
The REFLEX II galaxy cluster catalogue: two-point statistics and cosmological implications

Andrés Balaguera Antolínez



München 2010

The REFLEX II galaxy cluster catalogue: two-point statistics and cosmological implications

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

vorgelegt von
Andrés Balaguera Antolínez
aus Bogotá D.C., Kolumbien

München, den 28 September 2010

Erstgutachter: Prof. Dr. Hans Böhringer

Zweitgutachter: Prof. Dr. Ralf Bender

Tag der mündlichen Prüfung: 28.02.2011

“Menos tu vientre,
todo es confuso.
Menos tu vientre,
todo es futuro fugaz
pasado baldío, turbio.
Menos tu vientre,
todo es oculto.
Menos tu vientre,
todo inseguro,
todo postrero,
polvo sin mundo.
Menos tu vientre,
todo es oscuro.
Menos tu vientre
claro y profundo.”
Miguel Hernández

Para Pilar y Celeste

Contents

Zusammenfassung	xiv
1 Introduction	1
2 Cosmological model and structure formation	7
2.1 Introduction	7
2.2 The standard cosmological model	7
2.3 Statistical properties of large-scale structure	11
2.3.1 Two point statistics	12
2.3.2 Growth of inhomogeneities	14
2.4 Distortions in the linear clustering pattern	16
2.4.1 Galaxy bias	16
2.4.2 Redshift space distortions	17
2.4.3 Geometrical distortions	19
2.4.4 Non-linear evolution	20
2.4.5 Other source of distortions	22
2.5 Probes for cosmological parameters	23
2.5.1 CMB	23
2.5.2 Galaxy clustering and BAO	23
2.5.3 Other probes	25
2.6 Galaxy clusters	25
2.7 Summary	29
3 Dark matter haloes: abundance, bias and clustering	31
3.1 Introduction	31
3.2 Dark matter halo mass functions	32
3.3 The dark matter-halo bias	34
3.4 Properties of dark matter haloes	37
3.5 The halo model for gravitational clustering	40
3.5.1 Introduction	40
3.5.2 Dark matter power spectrum	42
3.5.3 Halo model for gravitational clustering	43
3.5.4 Clustering of galaxy clusters	45
3.6 Halo bias and two-point statistics from N -body simulations	48

3.6.1	The L-BASICC II N -body simulations	48
3.6.2	Halo power spectrum	51
3.6.3	Marked statistics with haloes	58
3.7	Summary	62
4	The REFLEX II galaxy cluster catalogue: sample and mock catalogues	63
4.1	Introduction	63
4.2	The REFLEX II galaxy cluster catalogue	63
4.2.1	REFLEX II sample	63
4.3	REFLEX II selection function	68
4.3.1	The REFLEX II X-ray luminosity function	68
4.3.2	Cosmological implications	70
4.3.3	Selection function	72
4.4	Construction of the REFLEX II mock catalogs	73
4.4.1	Mass-X-ray luminosity relation	74
4.4.2	Systematics	79
4.4.3	Properties of the mock catalogues	81
4.5	Summary	83
5	Power spectrum analysis of the REFLEX II sample	85
5.1	Introduction	85
5.2	The measurement of $P(k)$	85
5.2.1	Covariance matrix	88
5.3	Understanding the REFLEX II power spectrum	92
5.3.1	Luminosity bias	92
5.3.2	Effective bias	96
5.4	Distortions induced in a flux limited sample	97
5.4.1	Modeling the shape of $P(k)$	100
5.5	Analysis of the REFLEX II power spectrum	101
5.5.1	Measurements	101
5.5.2	Amplitude of the REFLEX II $P(k)$	105
5.5.3	Shape of the REFLEX II $P(k)$	106
5.6	Correlation function	107
5.7	Summary	109
6	Conclusions	113
A	Spherical collapse model	117
B	The halo model for gravitational clustering	123
C	The power spectrum estimator	137

List of Figures

2.1	Cosmological functions for different cosmological parameters	11
2.2	Description of the linear matter power spectrum	15
2.3	Power spectrum for different cosmologies	20
2.4	Linear matter power spectrum	24
3.1	Halo mass function	38
3.2	NFW and Einasto density profiles	40
3.3	Halo model	41
3.4	Halo model power spectrum and correlation function	43
3.5	Prediction of cluster luminosity function	45
3.6	Halo abundance as a function of redshift	47
3.7	Halo mass function. Comparisson	48
3.8	L-BASSIC II mass function at different redshifts	49
3.9	Absolute halo bias	51
3.10	Halo power spectrum	53
3.11	L-BASICC II: covariance matrix of the halo power spectrum	54
3.12	L-BASICC II: power spectrum variance	55
3.13	Absolute halo bias	56
3.14	L-BASSIC II Power spectrum	57
3.15	Marked correlation function	59
3.16	Bias factors in the marked-correlation function	60
3.17	BAO in the halo marked statistics	61
4.1	Aitoff projection of the REFLEX II galaxy clusters in equatorial coordinates	64
4.2	Propertie of the REFLEX II sample	65
4.3	Redshift distribution	66
4.4	Luminosity-redshift diagram for the REFLEX II sample (points).	67
4.5	Luminosity function of the REFLEX II sample	68
4.6	Confidence levels for the parameterization of the X-ray luminosity function	70
4.7	Luminosity function compared with predictions from the Λ CDM cosmological model	71
4.8	Mass- X-ray luminosity-mass for the REFLEX II mock catalogues	74
4.9	Testing Kaiser approximation	77
4.10	Distribution of luminosities around the mean mass luminosity	78

4.11	Luminosity function	79
4.12	Distribution of halo masses in the REFLEX II mock catalogues	80
4.13	Distribution of luminosities around the mean mass luminosity	81
4.14	Some properties of the mock catalogues	82
5.1	REFLEX II window function	86
5.2	Mean power spectrum from the REFLEX II mock catalogues	87
5.3	Power spectrum of mock catalogues	88
5.4	Correlation matrix from the REFLEX II mock power spectra	89
5.5	Effective volume probed by the REFLEX II sample	89
5.6	Bin-averaged variance of the REFLEX II power spectrum	90
5.7	Probability distribution of the 100 mock power spectra	91
5.8	Skewness and Kurtosis from the REFLEX II power spectra	92
5.9	Luminosity bias from the L-BASICC II simulation as a function of wavenumber	93
5.10	Luminosity bias from the L-BASICC II simulation	94
5.11	Absolute luminosity bias	95
5.12	Effective luminosity bias measured from the REFLEX II mock catalogues	97
5.13	Mean squared luminosity bias	98
5.14	Comparison between power spectrum estimators	99
5.15	Q -model description of the mock power spectrum	100
5.16	REFLEX vs REFLEX II power spectrum	102
5.17	REFLEX II power spectrum. Comparison between data and mock catalogues.	103
5.18	REFLEX II power spectrum for three different minimum luminosities	105
5.19	Relative luminosity bias as a function of scale in the REFLEX II sample	106
5.20	Relative luminosity bias measured from the REFLEX II sample	108
5.21	Best-fitting Q -model for the REFLEX II mock power spectrum	109
5.22	Correlation function of the REFLEX II sample	110
B.1	HOD	129
B.2	Galaxy power spectrum from the halo model	130
B.3	Scale dependent mean density	135
B.4	Power spectrum with halo exclusion	136
C.1	Shell-averaged window function	143

List of Tables

2.1	Cosmological parameters	10
3.1	Fitting formulae for the halo mass function	34
4.1	Minimum X-ray luminosities used to define sub-samples of the REFLEX II catalogue	67
4.2	Fitting of the REFLEX II luminosity function	69
4.3	Description of the mock catalogues	83
4.4	Content of the synthetic catalogue	84
5.1	Recovering the growth index	96
5.2	Numerical values of the REFLEX II power spectrum	107
5.3	Correlation function	111

Abstract

In this work I present the characterization of the spatial distribution of X-ray galaxy clusters, based on the measurements of the cluster power spectrum. The analysis is developed on the new ROSAT-ESO Flux-Limited X-Ray (REFLEX II) galaxy cluster catalogue. This new sample extends the flux limit of the original REFLEX to 1.8×10^{-12} erg s⁻¹ cm⁻², yielding a total of 911 clusters with $\geq 94\%$ completeness in redshift follow-up. The X-ray luminosity function was measured and parameterized in order to be used as cosmological probe. The statistical analysis is complemented by creating a set of 100 REFLEX II-like mock catalogues, built from a suite of large volume Λ CDM N -body simulations (L-BASICC II). We have characterized the main properties of this set of simulations, such as the halo abundance, the halo bias, the halo clustering and the mass-X-ray luminosity relations, all these required to properly understand the observed clustering signal of X-ray galaxy clusters. The mock catalogues have been designed to reproduce the observed X-ray luminosity function by calibrating the mass-X ray luminosity relation. Our measurements of the X-ray cluster power spectrum are in agreement with predictions from the Λ CDM cosmological model and show the expected increase in the amplitude of the power spectrum with increasing X-ray luminosity. The better statistics of the REFLEX II sample allowed us to explore the issue of luminosity bias in some detail. On large scales it was observed a scale-independent relative luminosity bias, which implies that the clustering of galaxy clusters, on those scales, is a scaled version of the clustering of dark matter. Moreover, this suggests that within the precision given by the surveyed volume of the REFLEX II catalogue, the shape of the measured galaxy cluster power spectrum is not affected by distortions induced due to the presence of clusters with different X-ray luminosities. We confirmed this result by implementing a luminosity dependent power spectrum estimator. The measured power spectrum is statistically compatible with a featureless power spectrum on scales $k > 0.01 h/\text{Mpc}$ and hence no statistically significant signal of baryonic acoustic oscillations has been detected. Similar conclusions are drawn from the analysis of the clustering in configurations space by means of the cluster correlation function. For the first time, a signature of non-linear evolution in the galaxy cluster power spectrum has been detected on scales $k \approx 0.15 h/\text{Mpc}$. We model the shape of the measured power spectrum by means of phenomenological parameterizations, which are useful for our measurements due the moderate volume probed by the survey. The measurements of cluster power spectrum described in this work are suitable to be used as probes for cosmological parameters.

