with the largest fraction of members is selected. A low value of  $f_c$  denotes a false positive detection of cluster, due to the fact that galaxies cannot be clearly associated to a single DM halo. If  $f_h$  falls below a threshold, the cluster is not detectable, because the number of identified galaxy halos within the cluster is too low. They obtained that for  $f_c = 0.3$  the purity is  $> 90\%$  for cluster  $N_g > 10$  and  $95 - 100\%$  pure for clusters with  $N_g > 20$ . On the other hand, for  $f_h = 0.3$ , the completeness of the sample is  $> 90\%$  above a mass of  $\sim 2 \times 10^{14} M_\odot$  and  $95 - 100\%$  complete above a mass of  $\sim 3 \times 10^{14} M_\odot$ .

### 3.3 MaxBCG cluster number counts

This Section introduces the first data set we need for the combined cosmological analysis presented here, namely the maxBCG cluster abundances. We also describe in detail the counts covariance error matrix which is composed of Poisson errors, sample variance and observational uncertainties.

#### 3.3.1 Cluster abundances

As already introduced in Section 2.3, the richness  $N_{\text{gal}}$  of a galaxy cluster is defined as the number of red galaxies within the radius  $R_{200}$  from the cluster centre. The cluster abundances we use in our analysis (see Table 3.1) are divided into nine richness bins in the range

$$11 \leq N_{\text{gal}} \leq 120 ,$$

which approximately corresponds to (Rozo et al. 2010)

$$7 \times 10^{13} h^{-1} M_\odot \leq M \leq 1.2 \times 10^{15} h^{-1} M_\odot .$$

The relatively low mass limit of this sample evolves weakly with redshift, resulting in a sample that is significantly large. The sample also includes five additional clusters with richness  $N_{\text{gal}} > 120$ , that we decide not to consider because of their negligible impact on the cosmological analysis. We extend the counts data used by Rozo et al. (2010) by including an additional bin at  $9 \leq N_{\text{gal}} \leq 11$  (Eduardo Rozo, private communication). However, we checked the results are not affected by this choice.

**Table 3.1:** MaxBCG counts data, binned into ten richness intervals: first column gives the mean  $N_{\text{gal}}$  value for each bin, while second one is the corresponding number of clusters.

$N_{\text{gal}}$	$\Delta N$
9.9	6070
12.8	5167
16.9	2387
21.4	1504
26.8	765
34.2	533
43.7	230
55.1	134
70.0	59
97.8	31

### 3.3.2 Counts covariance matrix

For the cosmological analysis we need to provide theoretical expectations of the counts covariance matrix  $\mathcal{C}_{\Delta N}$  between various bins, which is given by

$$\begin{aligned} \mathcal{C}_{\Delta N} &= \langle (N_i - \bar{N}_i) (N_j - \bar{N}_j) \rangle \\ &= [(C^{\text{Poisson+Obs}}|_{\Delta N})^{-1} + (C^{\text{Sample}}|_{\Delta N})^{-1} + (C^{\text{Photoz}}|_{\Delta N})^{-1}]^{-1} . \end{aligned} \quad (3.1)$$

The partial error matrices include the intrinsic error estimates, such as Poisson errors and sample variance, and observational systematics, such as photometric redshifts errors and uncertainties in purity and completeness of the sample. Following the prescription of Rozo et al. (2004, 2007),  $C^{\text{Poisson+Obs}}|_{\Delta N}$  represents the Poisson error matrix, including the uncertainties on purity and completeness, while  $C^{\text{Sample}}|_{\Delta N}$  represents the sample variance matrix and  $C^{\text{Photoz}}|_{\Delta N}$  the photometric redshift errors. In the top panel of Fig. 3.1 we show the counts data and errors, together with the predicted counts for a selection of different cosmologies. The red solid line represents the best-fit model to our full data set, while in green dotted-dashed and in blue dashed we represent the predicted counts for  $\Omega_m = 0.41$ ,  $\sigma_8 = 0.65$  and  $\Omega_m = 0.15$ ,  $\sigma_8 = 0.94$ , respectively. From this plot we deduce that a Universe with higher  $\Omega_m$  has more clusters at low richness. At the same time, a lower  $\sigma_8$  implies a suppression of the high richness (mass) counts. Vice versa, a low  $\Omega_m$  value brings less low richness objects and a higher  $\sigma_8$  allows for more high richness (mass) counts.

### Poisson errors

When counting a discrete number of events in cells, the first associated uncertainty is the **Poisson error**. Therefore, if we count clusters in bins, the probability of observing  $\Delta N$  clusters in a bin of given redshift and richness, is determined by the following Poissonian distribution:

$$P(\Delta N|N) = \frac{N^{\Delta N}}{\Delta N!} e^{-N} , \quad (3.2)$$

where  $N = \langle \Delta N \rangle = \text{Var}[\Delta N]$  is positive. This distribution is uni-modal, centred roughly on the mean and exhibits positive skew which decreases as the mean increases. Let us assume now that the mean  $N$  fluctuates in different bins as

$$N_i = \int d^3x W_i(\vec{x}) n(\vec{x}, z_i) , \quad (3.3)$$

where  $n(\vec{x}, z_i)$  is the spatial number density and  $W_i(\vec{x})$  is the usual top-hat window function. The sample averaged number counts is

$$\bar{N}_i = \bar{n}_i V_i , \quad (3.4)$$

where  $V_i = \int d^3x W_i(\vec{x})$  and  $\bar{n}_i$  is the averaged cluster density given by Eq. (2.28). Thus, the Poisson error contribution to the covariance matrix is

$$C_{ij}^{\text{Poisson}}|_{\Delta N} = \langle (N_i - \bar{N}_i) (N_j - \bar{N}_j) \rangle^{\text{Poisson}} = \delta_{ij} \bar{N}_i , \quad (3.5)$$

where  $\text{Var}[\bar{N}_i] = \bar{N}_i$ .

### Purity and completeness

For a correct analysis, we need to take into account for **purity** and **completeness** issues. For the cosmological analysis presented here, we assume 100% purity and completeness, adding a 5% uncertainty in quadrature (Rozo et al. 2010). This means adding a parameter  $\lambda = \bar{\lambda} \pm \sigma_\lambda = 1.00 \pm 0.05$  which transforms  $\bar{N}_i$  to  $\lambda \bar{N}_i$ . According to error propagation formulae, the variance of  $\lambda \bar{N}_i$  is given by:

$$\text{Var}[\lambda \bar{N}_i] = \left( \frac{\partial \lambda \bar{N}_i}{\partial \lambda} \right)^2 \sigma_\lambda^2 + \left( \frac{\partial \lambda \bar{N}_i}{\partial \bar{N}_i} \right)^2 \sigma_{\bar{N}_i}^2 = \bar{N}_i^2 \sigma_\lambda^2 + \lambda^2 \bar{N}_i . \quad (3.6)$$

Thus the Poisson error matrix diagonal elements change as:

$$C_{ij}^{\text{Poisson+Obs}}|_{\Delta N} = \langle (N_i - \bar{N}_i) (N_j - \bar{N}_j) \rangle^{\text{Poisson+Obs}} = \delta_{ij} (\bar{N}_i^2 \sigma_\lambda^2 + \bar{N}_i) . \quad (3.7)$$

### Sample variance

Another important source of uncertainty is the cell-to-cell variance due to large-scale structure (Hu & Kravtsov 2003; Lima & Hu 2004), usually called **sample variance** (or cosmic variance). In fact, observational estimates of number densities of a clustered population in finite volumes are subject to uncertainties which exceed the Poisson noise, arising from the underlying large-scale density fluctuations. If the typical clustering scale of the observed objects is much smaller than the sample volume, then cosmic variance is negligible. On the contrary, if the observed volumes are smaller than the clustering scale, one might count more or less objects with respect to the average number: in this case, cosmic variance becomes dominant and should be taken into account. For example, cosmic variance is relevant for deep galaxies surveys, because galaxies at high redshifts are more strongly clustered than dark matter compared to the local Universe (Kauffmann et al. 1999b). To check if a survey is independent of cosmic variance and accurately samples the Universe, one can compare the variation between the number counts for samples of different angular sizes which mimic different survey fields. In the case of strong cosmic variance, the number counts depend on where the samples are located in the sky and differ significantly for each region. If instead the counts agree while approaching the total size of the survey, the cosmic variance can be neglected. For clusters, we can recast the counts of Eq. (3.3) as

$$N_i = \int d^3 x W_i(\vec{x}) (\bar{n}_i + \delta n_i) = \int d^3 x W_i(\vec{x}) [\bar{n}_i (1 + b_i \delta_i)] , \quad (3.8)$$

where  $\delta_i \equiv \delta(\vec{x}, z_i)$  is the underlying linear density field and  $b_i$  is the average bias given by Eq. (2.40). Note that the index  $i$  identifies unique cells in redshift, angle and observable mass. The covariance matrix between cluster counts found in each bin due to sample variance is given by

$$\begin{aligned} C_{ij}^{\text{Sample}}|_{\Delta N} &= \langle (N_i - \bar{N}_i) (N_j - \bar{N}_j) \rangle^{\text{Sample}} = \\ &= b_i b_j \bar{n}_i \bar{n}_j \int d^3 x \int d^3 x W_i(\vec{x}) W_j(\vec{x}) \langle \delta_i \delta_j \rangle = \\ &= b_i b_j \bar{n}_i \bar{n}_j \int \frac{d^3 k}{(2\pi)^3} W_i(\vec{k}) \widehat{W}_j(\vec{k}) P(k) , \end{aligned} \quad (3.9)$$

where  $P(k)$  is the linear power spectrum and  $\widehat{W}_j(\vec{k})$  is the Fourier transform of  $W_j(\vec{k})$ . In our cosmological analysis, however, the survey volume is big enough ( $\sim 7,500 \text{ deg}^2$ ) to have a negligible sample variance contribution.

### Photometric redshift errors

The observational uncertainties associated to the photometric redshifts estimates are described by the probability distribution  $p(z_c|z_h)dz_c$ , where  $z_c$  is the photo-z estimate and  $z_h$  is the true halo redshift. This probability depends on the richness measure because the number of galaxies contributing to the photometric estimate increases with  $N_{\text{gal}}$ . However, systematics can reduce the sensitivity to cluster richness. We have found that the photo-z errors have a negligible impact on the number counts analysis presented here, so we neglect their contribution  $C^{\text{Photoz}}|_{\Delta N}$  in the number counts covariance matrix.

## 3.4 MaxBCG cluster masses

This Section describes the estimation of the mean cluster masses and errors from the maxBCG catalogue by the WL technique. We also specify the nature of the mass-richness relation and constraints on its scatter from previous works. Finally, we introduce the second data set we need for our analysis, namely the total cluster masses together with their covariance error matrix.

### 3.4.1 Mean cluster masses from weak lensing observations

Sheldon et al. (2009) first measured the weak lensing effect from maxBCG galaxy clusters. By stacking the clusters in narrow richness bins, the signal-to-noise was improved considerably compared to the measurement of the profile of an individual cluster. In this way, they could measure the cluster average weak lensing shear profile. These profiles were then converted to surface mass density contrast profiles (see Eq. 2.8) for different luminosity and richness bins, by means of the photometric redshift distribution of the background galaxies, estimated with the methods of Lima et al. (2008) and Oyaizu et al. (2008). Johnston et al. (2007b) fit the resulting profiles and reconstructed the mean 3D cluster density and mass profiles: this allows one to estimate the mass and the concentration of clusters in a given redshift bin. For this reconstruction a Navarro-Frenk-White profile (Navarro et al. 1997) for the cluster density was assumed (see Eq. 2.4). They found for the whole sample, the following mean mass-richness relation:

$$M_{200}(N_{\text{gal}}) \simeq 8.8 \times 10^{13} h^{-1} M_{\odot} (N_{\text{gal}}/20)^{1.28}, \quad (3.10)$$

where  $M_{200}$  is the mass contained within the radius  $R_{200}$ . Mandelbaum et al. (2008b) used a different method for estimating the redshift of the background sources, which corrected the biased mass estimates of Sheldon et al. (2009) upwards by a factor of 1.18, as described also in Rozo et al. (2009). Mandelbaum et al. (2008a) performed a similar but independent weak lensing analysis of the maxBCG clusters, but using a slightly different overdensity threshold for the mass definition (180 instead of 200), which added only a 2% correction. Applying this correction, the slopes of Johnston et al. (2007b) and Mandelbaum et al. (2008a) agreed almost perfectly, but the latter found a mean mass difference of approximately 6% with respect to Johnston et al. (2007b). Rozo et al. (2010) included this discrepancy by introducing an offset factor  $\beta$  with a suitable chosen prior, as we will describe in Chapter 4.

For the cosmological analysis presented here we follow Rozo et al. (2010) and fit for the mass-richness relation using the Johnston et al. (2007b) data and their errors (see Table 3.2). We use five richness bins for this, in the range  $12 \leq N_{\text{gal}} \leq 300$ , plus another extra bin at  $9 \leq N_{\text{gal}} \leq 12$  (Eduardo Rozo, private communication). The bottom panel of Fig. 3.1 shows the mean weak lensing mass estimates with errors, together with the theoretical expectations for a selection of different cosmologies. The red solid line represents the best-fit model to our full data set (counts, total masses and power spectrum). The green dotted-dashed (blue dashed) curve represent the theoretical predictions for  $\Omega_m = 0.41$ ,  $\sigma_8 = 0.65$  ( $\Omega_m = 0.15$ ,  $\sigma_8 = 0.94$ ). From this plot we can deduce that in a Universe with low  $\Omega_m$ , structures form early, thus being more numerous at high  $z$ . In addition to this, they have more time to accrete, implying they reach higher masses. On the contrary, if  $\Omega_m$  is high, structures form late, resulting in less structures at high  $z$ . Furthermore, they have less time to accrete and thus they will end up with lower mean masses.

Note that in general the estimated WL mass of a galaxy cluster depends on the underlying cosmological model. For the analysis presented here, this dependency is through the angular diameter distance, which is modified by the total matter density  $\Omega_m$ . In order to estimate the size of this cosmology dependence, we placed a galaxy cluster with mass  $M = 1.1 \times 10^{15} h^{-1} M_\odot$  at redshift  $z = 0.2$  and produced a mock catalogue of sheared background galaxies. From this catalogue we estimated the mass of the cluster by fitting to a NFW profile. We found that, if we allow  $\Omega_m$  to change within the  $1\text{-}\sigma$  level of our best fit cosmology, the mass varies within 5%. However, we allow for an uncertainty in the mass estimation with the offset factor  $\beta$  with a prior width of 6%. Hence any change due to a different  $\Omega_m$  is completely degenerate with the  $\beta$  parameter, which we assume does not depend on cosmology.

**Table 3.2:** MaxBCG mean mass data, binned into six richness intervals. From left to right, we have respectively: the mean  $N_{\text{gal}}$  value for each bin, the corresponding abundances, the WL mean mass and relative error on it.

$N_{\text{gal}}$	$\Delta N$	$\bar{M}[10^{14} M_{\odot}]$	$\sigma_M/\bar{M}$
10.4	7986	0.9659	0.11
14.7	5651	1.3	0.08
21.6	2269	1.986	0.10
32.6	1021	3.851	0.09
54.0	353	5.482	0.12
145.9	55	13.04	0.17

### 3.4.2 Mass-richness scaling relation

In order to perform a cosmological analysis, we need to make some assumptions on the scaling relation between the true mass of a cluster  $M$  and its observed richness  $N_{\text{gal}}$ . Following Johnston et al. (2007b) and Rozo et al. (2010), we assume the scaling relation to be a power law in mass,

$$\ln M = \ln M_{200|20} + \alpha_N \ln(N_{\text{gal}}/20) \quad , \quad (3.11)$$

with  $M_{200|20}$  the mass of a cluster with 20 member galaxies within a radius of  $R_{200}$  and  $\alpha_N$  the slope of the relation. This provides the mean of the distribution between  $N_{\text{gal}}^{\text{obs}}$  and  $M$ , given by Eq. (2.25). We fit this relation by fixing two pivot points in mass

$$M_1 = 1.3 \times 10^{14} M_{\odot} \quad \text{and} \quad M_2 = 1.3 \times 10^{15} M_{\odot} \quad ,$$

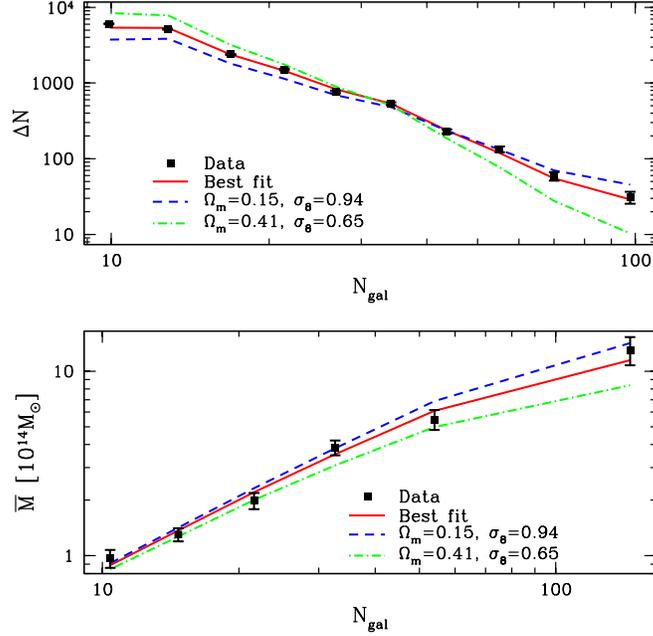
while the corresponding richness values

$$\ln N_1 \equiv \ln N_{\text{gal}}|M_1 \quad \text{and} \quad \ln N_2 \equiv \ln N_{\text{gal}}|M_2$$

are kept as free parameters. Note that we use  $\sigma_{\ln N_{\text{gal}}^{\text{obs}}|N_{\text{gal}}} = \sigma_{\ln N_{\text{gal}}^{\text{obs}}|M}$ , which holds because  $p(N_{\text{gal}}|M)$  is a delta function, according to Eq. (2.25). The statistical scatter around the scaling relation is assumed to be constant with redshift and mass for individual clusters. To get an estimate of this quantity is not trivial; however, by demanding consistency between the X-ray and weak lensing measurements, Rozo et al. (2009) found

$$\sigma_{\ln M|N_{\text{gal}}^{\text{obs}}} = 0.45_{-0.18}^{+0.20} \quad (95\% \text{c.l.}),$$

which is the scatter in mass given the richness. Specifically, Rozo et al. (2009) used observational constraints from maxBCG clusters on the mean  $M - N_{\text{gal}}$  relation,



**Figure 3.1:** *Top panel:* MaxBCG cluster counts data (black points) and theoretical predictions for a choice of different cosmologies (without primordial non-Gaussianity). The red solid line represents the best-fit model to our full data set (counts, total masses and power spectrum). In green dotted-dashed and in blue dashed we represent the predicted counts for  $\Omega_m = 0.41$ ,  $\sigma_8 = 0.65$  and  $\Omega_m = 0.15$ ,  $\sigma_8 = 0.94$ , respectively. *Bottom panel:* Mean masses of maxBCG clusters (black points) and theoretical predictions for different cosmologies, as above.

from weak lensing measurements, and on the mean and scatter of the  $L_X - N_{\text{gal}}$  relation, from X-ray measurements. The former analysis was done by Johnston et al. (2007a,b); Sheldon et al. (2009); Mandelbaum et al. (2008a,b). Rykoff et al. (2008) measured instead the cluster X-ray luminosity and scatter by stacking the RASS (Voges et al. 1999) photon maps, centred on maxBCG clusters, in narrow richness bins. In addition to these data sets, they also placed priors on the  $L_X - M$  relation from the constraints obtained by Vikhlinin et al. (2009), based on clusters from the 400 days cluster X-ray survey (Burenin et al. 2007). By combining these data, they obtained a constraint of the scatter in mass which is consistent with other analysis (e.g. Becker et al. 2007).