Zusammenfassung

In dieser Arbeit wird die großskalige Struktur von Galaxienhaufen aus dem ROSAT-ESO Flux-Limited X-Ray (REFLEX II) Katalog untersucht. Aufgrund der im Vergleich zum ursprünglichen REFLEX Katalog schwächeren Grenzhelligkeit von 1.8×10^{-12} er/s/cm⁻² konnte eine Gesamtanzahl von 911 Haufen gefunden werden; für mehr als 84% von diesen existieren spektroskopische Rotverschiebungen. Die Röntgenleuchtkraftfunktion wurde gemessen und parametrisiert, um sie zur Bestimmung kosmologischer Parameter verwenden zu können. Zur weitergehenden statistischen Analyse dieses neuen *Samples* wurden 100 REFLEX II *Mock*-Kataloge aus großen Λ CDM *N-body* Simulationen (L-BASICC II) hergestellt, und Messgrößen wie die Massenfunktion, den *Bias*, die großskalige Struktur und die Masse-Röntgenleuchtkraft-Beziehung der Dunklen Halos charakterisiert, die benötigt werden, um die beobachtete großskalige Struktur der Galaxienhaufen zu verstehen. Unter Verwendung der Masse-Röntgenleuchtkraftbeziehung wurden die *Mock*-Kataloge so konstruiert, dass die beobachtete Röntgenleuchtkraftfunktion reproduziert wird. Die gemessenen Leistungsspektren der Röntgenhaufen stimmen mit Vorhersagen der Λ CDM-Kosmologie überein und zeigen das erwartete Anwachsen der Amplitude des Leistungsspektrums mit wachsender Röntgenhelligkeit, was eine Folge der direkten Beziehung zwischen der beobachteten Leuchtkraft und der Gesamtmasse des Haufens ist. Die bessere Statistik des REFLEX II *Samples* erlaubt es, den Leuchtkraft *bias* mit hoher Genauigkeit zu untersuchen. Der beobachtete relative Leuchtkraft *bias* entspricht einem skalenunabhängigen Halomassen*bias*, in Übereinstimmung mit Befunden aus *N-body* Simulationen. Dies lässt darauf schließen, dass die Form des Leistungsspektrums der REFLEX II Galaxienhaufen innerhalb der durch das Beobachtungsvolumen gegebenen Genauigkeit nicht durch den Einfluss des beobachteten Populationsmixes verzerrt wird. Dies wird durch die Verwendung eines leuchtkraftunabhängigen Schätzers für das Leistungsspektrums bestätigt. Im Rahmen der Messgenauigkeit ist das gemessene Leistungsspektrum auf Skalen von $k > 0.01h$ Mpc⁻¹ vereinbar mit einem glatten Leistungsspektrum, weshalb kein statistisch signifikantes Signal der Baryonischen Akustischen Oszillationen detektiert werden kann. Ähnliche Schlussfolgerungen lassen sich aus der Analyse der großskaligen Struktur im Konfigurationsraum mithilfe der Korrelationsfunktion ziehen. Es kann jedoch gezeigt werden, dass das REFLEX II Leistungsspektrum auf Skalen von $k > 0.15h$ Mpc⁻¹ Signaturen von nichtlinearer Entwicklung aufweist. Die Form des gemessenen Leistungsspektrums wurde mithilfe phänomenologischer Parametrisierungen modelliert, welche aufgrund des überschaubaren Volumens der Himmelsdurchmusterung benutzbar sind.

Introduction

Within the last decade, three foundational observational probes have been well recognized as opening up observational windows to determine some of the most relevant features of our Universe. These experiments are: observations of the temperature fluctuations of the cosmic microwave background radiation (CMB), recently measured with high precision by the Wilkinson Microwave Anisotropy Probe (*WMAP*) satellite (Spergel et al., 2007; Komatsu et al., 2010); the Hubble diagram inferred from the Type Ia supernovae (SNe) observations (e.g. Perlmutter et al., 1998; Riess et al., 2004) and the measurement of the large scale structure (LSS) of the Universe as traced by the spatial distribution of galaxies (Percival et al., 2002; Tegmark et al., 2004; Sánchez et al., 2006; Percival et al., 2007a, 2010; Reid et al., 2009). With the completion of large redshift surveys, such as the Two-degree Field Galaxy Redshift Survey (2dFGRS)¹ and the Sloan Digital Sky Survey (SDSS)² it has been possible to improve the level of accuracy of LSS studies.

The recent detection of the baryon acoustic oscillations (BAO, e.g. Eisenstein et al., 2005; Gaztanaga et al., 2008; Sánchez et al., 2009; Percival et al., 2010) in redshift surveys have been key to this progress. Over the last few years these observations have established the concordance cosmological model, based on a flat space-time in a current phase of accelerated expansion due to the presence of a dominating dark energy component which exerts a negative pressure (i.e. a negative equation of state). Moreover, the equation of state of the dark energy has been shown to be compatible with the once rejected Einstein's cosmological constant Λ .

The advent of the era of precision cosmology has provided accurate measurements ($\sim 2\%$ precision) on the most relevant parameters describing the standard cosmological model. Such scenario, instead of providing a complete understanding of the Universe, had raised fundamental questions (what is the origin of dark energy, is there a dark-matter particle?). These are tried to be answered not only from a cosmological perspective, but from an overall view covering all fields of physics, ranging from quantum field theory to general relativity. Such current status leaves us with the picture of a Universe dominated by an unknown weird fluid with negative pressure. Major investments of scientists and money have been willing to develop large galaxy surveys that lead to the determination of the properties of dark energy, but still a great theoretical breakthrough must be done in order to understand its physical origin (e.g. Copeland et al., 2006). Following the dark energy, in the second position of

¹<http://msowww.anu.edu.au/2dFGRS/>

²<http://www.sdss.org/>

dominance we find a less weird fluid, the dark matter. The less strangeness of dark matter lies in the fact that it interacts gravitationally with the rest of the known Universe. From an astronomical point of view, this makes it suitable to be detected by indirect methods such as weak lensing (e.g Clowe et al., 2004). From a theoretical point of view, dark matter can find a place within the already complex particle physics scenario (e.g. Greiner and Muller, 1989). In order to translate this to something more than a beautiful idea, a great amount of effort has been placed in designing and building detectors and probes for dark matter particles (e.g. Bernabei et al., 2008).

Adding up these two contributions, we finish with $\sim 96\%$ of the total amount of matter-energy in the Universe, which in other words means that only 4% of the Universe is made of material we certainly know: galaxies, gas, stars, planets: in summary, atoms. We can be tempted to think that modern science is still rolling over the wheels of geocentrism, in the sense that we (the human species) dare to consider and call *strange* almost the 75% of the energy content of the Universe, while being ourselves (and all we can see), the real strange 4% minority. Summarizing, this is the so-called Λ CDM cosmological model.

In the search for a better understanding of the physical processes that generated the Universe we observe, galaxy clusters have played a relevant role and their study will provide, in the near future, an accurate description of the Universe, comparable to that obtained by current galaxy surveys (e.g. Vikhlinin et al., 2009a). The first astronomical studies concerning galaxy clusters date back to the estimates of the mass of a galaxy cluster carried out by Zwicky (1933) on the Coma cluster³, revealing for the first time the need for an invisible component, the dark matter. Nowadays galaxy clusters are still the best scenario where the existence of dark matter can be confirmed (e.g. Clowe et al., 2004). The compilation of galaxy clusters such as the Abell and ACO catalogues (Abell, 1958; Abell et al., 1989; Ebeling and Maddox, 1995) provided a valuable tool to explore the optical properties of galaxy clusters such as their richness⁴ and also led to the first characterization of the large scale structure with galaxy clusters (e.g. Peebles, 1980). With the advent of the satellite era, spatial missions such as the *Uhuru*⁵ or *Einstein*⁶ satellites provided an opening window for X-ray astronomy and showed (among other interesting aspects of X-ray astronomy) that galaxy clusters are active X-ray emitters. Such scenario demanded the existence of a hot intra-cluster plasma in thermal equilibrium generating X-ray photons via bremsstrahlung radiation. This picture was confirmed with the posterior detection of an iron (thermal) emission-line, which in turns provided a suitable scenario to study galaxy evolution (a complete description of X-ray galaxy clusters has been given by Sarazin 1988).

Today we refer to galaxy clusters as the largest bounded structures in the Universe. They are associated with the highest peaks in the (appropriately filtered) matter density field and are recognized as biased tracers of the underlying matter distribution (e.g. Bardeen et al., 1986). Typical galaxy clusters possess masses of the order of $10^{14}M_{\odot}$, radii of $\sim 2\text{Mpc}$ and can host around 10^3 galaxies. The skeleton of galaxy clusters consist in an invisible dark matter halo accounting for $\sim 90\%$ of the cluster total mass. The visible part accounting for the remaining fraction is embodied in hot gas and galaxies (stars, dust etc). The formation of galaxy clusters involve several steps, starting from the collapse of dark matter haloes, followed

³The Coma cluster contains ≈ 1000 identified galaxies at redshift $z = 0.025$, located in the constellation of Coma Berenices

⁴Defined in terms of the magnitude of the 10^{th} brightest galaxy in the cluster

⁵<http://heasarc.gsfc.nasa.gov/docs/uhuru/uhuru.html>

⁶<http://heasarc.gsfc.nasa.gov/docs/einstein/heao2.html>

by the occupation of such haloes by baryons and the formation of galaxies. Once formed, they are driven again by gravity and merge to compose the largest structures that can be approximately considered in virial equilibrium. Their deep potential wells make them the largest astrophysical laboratories in the Universe, where the combination of gravitation and baryonic physics has been intensively studied through the analysis of cluster properties such as scaling relations (e.g. Reiprich and Böhringer, 2002a; Pratt et al., 2009; Mantz et al., 2010), density profiles (e.g. Makino et al., 1998a; Croston et al., 2008), pressure profiles (e.g. Arnaud et al., 2009), baryon fractions (e.g. Giodini et al., 2009), etc. The abundances of galaxy clusters determined by their luminosity function (e.g. Böhringer et al., 2002) can also be used to constrain parameters like the matter content in the universe Ω_m , and the amplitude of the density fluctuations characterized by σ_8 , the root-mean-squared linear perturbation theory variance in spheres of radius 8 Mpc h^{-1} (e.g. Schuecker et al., 2003b). The spatial distribution of galaxy clusters, characterized mainly by its two-point statistics, provides as well useful information about the cosmological model of the Universe. The shape of this measurement is particularly sensitive to the parameter combination $\Omega_m h$, and the amplitude of the galaxy cluster power spectrum provides information related to the dark-matter halo clustering and the mass-X ray luminosity relation. In a sense, theoretical models involving dark-matter halo clustering and the baryon physics in clusters can be used as reliable frameworks to explain the observed clustering strength (e.g. Moscardini et al., 2000; Hütsi and Lahav, 2008), something that can not be easily reached when the clustering of galaxies is studied.