For our cosmological analysis we need to place a prior on the converse scatter,  $\sigma_{\ln N_{\text{gal}}^{\text{obs}}|M}$ . The scatter and the converse scatter can be readily related to each other by invoking the relation of Eq. (3.11), which results in

$$\sigma_{\ln M|N_{\text{gal}}^{\text{obs}}} = \alpha_N \sigma_{\ln N_{\text{gal}}^{\text{obs}}|M} .$$

### 3.4.3 Cluster total masses

In our combined analysis, we do not use the WL cluster mean masses as an additional data set to the cluster abundances. Instead, we use the cluster total masses  $\Delta N \bar{M}$ , i.e. the product of counts and mean mass in the corresponding richness bins. The covariance matrix of the total masses is calculated in an analogous way as done for counts (see Eq. 3.1):

$$\mathcal{C}_{\Delta N \bar{M}} = \langle (N_i \bar{M} - \bar{N}_i \bar{M}) (N_j \bar{M} - \bar{N}_j \bar{M}) \rangle . \quad (3.12)$$

Note that when introducing the uncertainty  $\lambda = \bar{\lambda} \pm \sigma_\lambda = 1.00 \pm 0.05$  on purity and completeness, as described for counts already (see also (Rozo et al. 2010)), the total masses change from  $\bar{N}_i \bar{M}$  to  $\lambda \bar{N}_i \bar{M}$ . According to error propagation formulae, the variance of  $\lambda \bar{N}_i \bar{M}$  is thus given by:

$$\begin{aligned} \text{Var}[\lambda \bar{N}_i \bar{M}] &= \left( \frac{\partial \lambda \bar{N}_i \bar{M}}{\partial \lambda} \right)^2 \sigma_\lambda^2 + \left( \frac{\partial \lambda \bar{N}_i \bar{M}}{\partial \bar{N}_i} \right)^2 \sigma_{\bar{N}_i}^2 + \left( \frac{\partial \lambda \bar{N}_i \bar{M}}{\partial \bar{M}} \right)^2 \sigma_{\bar{M}}^2 = \\ &= (\bar{N}_i \bar{M})^2 \sigma_\lambda^2 + \lambda^2 \bar{M}^2 \bar{N}_i + \bar{N}_i^2 \sigma_{\bar{M}}^2 . \end{aligned} \quad (3.13)$$

The Poisson error matrix diagonal elements then become:

$$C_{ij}^{\text{Poisson+Obs}}|_{\Delta N \bar{M}} = \delta_{ij} [(\bar{N}_i \bar{M})^2 \sigma_\lambda^2 + \bar{M}^2 \bar{N}_i + \bar{N}_i^2 \sigma_{\bar{M}}^2] . \quad (3.14)$$

## 3.5 MaxBCG cluster power spectrum

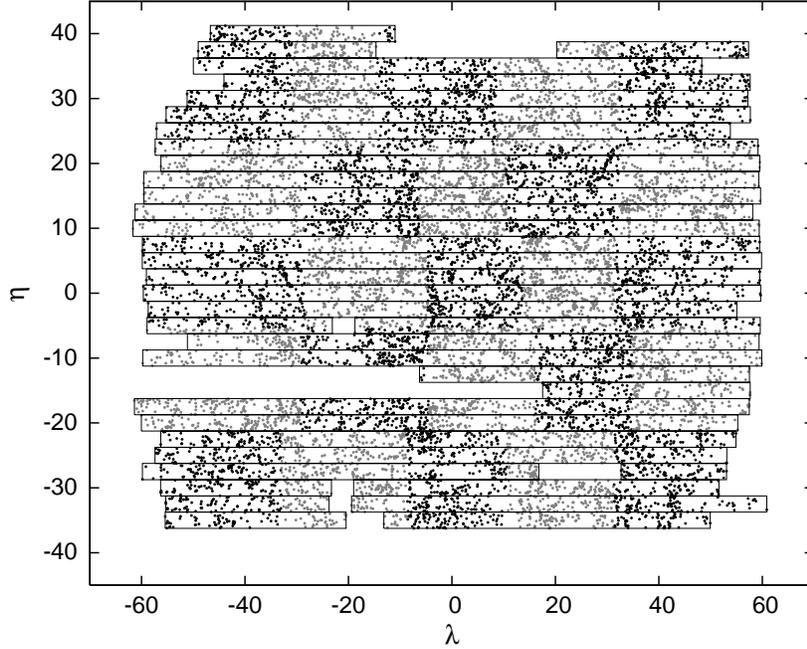
This Section details the measurements of the third data set we need, namely the maxBCG cluster power spectrum and corresponding covariance matrix.

### 3.5.1 Cluster power spectrum calculation

We will now describe the calculation of the redshift-space power spectrum of the maxBCG sample, as measured by Hütsi (2010). For the full details of the power spectrum measurement, along with systematics tests, we also refer the reader to Hütsi (2006a,b). From the SDSS DR5 full data set, the three southern stripes are removed: Fig. 3.2 shows angular distribution of the remaining clusters of galaxies, in the Sloan survey coordinate system (Stoughton et al. 2002), together with the reconstructed angular mask. The power spectrum calculation is then based on these 12,616 galaxy clusters, with redshift  $0.1 \leq z \leq 0.3$ , over  $\sim 6800 \text{ deg}^2$  and  $\sim 0.4 h^{-3} \text{ Gpc}^3$  of comoving volume. The distribution is divided into 25 regions in the latitude-longitude plane and in 3 redshift slices: in total, the divisions are 75, with  $\sim 168$  objects each. The selection function of the survey is given by  $n(\mathbf{r}) = n(\hat{\mathbf{r}})n(r)$ , where  $r = |\mathbf{r}|$ ,  $\hat{\mathbf{r}} = \mathbf{r}/r$ ,  $n(r)$  is the radial selection function, obtained from the redshift distribution of clusters assuming a flat  $\Lambda$ CDM model with  $\Omega_m = 0.27$ , and  $n(\hat{\mathbf{r}})$  the angular selection function, which is 1 (0) if the point is inside (outside) the mask.

The power spectrum has been calculated with the direct Fourier method (FKP, Feldman et al. 1994), which is appropriate for large  $k$ -modes (Tegmark et al. 1998), where the Fast Fourier Transforms (FFTs) were used to speed up instead of sums. This method actually yields the pseudo-spectrum, namely the measurement products are convolved with the window function of the survey. We take this into account when modelling the theoretical spectra in our analysis. More precisely, the following steps were followed:

1. The survey selection function  $n(\mathbf{r})$  was represented using a random catalogue with 100 times more points than maxBCG sample.
2. The density field has to be defined on a regular grid, with cells of finite size, thus we have to take into account for a smoothing effect. Jing (2005) connected the results from FFTs and those from Fourier summations, introducing a mass assignment window to construct the overdensity grid: the latter was the Triangular Shaped Cloud (TSC) filter (Hockney & Eastwood 1988), which can be



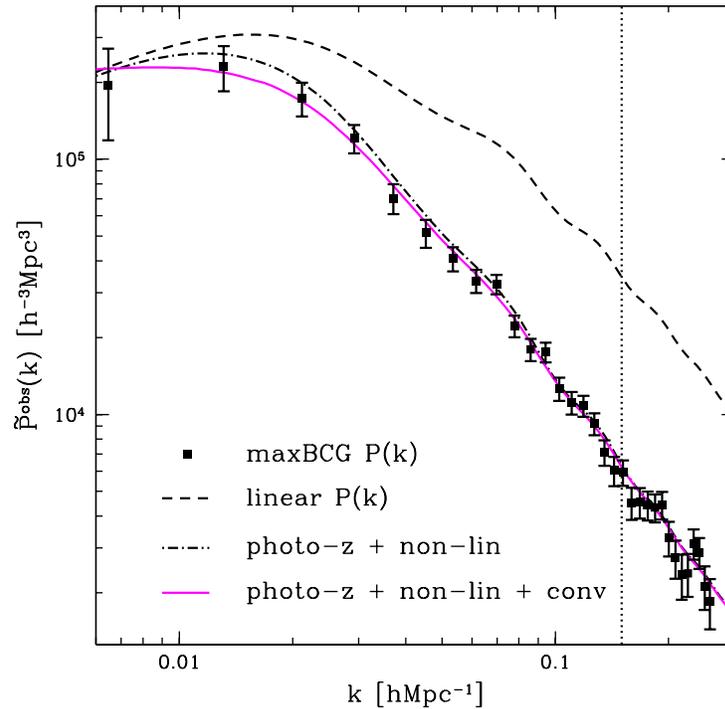
**Figure 3.2:** The angular distribution of the selected maxBCG clusters (after the removal of the southern stripes). The coordinates are taken from the Sloan survey convention:  $\mu$  is the latitude and  $\lambda$  is the longitude. The black-grey pattern represents the division used in the jackknife error calculation. The complete angular mask is given by the union of the rectangles. Credit: Hütsi (2010).

written in Fourier space as

$$\mathcal{W}(\mathbf{k}) = \left\{ \left[ \prod_{i=1}^3 \sin \left( \frac{\pi k_i}{2k_N} \right) \right] / \left[ \prod_{i=1}^3 \frac{\pi k_i}{2k_N} \right] \right\}^3, \quad (3.15)$$

where  $\mathbf{k} = (k_1, k_2, k_3)$ .

3. The gridded overdensity field  $\delta_g(\mathbf{k})$  was transformed to Fourier space using the FFTs.
4. The raw 3D power spectrum was estimated by taking the modulus squared of the previous, i.e.  $P_{\text{raw}} = \langle |\delta_g(\mathbf{k})|^2 \rangle$ .



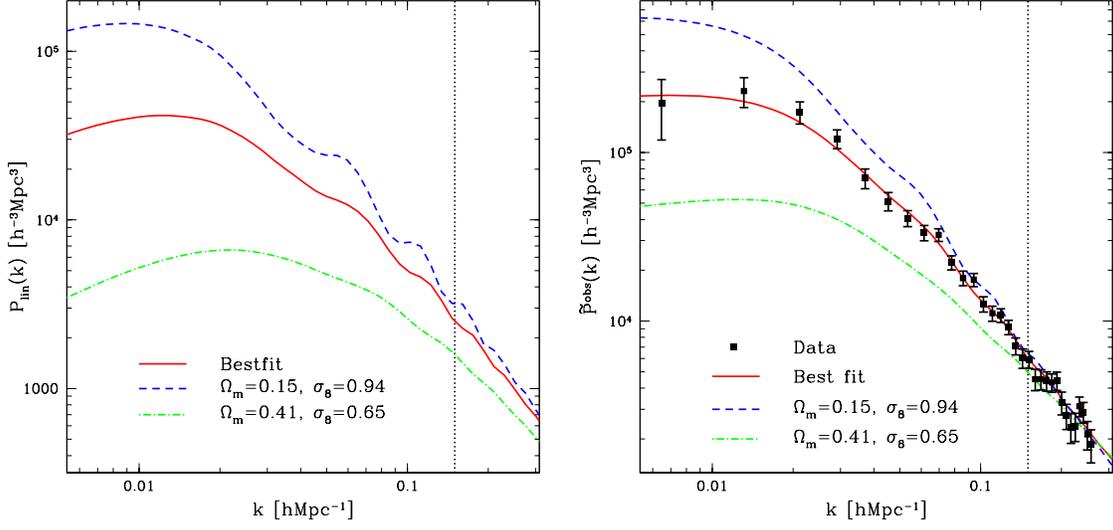
**Figure 3.3:** The redshift-space power spectrum of maxBCG cluster sample (black points) and errors. The dashed line corresponds to the linear matter power spectrum, dotted-dashed and solid magenta lines represent the final best fitting model in Hütsi (2010), without survey window convolution and with, respectively. Credit: Hütsi (2010).

5. When applying the TSC method, the shot noise term takes the form

$$P_{\text{shot}} = \frac{1}{N} \prod_{i=1}^3 \left[ 1 - \sin^2 \left( \frac{\pi k_i}{2k_N} \right) + \frac{2}{15} \sin^4 \left( \frac{\pi k_i}{2k_N} \right) \right]. \quad (3.16)$$

The latter was subtracted from the raw 3D power spectrum.

6. It was checked that the aliasing effects due to the finite grid size were negligible for the measurements, and were nonetheless corrected with the iterative method by Jing (2005) to obtain finally the angle averaged pseudo-spectrum, i.e. with smoothing effect of the TSC deconvolved.



**Figure 3.4:** The different lines represents the theoretical models for best fit (red solid), for  $\Omega_m = 0.15$ ,  $\sigma_8 = 0.94$  (blue dashed) and for  $\Omega_m = 0.41$ ,  $\sigma_8 = 0.65$  (green dotted dashed), while data are in black. The dotted line at  $k = 0.15$  represents the data cut to linear scales. Top panel: linear matter power spectrum. Bottom panel: full non-linear observed power spectrum.

### 3.5.2 Cluster power spectrum covariance matrix

The uncertainties on the power spectrum measurements have been estimated by Hütsi (2010) with three different methods, shown to be comparable:

- i) with the original FKP theoretical prescription, which assumes Gaussian errors from cosmic variance and a shot noise, neglecting redshift space distortions;
- ii) with a *jackknife* method (Lupton 1993), implemented by dividing the survey into the 75 regions with  $\sim 168$  clusters each;
- iii) with a Monte Carlo method, based on the fiducial  $\Lambda$ CDM cosmology, in which 1000 mock realisations of the maxBCG survey were generated, including redshift-space distortions and photo-z errors.

In our analysis we used only the Monte Carlo covariance matrix, so that we can write the covariance matrix  $\mathcal{C}_P$  for the cluster power spectrum as

$$\mathcal{C}_P = R \cdot \sigma_P^2, \quad (3.17)$$

where  $R$  is the correlation matrix, namely the covariance with diagonal elements equal to unity and off-diagonal ones showing the correlation between the errors, and  $\sigma_P^2$  is the variance of  $P(k)$ .

The power spectrum measurements and errors are reported in Table 3.3 and shown in Fig. 3.3 (black points). The dashed line represents the linear power spectrum. The dotted dashed curve is the theoretical model (see Eq. 2.44), including the photo- $z$  smoothing and the non-linearities, without the survey window convolution. Finally, solid magenta line shows the best fitting model convolved with the survey window. Note that to take into account data in the quasi-linear regime only, we restrict ourselves to scales larger than (wavenumbers smaller than)  $k_{\max} = 0.15 h \text{ Mpc}^{-1}$ .

We show in Fig. 3.4 the linear matter power spectrum  $P_{\text{lin}}(k)$  and the full halo power spectrum  $\tilde{P}^{\text{obs}}(k)$  from Eq. (2.48) respectively, as a function of  $k$ . The different lines correspond to the theory curves for our combined best-fit cosmological model (red solid) and for two other models ( $\Omega_m = 0.15$ ,  $\sigma_8 = 0.94$  for the blue dashed line and  $\Omega_m = 0.41$ ,  $\sigma_8 = 0.65$  for the green dotted dashed curve), chosen to be at the  $2\sigma$  limit of the marginalised  $\Omega_m - \sigma_8$  contour, compared with data and errors from Hütsi (2010) in black. We are assuming Gaussian initial condition and fixing all the other model parameters to our best-fit values.

### 3.6 The counts-clustering off-diagonal covariance

We estimate now the off-diagonal terms of the covariance matrix between the clustering and the binned number distributions  $\Delta N$  of the maxBCG clusters. For simplicity, instead of the power spectrum we use, as a clustering estimator here, the projected two-point angular correlation function  $w(\theta)$  of objects at a given angular distance  $\theta$  (Peebles 1980). The angular separation  $\theta$  is usually defined as

$$\cos(\theta) = \hat{n}_1 \cdot \hat{n}_2 ,$$

where  $\hat{n}_1$  and  $\hat{n}_2$  are the unit vectors pointing to the objects locations. So the projected correlation function can be written as

$$w(\theta) \equiv \langle \delta_h(\hat{n}_1) \delta_h(\hat{n}_2) \rangle , \quad (3.18)$$

where  $\delta_h(\hat{n})$  is the halo (cluster) projected overdensity in a direction  $\hat{n}$  and the average is carried over all pairs at an angular distance  $\theta$ . We use the *jackknife* technique as

follows: we split the maxBCG footprint into 100 equal-area jackknife regions using HEALPIX<sup>21</sup> (Górski et al. 2005) and populate the full footprint with 50 random points for each maxBCG cluster, to reduce shot noise. A common approach for estimating the correlation function of objects is to count pairs and to compare the data with an approximately ten times bigger random (unclustered) sample with same redshift distribution and angular coverage. It is then possible to define some estimators which combine data and random objects as number of data-data (DD), random-random (RR) and data-random (DR) pairs. We use the Landy-Szalay-estimator (Landy & Szalay 1993), given by:

$$w_{\text{LS}}(\theta) = \frac{DD - 2DR + RR}{RR} . \quad (3.19)$$

We also bin the number of clusters within six equal-width bins in  $\log_{10}$  space. We iteratively remove and replace each jackknife region and calculate the number histogram and correlation function at each iteration. In general, the covariance matrix  $\mathcal{C}^{\text{JK}}$  between measured statistics  $x = x(\alpha)$  and  $y = y(\beta)$  can be estimated from  $N$  jackknives using (see e.g. Efron 1982):

$$[\mathcal{C}^{\text{JK}}(x_i, y_j)]_{\alpha, \beta} = \frac{N-1}{N} \sum_{k=1}^N (x_{-i}^k - \bar{x}_i)_\alpha (y_{-j}^k - \bar{y}_j)_\beta , \quad (3.20)$$

where  $x_{-i}$  ( $y_{-j}$ ) is the statistic with jackknife region  $i$  ( $j$ ) removed, and  $\bar{x}_i$  ( $\bar{y}_j$ ) is the average value of all  $x_{-i}$  ( $y_{-i}$ ). We note that typically, but not necessarily,  $x$  and  $y$  are the same statistic. We compare the square root of the diagonal elements of the  $w(\theta)$  covariance matrix, i.e.

$$\mathcal{C}_{w(\theta)}^{\text{JK}} \equiv \mathcal{C}^{\text{JK}}[w(\theta), w(\theta)] , \quad (3.21)$$

with the error expected from Poisson counting statistics and find agreement with the theoretical expectations (as described by e.g., Ross et al. 2009). Moreover, the diagonal elements of the counts covariance matrix, i.e.

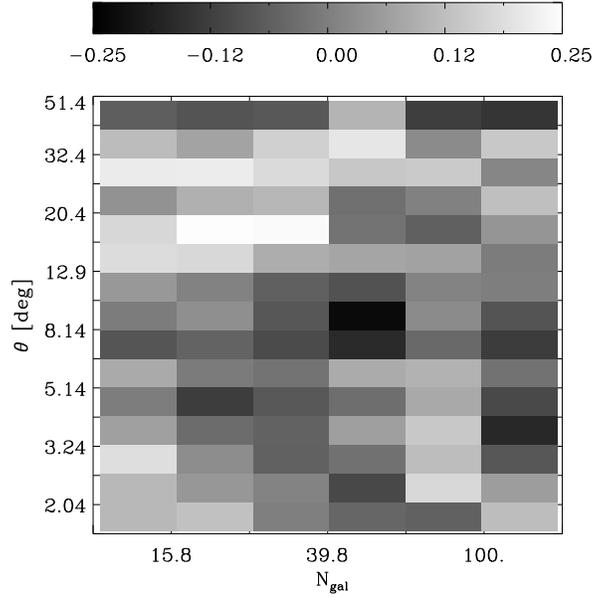
$$\mathcal{C}_{\Delta N}^{\text{JK}} \equiv \mathcal{C}^{\text{JK}}(\Delta N, \Delta N) , \quad (3.22)$$

are approximately Poissonian, independently of the number of jackknives used, as we would expect. Finally, the off-diagonal terms of the normalised counts-clustering covariance, i.e.

$$\mathcal{C}_{\Delta N, w(\theta)}^{\text{JK}} \equiv \mathcal{C}^{\text{JK}}[\Delta N, w(\theta)] , \quad (3.23)$$

---

<sup>21</sup><http://healpix.jpl.nasa.gov/>



**Figure 3.5:** The off-diagonal elements of the normalised covariance matrix of the correlation function  $w(\theta)$  and the histogram distribution of  $\Delta N$ , as calculated using the jackknife technique.

are shown in Fig. 3.5. Note that the average value of the off-diagonal terms is  $-0.03 \pm 0.10$  ( $1\sigma$ ), which is consistent with zero. We observe that as the number of jackknives increases, the mean of the average value of the off-diagonal terms approaches zero.

We also compare the magnitude of the off-diagonal terms obtained from the maxBCG clusters with simulated clusters from the Millennium Simulation (see e.g. Springel et al. 2005; Lemson & Virgo Consortium 2006). Specifically, we join the light-cone table of Henriques2012a.wmap1.BC03\_AllSky\_00 (Henriques et al. 2012; Guo et al. 2011) with the halo-tree table MPAHaloTrees..MHALo. We apply the same redshift and survey footprint constraints to mimic the maxBCG sample and calculate the correlation function and histogram distribution of  $\Delta N$ . We find the data and simulations agree closely: e.g., for 100 jackknives the mean and  $1\sigma$  error of the off-diagonal terms are  $0.00 \pm 0.10$  from the simulations. From these tests we conclude that our choice of ignoring the off-diagonal covariance matrix between clustering and number counts is a reasonable approximation.

**Table 3.3:** MaxBCG power spectrum data binned into  $k$ -intervals. From left to right: the mean  $k$  value for each bin, corresponding observed power spectrum value and standard deviation on it. Note that we restrict to the quasi-linear regime only, thus we will use in our cosmological analysis only the first 18  $k$ -bins, corresponding to data up to  $k_{\max} = 0.15 [h \text{ Mpc}^{-1}]$ .