X-ray detected galaxy clusters display a number of advantages compared to optically detected galaxy clusters and galaxy surveys. First, diffuse emission in X-rays associated mainly to thermal bremsstrahlung ensures us that we are witnessing deep potential wells associated to the high mass configurations. Note however that X-ray emission is not a warranty that the system is in equilibrium or in a relaxed state (e.g. Mulchaey, 2000). Second, if gravitation is the only interaction ruling the physical processes in galaxy-clusters, then galaxy clusters of a certain mass can be thought of as scaled copies of clusters with different masses (or radius, luminosities, temperature, etc) and the inferred X-ray luminosity (or temperature) can be directly associated to the total mass of the clusters by simple scaling relations. Third, being the emissivity associated to bremsstrahlung radiation proportional to the squared of electron density (e.g. Padmanabhan, 1993a), the emission in X-ray is more concentrated than the galaxy distribution observed in the optical band. Last, X-ray cluster catalogues might full-sky coverage simply because of being forced to be developed in space. Due to this example of advantages, the potential of galaxy clusters detected in the X-ray band has been recognized and great efforts have been made to construct large samples of X-ray galaxy clusters (e.g. Truemper, 1993; Ebeling et al., 1998; Böhringer et al., 2004). This work is based on one of such huge efforts: the REFLEX sample (Böhringer et al., 2004).

In this thesis we characterize the spatial inhomogeneous distribution of X-ray detected galaxy clusters by means of the galaxy cluster (gravitational) power spectrum $P(k)$. The relevance of the power spectrum has been widely recognized as a primordial statistical tool (Peebles, 1980) due to the amount of cosmological and astrophysical information it encodes (e.g. Feldman et al., 1994; Eisenstein and Hu, 1998; Einasto et al., 1999; Schuecker, 2005; Hamilton, 2005). A hand-wave definition of this quantity can be given as to be the amount of clustering power that gravity provides on spatial scales λ characterized by a wavenumber $k \sim 1/\lambda$. When we mention power spectrum, we might be talking about matter, galaxies or galaxy clusters. In general, saved the physical properties characterizing galaxies and clusters (some of which determine the strength of their clustering signal), the shape of the

power spectrum is basically determined by the way in which fluctuations in the underlying matter density distribution evolved in time due to gravitational interaction, in a homogeneous background determined by the so-called Friedman-Robertson-Walker Universe. On very large scales the slope of the power spectrum contains information related to primordial fluctuations, closely related to inflationary models (e.g. Liddle and Lyth, 2000). The scale associated to the horizon at the time of the matter-radiation equality (i.e. the largest distance a photon could travel from the Big-bang until that time) is remarkably imprinted in the shape of the power spectrum as a maximum whose value is sensitive to the content of matter. On scales smaller than the horizon, the shape of the power spectrum is dominated by the evolution of perturbations arising from causal connection. It is of special relevance that the presence of the BAO, associated to the comoving distance traveled by a sound-like perturbation in a radiation dominated fluid before the recombination epoch. Although this feature might be better understood in configuration space with the correlation function, its corresponding behavior in Fourier space has been theoretically modeled (e.g. Hu and Sugiyama, 1996; Eisenstein and Hu, 1998) and measured (e.g. Meiksin et al., 1999; Eisenstein et al., 2005; Percival et al., 2010), setting the BAO as a standard ruler useful to constrain cosmological parameters (e.g. Glazebrook and Blake, 2005; Wagner et al., 2008).

The statistical analysis of the spatial distribution of galaxy cluster dates back to Bahcall and Soneira (1983) and Klypin and Kopylov (1983) with the measurements of the two-point correlation function from rich clusters based on the Abell/ACO catalogue. The power spectrum from the same sample was measured by Peacock and West (1992) followed by the measurements of Einasto et al. (1997) and Retzlaff et al. (1998). The shape of the power spectrum determined from such measurements confirmed the predictions from a cosmological model with dark matter and was compatible with being a rescaled version of the observed galaxy power spectrum (e.g. Tadros and Efsthathiou, 1996; Tadros et al., 1998). Despite the good agreement found on intermediate scales, these measurements also displayed features that were not predicted by the standard cosmological model, such as the detection of periodicity patterns in the cluster distribution represented by a peak at very large scales of $\lambda \sim 120$ Mpc (Einasto et al., 1997). Only the use of larger samples with well-defined selection functions and supported by numerical simulations (e.g. Retzlaff et al., 1998) revealed that the significance of such feature was strongly dominated by cosmic variance and thus very large volumes are required to properly sample such large scales. These conclusions were derived a few years later with the first measurement of the power spectrum of X-ray galaxy clusters (e.g. Schuecker et al., 2001; Zandivares et al., 2001) and is also confirmed by the present work.

Concerning the detection of BAO in the cluster power spectrum, a first detection was claimed by Miller et al. (2001) based on the measurements of the power spectrum of the Abell/ACO catalogue by Miller and Batuski (2001). Interestingly, the sample we will study contains approximately the same number of objects as used by Miller et al. (2001) and samples a larger volume. However we do not expect to obtain a statistically significant signal of BAO from our sample (Schuecker et al., 2001), mainly due to the limited volume sampled by the survey. The detection claimed by Miller et al. (2001) did not transcend as the first clear signature of BAO in the distribution of luminous matter as was indeed the case of the detection of the acoustic peak in the spatial galaxy distribution by Eisenstein et al. (2005). Recently Hütsi (2010) reported a 2σ detection of BAO in the maxBCG (photometric) galaxy cluster survey (Koester et al., 2007), which samples a larger volume (approximately one order of magnitude bigger) than the largest X-ray galaxy cluster sample, the REFLEX catalogue.

The *Röntgen Satellite* (ROSAT) mission (1990 – 1999) represents a breakthrough in the field of X-ray astronomy, with a legacy of twenty years of research on cosmology and astrophysics. The ROSAT All Sky Survey (RASS) observations (Truemper, 1993; Voges et al., 1999) used for the construction of galaxy-cluster catalogues both in the northern (NORAS) (Böhringer et al., 2000) and in the southern hemisphere, the ROSAT-ESO Flux Limited X-ray (REFLEX) catalogue (Böhringer et al., 2004). The REFLEX catalogue was, to date, the largest statistically complete X-ray detected cluster sample, only overcome by itself with the advent of the REFLEX II sample. The clustering properties of this survey were analyzed by means of the power spectrum (Schuecker et al., 2001), the cluster correlation function (Collins et al., 2000), cluster-galaxy cross-correlation functions (Sánchez et al., 2005) and Minkowski functionals (Kerscher et al., 2001). Sub-samples of the REFLEX catalogue complemented by detailed follow-up observations have been used to constrain cluster scaling relations (e.g. Reiprich and Böhringer, 2002a; Ortiz-Gil et al., 2004; Stanek et al., 2006; Pratt et al., 2009; Mantz et al., 2010). The completion of the redshift measurements for galaxy-clusters in the RASS sample (NORAS+REFLEX) is planned to be carried out during the next two-years. This effort will provide the largest volume ever probed to date by X-ray surveys. The analysis of such sample will represent the end of the ROSAT era concerning LSS analysis and will become the preamble for the advent of the scientific results that are planned to be obtained with the launch of the *eROSITA* mission.

This thesis

The final aim of this work is to characterize the large scale structure of the Universe by the measurement of the two-point statistics of galaxy clusters based on the new ROSAT-ESO Flux-limited (REFLEX II) galaxy cluster survey. In an attempt to be brief and at the same time self-consistent, this thesis has been divided in four chapters that are designed to explain the different components of the title of this work, namely, cosmology, galaxy clusters and two-point statistics.

The overall content of this work lies in a cosmological context, where distances are measured in units of Megaparsecs (Mpc)⁷, masses given in units of billion times the mass of the sun and temperatures ranging from the cold Universe at 2.7 K to the inner regions of galaxy clusters up to 10⁸ K. The ranges in temperatures and therefore in energies marks perhaps a border between cosmology and astrophysical processes. We will not dig deeper into these interesting aspects, but will be aware in their importance in our analysis.

A summarized content of the chapters is given as follows. In Chapter 2 we introduce the standard cosmological model. We briefly describe the basics of the statistical analysis of large scale structure by introducing the two-point statistics. We comment how these measurements can be used as probes for cosmological models and thereby how are they used to set constraints on cosmological parameters. We also present some of the the most relevant experiments and probes from where the most precise measurements of cosmological parameters have been derived. For the statistical analysis we are mainly focused on a spectral decomposition, i.e, we concentrate on the properties of large-scale structures in Fourier space. Therefore we put less attention to the analysis based on the correlation function, developed in configuration space, though some results will be briefly shown and commented.

Chapter 3 is dedicated to introduce and study basic properties of dark matter haloes such

⁷1Mpc= 10⁶pc = 3.26 × 10⁶ light-years.

as their abundance and their distribution with respect to the underlying dark matter. We briefly introduce the theoretical context where the halo abundance, halo-matter bias and halo clustering is explored. We finally we use N -body simulations on dark matter and dark matter haloes from which we can directly measure these aforementioned properties. This chapter then provides a suitable basis from where the observations on the abundance and the clustering of galaxy clusters can be properly understood.

Chapter 4 introduces the REFLEX II galaxy cluster survey, from which our analysis on clustering will be extracted. We show some of the most relevant properties of the sample with a brief explanation of the machinery used for its construction. We measure the REFLEX II luminosity function (i.e, the cluster abundance as a function of X-ray luminosity). Based on this measurement and implementing N -body simulations, we construct a set of mock catalogues. The construction of these mock catalogues is important, since it allows us to develop a more realistic treatment of the statistical methods involved in this project. These catalogues contain underlying information concerning the galaxy cluster luminosity function and mass-X ray luminosity relation.

Chapter 5 concentrates on the two-point statistical analysis based on the REFLEX II sample. We describe the methods implemented, namely, the description of the power spectrum estimator and the construction of the survey window function. We explore the clustering properties of our suit of mock catalogues in order to validate them as suitable source to extract the power spectrum covariance matrix. We explore the behavior of the clustering strength as a function of the X-ray luminosity and with the help of our mock catalogues we determine theoretical models to describe the full shape and amplitude of the measured clustering. We briefly discuss the results concerning the measurements of the cluster-correlation function and show that it is in agreement with the respective measurements in Fourier space. We finally make some attempts to use the information encoded in our measurements to extract constraints on some relevant cosmological parameters. We end with a summary in Chapter 6.

Convention

The analysis presented here uses distances of galaxy clusters written in terms of their redshift, determined from individual redshift measurements of galaxies in each galaxy cluster. In order to determine the cluster luminosities from the measured fluxes and to determine comoving distances to these objects, a fiducial cosmological model is required. Our fiducial cosmological model consist in a flat Λ CDM Universe with a matter energy-density parameter of $\Omega_{\text{mat}} = 0.25$, a dark energy equation of state $w = -1$ and a dimensionless Hubble parameter of $h = 0.7$, such that the Hubble constant H_0 is given in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We refer by L_X to the X-ray luminosity as that in the *ROSAT* energy band $[0.1 - 2.4] \text{ keV}$ and, whenever it is not explicitly written, its units are $10^{44} \text{ erg s}^{-1} h^{-2}$. For bolometric luminosities we implement the notation $L_{X \text{ bol}}$.

The estimation of parameters from observations described along this work and the error analysis makes use of the Markov Chain Monte Carlo (MCMC) method (e.g. Verde, 2009; Heavens, 2009).

Cosmological model and structure formation

2.1 Introduction

We dedicate this chapter to introduce the basic concepts and terminology of the modern cosmology and the statistical analysis of large scale structures. In Section 2.2 we introduce the standard cosmological model and the latest constraints on cosmological parameters. In Section 2.3 we introduce the basic tools for developing statistical analysis of large scale structure, mainly the so-called two-point statistics: the correlation function and its Fourier transform, the power spectrum. These quantities are fairly well modeled by linear theory of gravitational collapse, and in Section 2.4 we present some physical effects that can lead to deviations from linear theory. We briefly describe of the probes from which precise measurements of cosmological parameters are currently extracted. Being this thesis focused on galaxy clusters, we end this chapter with a description of the main properties of these objects.

2.2 The standard cosmological model

The standard cosmological model is based on the assumption that on large scales, the Universe is isotropic, homogeneous, and evolves in time according to the Einstein-Hilbert field equations and the Friedmann-Robertson-Walker metric (e.g. Peebles, 1980; Padmanabhan, 1993b; Peacock, 1999; Carroll, 2003; Dodelson, 2004). Provided with a perfect fluid matter-energy tensor, the field equations and the FRW line element generates a set of differential equations encoding the kinematics of the Universe. From this set of equations is derived the so-called Friedman equation, written as

$$H(z) \equiv \frac{\dot{a}}{a} = H_0 \left(\Omega_{\text{cdm}}(1+z)^3 + \Omega_{\text{rad}}(1+z)^4 + \Omega_{\text{DE}}(1+z)^{3(1+\omega)} + \Omega_k(1+z)^2 \right)^{\frac{1}{2}} \equiv H_0 E^2(z), \quad (2.1)$$

where $a(t)$ is the scale factor¹ and $z = \frac{1}{a(t)} - 1$ is the cosmological redshift. Here Ω_i represents the contribution of the different components of matter-energy in the universe relative to the

¹The scale factor $a(t)$ links comoving distances r with physical distances x , via $x = ar$.

critical density $\rho_{\text{crit}}(z) = 3H^2(z)/8\pi G_{\text{N}}$, accounting for representing the CDM component Ω_{cdm} , the dark energy component Ω_{DE} , the contribution of radiation (photons plus neutrinos) Ω_{rad} and the curvature Ω_{k} . The time evolution of each component is strongly determined by its equation of state w_i , defined as the ratio of the pressure to the energy density $w = p/\rho$. Equation (2.1) assumes no interaction between different components (apart from gravitation) and a time-independent equation of state for all components, with $w_{\text{DM}} = 0$ for matter and $w_{\text{R}} = 1/3$ for radiation. Concerning the dark energy fluid, the Friedman equation with time-dependent equation of state $w_{\text{DE}}(z)$ is simple to derive. In that case, the rate of acceleration is obtained by replacing

$$w_{\text{DE}} \rightarrow w_{\text{eff}}(z) = \frac{1}{\ln(1+z)} \int_0^z \frac{1+w(z)}{1+z} dz - 1. \quad (2.2)$$

The Friedman equation (2.1) allows us to determine the behavior of the scale factor as a function of cosmic time as well as a function of the energy density parameters Ω_i . At very early times, radiation was the dominating component leading to an acceleration phase. During this epoch, the horizon grew as $d_{\text{hor}} \sim a$ (for $k = 0$) such that in this regime, regions that were causally disconnected regions couldn't come to causal connection (giving rise to the so-called the horizon problem). The transition from a radiation-dominated to a matter-dominated phase occurred at a redshift given by $z_{\text{eq}} \approx 2.519 \times 10^4 \Omega_{\text{m}} h^2$ after which the matter component started to rule the expansion. During this period, the process of recombination took place ($z \approx 1090$) giving origin to the first hydrogen and helium atoms. The Universe became transparent and the information from the last scattering surface is what we observe now as the CMB photons. This age was followed by the so-called “dark ages”, a period in the cosmic evolution where the only electromagnetic signal emitted was the one associated with the 21cm spin line of neutral hydrogen. The process of structure formation reached its non-linear regime during this epoch creating the first compact objects with deep potential wells where baryons fell and emitted radiation capable of re-ionizing the surrounding hydrogen clouds (e.g. Longair, 2005). Hierarchical structure formation took place, by forming the first stars, globular clusters, compact objects like quasars, and galaxies, all these embedded within dark matter haloes, which merged to finally form galaxy clusters. The contribution of dark energy started to be comparable with the contribution of dark matter at a redshift $z_{\text{cv}} \sim 0.33$ (using $w = -1$ for the dark energy fluid). In this phase the horizon scaled proportional to $a^{1/2}$, i.e., with a lower rate than in the radiation dominated epoch. As a consequence, regions that were causally disconnected at the beginning of the radiation dominated epoch can now enter the horizon of other regions and come in causal connection. Finally, the expansion in a non-flat universe will be dominated by a the curvature with $a \propto t$. In such scenario, the sign of the curvature might lead to an eternal expansion or an unavoidable re-collapse. Notice that in a flat universe, this scenario can be also reached by a dark energy fluid with an equation of state $0 > w > -1$ (expansion) or re-collapse (or big-rip) for values $w < -1$.

The energy content not only affects the kinematics of the homogeneous universe, but also the measurements of distances. The comoving (line-of-sight) separation between the observer at $z = 0$ and an object at redshift z in a homogeneous universe is determined from the FRW metric and depends on the Hubble parameter $H(z)$:

$$r(z) = c \int_0^z \frac{dz'}{H(z')}. \quad (2.3)$$

where c is the speed of light. The fact that the measurements of distances depend on the energy content and its parameterizations has allowed pure geometrical measurements (i.e.

separations of galaxies, observed radius of galaxy clusters) to be used as probes to constrain the parameters of the cosmological model. This is developed through the determination of the (comoving) angular diameter distance $D_A(z)$ given by

$$D_A(z) = \frac{c}{H_0 \sqrt{|\Omega_k|}} f_k \left(H_0 \sqrt{|\Omega_k|} \int_0^z \frac{dz}{H(z)} \right), \quad (2.4)$$

where $f_k(x) = \sinh(x)$, x , $\sin(x)$ for $k = -1, 0, +1$. The determination of the luminosity from the observed flux of a distant source depends also on the cosmological model; the relation between flux and luminosity defines the luminosity distance as the distance to a source with luminosity L that has been observed with a flux F : $D_L(z) = \sqrt{L/4\pi F}$. The luminosity distance is related to the angular diameter distance via $D_L(z) = (1+z)^2 D_A(z)$ (e.g. Hogg, 1999).

Based on General Relativity and standard particle physics, the standard cosmological model can predict a wide range of astronomical observations (e.g. CMB or the large-scale distribution of galaxies and clusters) with high precision, introducing only six variables: the density parameter for cold dark matter Ω_{cdm} , dark energy Ω_{de} , baryons Ω_{ba} , the amplitude and spectral index of primordial fluctuations (σ_8, n_s) and the optical depth τ associated to the re-ionization history of the universe. Other parameters are derived from this set (e.g. curvature, the age of the Universe). Nevertheless, astronomical observations are still bringing to stage some physical phenomena that can not be satisfactorily explained within this model. The commonly mentioned problems of the standard model are 1) the cosmological constant problem: what is its origin, why is it so small but still $\neq 0$ (e.g. Weinberg, 1989), 2) the horizon problem: why causally disconnected regions in the Universe display the same average temperature? ², 3) the fine tuning problem: why does the theory need to fine-tune initial conditions to explain a flat Universe? Inflation theory (e.g. Liddle and Lyth, 2000) has provided theoretical grounds to remove these problems by introducing a phase of exponential acceleration before the radiation dominated epoch; such extreme acceleration led regions that were originally in causal contact to be observed today as physically disjoint patches in the sky. Current astronomical probes such as the CMB can constrain inflationary parameters (Komatsu et al., 2009). Nevertheless the road to solving the cosmological constant problem and the origin of dark energy is long. Nowadays the most ambitious astronomical experiments, mainly galaxy redshift surveys (e.g. BOSS³ and PanSTARRS⁴) are being designed to explore the behavior of the dark energy fluid as a function of redshift and generate the most precise measurements of the dark energy equation of state.

The current state of acceleration of the Universe has been established by observations of the peak brightness of distant SNe Ia (e.g. Perlmutter et al., 1998; Garnavich et al., 1998; Perlmutter et al., 1999; Riess et al., 2004; Spergel et al., 2007). Within the last decade direct measurements of acceleration have reached unprecedented precision thanks to the increasing number of SNe Ia detected and improved techniques (e.g. Wang, 2000). From the point of view of General Relativity, such an accelerated state implies that the strong energy condition (e.g. Carroll, 2003) must be violated by the total amount of energy density in the universe. This compelled cosmologists to introduce a dark energy fluid with a negative equation of state and with a dominating budget in the total mass-energy of the Universe. At the present

²if two regions located at the redshift characterized by the last-scattering surface of CMB photons are observed to be separated by an angular scale of $\Delta\theta \sim z_{\text{CBM}}^{-1/2} \approx 2^\circ$, then these two regions are not in causal contact at that redshift.