$k[h \text{ Mpc}^{-1}]$	$\tilde{P}^{\text{obs}}[10^4 h^{-3} \text{ Mpc}^3]$	$\sigma_P[10^3]$
0.0065	19.470	76.180
0.0131	23.100	46.270
0.0212	17.340	26.140
0.0292	12.060	15.320
0.0371	7.0390	9.4570
0.0452	5.1410	6.4090
0.0534	4.0760	4.4630
0.0616	3.3410	3.4820
0.0698	3.2410	2.8120
0.0779	2.2220	2.1470
0.0860	1.7980	1.8040
0.0942	1.7570	1.5610
0.1024	1.2610	1.3170
0.1105	1.1110	1.1260
0.1188	1.0820	1.0180
0.1270	0.9194	0.9048
0.1351	0.7111	0.8153
0.1433	0.6027	0.7900
0.1514	0.5924	0.6750
0.1596	0.4517	0.6517
0.1678	0.4524	0.6199
0.1760	0.4414	0.5735
0.1841	0.4318	0.5392
0.1923	0.4423	0.5494
0.2005	0.3285	0.5095
0.2087	0.2738	0.4683
0.2169	0.2354	0.4751
0.2251	0.2384	0.4552
0.2332	0.3122	0.4222
0.2414	0.2882	0.4019
0.2496	0.2127	0.4108
0.2578	0.1852	0.4121

# Chapter 4

## Cosmological analysis

This Chapter is entirely dedicated to the statistical analysis of data for cosmological parameters estimation. The standard approach for extracting cosmological parameters from observational data sets makes use of maximum likelihood methods, based on the Bayes Theorem. If the dimensionality of the parameter space is very large, there are ways to study the likelihood surface efficiently, sampling the space more densely where the likelihood is high: these are the Monte Carlo Markov Chain methods (MCMC), a combination of Monte Carlo methods and the advantages of the Markov chains. Our primary MCMC tool for cosmological analysis is the COSMOMC software, for which we provide here a description. In particular we focus on the results of Rozo et al. (2010) and Zu et al. (2012), who derived cosmological constraints from the maxBCG cluster sample, using cluster counts and weak lensing mass estimates. The goal of this project is to extend their analysis to fully include the clustering information: we show that the inclusion of the cluster power spectrum significantly improves the cosmological constraints, and also reduces the degeneracies between the scaling relation nuisance parameters. We present the improved cosmological results obtained in this way from the maxBCG cluster counts, weak lensing masses and power spectrum. As an interesting application, we present the constraining power of these data on the amount of primordial non Gaussianity of the initial density perturbations.

### 4.1 Parameter estimation

One of the main purposes in cosmology is to estimate the parameters describing a model of the Universe and their confidence limits, on the basis of a set of observational data. This is a non-trivial task when more than few parameters are involved. In this

Section, we first introduce the importance of Bayesian statistics, which is based on the interpretation of probability as a conditional measure of uncertainty. We then consider the posterior calculation in case of Gaussian ( $\chi^2$  statistics) and Poissonian ( $C$ -statistics) distributions. Finally we introduce some useful concepts in parameters estimation, such as confidence regions and marginalisation.

### 4.1.1 Bayes theorem

Suppose that  $D$  is some set of data,  $\Theta = (\theta_1, \dots, \theta_m)$  is the  $m$ -dimensional unknown parameter vector of a certain model, then the Bayes theorem states that

$$P(\Theta|D) = \frac{P(D|\Theta) P(\Theta)}{P(D)} = \frac{P(D|\Theta) P(\Theta)}{\int P(D|\Theta) P(\Theta) d\Theta}, \quad (4.1)$$

where  $P(\Theta|D)$  is the multivariate **posterior** distribution, namely the inferred joint probability distribution of the parameters, after considering the data. This is achieved by calculating the **likelihood**  $P(D|\Theta) \equiv \mathcal{L}(D|\Theta)$ , which is the conditional probability of obtaining the observed data given a set of input parameter values, and by setting a **prior** probability on the parameters themselves,  $P(\Theta)$ . The likelihood tells us about the mismatch between data and theoretical predictions. The prior, instead, represents literally the *a priori* knowledge on the distribution of the expectation values of the parameters, before considering any measurement. The normalization constant  $P(D)$  plays a key role in the model selection and it is called **evidence** or marginal likelihood: from Eq. (4.1) it can be seen that is equal to the integral over  $\Theta$  of the likelihood times the prior. Throughout our analysis, however,  $P(D) = 1$ . The general problem of posterior estimation is then to calculate an integral: compressing the posterior distribution to a set of samples can save a vast amount of computational time, if compared to the integration over the  $m$ -dimensional parameter space. This can be done efficiently by a *Monte Carlo* integration method. This sampling method is very convenient when calculating quantities, like moments, which typically are the expectation values of a function of the model parameters.

### 4.1.2 Gaussian $\chi^2$ statistics

When comparing a theoretical model to real data, one has to determine the best fit model to the data and the corresponding errors. The  $\chi^2$  statistics is a measure of how good the fit of the model to the data is. The best fit condition is fulfilled by the model which minimises the  $\chi^2$ . This is also equivalent to maximise the likelihood

that the observations were produced by the theory we fit, when the samples are described by a Gaussian distribution. We assume then the following:  $\Theta$  represents the parameters of the model we need to estimate,  $x_i$  and  $\mu_i \equiv \mu_i(\Theta)$  are respectively the  $n$  observed data and corresponding theoretical predictions in some bin  $i$ ,  $\sigma_i$  the errors associated to the data. We also assume that the probability of observing a given value  $x_i$  is its Gaussian probability

$$P_i \propto \exp \frac{(x_i - \mu_i)^2}{2\sigma_i^2} . \quad (4.2)$$

If data are uncorrelated, then the probability of observing a set of samples is the product of the single probabilities over all the bins. The logarithm of the latter is the  $\chi^2$ :

$$\chi^2 \equiv \ln \prod_{i=1}^n P_i = \sum_{i=1}^n \frac{[x_i - \mu_i]^2}{\sigma_i^2} . \quad (4.3)$$

Finally, the posterior probability, proportional to the likelihood assuming a flat prior, is then given by

$$P(\Theta|D) \propto \mathcal{L}(\Theta|D) \propto \exp \left[ -\frac{1}{2}\chi^2 \right] , \quad (4.4)$$

from which is clear that a maximum likelihood corresponds to the minimum value of  $\chi^2$ . In the case of correlated data, the modified  $\chi^2$  equation is

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^n [x_i - \mu_i]^T \{C^{-1}\}_{ij} [x_j - \mu_j] , \quad (4.5)$$

where

$$C_{ij} = \langle (x_i - \mu_i) (x_j - \mu_j) \rangle \quad (4.6)$$

is the covariance error matrix, describing the covariance between the data. The latter can be also decomposed into the product of the correlation matrix  $R_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$ , i.e. a normalised version of the covariance matrix, and the diagonal elements  $C_{ii} = \text{Var}(x_i) = \sigma_i^2$ .

We can also define the **degrees of freedom** as  $\nu = n - m$ , where  $n$  the number of independent data points and  $m$  is the number of parameters. In addition to this, one can use the so-called *reduced chi-square*, which is given by  $\chi_{\text{red}}^2 = \chi^2/\nu$ . A useful practical rule is that a good fit is achieved for a reduced chi-square equal to unity, i.e.  $\chi^2 \sim \nu$ .

The binned data sets we use in our analysis (see Table 3.1) show that each bin contains a sufficient number of observations to be approximated by the Gaussian limit of the Poisson distribution. In the case of the cluster power spectrum, the Gaussian approximation is standard. Thus we restrict to the Gaussian picture hereafter. In addition to this, our data are correlated and we use the covariance matrices already defined.

### 4.1.3 $C$ -statistics

The probability distribution for bins with a small number of observations is Poissonian and the  $\chi^2$  statistics is not appropriate anymore. In fact, if the counts in a bin are less than  $\sim 10$ , the Poisson distribution differs significantly from the Gaussian one. In this case one can use the  $C$ -statistics (Cash 1979):

$$P_i = \frac{\mu_i^{x_i} e^{-\mu_i}}{x_i!}, \quad P = \prod_{i=1}^n P_i, \quad (4.7)$$

where  $x_i$  is the observation and  $\mu_i$  the theoretical expectation in the bin  $i$ . The quantity which needs to be minimised to obtain the best fit model is then:

$$\chi_{\text{cash}}^2 = -2 \ln P = -2 \sum_{i=1}^N (x_i \ln \mu_i - \mu_i - \ln x_i!) \approx -2 \sum_{i=1}^N (x_i \ln \mu_i - \mu_i). \quad (4.8)$$

### 4.1.4 Confidence regions and marginalisation

Once the best fit parameters are obtained, the accuracy of estimates can be visualised by showing the confidence limits (or region) around them. This can be done by selecting a region in the  $m$ -dimensional parameter space, around the best fit value, which encloses a given percentage of the probability distribution. One should thus integrate the likelihood surface and compute where a certain percentage of points lie. Typically this is obtained by considering compact regions of constant  $\chi^2$  boundaries. However, in case the  $\chi^2$  has more than one minimum, a non-connected confidence region is more appropriate. For multi-variate Gaussian distributions the regions of constant  $\chi^2$  are ellipsoidal regions. The usually defined confidence levels are the 68.3%, 95.4% and 99.5%, which correspond to the conventional 1, 2 and  $3\sigma$ , in the case of Gaussian distributions.

Knowing the joint posterior probability distribution of  $m$  parameters, i.e.

$$P(\Theta|D) \equiv P(\theta_1, \dots, \theta_m|D) \equiv P(\theta_1 = t_1, \dots, \theta_m = t_m|D) , \quad (4.9)$$

one could be interested into the probability distribution of a single parameter, for example  $\theta_1$ , regardless of the others,  $\theta_2, \dots, \theta_m$ . In this case, one can derive the **marginal probability** by integrating over all the parameters, except the one of interest. Mathematically, this translates into

$$P(\theta_1|D) = \int d\theta_2 \dots \int d\theta_m P(\theta_1, \dots, \theta_m|D) . \quad (4.10)$$

This is the simplest 1D case, where the marginalised distribution of a parameter is given by the maximum (or the mean) of the distribution, together with confidence limits. In cosmological analysis, however, the marginalisation is often used to obtain 2D posterior distributions for various combination of parameter pairs  $\theta_1, \theta_2$ , which can show possible degeneracies. This translates simply into:

$$P(\theta_1, \theta_2|D) = \int d\theta_3 \dots \int d\theta_m P(\theta_1, \theta_2, \dots, \theta_m|D) . \quad (4.11)$$

We include few 2D marginalised posterior probabilities contour plots of our results, together with their interpretation.

## 4.2 Sampling methods

Bayesian methods are often used to compare physical models to data. The parameters in different models are most easily constrained by sampling from the posterior distribution. The sampling methods scale at best linearly with the dimension of the parameter space, compared to integration which is computationally more expensive. The most common method for sampling from a general distribution in high-dimensions is the Markov Chain Monte Carlo (MCMC), based on the Monte Carlo approximation method and on the useful properties of the Markov Chains, which we both revise in this Section. MCMC generates a set of points in the parameter space, which have the same distribution as the target posterior distribution, by computing the Bayesian likelihood at random points selected using the Metropolis Hastings algorithm. Then, from the correlated samples of the posterior distribution it is possible to derive other significant statistical quantities.

### 4.2.1 Markov chains

Let us assume that  $X$  is a random variable and  $X_t$  is its value at time  $t$ . We define  $X$  as a **Markov process** if the probability distribution at the present status,  $X_t$ , depends only on the previous status,  $X_{t-1}$ , and not on the entire history:

$$P(X_t = x_t | X_0 = x_0, \dots, X_{t-1} = x_{t-1}) = P(X_t = x_t | X_{t-1} = x_{t-1}) . \quad (4.12)$$

A **Markov chain**  $\{X_t\}$  is a sequence of random variables  $X_0, X_1, \dots, X_t$  generated by a Markov process. If the probability is regular, the chain will gradually forget about the initial state  $X_0$  and approach an invariant distribution after a sufficient number of steps. In fact, the main property of Markov chain is *stationarity*: there exists a probability distribution  $\phi(x)$  such that if  $X_t \sim \phi(x)$ , then  $X_{t+1} \sim \phi(x)$ . In Fig. 4.1 we show an example of a sequence which approaches an asymptotic stationary distribution. If a chain starts in a random position of the parameter space, it will take some time (few hundreds of steps) to equilibrate. The term *burn-in period* refers to those iterations at the beginning of the chain, which are still far from converging to the invariant distribution: these chain steps are usually discarded, before computing the chain analysis. It could happen that a Markov chain looks like it has reached its invariant distribution when it has not: this is called *pseudo-convergence*. This happens when there are regions of the parameter space which are weakly connected by the chain, so that the chain remains there for many iteration. In the opposite case of a *good mixing*, the true convergence is reached after a sufficient number of steps.

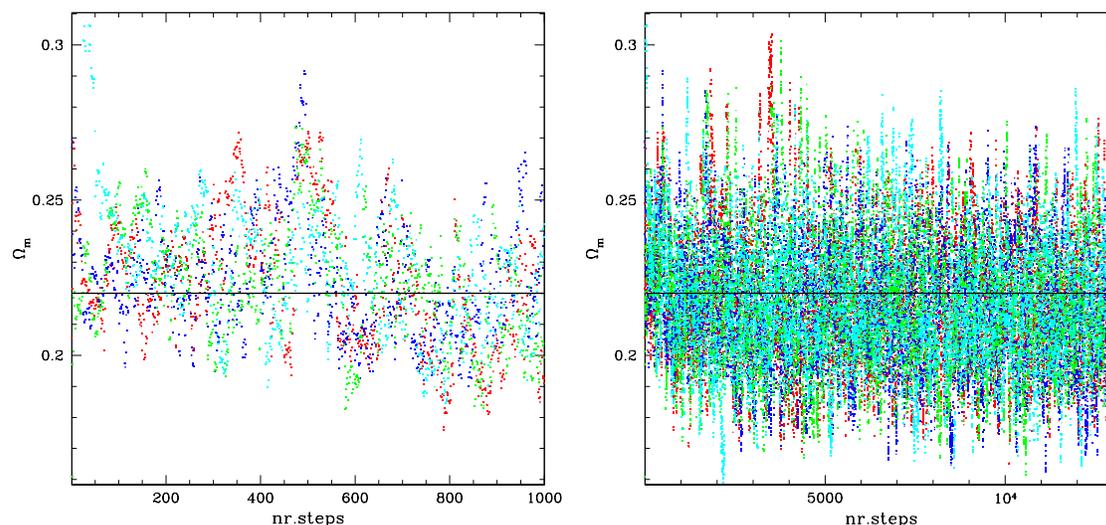
### 4.2.2 Monte Carlo methods

The **Ordinary Monte Carlo** (OMC) is based on the idea of using independent and identically distributed (IID) simulations  $X_1, \dots, X_n$  of a random process  $X$  to approximate the expectation

$$\mu = \langle h(X) \rangle = \int h(x) \pi(x) dx , \quad (4.13)$$

where  $h$  is a real-valued function and  $\pi(x)$  is the common distribution of the variables. The OMC method works by selecting  $n$  samples  $X_i$  from  $\pi(x)$  and then approximating Eq. (4.13) by the sample average of the  $h(X_i) = Y_i$ , having  $\langle Y_i \rangle = \mu$  and variance  $\text{Var}[Y_i] = \sigma^2$ :

$$\hat{\mu}_n \simeq \langle h(X) \rangle = \frac{1}{n} \sum_{i=1}^n h(X_i) = \frac{1}{n} \sum_{i=1}^n Y_i . \quad (4.14)$$



**Figure 4.1:** Four MCMC chains (blue, red, green, cyan) from our runs, which are approaching a stationary distribution for the  $\Omega_m$  parameter. The horizontal line represents the best fit value,  $\Omega_m = 0.22$ , while the marginalised mean value and  $1\sigma$  error are  $\Omega_m = 0.215 \pm 0.02$ . Left panel: first 1000 steps (burn-in). Right panel: full chains. Note that the number of parameters being varied in total here is nine, which justifies the quite significant spread of the chains.

The Central Limit Theorem (CLT) states that the mean of  $n$  independent random variables, with the same mean and variance, is well-approximated by a normally distributed random variable, for a large enough  $n$  and regardless of the variables distribution. The mean of this normal distribution is equal to the mean of random variables. The variance is instead equal to the variance of the random variables divided by the sample size. According to the CLT then,  $\hat{\mu}_n \sim \mathcal{N}(\mu, \sigma^2/n)$ .

### 4.2.3 MCMC methods

MCMC methods are used to simulate non-standard, complex multivariate distributions. The most important of these methods is the **Metropolis-Hastings** (M-H) algorithm, created by Metropolis & Ulam (1949); Metropolis et al. (1953) and improved by Hastings (1970). According to the previous definitions, a MCMC is an OMC where the IID simulations are replaced by a Markov Chain process with a certain equilibrium distribution. Following the notation in the previous sections, one needs to create a Markov chain which has  $\pi(x)$  as asymptotic invariant distribution.

Assuming that we have  $m$  burn-in steps, we are then able to estimate  $\langle h(X) \rangle$  as (see Eq. 4.14)

$$\hat{\mu}_n^* = \frac{1}{n-m} \sum_{i=m+1}^n h(X_i). \quad (4.15)$$

One of the advantages of the MCMC methods is that the number of iterations required for convergence roughly scales linearly with the number of parameters. However as the number of parameters in the models grows, the computational cost can be quite challenging. There is furthermore a dependence on the nature of the posterior distribution itself: longer chains need to be constructed when the posterior departs from a Gaussian distribution. In this Section, we revise only the parts concerning the MCMC methods which are relevant for our work. For a more detailed description, we refer to: MacKay (2003); Gamerman & Lopes (2006); Neal (1993); Gilks (1999); Hobson et al. (2010); Geyer (2011).

### Metropolis-Hastings algorithm

The most interesting aspect is to understand how the Metropolis-Hastings algorithm generates a sequence of correlated variables from a certain distribution. This can be summarised into the following steps.

- i) Set any initial random value  $X_0$ .
- ii) Sample a candidate point  $X^*$  from a conditional probability density given the current state  $X_t$ , denoted  $\psi(\cdot|X_t)$ . The latter can have any form, but the final stationary distribution of the chain will be  $\pi(x)$ .
- iii) Calculate the Hastings ratio defined as

$$r(X^*|X_t) = \frac{\pi(X^*)\psi(X|X^*)}{\pi(X)\psi(X^*|X)}. \quad (4.16)$$

- iv) Sample a Uniform variable  $U$  in  $(0, 1)$ .
- v) Define the acceptance probability for the state  $X^*$  as

$$a(X^*|X_t) = \min [1, r(X^*|X_t)] = \min \left[ 1, \frac{\pi(X^*)}{\pi(X)} \right]. \quad (4.17)$$

The last equivalence comes from the fact that this algorithm typically considers only symmetric proposals:  $\psi(X^*|X) = \psi(X|X^*)$ . Then, if  $U \leq a(X^*|X_t)$ , the

candidate point is accepted and the next state becomes  $X_{t+1} = X^*$ . Otherwise, the candidate point is rejected and the chain does not move, i.e.  $X_{t+1} = X_t$ . This step is usually called *Metropolis rejection*.

vi) Increment  $t$  and start again from step ii).

Note that in general, the choice of the proposal density  $\psi(\cdot|X_t)$  is relevant for the algorithm efficiency: for a quicker convergence, it is more appropriate to choose one with a shape which is similar to the posterior. Similarly, the choice of its covariance matrix is important. A tight distribution will cause slow mixing and high acceptance rate, while a wide distribution will result in low acceptance and no movement of the chain, hence resulting in slow mixing as well.

### Convergence

The next question to address is about how long do we need to run the chain to be sure that it has converged. Having sampled sufficiently enough points in the parameter space the MCMC chain equilibrates to the target distribution and it is said to have reached convergence. In practice, convergence may be tested through one of the many convergence tests.