³<http://www.sdss3.org/cosmology.php>

⁴<http://pan-starrs.ifa.hawaii.edu/public/science-goals/galaxies-cosmology.html>

Parameter	Symbol	Numerical value
Hubble parameter	h	$0.706^{+0.017}_{-0.017}$
Curvature parameter	Ω_k	$-0.0018^{+0.0054}_{-0.0054}$
Dark energy density parameter	Ω_{vac}	$0.740^{+0.013}_{-0.013}$
Matter density parameter	Ω_{mat}	$0.260^{+0.013}_{-0.013}$
Dark energy EOS	Ω_{vac}	$-0.965^{+0.078}_{-0.081}$
Mass Dispersion	σ_8	$0.775^{+0.034}_{-0.033}$
Spectral index	n_s	$0.962^{+0.012}_{-0.012}$
CMB temperature	T_{cmb}/K	2.728
Age of the Universe	t_0/Gyr	$13.78^{+0.26}_{-0.26}$

Table 2.1: Constraints on cosmological parameters derived from the joint data analysis from the Cosmic background radiation (measured by *WMAP*), large-scale spatial distribution of luminous red galaxies from SDSS and the Hubble diagram inferred from measurements of Type Ia supernova. Taken from Sánchez et al. (2008).

epoch, no physical theory has provided a reliable model for dark energy and its properties. Theoretical models to explain the cosmic acceleration have been widely discussed in the literature (e.g. Copeland et al., 2006). Among others, quintessence models (e.g. Wetterich, 2001; Wang, 2000; Peebles and Ratra, 2003), Chaplygin gas model (Kamenshchik et al., 2001), decaying vacuum energy density (e.g. Borges and Carneiro, 2005). Attempts to avoid the dark energy fluid and explain the cosmic acceleration as a result of General Relativity have also been proposed as $f(R)$ gravity (Carroll et al., 2004) and the concept of back-reaction (e.g. Buchert, 2001). The cosmological constant (with $w = -1$) is the immediate candidate for this dark energy fluid, being compatible not only with the SNe Ia observations, but also with the temperature fluctuation analysis in the CMB and the large-scale structure analysis of galaxy clustering (see table 2.1). Figure 2.1 shows some fundamental cosmological functions as a function of the redshift and how they change with the cosmological parameters.

Currently the consensus in modern cosmology is to consider Λ CDM as the standard cosmological model, being this description of the Universe most compatible with the astronomical evidences. Within this model the presence of dark energy has had to be accepted to explain the cosmic acceleration and account for $\sim 75\%$ of the total matter-energy in the Universe, together with a cold dark matter and baryonic component, responsible for structure formation and accounting for $\sim 25\%$ of the total budget. The constraints on cosmological parameters have achieved a high precision in the last decade due to the advent of large galaxy redshift surveys as 2dFGRS, SDSS, the increasing resolution in CMB probes as WMAP and the increasing number of SNe Ia detections based on deep observations. Future galaxy surveys such as BOSS⁵, PanSTARRS⁶ and Hetdex promise to achieve constraints on cosmological parameters to the 1% level of accuracy. In table 2.1 we show some of the constrains on cosmological parameters based on SNe Ia data together with the CMB angular power spectrum and the measurement of the two-point correlation function from the Luminous Red Sample (LRG) of galaxies from the SDSS database.

⁵<http://www.sdss3.org/cosmology.php>

⁶<http://pan-starrs.ifa.hawaii.edu/public/science-goals/galaxies-cosmology.html>

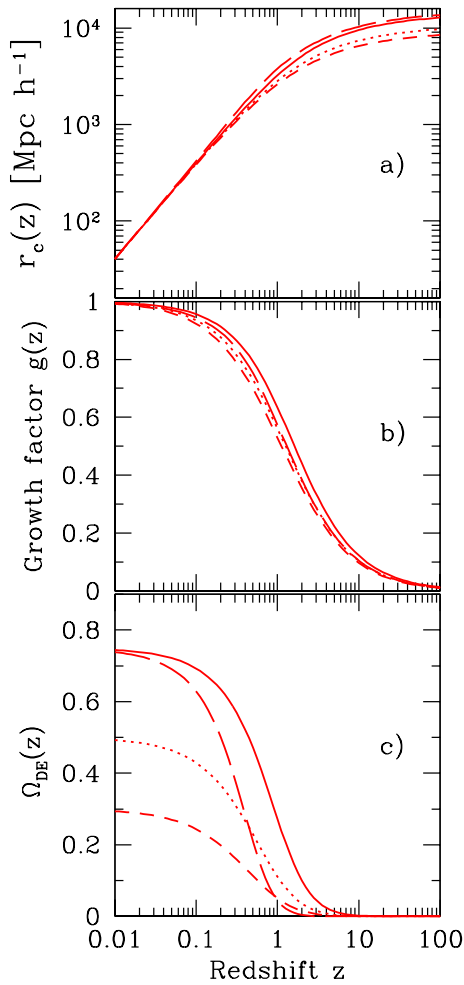


Figure 2.1: Cosmological functions for different values of the matter content Ω_m , the dark energy content Ω_{DE} and the dark energy equation of state w : $(\Omega_m, \Omega_{\text{DE}}, w) = (0.25, 0.75, -1)$ (solid line), $(0.5, 0.5, -1)$ (dotted line), $(0.7, 0.3, -1)$ (dashed line) and $(0.25, 0.75, -2)$ (long-dashed line). Panel a) shows the comoving distance $r_c(z)$, which only is affected by the different cosmological parameter at large redshifts, reducing the the Hubble low at low values of z . Panel b) shows the growth factor, the growing solution of the evolution equation for cold-dark matter perturbations. Panel c) shows the redshift dependence of the energy content provided by the dark energy fluid.

2.3 Statistical properties of large-scale structure

The basic model to describe the statistics of the large-scale structure of the universe is founded on the assumption that the distribution of matter and galaxies can be associated with a random realization of a Poisson process. The main features of such realizations, like the fluctuations about the mean density,

$$\delta(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t) - \langle \rho(\mathbf{r}, t) \rangle}{\langle \rho(\mathbf{r}, t) \rangle}, \quad (2.5)$$

depend on the initial conditions and the way that those initial fluctuations evolve with time. The LSS analysis is developed in configuration space or in Fourier space. In the latter, the statistical analysis is based on a plane-wave expansion of the field $\delta(\mathbf{r}, t)$, leading to a fair description once the volume from which the perturbations are measured is large enough to be a fair sample of the Universe. In this limit, an infinite universe can be re-constructed by replicating a fair volume with periodic boundary conditions. However, real galaxy and cluster surveys are not regions satisfying boundary conditions, but portions of the sky (or the full sky) covering a certain range of redshift. We can therefore use another basis for our spectral decomposition that is better adjusted to the geometry of the survey. This has been done in galaxy redshift surveys using the Fourier-Bessel expansion (e.g Fisher et al., 1996; Heavens and Taylor, 1994, 1997; Tadros et al., 1999; Percival et al., 2004a; Zhan et al., 2006). Also,

the Laguerre polynomials and logarithmic spherical waves (Hamilton and Culhane, 1996) have been implemented in the spectral decomposition, motivated by physical grounds (e.g. redshift distortions) or intrinsic features of the galaxy survey.

Gaussian initial conditions have been assumed for a long time in the literature, based mainly on the central limit theorem for random fluctuations (e.g. Reif, 1965). In the linear regime of perturbations ($\delta \ll 1$) Fourier modes evolve independently and Gaussianity persists along the time evolution and therefore is enough to characterize the statistical properties of LSS with its two-point statistics. Nevertheless, as long as the perturbations evolve in time, they become non-linear at some stage and thus generate non-Gaussianities in the matter and galaxy distribution. Although a complete description of the probability distribution for non-Gaussian fluctuations has not been yet determined from first principles, numerical simulations have shown that the probability distribution function for dark matter perturbations in the non-linear regime can be fairly well modeled by a log-normal distribution (Coles and Jones, 1991; Kayo et al., 2001).

A statistical description of the large-scale structure of the Universe is based on the measurement of fluctuations δ above the mean overdensities averaged over a large ensemble of realizations, as defined in equation (2.5). Using N -body simulations we have the opportunity to create such a large set (typically $\approx 10^2$) of realizations with the same initial conditions and thus we can properly talk about ensemble averages. However, in real life, as only one realization (the observed Universe) is available, the ensemble average must be modified by a volume average such that both concepts would lead to the same unbiased statistics if the volume sampled by a galaxy or galaxy cluster survey is large enough to be treated as a fair sample of the whole Universe, following the notion of ergodicity.

In this thesis we will only concentrate on the two-point statistics of the inhomogeneous distribution of galaxy clusters. Higher statistics (e.g. Matarrese et al., 1997; Verde and Heavens, 2001; Smith, 2008) are beyond of the scope of this thesis. These high-order statistics are currently measured from N -body simulations and large galaxy redshift surveys and provide relevant information concerning non-Gaussianities and the distribution of galaxies and clusters with respect to the underlying matter distribution.

2.3.1 Two point statistics

Let us briefly introduce the most relevant statistical tools for the large scale structure analysis. In a Poisson random process with no correlations the probability of finding an object at the volume element δV is just the product of the number density of those objects in the volume times the volume they occupy δV , $\delta P = \langle \rho \rangle \delta V$. Hence, as the events in a Poisson distributions are independent of each other, the joint probability of finding a galaxy at \mathbf{r}_1 and another galaxy at \mathbf{r}_2 is just the product of individual probabilities: $\delta P(1, 2) = \delta P_1 \delta P_2 = \langle \rho \rangle^2 \delta V_1 \delta V_2$. In the presence a physical interaction, correlations may appear at different scales. For a two-point process, the two-point correlation function will $\xi(\mathbf{r})$ is defined as

$$\xi(\mathbf{r}_1 - \mathbf{r}_2) \equiv \langle \delta(\mathbf{r}_1) \delta(\mathbf{r}_2) \rangle. \quad (2.6)$$

This function represents a measure of the excess of probability for having a pair of objects occupying volumes $\delta V_{1,2}$ separated by a distance $|\mathbf{r}_1 - \mathbf{r}_2|$ under the effects of gravitation (e.g. Peebles, 1973; Hamilton, 1993) with respect to the probability associated with pairs in a random distribution separated by the same distance:

$$dP_{12} = \langle \rho \rangle^2 (1 + \xi(\mathbf{r}_1 - \mathbf{r}_2)) \delta V_1 \delta V_2. \quad (2.7)$$

The power spectrum of fluctuations $P(k)$ is defined as the Fourier transform of the two-point correlation function. It can be shown (e.g. Padmanabhan, 1993b) that the power spectrum corresponds to the dispersion of the probability distribution function associated with fluctuations in the density field in Fourier space:

$$\langle \delta(\mathbf{k})\delta^*(\mathbf{k}') \rangle = (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') P(k). \quad (2.8)$$

The power spectrum only depends on the magnitude of the wave vector, following the assumption of statistical homogeneity on large-scales. This allows us to write the Fourier transform of the correlation function and its inverse relation as a zeroth-order Hankel transform (Hamilton, 2000):

$$P(k) = 4\pi \int_0^\infty \xi(r)r^2 j_0(kr)dr, \quad \xi(r) = \int_0^\infty d \ln k \Delta^2(k) j_0(kr), \quad (2.9)$$

where $j_0(x) = \text{sinc}(x)$ is the zeroth-order spherical Bessel function. The quantity $\Delta^2(k) \equiv P(k)k^3/2\pi^2$ is a dimensionless power spectrum, corresponding to the contribution of power in logarithmic bin in wavenumber to the correlation function.