The simplest method to solve the convergence issue is by running multiple chains in parallel, with *over-dispersed* initial values and compare the estimates  $\hat{\mu}_n^*$ . A fundamental problem is that there will always be regions of the target distribution that have not been covered by the finite chain. Over-dispersion can be obtained after running a single chain and use the variance from the chain itself. Let us assume we are interested in a quantity  $\phi$  from the chain, e.g. some parameters or any function of them. Let us further assume that we run  $m$  parallel sequences of length  $n$  and label the quantities  $(\phi_{ij})$ ,  $j = 1, \dots, n$  and  $i = 1, \dots, m$ . We hence compute two quantities: the between-sequence variance  $V$  and the within-sequence variances  $W$ , i.e.

$$V = \frac{n}{m-1} \sum_{i=1}^m (\bar{\phi}_i - \bar{\phi})^2 \quad , \quad W = \frac{1}{m} \sum_{i=1}^m s_i^2 \quad , \quad (4.18)$$

where

$$\bar{\phi}_i = \frac{1}{n} \sum_{j=1}^n \phi_{ij} \quad , \quad \bar{\phi} = \frac{1}{m} \sum_{i=1}^m \bar{\phi}_i \quad , \quad s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (\phi_{ij} - \bar{\phi}_i)^2 \quad . \quad (4.19)$$

$W$  is the *average variance* of all the chains, while  $V$  measures the *variance of the averages* of the chains. Note that the between-sequence variance  $V$  contains a factor

$n$  because it is based on the variance of the within-sequence means,  $\bar{\phi}_i$ , each of which is an average of  $n$  values  $\phi_{ij}$ . An overestimate of the variance of  $\phi$  in the target distribution is

$$\text{Var}^*(\phi) = \frac{n-1}{n} W + \frac{1}{n} V. \quad (4.20)$$

Further,  $W$  is an underestimate of the target variance  $\text{Var}(\phi)$ , because individual chains do not have time to cover the entire target distribution. However it holds that:

$$\lim_{n \rightarrow \infty} \text{Var}^*(\phi) = \lim_{n \rightarrow \infty} W = \text{Var}(\phi). \quad (4.21)$$

Convergence can now be established by monitoring

$$\sqrt{R} = \sqrt{\frac{\text{Var}^*(\phi)}{W}}, \quad (4.22)$$

which approaches 1 at convergence, in the so-called **(R-1) test**. Note that there are many other convergence criteria (see for example Cowles & Carlin 1996), which however we do not consider in our work.

### 4.3 The Cosmological Monte-Carlo

The MCMC method in a cosmological context has been developed by the Cosmological Monte-Carlo (COSMOMC<sup>1</sup>) publicly available parameter estimation tool. This software has been developed by Lewis & Bridle (2002) and it is supported by the software forum<sup>2</sup>. It includes two main Fortran 90 programs: a MCMC code for exploring cosmological parameter space (`cosmomc*`) and a code for analysing Monte-Carlo samples and importance sampling (`getdist*`). COSMOMC is originally based on the M-H sampler, but there are options for other sampling and methods for exploring fast/slow parameter space. It contains the Boltzmann code CAMB<sup>3</sup>, i.e. the Code for Anisotropies in the Microwave Background by Lewis et al. (2000) for calculating theoretical matter power spectrum and CMB anisotropies, based on CMBFAST<sup>4</sup> by Seljak & Zaldarriaga (1996). Useful overviews can be found also in Lewis & Bridle (2006); Lewis (2011). The data sets available include WMAP7, SN-Ia, BAO and HST. The last version contains also Planck data via a specific Planck likelihood code.

<sup>1</sup><http://cosmologist.info/cosmomc/>

<sup>2</sup><http://cosmocoffee.info/>

<sup>3</sup><http://camb.info/>

<sup>4</sup><http://cmbfast.org>

### How does it work?

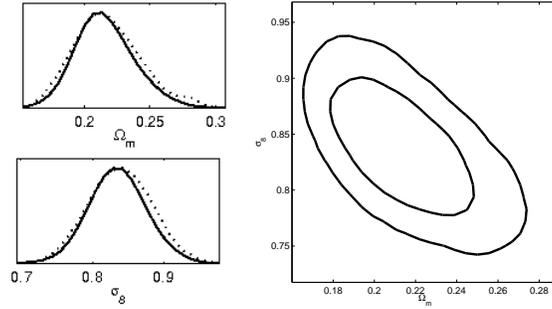
The `cosmomc*` algorithm produces chain files, containing all the accepted set of parameters: the first column gives the sample weight, the second the likelihood and from third on all the parameters (primary and derived ones). The program takes as inputs the central values, ranges and standard deviations of the parameters, from an input file called `params.ini`. The default primary parameters are:  $\Omega_b h^2$ ,  $\Omega_c h^2$ ,  $\Omega_k$ ,  $\log(10^{10} A_s)$ ,  $w$ ,  $\tau$ ,  $n_s$ ,  $n_t$ ,  $\theta$ ,  $m_\nu$ ,  $n_\nu$ ,  $n_{\text{run}}$  and  $r = A_t/A_s$ . The derived ones are instead:  $H_0$ ,  $t_0$  and  $\Omega_\Lambda$ . Priors on the fundamental parameters have to be assigned: the final results should not depend on this choice. This helps the program to estimate a proposal distribution: if one provides an input covariance matrix with the parameters correlation, the estimation procedure can be significantly improved. It is also possible to get an estimated covariance matrix for the best fit model, as well as to learn the proposal distribution from the covariance with the MPI option (run on different cluster nodes). The process will stop when convergence is reached and the diagnostics for multiple chains is done by the (R-1) test (see Eq. 4.22), using the second half of each chain. Note that varying some parameters can be much more computationally expensive (the so-called *slow* parameters) than varying others (the so-called *fast* parameters).

### Chains analysis

Once the convergence is reached and the burn-in steps removed, the `getdist*` analyses the chains and generates the following outputs:

- i) `.likestats`, which contains the best fit sample model and likelihood, as well as the limits from the extremal values of the posterior distribution;
- ii) `.margestat`, which gives mean, standard deviation and marginalised limits for all the parameters;
- iii) `.covmat`, which is a covariance proposal matrix for generating other chains;
- iv) Matlab or Python files for generating marginalised plots (1D, 2D, 3D).

Generally, it is useful to consider both the likelihood and marginalised distributions. The likelihood function is used to estimate the best fit model parameters as those which maximise the likelihood itself, or equivalently the logarithm of the likelihood. The marginalised distribution instead gives the projected shape of probability density in one or two dimensions only. We show in Fig. 4.2 an example of 1D and 2D marginalised posteriors for the parameters  $\Omega_m$  and  $\sigma_8$ , obtained from one of our runs.



**Figure 4.2:** 1D and 2D marginalised posterior probabilities for parameters  $\Omega_m$  and  $\sigma_8$ , for same run used in Fig. 4.1. The best fit values of the parameters are  $\Omega_m = 0.22$  and  $\sigma_8 = 0.83$ , while the marginalised mean and  $1\sigma$  error are  $\Omega_m = 0.215 \pm 0.02$  and  $\sigma_8 = 0.84 \pm 0.04$ . Left panel: the solid lines show the fully marginalised posterior, the dotted lines show the mean likelihood of the samples. For Gaussian distributions they should be the same. However, for skewed distributions, or if chains are poorly converged, they will not be. Right panel: 2D marginalised posterior probability for the pair  $\Omega_m - \sigma_8$ . The external contour corresponds to the 95% confidence level, while the internal contour corresponds to the 68% c.l.

### Importance sampling

This program also does post-processing on chains, like **importance sampling**: this is a quick way to re-weight the results for different priors, new data sets or small theoretical corrections. More precisely, according to the previous notation, the expectation value of a function  $h$  under the distribution  $\pi(x)$  (see Eq. 4.13) can be converted into:

$$\langle h(X) \rangle_\pi = \int h(x) \pi(x) dx = \int h(x) \frac{\pi(x)}{\pi'(x)} \pi'(x) dx = \langle h(x) \frac{\pi(x)}{\pi'(x)} \rangle_{\pi'}, \quad (4.23)$$

where  $\pi'(x)$  is another distribution.

## 4.4 Combined maxBCG analysis

The purpose of our analysis is to constrain cosmology and fit for the richness-mass relation and scatter around it, simultaneously. To do this, we perform a Bayesian parameter estimation by running an MCMC, using M-H sampler with a modified version of COSMOMC. In this Section we describe the set up of the runs, namely the code which produces the theory, the data sets we need and the parameters we vary. We also list the different runs, which are characterised by different combinations of data sets and free parameters.

### The code

We work on a Fortran 90 code, which implements the standard cosmological framework and cluster statistics, first developed by Battye & Weller (2003). In particular, the code provides calculation of the mass function and cluster number counts, by linking to CAMB and solving perturbation equation. The code has been adapted to the data sets we want to use, by applying some crucial modifications. In particular, we need first to introduce the mass-observable relation and convert functions of the true mass into functions of the real observable (optical richness). Secondly, we have to implement the bias and the observed power spectrum theoretical expectation. Finally, the coding of primordial non-Gaussianity routines is necessary to introduce  $f_{\text{NL}}$  parameter in the analysis. We then need to link properly the modified code to COSMOMC in order to compare theory and real data for the parameter estimation, by means of the  $\chi^2$  statistic.

### The data sets

As already introduced, we use the following data sets from the maxBCG sample, integrated over all the redshift range:

- (i) cluster number counts  $\Delta N$ , divided into 10 richness bins;
- (ii) total mass of clusters  $\Delta N \bar{M}$ , divided into 6 richness bins;
- (iii) cluster power spectrum  $\tilde{P}^{\text{obs}}$ , divided into 18  $k$ -bins.

We also include the corresponding covariance matrices. In addition to the cluster data, we also use the CMB power spectra from WMAP7 (Larson et al. 2011), in the cases specified below.

## The parameters

We assume a flat  $\Lambda$ CDM cosmological model. In Table 4.1 we list all the parameters of the analysis, including their flat or Gaussian priors. The main parameters we consider in our analysis are:

- i) standard cosmological parameters: CDM energy density  $\Omega_c$ , amplitude of primordial perturbations  $\log(10^{10}A_s)$ , primordial non-Gaussianity parameter  $f_{\text{NL}}$ ;
- ii) scaling relation parameters:  $\ln N_1$ ,  $\ln N_2$  and the scatter  $\sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}$ ;
- iii) nuisance parameters: weak lensing mass bias  $\beta$ , power spectrum non-linear correction  $q_{\text{NL}}$ , photo-z errors  $\sigma_z$ , scatter on bias  $B$ .

When using cluster data alone, we fix the Hubble parameter  $h = 0.7$ , primordial spectral index  $n_s = 0.96$  and baryon density  $\Omega_b = 0.044$ , as these parameters are not easily constrained in this case. When adding the external CMB data instead, we relax the assumptions on  $h$ ,  $n_s$ ,  $\Omega_b$  and we add the optical depth  $\tau$  as well as the amplitude of CMB SZ template  $A_{\text{sz}}$ . Note that for us the total matter energy density  $\Omega_m$  and the power spectrum normalization  $\sigma_8$  are always derived parameters.

## The runs

In order to understand the constraining power of each data set, we estimate the high-dimensional posterior parameter distributions in the following cases:

1. **counts only**: 6 free parameters  $[\Omega_c, \log(10^{10}A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta]$ , with counts and masses data sets;
2. **counts with  $f_{\text{NL}}$** : 7 free parameters  $[\Omega_c, \log(10^{10}A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, f_{\text{NL}}]$ , with counts and masses data sets;
3. **counts+ $P(\mathbf{k})$** : 9 free parameters  $[\Omega_c, \log(10^{10}A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, q_{\text{NL}}, \sigma_z, B]$ , with cluster counts, masses and cluster power spectrum;
4. **counts+ $P(\mathbf{k})$  with  $f_{\text{NL}}$** : 10 free parameters  $[\Omega_c, \log(10^{10}A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, q_{\text{NL}}, \sigma_z, B, f_{\text{NL}}]$ , with cluster counts, masses and cluster power spectrum;
5.  **$P(\mathbf{k})$  only**: 9 free parameters  $[\Omega_c, \log(10^{10}A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, q_{\text{NL}}, \sigma_z, B]$ , with cluster power spectrum only;

6. **CMB only**: 7 free parameters  $[\Omega_b, h, \tau, n_s, A_{sz}, \Omega_c, \log(10^{10} A_s)]$ , with CMB data only;
7. **CMB+clusters**: 14 free parameters  $[\Omega_b, h, \tau, n_s, A_{sz}, \Omega_c, \log(10^{10} A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, q_{\text{NL}}, \sigma_z, B]$ , with CMB and all cluster data;
8. **CMB+clusters with  $f_{\text{NL}}$** : 15 free parameters  $[\Omega_b, h, \tau, n_s, A_{sz}, \Omega_c, \log(10^{10} A_s), \ln N_1, \ln N_2, \sigma_{\ln M|N_{\text{gal}}^{\text{obs}}}, \beta, q_{\text{NL}}, \sigma_z, B, f_{\text{NL}}]$ , with CMB and all cluster data.

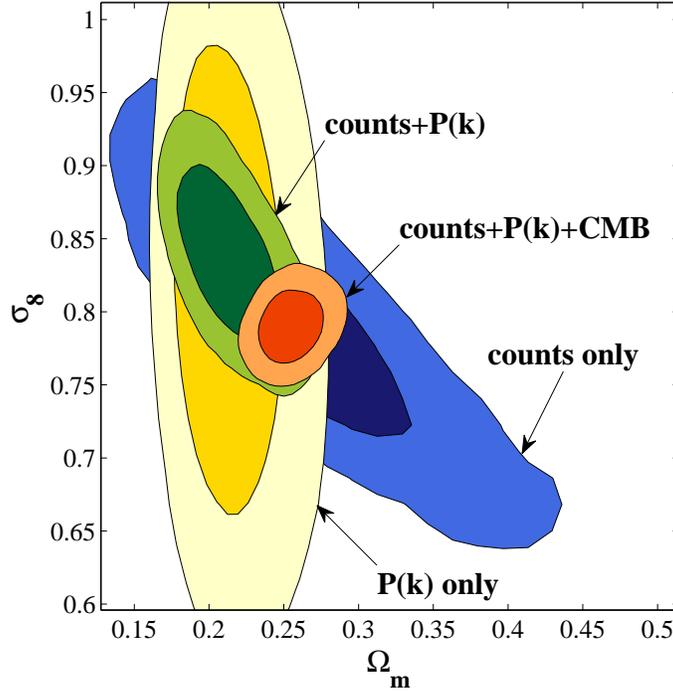
The addition of the  $f_{\text{NL}}$  parameter significantly affects the computational time because of the complexity of the calculations (integrations mainly). Obviously, the runs including the CMB data set take longer to converge because more parameters are varying simultaneously. Typically, we run chains of the order of 20,000 steps, reaching a convergence of  $R-1 \leq 0.03$ .

## 4.5 Results

We summarise our results in Table 4.2, and we show in Figs. 4.3, 4.6, 4.7, 4.9 and 4.10 the 2D 68% and 95% marginalised confidence regions for different pairs of parameters in our analysis. The colour scheme is the same for all figures: blue contours refer to runs with counts and WL mean masses data only, green contours include in addition the cluster power spectrum data, while orange contours also include CMB data.

### 4.5.1 $\Omega_m - \sigma_8$ contours

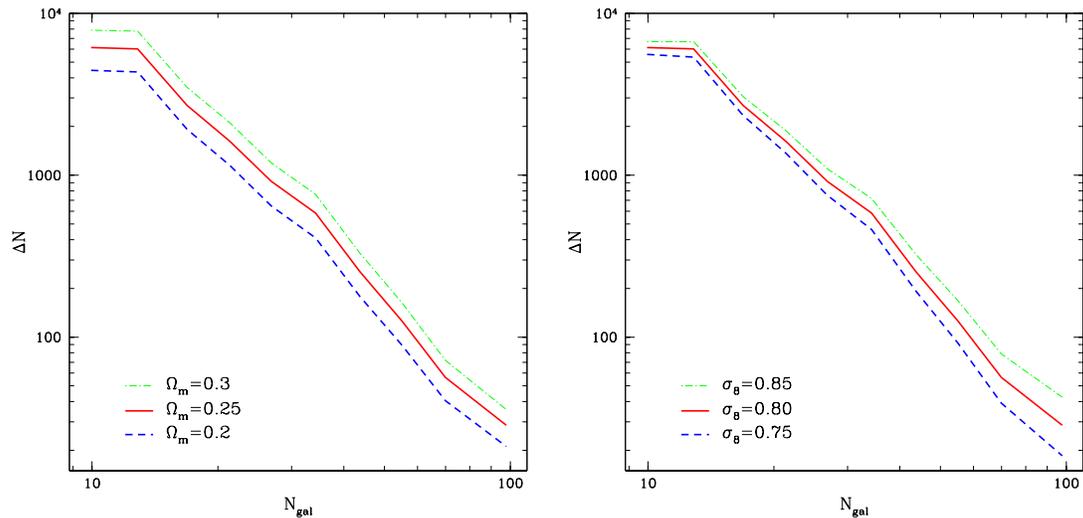
The joint constraint in the  $\Omega_m - \sigma_8$  plane in Fig. 4.3 displays the typical degeneracy from cluster counts: the counts increase with increasing  $\Omega_m$  and  $\sigma_8$  values, hence any increase in  $\Omega_m$  must be balanced by a decrease in  $\sigma_8$  (and vice versa), to keep the abundances at the observed values. This behaviour is also confirmed by Fig. 4.4, showing the dependence of the richness binned number counts on  $\Omega_m, \sigma_8$  parameters. We already discuss this dependence for the mass function and in Fig. 3.1, but it is worth to recall it here. In both panels, the red curve corresponds to the best fit of the run **counts only**:  $\Omega_m = 0.25$  and  $\sigma_8 = 0.80$ . In the left panel, for fixed  $\sigma_8 = 0.80$ , we see that the predicted abundances increases for increasing value of  $\Omega_m$ . In the right panel, for fixed  $\Omega_m = 0.25$ , it is evident that any increase of  $\sigma_8$  leads to an increase of counts more evident at high richness than the increase at low richness. Thus, if one of the two parameters increases, the other one has to decrease to match the observed counts.



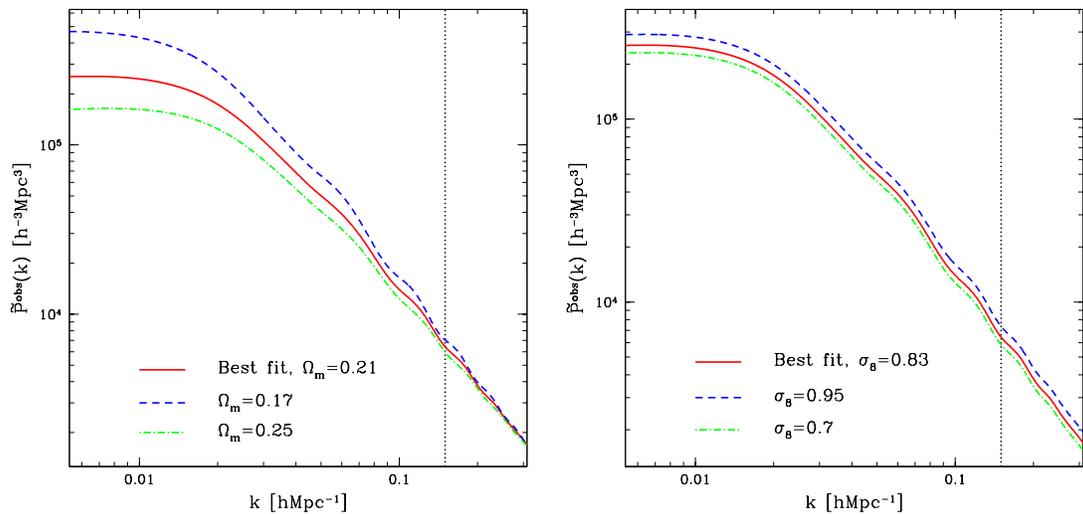
**Figure 4.3:** Marginalised posterior probability distributions on the parameters  $\Omega_m - \sigma_8$  for the runs using **counts only** (blue), **counts+ $P(k)$**  (green) and **counts+ $P(k)$ +CMB** (orange), at 68% and 95% confidence levels. The yellow contours show the joint constraints in the case of  $P(k)$  data only.

The constraints on individual parameters with counts and masses only are consistent with Rozo et al. (2010), and we find  $\Omega_m = 0.25 \pm 0.06$ ,  $\sigma_8 = 0.80 \pm 0.06$  ( $1\sigma$  errors throughout), while the errors are improved by a factor between 1.5 and 3, depending on the parameter, when adding the maxBCG power spectrum: in this case we obtain  $\Omega_m = 0.215 \pm 0.022$ ,  $\sigma_8 = 0.84 \pm 0.04$ . Combining these results with the CMB data, the constraints then shrink to  $\Omega_m = 0.255 \pm 0.014$  and  $\sigma_8 = 0.790 \pm 0.016$ : the contribution of the CMB tightens the errors by a further factor of two.

As an interesting comparison, we also show the joint constraints for the case of  $P(k)$  data only (yellow contours in Fig. 4.3), with a prior on the scaling relation parameters: the degeneracy direction is complementary to that of the counts. It is evident that  $P(k)$  essentially constraints only the  $\Omega_m$  parameter, because of the degeneracy with  $\sigma_8$ , as we would expect by definition. To have a better understanding



**Figure 4.4:** Cluster number counts for different values of  $\Omega_m$  (left panel) and  $\sigma_8$  (right panel). The red curve corresponds to the best fit of the run **counts only**:  $\Omega_m = 0.25$  and  $\sigma_8 = 0.80$ . In the left panel, we fix  $\sigma_8 = 0.80$  and give  $\Omega_m$  the values 0.2 (blue dashed), 0.25 (red solid), 0.3 (green dotted-dashed). In the right panel, we fix  $\Omega_m = 0.25$  and assign to  $\sigma_8$  the values 0.75 (blue dashed), 0.80 (red solid), 0.85 (green dotted-dashed).



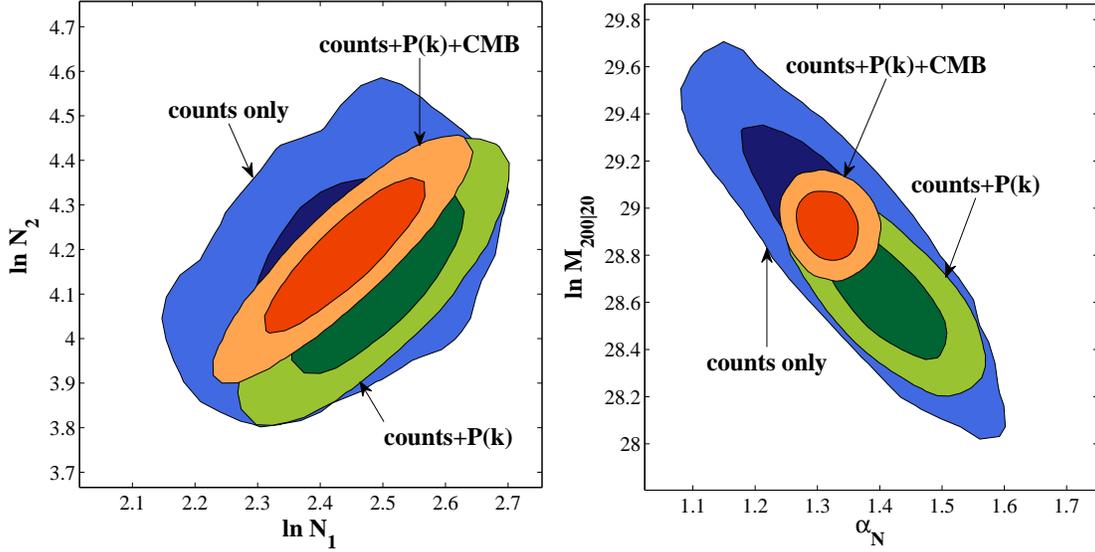
**Figure 4.5:** Full observed power spectrum for different values of  $\Omega_m$  (left panel) and  $\sigma_8$  (right panel). The red curve corresponds to the best fit of the run  $\mathbf{P}(\mathbf{k})$  **only**:  $\Omega_m = 0.21$  and  $\sigma_8 = 0.83$ . The other curves select values at approximately  $1\sigma$  level of the marginalised posterior distribution. In the right panel, we fix  $\sigma_8 = 0.80$  and give  $\Omega_m$  the values 0.17 (blue dashed), 0.21 (red solid), 0.25 (green dotted-dashed). In the right panel, we fix  $\Omega_m = 0.21$  and assign to  $\sigma_8$  the values 0.95 (blue dashed), 0.83 (red solid), 0.7 (green dotted-dashed).

of the power spectrum dependence on these parameters, we show the full observed power spectrum for different values of  $\Omega_m$  and  $\sigma_8$  in Fig. 4.5. In both panels, the red curve corresponds to the best fit of the run  **$\mathbf{P}(\mathbf{k})$  only**:  $\Omega_m = 0.21$  and  $\sigma_8 = 0.83$ . The left panel shows the power spectrum varying  $\Omega_m$ , at fixed  $\sigma_8 = 0.83$ , while the right panel shows the power spectrum varying  $\sigma_8$ , at fixed  $\Omega_m = 0.21$ . We notice that if we increase (decrease)  $\Omega_m$ , the peak of the power spectrum decreases (increases) while also being shifted to higher (lower) values of  $k$ , while  $\sigma_8$  simply changes the overall normalisation. In our case, the size of the posterior on  $\sigma_8$  depends on the assumptions made on the cluster bias: allowing for a completely free bias would cause a complete degeneracy with  $\sigma_8$ . However here the degeneracy is partially broken because we are assuming that the bias is centred around the predicted values from the mass function, allowing only for deviations from this (parametrised by the scatter  $B$ ), whose amplitude is limited by the prior on  $B$ .

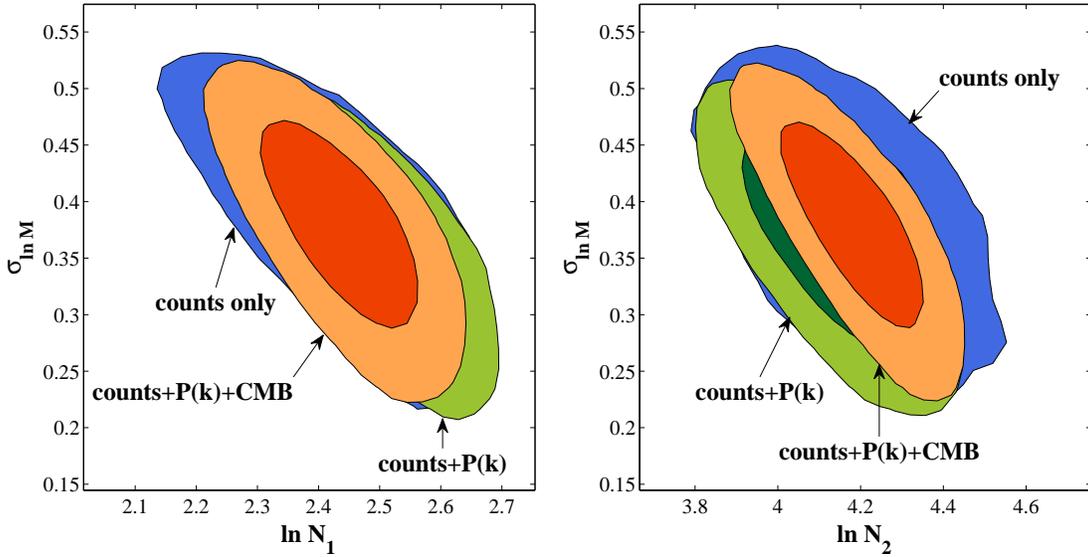
### 4.5.2 Scaling relation parameters contours

In Fig. 4.6, left panel, we show the marginalised posterior probability contours of the scaling relation parameters  $\ln N_1 - \ln N_2$ . These two parameters correspond to the two pivot points in (log) mass  $\ln M_1, \ln M_2$ , which define a linear scaling relation in the  $\ln N_{\text{gal}} - \ln M$  plane, with slope  $\alpha_N$ , intercept  $\ln M_{200|20}$  and scatter  $\sigma_{\ln M}$ .

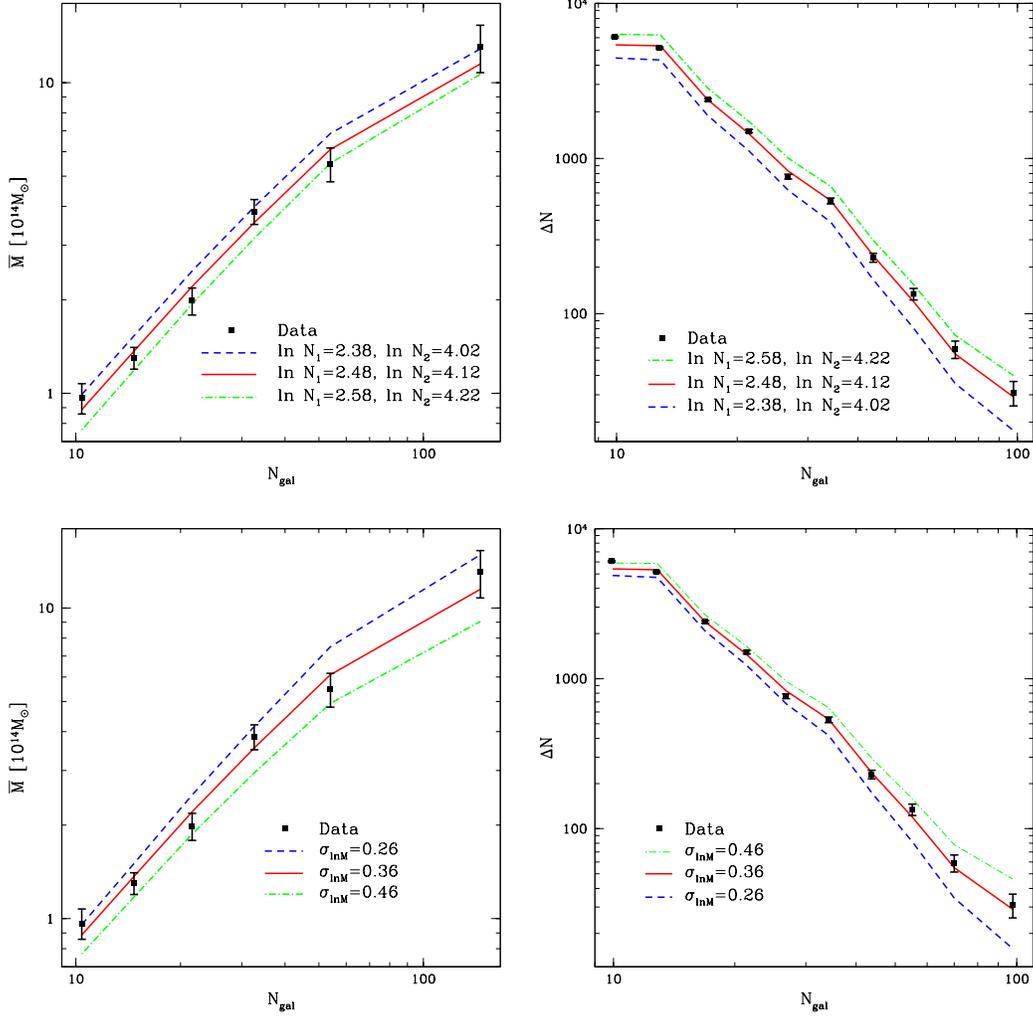
From Fig. 4.8 we can better understand the degeneracy between these two parameters. The left panels show the mean masses estimates in richness bins, while the right panels show the counts in richness bins. In all plots the best fit is shown by a red solid curve, while any increase or decrease of scaling relation parameters or scatter is plotted in green dotted-dashed or blue dashed line, respectively. Note that we do not plot the linear relation across the two points  $(\ln N_1, \ln M_1), (\ln N_2, \ln M_2)$ , but instead we connect the mass estimates in each richness bin. This can give us an intuition on the approximate scaling relation behaviour. If we fix the two pivot points in mass, as  $\ln N_1$  increases, so does  $\ln N_2$ , to keep the slope of the linear relation at the observed value: this is shown by the green dotted-dashed curve in the top left panel of Fig. 4.8. At the same time, the intercept  $\ln M_{200|20}$  decreases, resulting in a parallel displacement of the linear scaling relation towards higher richness values. We also notice that for fixed  $\ln N_{\text{gal}}$ , the corresponding mass has now decreased. On the other hand, the blue dashed line displays the case of a decreasing  $\ln N_1$ , and thus  $\ln N_2$ , and an increasing intercept.



**Figure 4.6:** Left panel: constraints on the scaling relation parameters  $\ln N_1 - \ln N_2$ . Note that  $\ln N_1 \equiv \ln N_{\text{gal}}|M_1$  and  $\ln N_2 \equiv \ln N_{\text{gal}}|M_2$ , where  $M_1 = 1.3 \times 10^{14} M_\odot$  and  $M_2 = 1.3 \times 10^{15} M_\odot$ . Right panel: constraints on the slope  $\alpha_N$  and intercept  $\ln M_{200|20}$  of the scaling relation. The runs used for this plot are: **counts only** (blue), **counts+P(k)** (green) and **counts+P(k)+CMB** (orange). The confidence levels are at 68% and 95%.



**Figure 4.7:** Left panel: constraints on the scaling relation parameters  $\ln N_1 - \sigma_{\ln M}$ . Right panel: constraints on the scaling relation parameters  $\ln N_2 - \sigma_{\ln M}$ . The runs used for this plot are: **counts only** (blue), **counts+P(k)** (green) and **counts+P(k)+CMB** (orange). The confidence levels are at 68% and 95%.



**Figure 4.8:** Predictions for mean masses (left panels) and number counts (right panels) varying scaling relation parameters. First row: for fixed scatter  $\sigma_{\ln M} = 0.36$ ,  $\ln N_1 \equiv \ln N_{\text{gal}}|M_1$ , with  $M_1 = 1.3 \times 10^{14} M_{\odot}$ , is equal to  $\{2.38, 2.48, 2.58\}$  and  $\ln N_2 \equiv \ln N_{\text{gal}}|M_2$ , with  $M_2 = 1.3 \times 10^{15} M_{\odot}$ , is equal to  $\{4.02, 4.12, 4.22\}$ , which we both plot in blue dashed, red (best fit) and green dotted-dashed respectively. Second row: for fixed  $\ln N_1 = 2.48$ ,  $\ln N_2 = 4.12$ ,  $\sigma_{\ln M}$  is equal to  $\{0.26, 0.36, 0.46\}$ , which we show in blue dashed, red (best fit) and green dotted-dashed respectively. The run used for these plots is **counts+P(k)**.

In order to keep the scaling relation to the observed value, then, the scatter plays a key role: even if the slope is kept constant by the increasing of both  $\ln N_1$  and  $\ln N_2$ , the decrease of the intercept is compensated by a decrease of  $\sigma_{\ln M}$ . This is confirmed by Fig. 4.7, showing the joint constraints  $\ln N_1 - \sigma_{\ln M}$  and  $\ln N_2 - \sigma_{\ln M}$ : any increase of  $\ln N_1$  or  $\ln N_2$  corresponds to a decrease of  $\sigma_{\ln M}$ . The converse is of course also true. Furthermore, the bottom left panel of Fig. 4.8 displays an analogous displacement of the predicted masses to higher  $N_{\text{gal}}$  values obtained by increasing the scatter (green dotted-dashed line) and to lower  $N_{\text{gal}}$  values obtained by decreasing the scatter (blue dashed line). As a consequence, the mean masses corresponding to the same richness result lower or higher, respectively.

Finally, from the right panels of Fig. 4.8 we can deduce another confirmation from the cluster number counts. If the scaling relation parameters or scatter increase, we will count more objects in general, thus the counts increase in each richness bin. This is motivated by the fact that if we have lower mass estimates, we automatically have more structures due to the shape of the mass function. The opposite arguments can also be formulated for a decrease in  $\ln N_1$ ,  $\ln N_2$  and  $\sigma_{\ln M}$ . As a consequence, again, any increase of  $\ln N_1$  and  $\ln N_2$  implies a decrease of  $\sigma_{\ln M}$ , to keep the abundances to the observed values.

Constraints on individual parameters using counts and masses are only compatible with Rozo et al. (2010) ( $\ln N_1 = 2.44 \pm 0.11$ ,  $\ln N_2 = 4.16 \pm 0.15$ ), while errors are reduced when adding the power spectrum, even if less significantly ( $\ln N_1 = 2.49 \pm 0.09$ ,  $\ln N_2 = 4.13 \pm 0.13$ ). Combining these results with the CMB data, the constraints are almost identical, as the CMB is not sensitive to the clusters scaling relation. Our constraints on the scaling relation scatter  $\sigma_{\ln M}$  are in agreement with Rozo et al. (2010), and they are not improved by the addition of power spectrum and CMB data.

We then calculated the likelihood contours on the derived parameters  $\alpha_N - \ln M_{200|20}$ : these are shown in the right panel of Fig. 4.6. By definition of slope and intercept, it is evident that any increase of  $\alpha_N$  implies a decrease of  $\ln M_{200|20}$  (and vice versa), motivating the degeneracy line of these two parameters. The marginalised mean values and  $1\sigma$  errors on individual parameters, using counts and masses only, are  $\alpha_N = 1.35 \pm 0.11$ ,  $\ln M_{200|20} = 28.85 \pm 0.33$ . When adding the power spectrum, the errors reduce to  $\alpha_N = 1.41 \pm 0.06$ ,  $\ln M_{200|20} = 28.64 \pm 0.17$ . Combining then with the CMB data, the constraints are further improved to  $\alpha_N = 1.32 \pm 0.03$ ,  $\ln M_{200|20} = 28.93 \pm 0.09$ .

### 4.5.3 $\log(10^{10}A_s) - \sigma_8$ contours

In Fig. 4.9, left panel, we show the degeneracy between the amplitude of primordial fluctuations  $\log(10^{10}A_s)$  and power spectrum normalization  $\sigma_8$ , being respectively a primary and a derived parameter of our analysis. The constraints on individual parameters with counts and masses only are  $\sigma_8 = 0.8 \pm 0.06$ ,  $\log(10^{10}A_s) = 3.2 \pm 0.5$ , and adding the maxBCG power spectrum they become  $\sigma_8 = 0.84 \pm 0.04$ ,  $\log(10^{10}A_s) = 3.5 \pm 0.25$ . Finally, the inclusion of the CMB data shrink the constraints to  $\sigma_8 = 0.079 \pm 0.016$  and  $\log(10^{10}A_s) = 3.06 \pm 0.03$ .

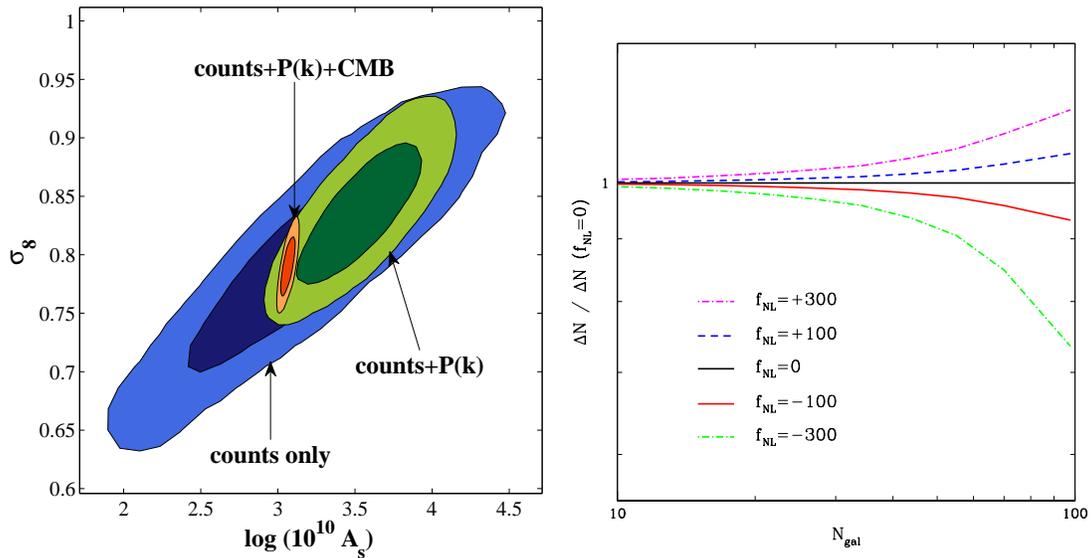
### 4.5.4 $f_{\text{NL}} - \Omega_m$ and $f_{\text{NL}} - \sigma_8$ contours

In Fig. 4.10 we show the constraints on the amount of primordial non-Gaussianity and its degeneracies with  $\Omega_m$  and  $\sigma_8$ . First, it is evident that when only counts and masses are used, the constraints on  $f_{\text{NL}}$  are weak as expected. In Fig. 4.9, right panel, we show how the relative counts with respect to the  $f_{\text{NL}} = 0$  vary for different  $f_{\text{NL}}$  values. For this theoretical expectations we use the run **counts with  $f_{\text{NL}}$** . It is clear that the difference between the curves is more relevant at large values of  $N_{\text{gal}}$ : higher  $f_{\text{NL}}$  corresponds to higher skewness, which implies more rare (massive) objects.