A key element of the statistical description of large-scale structures is the concept of filtering. Instead of describing the evolution of matter perturbations at a certain position, we would rather describe the statistical properties behavior of the matter fluctuations filtered (or smoothed) with a window function $W(\mathbf{r}; R)$ on a certain scale R . The smoothed overdensity is a convolution:

$$\delta(\mathbf{x}; R) = \int \delta(\mathbf{y}) W(\mathbf{x} - \mathbf{y}; R) d^3y = \int \frac{d^3k}{(2\pi)^3} \delta(\mathbf{k}) W(k; R) e^{-ik \cdot \mathbf{x}}, \quad (2.10)$$

where $W(k, R)$ is the Fourier transform of the filter function. The root mean square of the matter fluctuation can be written as

$$\sigma^2(R) = \langle \delta(\mathbf{x}, R)^2 \rangle = \int_{-\infty}^{\infty} d \ln k \Delta^2(k) W^2(k; R). \quad (2.11)$$

The selection of the smoothing function is in principle arbitrary. Some functions are selected because they yield simple analytical expressions and possess a well-defined volume (Bond et al., 1991). The most common filters are the top-hat $W(x, R) = \theta(x - R)$ (defining a volume $V = \int d^3x W(x, R) = (4/3)\pi R^3$) and a Gaussian filter which also yields a well-defined volume. Sometimes a sharp k -space filter $W(k, R) = \theta(k - k_s)$ where k_s is a cut-off wavenumber is also considered (Lacey and Cole, 1993), leading to a volume which is not well defined. Despite these irregularities and arbitrariness in the selection of the sharp-filter function in Fourier space, the relation $M = (4/3)\pi \bar{\rho} R^3$ is often used when transforming from the filter scale R to masses. The comoving scale $R \approx 8 \text{Mpc } h^{-1}$ at which $\sigma^2 \approx 1$ can be associated with a structure with a mean density given by the background density at the present time. The corresponding mass of such object is $M \approx 6 \times 10^{14} M_\odot$, which corresponds to the typical mass of a galaxy cluster.

In this thesis we will mainly concentrate on the Fourier analysis of the inhomogeneous distribution of galaxy clusters and few attempts will be made at an analysis in configuration space with the correlation function (e.g. Peebles, 1980; Hamilton, 1992; Landy and Szalay, 1993). The analysis of the correlation function has been developed in parallel though, and we will briefly mention some results but no detail on the estimators or theoretical modeling will be presented.

2.3.2 Growth of inhomogeneities

The current Λ CDM paradigm establishes that the main ingredient for structure formation is a pressurless cold dark matter (CDM) composed of particles which decoupled from radiation at redshifts when these particles were non-relativistic (White et al., 1984). Despite being a pressurless fluid, it can be shown directly from the coupled set of Boltzmann-Poisson equations that on large-scales such a fluid is unstable under gravitational perturbations (Binney and Tremaine, 1987). The evolution equation for CDM perturbations in the linear regime can be written as

$$\ddot{\delta}_{\text{cdm}}(t) + 2H(t)\dot{\delta}_{\text{cdm}}(t) = \frac{3}{2}H(t)^2\Omega_{\text{cdm}}(t)\delta_{\text{cdm}}(t). \quad (2.12)$$

where $\dot{\delta} = d\delta/dt$. The solution of these differential equations at different epochs of the universe will help us to understand the patterns in large-scale structure we observe today. As an example, during the radiation dominated epoch, CDM perturbations inside the horizon evolved according to the Meszaros effect (e.g. Peacock, 1999), represented by an evolution with constant amplitude followed by a slow logarithmic growth arising when the redshift approaches the decoupling (matter-radiation) time. This suppression of growth is caused by the fast acceleration of the universe during these epochs. During the matter dominated epoch the CDM fluctuations evolve with a growing solution given by Heath (1977) and generalized to a cosmological model with dark energy (e.g. Carroll et al., 1992; Percival, 2005; Wang, 2000).

During the radiation dominated epoch, adiabatic perturbations in the baryonic fluid follows the CDM solution and hence no relevant growth of baryonic structures occurs. Therefore the perturbations in the baryonic component (or in the photon density) are a good tracers of the anisotropies in the temperature field during the early ages of the universe, which are however impossible to observe. As the Universe passes from radiation domination to matter domination, CDM fluctuations are allowed to collapse; diffusion of dark matter particles from forming overdensities take place and suppresses the growth of structures on very small scales (e.g. Padmanabhan, 1993b) though. Meanwhile baryonic perturbations in their attempt to form bound structures, are forced, following the Jeans criteria, to execute acoustic oscillations due to the remaining pressure exerted by the photon background (Hu and Sugiyama, 1996). Baryons and photons are then maintained in equilibrium mainly by Thomson scattering until the recombination redshift ($z_* \approx 1500$), when the temperature of the baryon-photon fluid drops enough to let baryons form hydrogen atoms. Directly after the redshift of decoupling, baryons can still be dragged by the escaping photons, preventing the collapse of baryonic structures (Silk, 1968) on scales $\lambda_s \approx 8.38$ Mpc (Eisenstein and Hu, 1998). This process ends at the so-called drag redshift z_d , when neither photon pressure or the baryon-drag can no longer prevent the collapse: the maximum comoving distance traveled by the baryonic perturbation $r_s(z)$ (the so-called sound horizon) can be written:

$$r_s(z) = \frac{c}{H_0} \int_0^{\frac{1}{1+z}} \frac{da}{a^2 H(a) \sqrt{1 + \mathcal{R}a}}, \quad (2.13)$$

where $\mathcal{R} \equiv 3\Omega_m/4\Omega_r$. Following the latest results from the *WMAP* mission, the sound horizon at recombination gives $r_s = 146.8 \pm 1.8$ Mpc. The acoustic oscillations are thus present in the baryonic power spectrum as series of damped oscillations (e.g. Percival et al., 2007a). In the correlation function, this is translated into an excess of galaxy pairs at the scale of the sound horizon (Eisenstein et al., 2005, e.g.). Similarly, the angular power spectrum of

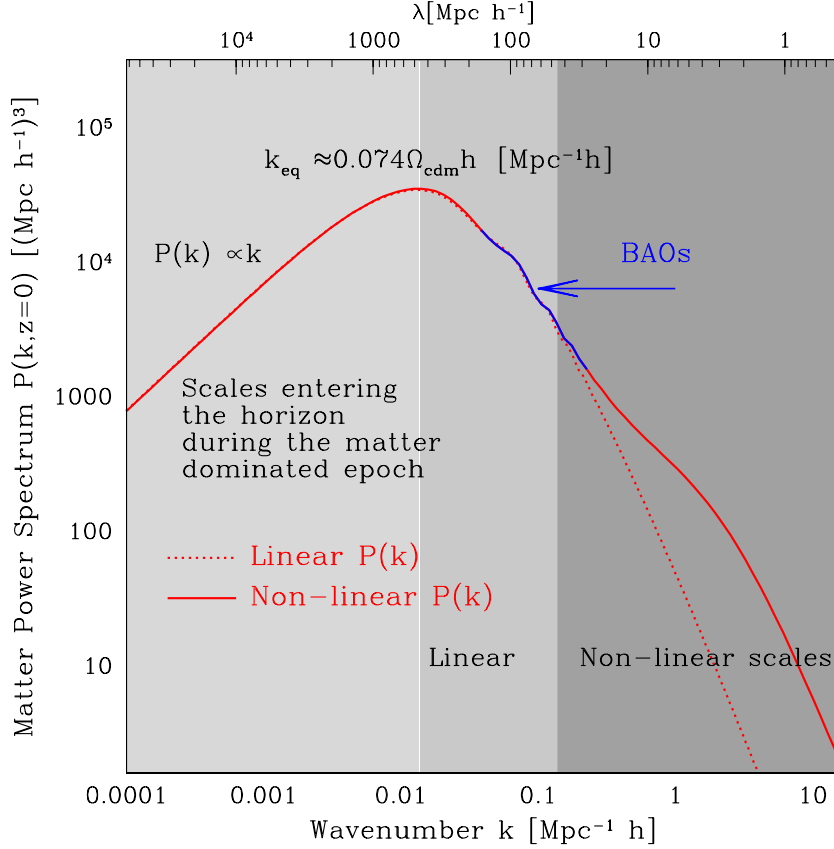


Figure 2.2: Linear matter power spectrum for $z = 0$, with $\Omega_{\text{cdm}} = 0.26$ and $h = 0.7$, and $k_{\text{eq}} = 0.0154$ normalized with $\sigma_8 = 0.9$. On large scales the power is increasing for those perturbations that enters the horizon during the matter dominated epoch, while on small scales we have the scales that enter the horizon during the radiation dominated epoch. In this part the power spectrum is a decreasing function of k because matter fluctuations that enter the horizon before matter-radiation equality are suppressed due to the Meszaros effect and the baryonic component is forced to oscillate displaying the acoustic peaks, with a decreasing amplitude due to Silk damping. Deviations from the linear power spectrum appear on small scales due to mode coupling of the Fourier modes, generating an increase in the power of small scale fluctuations.

temperature fluctuations will contain this acoustic signature; nevertheless, the acoustic scale is slightly different, since the temperature fluctuation will account for the sound horizon at the decoupling time, while the distribution of baryons (i.e., galaxies) encodes the sound-horizon at the drag redshift.

After decoupling, photons can freely propagate and baryonic perturbations are allowed to collapse and fall within the dark matter potential wells to create the first stars, globular clusters and galaxies, establishing the bottom-up hierarchical scenario of structure formation. This mechanism then guaranties that the tiny baryonic fluctuations (proportional to $\delta T/T$ at $z \leq z_{\text{dec}}$) can grow to generate the large fluctuations that we observe today. This also predicts the presence of BAO signatures in the CDM power spectrum, though with smaller amplitude.

The wavelength-dependent evolution of the matter perturbations are determined from the so-called transfer function $T(k)$. The linear matter power spectrum can be obtained from the transfer function as $P(k, z) = A k^{n_s} T(k)^2 g^2(z)$, where n_s is the spectral index ruling the behavior of

the primordial power spectrum Ak^n and $g(z)$ represents the growing solution $D_+(z)$ of the matter perturbation equation, normalized to be $g(z=0) = 1$ (see Fig. 2.1). The transfer function, defined as $T(k, z) \equiv (\delta(\mathbf{k}, z)D_+(z_i))/(\delta(\mathbf{k}, z_i)D_+(z))$, describes the transition of matter perturbations from super-horizon evolution (in the radiation dominated phase) to sub-horizon evolution. It basically depends on the solution of Boltzmann equation (Huang, 1987; Dodelson, 2004) for the linear evolution of the phase space density of each type of matter energy in the Universe and the way it is coupled with the rest of the forms of matter-energy. Different shapes of the transfer function can be fitted for the different components of the total energy density of the universe based on the results from Boltzmann codes like CMBFAST (Seljak and Zaldarriaga, 1996) and CAMB (Lewis et al., 2000). Parameterizations from Bardeen et al. (1986) and the fitting formulae of Eisenstein and Hu (1998) are used in the literature, though these have approximately $\leq 10\%$ deviations with respect to the exact solutions (e.g. Sánchez et al., 2008).

The shape of the linear matter power spectrum is determined by the way perturbations in the different fluid components evolved during the different epochs in the evolution of the Universe and how they compared to the horizon scale. On sub-horizon scales, a good understanding of the physics of gravitational instabilities can be explored from a Newtonian analysis of perturbations under the assumption that the processes leading to the formation of self-gravitational bound structures does not imply strong gravitational fields or intense velocity fields (e.g. Padmanabhan, 1993b).

The determination of the amplitude of the matter power spectrum is usually translated to the determination of the mass dispersion measured on a sphere of radius $8\text{Mpc } h^{-1}$, following Equation (2.11). The clustering analysis focused on the amplitude of the power spectrum contains degeneracies between cosmological parameters, especially on σ_8 and Ω_{mat} . This is due to the fact that these parameters influence the strength of the matter fluctuations in the same way. Some novel statistics and new probes like galaxy clusters have been introduced to break these degeneracies (e.g. Schuecker et al., 2003a; White and Padmanabhan, 2009; Xu et al., 2010). The degeneracy between these parameters is found as well in the study of cluster abundance (e.g. Reiprich and Böhringer, 2002b).

2.4 Distortions in the linear clustering pattern

The observed clustering pattern of galaxies and galaxy clusters display noticeable differences with respect to the linear theory matter power spectrum. In this section we describe some of the most relevant sources of distortions.

2.4.1 Galaxy bias

Light does not fairly trace the matter distribution in the Universe. The phenomena of galaxies and galaxy cluster as biased tracers has been widely recognized and yet a complete modeling is hard to be obtained from the theoretical point of view. This is mainly due to the highly non-linear processes taking place in the assembling and clustering of dark matter haloes and galaxies.

Galaxy bias can be recognized as an increment of the clustering strength of a galaxy population with respect to the clustering of the underlying dark matter distribution. The simplest scenario for galaxy bias consists of a deterministic scale independent bias $\delta_{\text{gal}}(z) = b(z)\delta_{\text{mat}}(z)$, which assumes that galaxies are fair tracers of the peculiar velocity field on large scales. Also,

the constant galaxy clustering hypothesis (Lahav, 2002) with $b(z) = g^{-1}(z)$ has been claimed to explain the observed clustering of galaxies. Similar models for a redshift dependent linear deterministic bias have been modeled by Tegmark and Peebles (1998) and Fry (1996). However galaxy bias has been observed to possess a non-linear behavior together with a stochastic component arising as a result of the different ways a galaxy can populate a dark matter halo as a function of the properties of galaxies and of haloes as well. Moreover, the halo-matter bias contributes with its own scatter, coming from the differences in halo masses, merger histories and formation redshifts, which makes galaxy bias a highly complicated problem to resolve. Clearly the best way is to first explore how dark matter haloes are biased with respect to the underlying matter distribution and then explore how galaxies are biased with respect to their hosting dark matter haloes (Bardeen et al., 1986; Kaiser, 1984; Frieman and Gaztanaga, 1994; Fry and Gaztanaga, 1993). The clustering of galaxies (and therefore galaxy bias) has been observed to be a function of intrinsic galaxy properties (e.g. Dressler, 1980), like their density or luminosity (e.g. Norberg et al., 2001).

2.4.2 Redshift space distortions

The distance to a galaxy is determined from its redshift through a cosmological model, following equation (2.3). Therefore the determination of separations can be affected by a mixture of physical effects, e.g. gravitational redshift, Doppler redshift, and observational effects such as photometric redshift errors. In a galaxy redshift survey, distortions on large scales are associated with the motion of the hosting dark matter halo towards matter overdensities, while on small scales, the distortions are mainly due to virialized motion of galaxies in dark matter haloes (the so called “finger-of-god” effect). Thus the observed distance to a galaxy with measured redshift z is the contribution of the distance corresponding to its cosmological redshift z_0 , plus the deviation along the line of sight due to its peculiar velocity in that component $u(r)$:

$$s = r(z_0) + \frac{u_r}{H(z_0)},$$

Following Dodelson (2004), in order to measure the peculiar velocity of a galaxy, we must have an estimate of its comoving distance r , which must be given by some redshift independent measurement (such like the determination of light-curves of Cepheid stars or Supernovae). Typical velocities of galaxies during their coherent in-fall towards clusters are of the order of ~ 500 km/s, leading to $\Delta r = v/H(z) \sim 5E(z)^{-1} \text{ Mpc } h^{-1}$. This also shows that the peculiar velocity field induces distortions in the distances of $\Delta r \propto v/H(z)$. In order to measure the peculiar motion of a galaxy, an experiment determining its comoving distance must generate results with a precision such that it is possible to distinguish the motion due to overdensities from the motion due to the Hubble expansion. The limiting precision the distance of a galaxy at some distance r must have in order to make such differentiation is $\sigma_r \sim 100(\Delta r/r)\% = (500/r)E(z)^{-1}\%$, such that the more distant the galaxy is, the higher the relative precision the experiment must achieve in order to decouple the two effects. A galaxy at redshift $z = 0.05$ implies $\sigma_r \approx 3\%$, rising up to $\sim 16\%$ for a galaxy at $z = 0.01$. In other words, given some precision from a redshift independent experiment σ , only galaxies located at distances $r \leq huE(z)^{-1}/\sigma \text{ Mpc } h^{-1}$ can be used to measure peculiar velocities.

When no measurement of peculiar velocities are available, the statistical analysis must be developed directly in redshift space. Transforming the basic quantities from real to redshift space requires some approximations. Assuming galaxy number conservation to leading order

in the peculiar velocity $u(r)$, it can be shown that, on large scales, the matter fluctuation in redshift space can be written to leading order in the peculiar velocity as (Kaiser, 1987; Hamilton, 1992):

$$\delta[\mathbf{s}(\mathbf{r})] \approx \delta(\mathbf{r}) - \frac{\partial u(r)}{\partial r}.$$

When transformed to Fourier space, the redshift space power spectrum can be written as (e.g. Kaiser, 1987; Heavens et al., 1998)

$$P^s(\mathbf{k}) = P^s(k, \mu_k) = \left(1 + \beta\mu(k)^2\right)^2 P^r(k), \quad (2.14)$$

where $\mu_k \equiv k_{\parallel}/k = \hat{\mathbf{r}} \cdot \hat{\mathbf{k}}$ is the cosine of the angle between the wave vector and the line-of-sight direction, k_{\parallel} is the projection of the wave vector in the line-of-sight direction ($k = (k_{\parallel}^2 + k_{\perp}^2)^{1/2}$) and $\beta \equiv f(z)/b(z)$, with $f(z) \equiv d \ln \delta / d \ln a$ as the growth index. Hence, in redshift space, the three dimensional power spectrum acquires an enhancement with respect to the power in real space, which depends on the direction of the wavenumber (note that the real power spectrum $P(k)$ is isotropic). For parallel directions and $\beta = 1$, one has $P^s(k) = 4P^r(k)$. Note also that $P^s(k_{\perp} = k, k_{\parallel} = 0, z) = P(k, z)$ i.e, the real power spectrum can be obtained by measuring the redshift power spectrum in the transverse direction, as can be seen in the Kaiser effect equation (2.14) by setting $\mu_k = 0$ (Hamilton and Tegmark, 2000; Tegmark et al., 2004). Nevertheless, it is not clear how determine, in Fourier space, what the tangential or the parallel direction with respect to the line-of- sight are.

It is interesting to note that the source of redshift distortions, mainly the peculiar velocity field, satisfies the linearized equation $\nabla \cdot \mathbf{u} \propto -\dot{a}\delta$. Only in the context of General Relativity, the proportionality constant is $\beta = f/b$ (e.g. Hamilton, 1992). For a matter dominated universe, it has been written as $f(z) = \Omega_{\text{cdm}}(z)^{0.6}$ Peebles (1980), fitted for a Λ CDM universe by Carroll et al. (1992) and extended by Wang (2000) for constant dark energy equation of state:

$$f(a) = \Omega_{\text{mat}}(a)^{\gamma(z)} \quad \gamma(a) = \frac{3(1 - \omega_x)}{5 - 6\omega_x} + \frac{3}{2} \frac{(2 - 3w_x)(1 - w_x)}{(5 - 6w_x)^3} (1 - \Omega_{\text{mat}}(a)), \quad (2.15)$$

with a 1% error with respect to the exact solution. This leads to a growth index for the Λ CDM cosmological model of $\gamma(z) = 0.55 + 0.011\Omega_x(z)$, leading a value $\gamma = 0.5580$ for $w_x = -1$. Given the current estimates of the dark energy equation of state, the variation of the exponent γ under variations of its value 0.55 is just $\delta\gamma = 0.154\delta w_x$. A 5% result in the measurement of w_x leads to a deviation of 10^{-4} in the value $\gamma = 0.55$. Therefore a deviation from the value 0.55 in a measurement of the exponent γ constitutes a smoking gun for alternative theories of gravity (e.g. Linder, 2008).

The small angle approximation provides a simple frame to treat redshift distortions. For wide angle surveys, this approximations does not hold anymore and a complete treatment must be carried out as developed by Szalay et al. (1998). Also, redshift distortions can affect the clustering pattern encoded in the angular correlation function, as explored by Nock et al. (2010).