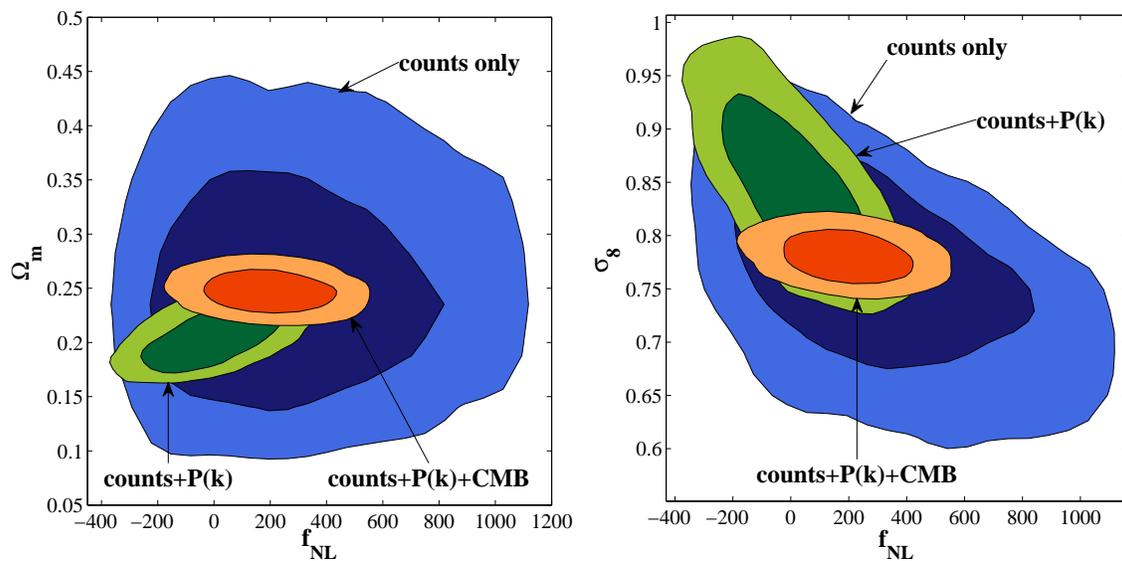
The situation improves when adding the cluster power spectrum: in this case, the constraints are tighter, and we observe a positive correlation between  $f_{\text{NL}}$  and  $\Omega_m$  and an anti-correlation with  $\sigma_8$ . To understand this behaviour, we can refer again to the power spectrum dependence on  $\Omega_m$  and  $\sigma_8$  parameters. Any increase of  $\Omega_m$  tilts the power spectrum, decreasing and shifting the peak to higher values of  $k$ , while any increase of  $\sigma_8$  shifts up the whole shape. The converse of this is also true. We have also understood that an increase in  $f_{\text{NL}}$  causes a boost in the power spectrum on large scales (small  $k$ ), so that  $\sigma_8$  needs to decrease to compensate a higher  $f_{\text{NL}}$ : this is exactly what is shown in the right panel of Fig. 4.10. In addition to this  $\Omega_m$  should increase to compensate a higher  $f_{\text{NL}}$ : this can be seen in the left panel.

We also see that the addition of the CMB power spectrum data improves the constraints on  $\Omega_m$  and  $\sigma_8$  and only indirectly reduces the bounds on  $f_{\text{NL}}$ , since PNG simply affects the higher-order statistics of the CMB.

Our constraints on PNG are  $f_{\text{NL}} = 12 \pm 157$  ( $1\sigma$ ) (without CMB) and  $f_{\text{NL}} = 194 \pm 128$  (with CMB), which are statistically compatible with zero and with each other. The shift in the mean between the two results is clear by looking at Fig. 4.10: the addition of the CMB favours lower values of  $\sigma_8$  (and higher values of  $\Omega_m$ ), thus



**Figure 4.9:** Left panel: marginalised posterior probability distributions on  $\log(10^{10} A_s) - \sigma_8$  parameters. The runs used are: **counts only** (blue), **counts+P(k)** (green) and **counts+P(k)+CMB** (orange), at 68% and 95% confidence levels. Right panel: counts predictions for  $f_{\text{NL}} = \{300, 100, 0, -100, -300\}$ , relative to the  $f_{\text{NL}} = 0$  case. The run used in this plot is **counts with  $f_{\text{NL}}$** .



**Figure 4.10:** Marginalised posterior probability distributions on the parameters  $f_{\text{NL}} - \Omega_m$  (left panel) and  $f_{\text{NL}} - \sigma_8$  (right panel). The runs used for these plots are **counts with  $f_{\text{NL}}$**  (blue), **counts+P(k) with  $f_{\text{NL}}$**  (green) and **CMB+clusters with  $f_{\text{NL}}$**  (orange), at 68% and 95% confidence levels.

shifting the favoured  $f_{\text{NL}}$  values in the process. While not competitive with results from the CMB bispectrum or from combined analyses of multiple galaxy surveys, it is interesting to find such constraints independently and for the first time with the clustering of optically selected galaxy clusters.

It is also worth mentioning the results on the  $B$  parameter, which was introduced to take into account the uncertainty in the bias expression derived from the mass function. As this parameter allows an arbitrary constant rescaling of the bias, it also has the desirable property of cancelling the effect of the scale-independent bias correction  $\delta b(f_{\text{NL}})$ , as described in Section 2.5. To check that the Gaussian prior we are imposing  $B = 1.0 \pm 0.15$  is large enough for both purposes, we made an additional run replacing it with a flat prior,  $B \in [0.0001, 5]$ . In this way, we obtain nearly unchanged results on  $f_{\text{NL}}$ . We also obtain no relevant degeneracies of  $B$  with any of the other parameters in the analysis.

Finally, since we restrict our analysis to near-linear scales by imposing the data cut at  $k_{\text{max}} = 0.15 h \text{ Mpc}^{-1}$ , we are not expecting strong constraints on  $q_{\text{NL}}$ . The constraints we found are indeed broad and in agreement within the errors with the results by Hütsi (2010), who found  $q_{\text{NL}} = 14.2 \pm 2.8$  when marginalising over three parameters only. We instead obtain  $q_{\text{NL}} = 26 \pm 10$  (without CMB) and  $q_{\text{NL}} = 14 \pm 6$  (with CMB).

**Table 4.1:** Parameters used in the analysis and their assumed priors (with or without the CMB data). In the prior columns a single number  $n$  stands for a fixed value,  $[a, b]$  stands for a flat prior,  $\mu \pm \sigma$  means a Gaussian prior of mean  $\mu$  and standard deviation  $\sigma$ .

Type	Symbol	Definition	Prior (no CMB)	Prior (+CMB)
Cosmology	$h$	Dimensionless Hubble parameter	0.7	[0.4, 0.9]
	$n_s$	Scalar spectral index	0.96	[0.5, 1.5]
	$\Omega_b$	Baryon energy density	0.04397	[0.01, 0.2]
	$\Omega_c$	Cold dark matter energy density	[0.1, 0.9]	[0.1, 0.9]
	$\log(10^{10} A_s)$	Amplitude of primordial perturbations	[0.1, 6.0]	[0.1, 6.0]
	$\tau$	Optical depth	0.09	[0.01, 0.125]
	$f_{\text{NL}}$	Primordial non-Gaussianity	[-900, 900]	[-900, 900]
Scaling relation	$\ln N_1 \equiv \ln N_{\text{gal}} M_1$	Richness at $M_1 = 1.3 \times 10^{14} M_{\odot}$	[1.0, 4.0]	[1.0, 4.0]
	$\ln N_2 \equiv \ln N_{\text{gal}} M_2$	Richness at $M_2 = 1.3 \times 10^{15} M_{\odot}$	[1.0, 4.0]	[1.0, 4.0]
	$\sigma_{\ln M N_{\text{gal}}^{\text{obs}}}$	Scatter	$0.45 \pm 0.1$	$0.45 \pm 0.1$
Nuisance	$\beta$	WL mass measurements bias	$1.0 \pm 0.06$	$1.0 \pm 0.06$
	$B$	Scatter on bias	$1.0 \pm 0.15$	$1.0 \pm 0.15$
	$q_{\text{NL}}$	Non-linear P(k) correction	[0.0, 50.0]	[0.0, 50.0]
	$\sigma_z$	Photo-z errors	[0, 120]	[0, 120]
	$A_{\text{sz}}$	Amplitude of CMB SZ template	1	[0, 2]
Derived	$\Omega_m$	Total matter energy density	—	—
	$\sigma_8$	Power spectrum normalization	—	—

**Table 4.2:** Marginalised mean values and  $1\sigma$  errors on the cosmological parameters, for the runs **counts only**, **counts with  $f_{\text{NL}}$** , **counts+ $P(k)$** , **counts+ $P(k)$  with  $f_{\text{NL}}$** , **CMB+clusters** and **CMB+clusters with  $f_{\text{NL}}$** . Note that  $\Omega_m$  and  $\sigma_8$  are derived parameters in our analysis.

Params	counts only		counts+ $P(k)$		clusters+CMB	
	no $f_{\text{NL}}$	+ $f_{\text{NL}}$	no $f_{\text{NL}}$	+ $f_{\text{NL}}$	no $f_{\text{NL}}$	+ $f_{\text{NL}}$
$\Omega_m$	$0.25 \pm 0.06$	$0.25 \pm 0.06$	$0.215 \pm 0.022$	$0.209 \pm 0.022$	$0.255 \pm 0.014$	$0.248 \pm 0.013$
$\sigma_8$	$0.80 \pm 0.06$	$0.77 \pm 0.07$	$0.84 \pm 0.04$	$0.85 \pm 0.05$	$0.790 \pm 0.016$	$0.780 \pm 0.016$
$\ln N_1$	$2.44 \pm 0.11$	$2.44 \pm 0.11$	$2.49 \pm 0.09$	$2.49 \pm 0.08$	$2.44 \pm 0.08$	$2.43 \pm 0.08$
$\ln N_2$	$4.16 \pm 0.15$	$4.15 \pm 0.15$	$4.13 \pm 0.13$	$4.11 \pm 0.12$	$4.19 \pm 0.11$	$4.15 \pm 0.11$
$\sigma_{\ln M}$	$0.38 \pm 0.06$	$0.38 \pm 0.06$	$0.36 \pm 0.06$	$0.37 \pm 0.06$	$0.378 \pm 0.059$	$0.38 \pm 0.06$
$\beta$	$1.00 \pm 0.06$	$1.01 \pm 0.06$	$1.01 \pm 0.06$	$1.01 \pm 0.06$	$1.01 \pm 0.06$	$1.00 \pm 0.06$
$q_{\text{NL}}$	-	-	$26 \pm 10$	$27 \pm 10$	$14 \pm 6$	$16 \pm 7$
$\sigma_z$	-	-	$46 \pm 12$	$42 \pm 8$	$43 \pm 10$	$31 \pm 5$
$B$	-	-	$1.07 \pm 0.13$	$1.01 \pm 0.15$	$1.19 \pm 0.11$	$1.00 \pm 0.14$
$f_{\text{NL}}$	-	$282 \pm 317$	-	$12 \pm 157$	-	$194 \pm 128$

# Chapter 5

## Clusters-galaxies cross correlation

The correlation function is one of the best statistical tools to measure quantitatively the clustering of objects in a sample. Correlation statistic methods are an important tool for relating galaxies and clusters to the underlying mass distribution, being clusters more strongly correlated in space than galaxies. Therefore, the two-point angular correlation function of galaxies and clusters efficiently quantifies the large-scale structure of the Universe. In this Chapter, we describe the measuring of the clustering properties of clusters and galaxies by using the two-point angular correlation function  $w(\theta)$ . As we have been focusing on optical data so far, the measurement of the 2D angular projected correlation function is to prefer to the 3D correlation function, because of the uncertainty on the photometric redshift estimates and the smoothing due to the relatively large radial distance errors. This part of the project is still in a preliminary, incomplete status. Here we introduce the measurement by pixelization, the theoretical expectation and the error estimates, in the case of cluster auto-correlation function only. The goal is to fully complete the estimations for the galaxies auto-correlation and for the clusters-galaxies cross-correlation, in order to investigate their constraining power on cosmological models.

### 5.1 Measurement by pixelization

Estimating the angular correlation function with pixelated analysis is similar to the approach used to calculate the CMB temperature anisotropies. The angular resolution of the particular instruments gives a natural limit for the maximum reasonable resolution used in the analysis. Here we describe the pixelization technique we employ, based on the HEALPIX software. We also present the mask calculation and the catalogues of clusters and galaxies we choose.

### 5.1.1 Pixelization method

The HEALPIX (Hierarchical Equal Area isoLatitude Pixelization of a sphere) pixelization scheme (Górski et al. 2005) is a software package available in C, C++, Fortran90, IDL, Java and Python, which is often used for CMB calculations and we use for the cross-correlation calculation. HEALPIX divides the spherical surface into pixels of same size. The  $N_{\text{side}}$  parameter is a measure of the map angular resolution and could be any power of 2. The total number of pixels is simply  $N_{\text{pix}} = 12 \times N_{\text{side}}^2$ . For example, the lowest resolution corresponds to  $N_{\text{side}} = 2^0 = 1$ , where the sphere is divided into  $N_{\text{pix}} = 12$  equivalent pixels. A single step to higher resolution (to the following power of 2) would divide each pixel into 4 subpixels, namely if  $N_{\text{side}} = 2$ , then  $N_{\text{pix}} = 12 \times 2^2 = 48$ . The isolatitude feature means that pixels are located at rings of constant latitude: this is fundamental to optimize the computational time of calculations with spherical harmonics. The ordering of pixels in a HEALPIX map can be *ring* or *nested*: we use the latter, because neighbouring pixels have similar indices, allowing for faster pair finding algorithms. An important issue when using a pixelization method is the damping of power on small scales ( $\theta < 1$  deg): this is because the true signal is smoothed over the area of each pixel, erasing small scale features, while on large scales the effect is not significant.

In our pixelated analysis, after comparing the following three different resolutions:

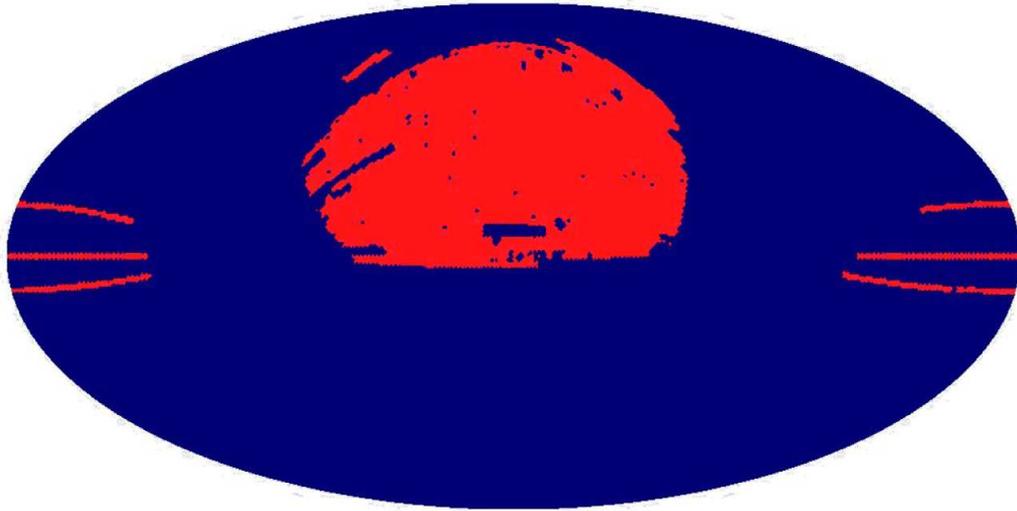
- $N_{\text{side}} = 64$ ,  $N_{\text{pix}} = 49,152$  (pixel size  $\sim 50$  arcmin);
- $N_{\text{side}} = 128$ ,  $N_{\text{pix}} = 196,608$ ;
- $N_{\text{side}} = 256$ ,  $N_{\text{pix}} = 786,432$  (pixel size  $\sim 13$  arcmin);

we rely on the highest resolution ( $N_{\text{side}} = 256$ ).

The angular distance  $\theta$  (in decimal degrees) between two points on a sphere  $P_1(\lambda_1, \phi_1)$  and  $P_2(\lambda_2, \phi_2)$ , where  $\lambda_{1,2}$  and  $\phi_{1,2}$  are respectively the latitudes and the longitudes (in radians), in the limit of relatively small pixels can be expressed as

$$\theta = \frac{180^\circ}{\pi} \tan^{-1} \left\{ \frac{\sqrt{\cos^2 \lambda_2 \sin^2(\phi_2 - \phi_1) + [\cos \lambda_1 \sin \lambda_2 - \sin \lambda_1 \cos \lambda_2 \cos(\phi_2 - \phi_1)]^2}}{\sin \lambda_1 \sin \lambda_2 + \cos \lambda_1 \cos \lambda_2 \cos(\phi_2 - \phi_1)} \right\}. \quad (5.1)$$

We use a linear binning in  $\theta$ , with  $\theta \in [0.2, 10.5]$  deg, divided into 20 bins.



**Figure 5.1:** Mask used for the calculation of angular correlation function with pixelated analysis, in equatorial coordinates.

### 5.1.2 The mask and the catalogues

We adopt a mock catalogue provided by Will Percival (private communication), consisting of about 5.2 million objects corresponding to the SDSS DR7 spectroscopic footprint, to create our *mask* (see Fig. 5.1). We use  $N_{\text{side}} = 256$  to determine which pixels are covered by the catalogue, downscale to  $N_{\text{side}} = 64$  and remove all pixels in the lower resolution that are not fully covered. The final mask is binary ( $m_i = 1, 0$ ). We then applied the mask to the following two catalogues:

- GMBCG (Gaussian Mixture Brightest Cluster Galaxy) largest publicly available cluster catalogue (Hao et al. 2010) from SDSS DR7 optical data, with  $z_{\text{photo}} < 0.55$ ,  $8 < N_{\text{gal}} < 140$  over  $8,200 \text{ deg}^2$  of comoving volume, with 55,431 objects;
- A volume limited sample of galaxies from the SDSS-II DR7 (Ben Hoyle, private communication), with  $z_{\text{photo}} < 0.55$ ,  $\Delta z_{\text{photo}} < 0.1$  and R-band absolute magnitude smaller than  $-21.5$ , with 9,489,664 objects .

## 5.2 Angular correlation function estimator

From Giannantonio et al. (2008) we employ the following estimator for the angular correlation function between clusters (c) and galaxies (g):

$$w(\theta)_{\text{pix}}^{\text{cg}} = \frac{1}{N_\theta} \sum_{i,j} m_i \left( \frac{c_i - \bar{c}}{\bar{c}} \right) m_j \left( \frac{g_j - \bar{g}}{\bar{g}} \right), \quad (5.2)$$

where  $N_\theta = \sum_{i,j} m_i m_j$  is the weighted number of pixel pairs separated by  $\theta$ ,  $c_i$  is the number of clusters inside pixel  $i$ ,  $g_j$  is the galaxies number inside pixel  $j$ ,  $m_{i,j}$  is the unique mask value for pixels  $i, j$ ,  $\bar{c}$  and  $\bar{g}$  are respectively the average numbers of clusters and galaxies per pixel (i.e. the total number of objects divided by the number of the non-empty pixels). In the case of clusters auto-correlation, we have then

$$w(\theta)_{\text{pix}}^{\text{cc}} = \frac{1}{N_\theta} \sum_{i,j} m_i \left( \frac{c_i - \bar{c}}{\bar{c}} \right) m_j \left( \frac{c_j - \bar{c}}{\bar{c}} \right), \quad (5.3)$$

while for the galaxies auto-correlation we have

$$w(\theta)_{\text{pix}}^{\text{gg}} = \frac{1}{N_\theta} \sum_{i,j} m_i \left( \frac{g_i - \bar{g}}{\bar{g}} \right) m_j \left( \frac{g_j - \bar{g}}{\bar{g}} \right). \quad (5.4)$$

## 5.3 Theoretical prediction

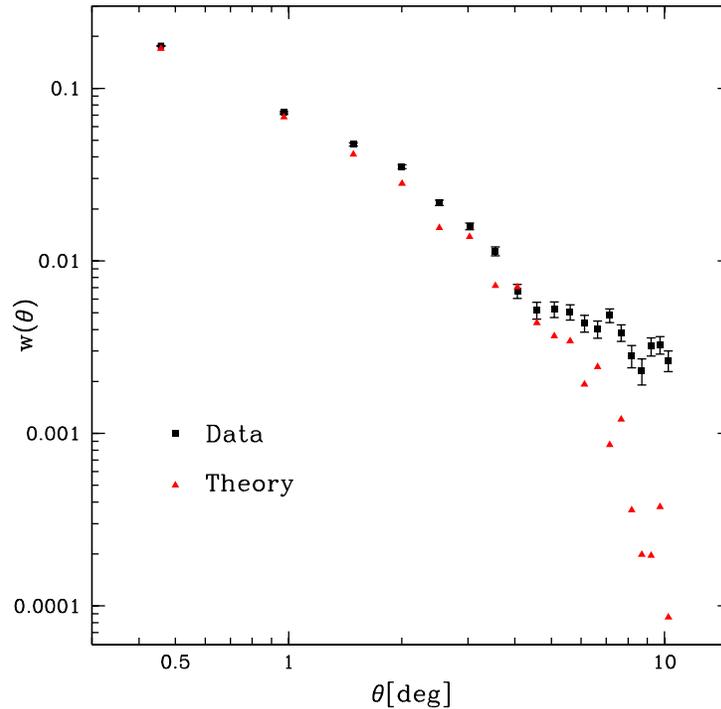
The numerical tool we use for calculating the theoretical correlation function is CAMB Sources<sup>1</sup>(Lewis & Challinor 2007) a Fortran 90 code similar to CAMB, CMBFAST (Seljak & Zaldarriaga 1996), CMBEASY (Doran 2005), CLASS (Lesgourgues 2011; Blas et al. 2011). It computes the angular power spectrum  $C_l$  of the CMB anisotropies as well as matter perturbations at different redshifts. We will calculate the  $C_l$ s of the matter density, given a certain redshift distribution  $d(z)$ , for both clusters and galaxies. The relation between the  $C_l$ s and  $w(\theta)$  is given by

$$w(\theta)_{\text{theory}} = \sum_{l \geq 0} \left( \frac{2l+1}{4\pi} \right) P_l(\cos \theta) C_l, \quad (5.5)$$

where  $P_l$  are the Legendre polynomials of degree  $l$ . The observed signal is then expected to be  $w(\theta) = b_{\text{eff}}^2 w(\theta)_{\text{theory}}$ , for a bias factor  $b_{\text{eff}}^{\text{cl}} \sim 2$  for clusters and  $b_{\text{eff}}^{\text{g}} \sim 1$  for galaxies. In Fig. 5.2 we show our preliminary results for the cluster auto-correlation function, together with theoretical predictions and errors estimates.

---

<sup>1</sup><http://camb.info/sources/>



**Figure 5.2:** Cluster auto-correlation function, with estimated errors and theoretical prediction. The angle  $\theta$  is given in degrees.

## 5.4 Error estimates

Having the theoretical  $C_l$ , we can use the following analytical prescription (Ross et al. 2011) for the covariance matrix of the correlation function:

$$\mathcal{C}_{\theta\theta'} = \frac{2}{f_{\text{sky}}} \sum_{l \geq 0} \frac{2l+1}{(4\pi)^2} P_l(\cos \theta) P_l(\cos \theta') \left( b_{\text{eff}}^2 C_l + \frac{1}{\bar{n}} \right)^2, \quad (5.6)$$

where  $f_{\text{sky}}$  is the number of observed steradians,  $\bar{n}$  is the number of object per steradian. Note that  $1/\bar{n}$  is the shot-noise contribution, while  $b_{\text{eff}}^2 C_l$  is the cosmic variance term.



# Chapter 6

## Conclusions

This dissertation was aimed to examine the contribution of clusters of galaxies in confirming the concordance  $\Lambda$ CDM cosmological model. After introducing the current cosmological framework and the main properties of galaxy clusters, we showed how these massive gravitationally bound systems in the Universe provide a unique test of cosmology and structure formation. Their number density and distribution are highly sensitive to cosmology and provide tight and robust constraints on cosmological parameters, complementing those which result from other observational probes. However, these constraints are critically dependent on our ability to estimate cluster masses. Our work highlighted in particular the power of optically selected cluster samples to produce precision constraints on cosmological parameters. This was achieved by adding the clustering information of galaxy clusters to complement the cluster number counts and observable-mass relation analyses.

The first achievement was to reproduce the cosmological constraints obtained in the analysis done by Rozo et al. (2010). In this analysis, they used cluster abundances in richness bins together with weak-lensing mass measurements of the SDSS maxBCG cluster catalogue to constrain cosmology and the richness-mass relation of the clusters. They assumed a flat Universe and they used the Tinker halo mass function for the computation. The richness-mass relation they adopted was a power-law in mass and they constrained the scatter around it by demanding consistency between X-ray and WL measurements. Their final cosmological results were further improved by a joint analysis with the WMAP DR5 data. We followed their theoretical modelling of cluster counts and cluster total masses, together with their corresponding covariance matrices. By adopting a Bayesian approach for deriving cosmological constraints from the sample, we obtained consistent constraints on the

matter density  $\Omega_m$ , the fluctuation amplitude  $\sigma_8$ , the scaling relation parameters  $\ln N_1$ ,  $\ln N_2$  and the scatter  $\sigma_{\ln M}$ .

An interesting development of this joint cosmological analysis was to fully include the redshift space power spectrum of the maxBCG clusters, calculated by Hütsi (2010). We followed this prescription to model the cluster power spectrum, including the effect of weak non-linearities and redshift space distortions and allowing for an arbitrary photometric redshift smoothing. We found that the inclusion of the power spectrum typically brings a factor between 1.5 and 3, depending on the parameter, improvement of the errors on  $\sigma_8$  and  $\Omega_m$ . Constraints on other parameters are also improved, even if less significantly. In addition to the cluster data, we also used the CMB power spectra from WMAP7, which further contributed to tighten the confidence regions by an additional factor of 2, mainly on  $\sigma_8$  and  $\Omega_m$  parameters.

Another peculiar application of this analysis was concerning the study of non-Gaussian initial conditions to constrain models of the early universe. MaxBCG galaxy clusters provided constraints on the amount of the local type primordial non-Gaussianity of the initial density perturbations which are statistically compatible with zero and consistent with the latest constraints. While these results are not competitive with those from combined analyses of multiple galaxy surveys or from the CMB bispectrum, we found such constraints independently and for the first time with the clustering of optically selected galaxy clusters.

All our MCMC analysis was performed using the Cosmological Monte-Carlo (COSMOMC) publicly available parameter estimation tool by Lewis & Bridle (2002). All the cosmological analysis implementation and results have been published in Mana et al. (2013). We are currently updating our results to include the Planck likelihood.

In the last Chapter of this thesis, we presented some preliminary calculations on the clustering of clusters and galaxies, by means of the two-point angular correlation function, estimated by a pixelization technique and theoretical modelling. This statistical quantity efficiently quantifies the large-scale structure of the Universe, thus providing additional constraining power on cosmological models. Our work can be seen as a proof of concept towards a full joint analysis of the LSS, consistently including both galaxies and clusters as dark matter tracers, to achieve the full potential of the upcoming galaxy surveys such as the Dark Energy Survey and the Euclid mission.

We conclude by suggesting two possible extensions to this work, which could result in promising results: the inclusion of massive neutrinos and the application to modified gravity models.

## Modified gravity

The accelerated expansion of the Universe can be explained by introducing modifications to gravity, instead of a dark energy component. A possible modification is the addition of a free function  $f(R)$  of the Ricci scalar  $R$  to the Einstein-Hilbert action. This leads to modified Einstein equations, containing terms with  $f(R)$  and differentials of this function with respect to  $R$  (i.e.  $f_R, f_{RR}$ ). Valid  $f(R)$  model should closely match the  $\Lambda$ CDM expansion history and satisfy local gravity constraints. Strong constraints on deviation from GR can be inferred from the large-scale structure: in fact, the increased growth of structure observed in  $f(R)$  gravity affects the large scales of the CMB temperature power spectrum. However, the most stringent constraints on  $f(R)$  models are due to the abundance of low-redshift ( $z < 0.15$ ) X-ray clusters (Schmidt et al. 2009). An interesting analysis was performed by Lombriser et al. (2012). They performed a MCMC study of  $f(R)$  gravity models which reproduce the  $\Lambda$ CDM expansion history using data from CMB (WMAP DR5), SNIa (Supernova Cosmology Project, SCP), BAO (SDSS DR7),  $H_0$  (SH0ES) and cluster abundances from the likelihood code of Seljak et al. (2013). They parametrize their solutions in terms of the Compton wavelength parameter, defined as

$$B = \frac{f_{RR}}{1 + f_R} R' \frac{H}{H'}, \quad (6.1)$$

where prime denote derivative with respect to  $\ln a$ : more precisely, they considered the value of  $B$  at  $\ln a = 0$ , i.e.  $B_0 \equiv B(\ln a = 0)$ . Standard gravity corresponds to  $B_0 = 0$ . The constraint on the Compton wavelength parameter  $B_0$  is essentially driven by the cluster abundances alone: they obtained  $100B_0 < 0.333$  (95% c.l.) from the cluster data alone and  $100B_0 < 0.11$  (95% c.l.) from the combined datasets.

These interesting results involving low redshift galaxy clusters suggest us a possible development direction of our research. It would be intriguing to test these  $f(R)$  gravity models with the maxBCG optical cluster sample.

## Massive neutrinos

Neutrino oscillation experiments have provided evidence that neutrinos have non-zero masses. In the standard picture, there are three neutrino species with a lower bound on the summed mass  $\sum m_\nu$  at 0.06 eV coming from solar and atmospheric oscillations observations. On the other hand, combined cosmological data, such as CMB, BAO, galaxy clustering and cluster mass function, provide an upper limit on the summed mass in the range  $\sum m_\nu < 0.3 - 0.8$  eV (95% c.l.). Neutrinos, in fact, produce visible effects on the background evolution and structure formation. More specifically, neutrinos decouple from the primordial plasma when they are still ultra-relativistic, becoming non-relativistic only after recombination, due to their small mass. During this period, their energy density contributes as radiation and changes both the expansion rate and the time of matter-radiation equality: a larger value of  $\sum m_\nu$  implies a larger value of the radiation density and therefore a postponed time of equality. These modifications appear in the matter power spectrum as a shift of the peak to larger scales. Moreover, since density fluctuations grow more efficiently on small scales after equality (i.e. in the matter-dominated era), the matter power spectrum is suppressed on small scales with respect to large scales. As a consequence, a suppression effect is also visible in the halo mass function. Analogous features are caused by non-relativistic neutrinos, which suppress density fluctuations on scales smaller than their free-streaming length.

Costanzi Alunno Cerbolini et al. (2013) performed a forecast analysis for a large photometric galaxy cluster survey (like Euclid survey) on neutrino properties constraints: in particular, the total neutrino mass and effective number of neutrino species. They combined cluster number counts and cluster power spectrum and used a MCMC method. They found that combining cluster data with CMB from Planck improves by more than an order of magnitude the constraint on neutrino masses, if compared to each probe used independently. On the other hand, galaxy clusters together with CMB and BAO can provide precise constraints on the sum of neutrino masses but these constraints depend on the calibration of the mass-observable relation (Rozo et al. 2013).

These studies motivate us to include the effect of massive neutrinos in our analysis, through the modifications on the cluster mass function and power spectrum. We will thus investigate the constraints on neutrinos properties in the near future.

# List of Figures

1.1	Energy densities of the Universe. . . . .	7
1.2	Scale factor as a function of time. . . . .	7
1.3	The original Hubble diagram. . . . .	9
1.4	Evolution of density perturbations. . . . .	13
1.5	Growth factor dependence on cosmological parameters . . . . .	16
1.6	Hubble diagram from SNIa. . . . .	20
1.7	SDSS LRGs galaxy cross correlation. . . . .	21
1.8	CMB map from Planck. . . . .	23
1.9	Temperature power spectrum from WMAP DR7 . . . . .	24
1.10	Composite image of galaxy cluster Abell 520. . . . .	27
1.11	Cluster mass function from Bahcall & Cen (1992). . . . .	30
1.12	Rotation curve of galaxy NGC 6503. . . . .	32
1.13	Constraints on $\Omega_m$ , $\sigma_8$ , $\Omega_\Lambda$ . . . . .	34
1.14	Constraints on $\Omega_m$ , $w$ and GR . . . . .	35
2.1	The gravitational lensing system representation. . . . .	41
2.2	Tinker mass function . . . . .	47
2.3	Mass function cosmological dependence on $\Omega_m$ and $\sigma_8$ . . . . .	49
2.4	High-z clusters sensitivity to cosmological models . . . . .	50
2.5	Mass functions comparison . . . . .	61
2.6	Mass dependence of the linear halo bias. . . . .	63
2.7	The effect of PNG on the cluster power spectrum. . . . .	64
3.1	MaxBCG counts and masses data and theoretical curves. . . . .	79
3.2	The angular distribution of the selected maxBCG clusters. . . . .	82
3.3	The redshift-space power spectrum of maxBCG sample. . . . .	83
3.4	Linear and full power spectrum. . . . .	84
3.5	The off-diagonal elements of $w(\theta)$ - $\Delta N$ covariance matrix. . . . .	87

---

4.1	Example of MCMC chains. . . . .	95
4.2	Example of 1D, 2D marginalised posteriors. . . . .	100
4.3	Marginalised posterior probability distribution on $\Omega_m - \sigma_8$ . . . . .	104
4.4	Number counts with varying $\Omega_m$ and $\sigma_8$ . . . . .	105
4.5	Power spectrum with varying $\Omega_m$ and $\sigma_8$ . . . . .	105
4.6	Constraints on scaling relation parameters. . . . .	107
4.7	Constraints on the scaling relation scatter. . . . .	107
4.8	Masses and counts varying scaling relation parameters. . . . .	108
4.9	Marginalised posteriors on $\log(10^{10} A_s) - \sigma_8$ . . . . .	111
4.10	Marginalised posteriors of $f_{\text{NL}} - \Omega_m, f_{\text{NL}} - \sigma_8$ . . . . .	111
5.1	Mask for the angular correlation function with pixelated analysis. . .	117
5.2	Cluster auto-correlation function with theoretical prediction. . . . .	119

# List of Tables

1.1	Evolution of energy densities components of the Universe. . . . .	6
1.2	Cosmological parameters of $\Lambda$ CDM model. . . . .	33
3.1	MaxBCG counts data. . . . .	73
3.2	MaxBCG mean mass data. . . . .	78
3.3	MaxBCG power spectrum data. . . . .	88
4.1	Parameters and priors used in the analysis. . . . .	113
4.2	Marginalised mean values and $1\sigma$ errors in the analysis. . . . .	114



# Acronyms

**ACT** Atacama Cosmology Telescope  
**AP** Alcock-Paczynski  
**BAO** Baryons Acoustic Oscillations  
**BCG** Brightest Cluster Galaxy  
**CAMB** Code for Anisotropies in the Microwave Background  
**CDM** Cold Dark Matter  
**CLT** Central Limit Theorem  
**CMB** Cosmic Microwave Background  
**COSMOMC** Cosmological Monte Carlo  
**DE** Dark Energy  
**DGP** Dvali, Gabadadze, Porrati model  
**DM** Dark Matter  
**DR** Data Release  
**FE** Friedmann Equations  
**FFTs** Fast Fourier Transforms  
**FKP** Feldman-Kaiser-Peacock  
**FLRW** Friedmann-Lemaître-Robertson-Walker  
**FT** Fourier Transformation  
**GMBCG** Gaussian Mixture Brightest Cluster Galaxy  
**GR** General Relativity  
**HEALPIX** Hierarchical Equal Area isoLatitude Pixelization of a sphere  
**HST** Hubble Space Telescope  
**ICM** Intracluster Medium  
**IID** Independent Identically distributed  
**JK** Jackknife  
 **$\Lambda$ CDM** Lambda Cold Dark Matter  
**LS** Landy-Szalay  
**LSS** Large Scale Structure  
**LV** LoVerde

**MCMC** Monte Carlo Markov Chain  
**M-H** Metropolis-Hastings  
**NFW** Navarro-Frenk-White density profile  
**OMC** Ordinary Monte Carlo  
**PNG** Primordial Non-Gaussianity  
**PS** Press-Schechter  
**RDS** Redshift Space Distortions  
**SDSS** Sloan Digital Sky Survey  
**SIS** Singular Isothermal Sphere  
**SNIa** Type Ia Supernovae  
**SO** Spherical Overdensity  
**SPT** South Pole Telescope  
**ST** Sheth-Tormen  
**SZ** Sunyaev-Zel'dovich  
**TSC** Triangular Shaped Cloud  
**WL** Weak gravitational Lensing  
**WMAP** Wilkinson Microwave Anisotropy Probe

# Bibliography

- Abazajian K. N., Adelman-McCarthy J. K., Agüeros M. A., et al., 2009, *ApJS*, 182, 543
- Abell G. O., 1958, *ApJS*, 3, 211
- Abell G. O., Corwin Jr. H. G., Olowin R. P., 1989, *ApJS*, 70, 1
- Acquaviva V., Bartolo N., Matarrese S., Riotto A., 2003, *Nuclear Physics B*, 667, 119
- Afshordi N., Tolley A. J., 2008, *Phys. Rev. D*, 78, 12, 123507
- Ahn C. P., Alexandroff R., Allende Prieto C., et al., 2012, *ApJS*, 203, 21
- Aihara H., Allende Prieto C., An D., et al., 2011, *ApJS*, 193, 29
- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Allen S. W., Evrard A. E., Mantz A. B., 2011, *ARA&A*, 49, 409
- Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, *MNRAS*, 383, 879
- Allen S. W., Schmidt R. W., Ebeling H., Fabian A. C., van Speybroeck L., 2004, *MNRAS*, 353, 457
- Angulo R. E., Hahn O., Abel T., 2013, *MNRAS*, 434, 1756
- Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046
- Bahcall N. A., Cen R., 1992, *ApJ*, 398, L81
- Bahcall N. A., Fan X., 1998, *ApJ*, 504, 1

- Bahcall N. A., Soneira R. M., 1983, *ApJ*, 270, 20
- Balaguera-Antolínez A., Sánchez A. G., Böhringer H., Collins C., Guzzo L., Phleps S., 2011, *MNRAS*, 413, 386
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291
- Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, *Phys. Rep.*, 402, 103
- Battye R. A., Weller J., 2003, *Phys. Rev. D*, 68, 8, 083506
- Becker M. R., McKay T. A., Koester B., et al., 2007, *ApJ*, 669, 905
- Bennett C. L., Larson D., Weiland J. L., et al., 2013, *ApJS*, 208, 20
- Benson B. A., de Haan T., Dudley J. P., et al., 2013, *ApJ*, 763, 147
- Binney J., Tremaine S., 1987, *Galactic dynamics*
- Blas D., Lesgourgues J., Tram T., 2011, *J. Cosmology Astropart. Phys.*, 7, 34
- Bogness N. W., Mather J. C., Weiss R., et al., 1992, *ApJ*, 397, 420
- Böhringer H., Schuecker P., Guzzo L., et al., 2004, *A&A*, 425, 367
- Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440
- Burenin R. A., Vikhlinin A., Hornstrup A., Ebeling H., Quintana H., Mescheryakov A., 2007, *ApJS*, 172, 561
- Burenin R. A., Vikhlinin A. A., 2012, *Astronomy Letters*, 38, 347
- Burkert A., 1995, *ApJ*, 447, L25
- Carlstrom J. E., Ade P. A. R., Aird K. A., et al., 2011, *PASP*, 123, 568
- Cash W., 1979, *ApJ*, 228, 939
- Chevallier M., Polarski D., 2001, *International Journal of Modern Physics D*, 10, 213
- Clowe D., Gonzalez A., Markevitch M., 2004, *ApJ*, 604, 596
- Cole S., Kaiser N., 1989, *MNRAS*, 237, 1127
- Cole S., Percival W. J., Peacock J. A., et al., 2005, *MNRAS*, 362, 505

- Coles P., Lucchin F., 1995, *Cosmology. The origin and evolution of cosmic structure*
- Collins C. A., Guzzo L., Bhringer H., et al., 2000, *Monthly Notices of the Royal Astronomical Society*, 319, 3, 939
- Costanzi Alunno Cerbolini M., Sartoris B., Xia J.-Q., Biviano A., Borgani S., Viel M., 2013, *J. Cosmology Astropart. Phys.*, 6, 20
- Cowles M. K., Carlin B. P., 1996, *Journal of the American Statistical Association*, 91, 434, 883
- Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, *MNRAS*, 423, 2279
- Dahle H., Pedersen K., Lilje P. B., Maddox S. J., Kaiser N., 2003, *ApJ*, 591, 662
- Dalal N., Doré O., Huterer D., Shirokov A., 2008, *Phys. Rev. D*, 77, 12, 123514
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Davis M., Peebles P. J. E., 1983, *ApJ*, 267, 465
- de Putter R., White M., 2005, *New A*, 10, 676
- Desjacques V., Jeong D., Schmidt F., 2011, *Phys. Rev. D*, 84, 6, 061301
- Desjacques V., Seljak U., 2010, *Advances in Astronomy*, 2010
- Desjacques V., Seljak U., Iliev I. T., 2009, *MNRAS*, 396, 85
- Doran M., 2005, *J. Cosmology Astropart. Phys.*, 10, 11
- Dunkley J., Komatsu E., Nolte M. R., et al., 2009, *ApJS*, 180, 306
- Dvali G., Gabadadze G., Porrati M., 2000, *Physics Letters B*, 485, 208
- Ebeling H., Edge A. C., Mantz A., et al., 2010, *MNRAS*, 407, 83
- Edge A. C., Stewart G. C., Fabian A. C., Arnaud K. A., 1990, *MNRAS*, 245, 559
- Efron B., 1982, *CBMS-NSF conference series in applied mathematics*, Soc. Ind. and App. Math., Philadelphia
- Einstein A., 1916, *Annalen der Physik*, 354, 769
- Eisenstein D. J., et al., 2005, *ApJ*, 633, 560

- Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605
- Estrada J., Sefusatti E., Frieman J. A., 2009, *ApJ*, 692, 265
- Evrard A. E., MacFarland T. J., Couchman H. M. P., et al., 2002, *ApJ*, 573, 7
- Fassbender R., Nastasi A., Böhringer H., et al., 2011, *A&A*, 527, L10
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
- Freeman K. C., 1970, *ApJ*, 160, 811
- Gamerman D., Lopes H. F., 2006, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman and Hall/CRC
- Garnavich P. M., Kirshner R. P., Challis P., et al., 1998, *ApJ*, 493, L53
- Geyer C. J., 2011, *Introduction to Markov chain Monte Carlo.*, 3–48, Boca Raton, FL: CRC Press
- Giacconi R., Murray S., Gursky H., Kellogg E., Schreier E., Tananbaum H., 1972, *ApJ*, 178, 281
- Giannantonio T., Porciani C., 2010, *Phys. Rev. D*, 81, 6, 063530
- Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, *MNRAS*, 422, 2854
- Giannantonio T., Ross A. J., Percival W. J., et al., 2013, arXiv:1303.1349
- Giannantonio T., Scranton R., Crittenden R. G., et al., 2008, *Phys. Rev. D*, 77, 12, 123520
- Gilks W. R., 1999, *Markov Chain Monte Carlo In Practice*, Chapman and Hall/CRC
- Gioia I. M., Henry J. P., Maccacaro T., Morris S. L., Stocke J. T., Wolter A., 1990, *ApJ*, 356, L35
- Gladders M. D., Yee H. K. C., 2005, *ApJS*, 157, 1
- Gladders M. D., Yee H. K. C., Majumdar S., et al., 2007, *ApJ*, 655, 128
- Górski K. M., Hivon E., Banday A. J., et al., 2005, *ApJ*, 622, 759
- Guo Q., White S., Boylan-Kolchin M., et al., 2011, *MNRAS*, 413, 101

- Guth A. H., 1981, *Phys. Rev. D*, 23, 347
- Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545
- Hamuy M., Maza J., Phillips M. M., et al., 1993, *AJ*, 106, 2392
- Hamuy M., Phillips M. M., Maza J., Suntzeff N. B., Schommer R. A., Aviles R., 1995, *AJ*, 109, 1
- Hansen S. M., McKay T. A., Wechsler R. H., Annis J., Sheldon E. S., Kimball A., 2005, *ApJ*, 633, 122
- Hao J., McKay T. A., Koester B. P., et al., 2010, *ApJS*, 191, 254
- Harrison E. R., 1970, *Phys. Rev. D*, 1, 2726
- Hastings W., 1970, *Biometrika*, 97–109
- Henriques B. M. B., White S. D. M., Lemson G., et al., 2012, *MNRAS*, 421, 2904
- Herschel F., 1785, *Phil. Trans.*, 75, 213
- Hinshaw G., Larson D., Komatsu E., et al., 2013, *ApJS*, 208, 19
- Hobson M. P., Jaffe A. H., Liddle A. R., Mukherjee P., Parkinson D., 2010, *Bayesian methods in cosmology*, Cambridge Univ. Press
- Hockney R. W., Eastwood J. W., 1988, Bristol, Hilger, 1988
- Hong T., Han J. L., Wen Z. L., Sun L., Zhan H., 2012, *The Astrophysical Journal*, 749, 1, 81
- Hotchkiss S., 2011, *J. Cosmology Astropart. Phys.*, 7, 4
- Hoyle B., Jimenez R., Verde L., Hotchkiss S., 2012, *J. Cosmology Astropart. Phys.*, 2, 9
- Hu W., Dodelson S., 2002, *ARA&A*, 40, 171
- Hu W., Kravtsov A. V., 2003, *ApJ*, 584, 702
- Hubble E., 1929, *Contributions from the Mount Wilson Observatory*, vol. 3, pp.23-28, 3, 23
- Hütsi G., 2006a, *A&A*, 449, 891

- Hütsi G., 2006b, *A&A*, 446, 43
- Hütsi G., 2006c, *A&A*, 459, 375
- Hütsi G., 2010, *MNRAS*, 401, 2477
- Jeans J. H., 1902, *Royal Society of London Philosophical Transactions Series A*, 199, 1
- Jenkins A., Frenk C. S., White S. D. M., et al., 2001, *MNRAS*, 321, 372
- Jing Y. P., 1999, *ApJ*, 515, L45
- Jing Y. P., 2005, *ApJ*, 620, 559
- Johnston D. E., Sheldon E. S., Tasitsiomi A., Frieman J. A., Wechsler R. H., McKay T. A., 2007a, *ApJ*, 656, 27
- Johnston D. E., Sheldon E. S., Wechsler R. H., et al., 2007b, *arXiv:0709.1159*
- Kaiser N., 1984, *ApJ*, 284, L9
- Kaiser N., 1987, *MNRAS*, 227, 1
- Kamionkowski M., 1998, *arXiv:astro-ph/9809214*
- Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999a, *MNRAS*, 303, 188
- Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999b, *MNRAS*, 307, 529
- Klypin A. A., Kopylov A. I., 1983, *Soviet Astronomy Letters*, 9, 41
- Koester B. P., McKay T. A., Annis J., et al., 2007a, *ApJ*, 660, 239
- Koester B. P., McKay T. A., Annis J., et al., 2007b, *ApJ*, 660, 221
- Komatsu E., Smith K. M., Dunkley J., et al., 2011, *ApJS*, 192, 18
- Kosowsky A., 2006, *New A Rev.*, 50, 969
- Kowalski M., Rubin D., Aldering G., et al., 2008, *ApJ*, 686, 749
- Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, *ApJS*, 111, 73

- Kuhlen M., Vogelsberger M., Angulo R., 2012, *Physics of the Dark Universe*, 1, 50
- Lacey C., Cole S., 1994, *MNRAS*, 271, 676
- Lahav O., Fabian A. C., Edge A. C., Putney A., 1989, *MNRAS*, 238, 881
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Larson D., et al., 2011, *ApJS*, 192, 16
- Lemson G., Virgo Consortium t., 2006, *arXiv:astro-ph/0608019*
- Lesgourgues J., 2011, *arXiv:1104.2932*
- Lewis A., 2011, *CAMB Notes*, <http://cosmologist.info/notes/CAMB.pdf>
- Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 10, 103511
- Lewis A., Bridle S., 2006, *CosmoMC Notes*, <http://cosmologist.info/notes/CosmoMC.pdf>
- Lewis A., Challinor A., 2007, *Phys. Rev. D*, 76, 8, 083005
- Lewis A., Challinor A., Lasenby A., 2000, *Astrophys. J.*, 538, 473
- Lima M., Cunha C. E., Oyaizu H., Frieman J., Lin H., Sheldon E. S., 2008, *MNRAS*, 390, 118
- Lima M., Hu W., 2004, *Phys. Rev. D*, 70, 4, 043504
- Lima M., Hu W., 2005, *Phys. Rev. D*, 72, 4, 043006
- Linder E. V., 2003, *Physical Review Letters*, 90, 9, 091301
- Linder E. V., 2005, *Phys. Rev. D*, 72, 4, 043529
- Linder E. V., Cahn R. N., 2007, *Astroparticle Physics*, 28, 481
- Lombriser L., Slosar A., Seljak U., Hu W., 2012, *Phys. Rev. D*, 85, 12, 124038
- Longair M. S., 2008, *Galaxy Formation*
- LoVerde M., Miller A., Shandera S., Verde L., 2008, *J. Cosmology Astropart. Phys.*, 4, 14

- Lupton R., 1993, *Statistics in theory and practice*
- Lynden-Bell D., 1967, *MNRAS*, 136, 101
- MacKay D. J. C., 2003, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, <http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html>
- Mahdavi A., Hoekstra H., Babul A., Sievers J., Myers S. T., Henry J. P., 2007, *ApJ*, 664, 162
- Majumdar S., Mohr J. J., 2003, *ApJ*, 585, 603
- Maldacena J., 2003, *Journal of High Energy Physics*, 5, 13
- Mana A., Giannantonio T., Weller J., Hoyle B., Hütsi G., Sartoris B., 2013, *MNRAS*, 434, 684
- Mandelbaum R., Seljak U., Hirata C. M., 2008a, *J. Cosmology Astropart. Phys.*, 8, 6
- Mandelbaum R., Seljak U., Hirata C. M., et al., 2008b, *MNRAS*, 386, 781
- Mantz A., Allen S. W., Ebeling H., Rapetti D., 2008, *MNRAS*, 387, 1179
- Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *MNRAS*, 406, 1759
- Maor I., Brustein R., Steinhardt P. J., 2001, *Physical Review Letters*, 86, 6
- Markevitch M., Gonzalez A. H., Clowe D., et al., 2004, *ApJ*, 606, 819
- Marriage T. A., Acquaviva V., Ade P. A. R., et al., 2011, *ApJ*, 737, 61
- Matarrese S., Verde L., 2008, *ApJ*, 677, L77
- Menanteau F., Sifón C., Barrientos L. F., et al., 2013, *ApJ*, 765, 67
- Messier C., 1784, *Connaissance des Temps* (Paris)
- Metropolis, Ulam, 1949, *Journal of the American statistical Association*, 44, 247, 335
- Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, *J. Chem. Phys.*, 21, 1087

- Mo H. J., Jing Y. P., White S. D. M., 1996, MNRAS, 282, 1096
- Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
- Moore B., Quinn T., Governato F., Stadel J., Lake G., 1999, MNRAS, 310, 1147
- Narayan R., Bartelmann M., 1996, arXiv:astro-ph/9606001
- Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
- Neal R. M., 1993, Probabilistic Inference Using Markov Chain Monte Carlo Methods, <http://cosmologist.info/Neal93>
- Oguri M., 2009, Physical Review Letters, 102, 21, 211301
- Oyaizu H., Lima M., Cunha C. E., Lin H., Frieman J., Sheldon E. S., 2008, ApJ, 674, 768
- Peebles P. J. E., 1980, The large-scale structure of the universe
- Peebles P. J. E., 1993, Principles of Physical Cosmology
- Percival W. J., Reid B. A., Eisenstein D. J., et al., 2010, MNRAS, 401, 2148
- Percival W. J., Reid B. A., Eisenstein D. J., et al., 2011, MNRAS, 417, 3101
- Perlmutter S., 2003, Physics Today, 56, 4, 040000
- Perlmutter S., Aldering G., Goldhaber G., et al., 1999, ApJ, 517, 565
- Pillepich A., Porciani C., Hahn O., 2010, MNRAS, 402, 191
- Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44
- Planck Collaboration, Ade P. A. R., Aghanim N., et al., 2011, A&A, 536, A8
- Planck Collaboration, Ade P. A. R., Aghanim N., et al., 2013a, arXiv:1303.5062
- Planck Collaboration, Ade P. A. R., Aghanim N., et al., 2013b, arXiv:1303.5076
- Planck Collaboration, Ade P. A. R., Aghanim N., et al., 2013c, arXiv:1303.5084
- Press W. H., Schechter P., 1974, ApJ, 187, 425
- Rapetti D., Allen S. W., Weller J., 2005, MNRAS, 360, 555

- Rapetti D., Blake C., Allen S. W., Mantz A., Parkinson D., Beutler F., 2013, MNRAS, 432, 973
- Rasia E., Tormen G., Moscardini L., 2004, MNRAS, 351, 237
- Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2
- Reichardt C. L., Stalder B., Bleem L. E., et al., 2013, ApJ, 763, 127
- Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716
- Riess A. G., Filippenko A. V., Challis P., et al., 1998a, AJ, 116, 1009
- Riess A. G., Macri L., Casertano S., et al., 2011, ApJ, 730, 119
- Riess A. G., Nugent P., Filippenko A. V., Kirshner R. P., Perlmutter S., 1998b, ApJ, 504, 935
- Rood H. J., 1974a, ApJ, 194, 27
- Rood II H. J., 1974b, ApJ, 188, 451
- Rosati P., della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
- Ross A. J., et al., 2013, MNRAS, 428, 1116
- Ross A. J., Percival W. J., Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 415, 2193
- Ross N. P., Shen Y., Strauss M. A., et al., 2009, ApJ, 697, 1634
- Rozo E., Dodelson S., Frieman J. A., 2004, Phys. Rev. D, 70, 8, 083008
- Rozo E., Rykoff E. S., Bartlett J. G., Evrard A. E., 2013, arXiv:1302.5086
- Rozo E., Rykoff E. S., Evrard A., et al., 2009, ApJ, 699, 768
- Rozo E., Wechsler R. H., Koester B. P., Evrard A. E., McKay T. A., 2007, arXiv:astro-ph/0703574
- Rozo E., Wechsler R. H., Rykoff E. S., et al., 2010, ApJ, 708, 645
- Rykoff E. S., McKay T. A., Becker M. R., et al., 2008, ApJ, 675, 1106
- Sachs R. K., Wolfe A. M., 1967, ApJ, 147, 73

- Sartoris B., Borgani S., Fedeli C., et al., 2010, MNRAS, 407, 2339
- Sato K., 1981, MNRAS, 195, 467
- Schmidt B. P., Suntzeff N. B., Phillips M. M., et al., 1998, ApJ, 507, 46
- Schmidt F., Kamionkowski M., 2010, Phys. Rev. D, 82, 10, 103002
- Schmidt F., Vikhlinin A., Hu W., 2009, Phys. Rev. D, 80, 8, 083505
- Schuecker P., Böhringer H., Collins C. A., Guzzo L., 2003, A&A, 398, 867
- Sefusatti E., Croce M., Desjacques V., 2012, MNRAS, 425, 2903
- Seljak U., Slosar A., Mandelbaum R., 2013, In preparation
- Seljak U., Zaldarriaga M., 1996, Astrophys. J., 469, 437
- Shandera S., Mantz A., Rapetti D., Allen S. W., 2013, J. Cosmology Astropart. Phys., 8, 4
- Sheldon E. S., Annis J., Böhringer H., et al., 2001, ApJ, 554, 881
- Sheldon E. S., Johnston D. E., Scranton R., et al., 2009, ApJ, 703, 2217
- Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
- Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
- Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
- Silk J., 1968, ApJ, 151, 459
- Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, J. Cosmology Astropart. Phys., 8, 31
- Smith S., 1936, ApJ, 83, 23
- Smoot G. F., Bennett C. L., Kogut A., et al., 1992, ApJ, 396, L1
- Spergel D. N., Bean R., Doré O., et al., 2007, ApJS, 170, 377
- Spergel D. N., Verde L., Peiris H. V., et al., 2003, ApJS, 148, 175
- Springel V., 2005, MNRAS, 364, 1105

- Springel V., White S. D. M., Jenkins A., et al., 2005, *Nature*, 435, 629
- Stoughton C., Lupton R. H., Bernardi M., et al., 2002, *AJ*, 123, 485
- Sunyaev R. A., Zeldovich Y. B., 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Szabo T., Pierpaoli E., Dong F., Pipino A., Gunn J., 2011, *ApJ*, 736, 21
- Tegmark M., Hamilton A. J. S., Strauss M. A., Vogeley M. S., Szalay A. S., 1998, *ApJ*, 499, 555
- Testi L., Hills R., Laing R., Stanghellini S., Wild W., 2010, *The Messenger*, 142, 17
- Tinker J., Kravtsov A. V., Klypin A., et al., 2008, *ApJ*, 688, 709
- Tinker J. L., Robertson B. E., Kravtsov A. V., et al., 2010, *ApJ*, 724, 878
- Tormen G., 1998, *MNRAS*, 297, 648
- Vanderlinde K., Crawford T. M., de Haan T., et al., 2010, *ApJ*, 722, 1180
- Vikhlinin A., Kravtsov A. V., Burenin R. A., et al., 2009, *ApJ*, 692, 1060
- Voges W., Aschenbach B., Boller T., et al., 1999, *A&A*, 349, 389
- Wang L., Steinhardt P. J., 1998, *ApJ*, 508, 483
- Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, *ApJ*, 646, 881
- Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, *ApJ*, 652, 71
- Weller J., Albrecht A., 2001, *Physical Review Letters*, 86, 1939
- Weller J., Albrecht A., 2002, *Phys. Rev. D*, 65, 10, 103512
- Wen Z. L., Han J. L., Liu F. S., 2009, *ApJS*, 183, 197
- White M., Kochanek C. S., 2002, *ApJ*, 574, 24
- White S. D. M., Davis M., Efstathiou G., Frenk C. S., 1987, *Nature*, 330, 451
- Wittman D., Margoniner V. E., Tyson J. A., Cohen J. G., Becker A. C., Dell'Antonio I. P., 2003, *ApJ*, 597, 218

- Wittman D., Tyson J. A., Margoniner V. E., Cohen J. G., Dell'Antonio I. P., 2001, *ApJ*, 557, L89
- Xia J.-Q., Baccigalupi C., Matarrese S., Verde L., Viel M., 2011, *J. Cosmology Astropart. Phys.*, 8, 33
- Xia J.-Q., Bonaldi A., Baccigalupi C., et al., 2010a, *J. Cosmology Astropart. Phys.*, 8, 13
- Xia J.-Q., Viel M., Baccigalupi C., De Zotti G., Matarrese S., Verde L., 2010b, *ApJ*, 717, L17
- Zeldovich Y. B., 1972, *MNRAS*, 160, 1P
- Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2012, *arXiv:1207.3794*
- Zwicky F., 1937, *ApJ*, 86, 217
- Zwicky F., Herzog E., Wild P., 1968, *Catalogue of galaxies and of clusters of galaxies*



# Curriculum vitae

## Personal

Date of birth 29 March 1985  
Place of birth Fossano (Cuneo)  
Nationality Italian

---

## Education

2010 - present Ph.D. student at University Observatory (München).  
Thesis topic: Cosmology with optically selected galaxy clusters.  
Supervisor: Prof. Dr. Jochen Weller

2009 Master degree in Mathematics at the University of Torino, Italy.  
Grade: 110/110.  
Thesis Topic: Singularities in Einstein's equations solutions for  
polytropic stars in Palatini  $f(R)$  theories.  
Supervisor: Prof. Dr. Lorenzo Fatibene

2007 Bachelor degree in Mathematics the University of Torino, Italy.  
Grade: 110/110.  
Thesis topic: Morphogenesis of fractal structures.  
Supervisor: Prof. Dr. Guido Magnano.

2004 High School Diploma.  
Grade: 100/100

---

## Publications

Mana et al. (2013) Mana A., Giannantonio T., Weller J., Hoyle B., Hütsi G., Sartoris B., 2013, MNRAS, 434, 684

---



# Danksagung

I owe plenty of thanks to all the people who have been at my side during the past three years here in Munich.

First of all, I would like to thank my supervisor, Jochen Weller, for the constant support, fundamental advices and constructive discussions. You always motivated me to achieve my scientific purposes, showing great enthusiasm for the subject and the results we obtained. You were also able to well understand my personality and the problems I encountered during my Ph.D. This really helped a lot to have a great working and human experience.

Secondly, I wish to warmly thank Tommaso and Ben for their great help and collaboration, and all my past and present group members (Marco, Stephen, Martin, Michi, Katarina, Laura, Florian, Ixandra) for the fruitful discussions and good time spent together. Thanks also to all professors and post-docs I had the pleasure to meet personally here or around the world, because I learned a lot from their lessons.

Very special thanks go far away, to my parents, without whom I would not be here now: no distance can ever keep us apart. *Grazie mamma e papi: senza di voi e ai vostri sacrifici, tutto questo non sarebbe stato possibile e ve ne sarò eternamente grata. Vi porto con me, perchè siete e sarete sempre la parte migliore di me.*

Special thanks go to Alessandro: you light up my days. *Grazie Ale per la tua infinita pazienza e comprensione, per il tuo sostegno morale nei momenti di sconforto, per la serenità che mi hai regalato in questi anni. Grazie soprattutto per avermi insegnato ad avere più fiducia in me stessa.*

Thank you to the people who made me spend a really good time in Munich: Manos, Luca, Andi, Annina, Viola, Michele, Giancarlo, Anna, Alexandra, Alex B., Alex K., Veronika, all the IMPRS students I had the chance to meet, all the USM people I have interacted with. Thanks also to Alex Saro and Fra for helping me with the corrections.

I also wish to thank my far away friends (Marzia, Al, Andre, Umbe, Paul, Clod) because you never forgot about your crazy friend up there, and the rest of my family (Paolo, Alex, Fabri, nonna Piera, nonna Pina, Piery, Laura, Davide, Fede, Vale, Andre, Marina, Ale, Giulia e Lory) for the long distance support.

And finally, thanks to myself, to my stubbornness, my heart, my mind, my arms, my tears, my passion.