Multipole expansion

In general one can write the galaxy power spectrum $P_{gg}^{s,r}(k, \mu_k)$ as a linear combination of Legendre Polynomials $\mathcal{P}_l(\mu)$ by developing a power expansion with respect to μ_k of the form $P^s(\mathbf{k}) = \sum_{\ell} f_{\ell}(\beta) P_{gg}^r(k) \mathcal{P}_{\ell}(\mu_k)$, where only the first three even indices contribute: the monopole

($\ell = 0$), quadrupole ($\ell = 2$) and a hexadecapole contribution ($\ell = 4$). The coefficients are given by

$$f_0(\beta) = 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2, \quad f_2(\beta) = \frac{4}{3}\beta + \frac{4}{7}\beta^2, \quad f_4(\beta) = \frac{8}{35}\beta^2,$$

The azimuthally-averaged power spectrum in redshift space is given by the monopole contribution $P^s(k) = P(k)f_0(\beta)$ which is only a function of β . Hence on large scales the increase in the power spectrum is scale-independent, providing a way to measure the parameter β (Hamilton, 1998). Note that when one wants achieve such measurement from the monopole contribution to the redshift power spectrum $P^s(k)$, one has to have a previous knowledge of the underlying real space power spectrum. Instead, what was usually done was to use the information contained in all the multipoles by measuring the three dimensional power spectrum $P^s(\mathbf{k})$ and use for instance the quadrupole-to-monopole ratio $Q = P_2^s(k)/P_0^s(k)$, which only depends on β . Translating the multipole expansion to configuration space is almost straightforward (e.g Hawkins et al., 2003).

Small-scale distortions

On small scales, the main distortions in redshift space are due to the motion of galaxies in virialized dark matter haloes. These distortions are modeled via the streaming model where the correlation function in real space is convolved with the pairwise velocity distribution function of galaxies (Peebles, 1980). Numerical simulations have shown how a Gaussian distribution or a Lorentzian distribution for the pair-wise velocity is not a good description of galaxies in dark matter haloes, though these are the two typical models that lead to analytical expressions (e.g. Scoccimarro, 2004). The first of these approximations describes galaxies with velocities drawn from a Maxwellian distribution function, such that the best one-dimensional pairwise velocity distribution function is an exponential function (Landy, 2002; Jing et al., 2002; Peebles, 1980). Measurements of the two-dimensional correlation function can thus give constraints on the parameter β and estimates of pair-wise velocity dispersions (e.g Guzzo et al., 2008; Hawkins et al., 2003). When translated to Fourier space, the streaming model generates a three dimensional power spectrum written as

$$P^s(\mathbf{k}, z) = P^s(k, \mu_k, z) = P(k, z) \left(1 + \mu_k^2 \beta(z)\right)^2 D(\mu_k, k, z), \quad (2.16)$$

where $k^2 = k_{\parallel}^2 + k_{\perp}^2$ and the factor $D(\mu_k, k, z)$ is the Fourier transform of the velocity distribution function. Analytical results can be obtained for the function $D(\mu_k, k, z)$ with Gaussian and exponential distributions (e.g. Lahav and Suto, 2004).

2.4.3 Geometrical distortions

When measuring the n -point statistics from a galaxy redshift survey, the choice of a set of cosmological parameters can lead to a distortion in the amplitude and shape of the clustering signal when compared to the predictions based on a fiducial cosmological model. This is the so-called Alcock-Paczynski effect (Alcock and Paczynski, 1979), mainly arising due to the fact that the parallel and transversal separations with respect to the line of sight do not transform in the same way from one cosmological model to another. If we assume a cosmological model characterized by a Hubble parameter $H^{\text{ref}}(z)$ and another characterized by $H(z)$, then parallel and transverse separation in a FRW universe transform as $x_{\parallel} = f_{\parallel}^{-1}(z)x_{\parallel}^{\text{fid}}$ and $x_{\perp} = f_{\perp}^{-1}(z)x_{\perp}^{\text{fid}}$

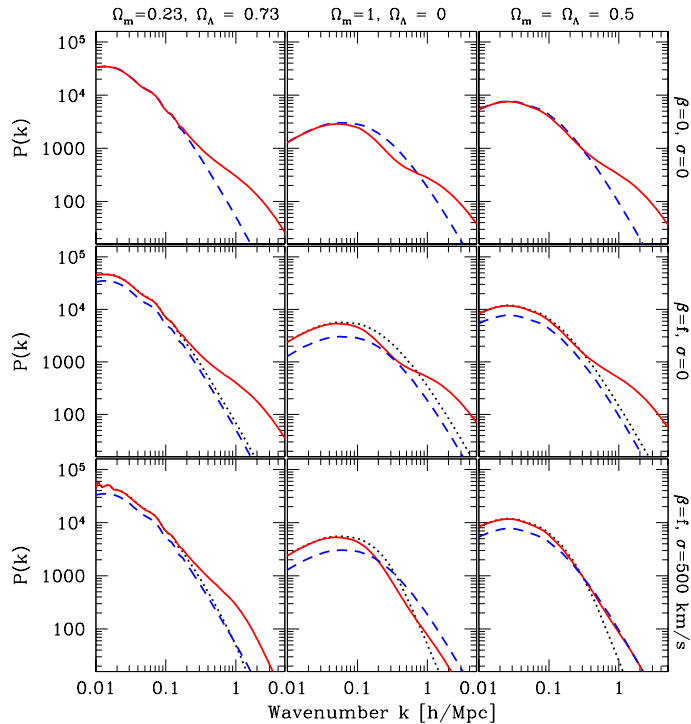


Figure 2.3: Behavior of the power spectrum for different cosmologies in real space (first row) and redshift space (second and third row). The first column represents the power spectrum with cosmological parameters close to the recent constraints (Sánchez et al., 2008). The distortions in redshift space have been assumed to be modeled with a Gaussian distribution.

where the distortion functions are given by:

$$f_{\perp}(z) = \frac{D_A^{\text{fid}}(z)}{D_A(z)}, \quad f_{\parallel}(z) = \frac{H(z)}{H^{\text{fid}}(z)},$$

and $D_A(z) = (1+z)^{-1}r(z)$ is the angular diameter distance. As a consequence, the Fourier modes and the amplitude of the power spectrum being determined from the measurements of distances are modified according to (Ballinger et al., 1996),

$$P(k, z) = \frac{1}{\alpha(z)^3} P^{\text{fid}}\left(\frac{k}{\alpha(z)}; z\right), \quad (2.17)$$

where the stretch factor $\alpha(z)$ is given by

$$\alpha(z) = [f_{\parallel}^2(z) f_{\perp}(z)]^{1/3} = \left[\frac{H(z)^2 D_A^{\text{fid}}(z)}{D_A(z) H^{\text{fid}}(z)^2} \right]^{1/3}. \quad (2.18)$$

In practice, the exponents in this expressions are only valid for a wide angle survey and not in pencil-beams observations. This has then to be taken into account when constraints on cosmological parameters are derived from the power spectrum, for it allows us to define a fiducial cosmology and through it generate a power spectrum with another set of parameters without having to determine measurements of two-point statistics of each new set of cosmological parameters. The Alcock-Paczynski effect has been used to constrain regions in the $\Omega_{\text{vac}} - \beta$ parameter space (e.g. Ballinger et al., 1996; Outram et al., 2004).

2.4.4 Non-linear evolution

On small scales, Fourier modes evolve in the highly non-linear regime at low redshifts. These non-linear effects lead to an increase in the power spectrum on small scales $k > 0.2 h \text{ Mpc}^{-1}$ with

respect to the amplitude expected from linear theory. Moreover, non-linearities can introduce shifts and distortions in the location and shape in the signal of BAO, which suggests that a good understanding of non-linear effects is imperative if the such feature is to be used as standard ruler in the LSS analysis.

A numerically-based attempt to model non-linear effects using two-point statistics is given by Hamilton et al. (1991) by means of the stable clustering hypotheses (Peebles, 1980). With the advent of large N -body simulations, it has been possible to describe the transition from linear to non-linear regimes by designing fitting algorithms. Among these procedures are the work developed by Peacock and Dodds (1994) and the Halo-Fit of Smith et al. (2003). Phenomenological parameterizations of the non-linear evolution of the galaxy power spectrum has been also constructed based on N -body simulations and tested against observations, such as the Q -model of Cole et al. (2005). In Chapter 5 we will implement this phenomenological model to describe the shape of the cluster power spectrum.

On the theoretical side, perturbation theory (e.g. Jain and Bertschinger, 1994; Scoccimarro and Frieman, 1996) provides a theoretical framework based on first principles (Euler equation and continuity equations) capable of accurately reproducing non-linear effects at intermediate redshifts $z > 1$ and has been successfully tested against numerical simulations (Jeong and Komatsu, 2009). The theory is based on a perturbative analysis of the complete set of conservation equations for matter, the matter distribution and its peculiar velocity fields. Under the assumption of a Λ CDM model of the Universe, the evolution equations for the matter perturbations allow for power-law solutions, leading to a power spectrum written as $P(k, z) = g^2(z)P(k, 0) + P_{22} + P_{13} \dots$, where the quantities P_{ij} correspond to terms of order $\mathcal{O}(\delta)^{i+j}$ in the density fluctuation, arising due to mode coupling. PT, being a perturbative approach, fails at low redshifts where the perturbative parameter, i.e. δ , goes above unity.

In the renormalized perturbation theory (Wyld, 1964; Crocce and Scoccimarro, 2006) a re-summation of the different non-linearities leads to an expansion whose perturbative parameter is not the perturbation amplitude. Instead, the re-summation is done such that when the expansion is truncated at a given scale, all non-linearities are already taken into account. This model has succeeded in reproducing the non-linear matter power spectrum measured from N -body simulations (e.g. Crocce and Scoccimarro, 2008; Smith et al., 2008), without the need for free parameters. The full-shape modeling of the measured correlation function from the LRG sample of the SDSS based on renormalized perturbation theory has led to tighter constraints on cosmological parameters (Sánchez et al., 2009). Non-linear effects must be taken into account in order to construct theoretical predictions capable of modeling the high-precision data expected from the next generation of cosmological probes. As will be discussed in Chapter 5, the power spectrum that we measured from a galaxy cluster catalogue will display non-linear evolution. Nevertheless, given the moderate volume probed by the survey we will explore (REFLEX, see Chapter 4), it will be sufficient to model non-linearities with a simple model such as the Q -model.

In Fig. 2.3 the matter power spectrum at redshift $z = 0$ is shown. The linear power spectrum is taken from the fitting formulae of (Eisenstein and Hu, 1998), while the non-linear power spectrum is calculated with the fitting formulae of Smith et al. (2003). In order to see how the shape of the power spectrum changes with the underlying cosmological parameters, we chose three sets: the concordance set which describes a flat universe with $\Omega_{\text{mat}} = 0.237$, $n = 0.98$, $\sigma_8 = 0.77$ and $h = 0.7$. The second model refers to of a matter dominated universe $\Omega_{\Lambda} = 0$ and the third contains $\Omega_{\text{mat}} = 0.5$. Only the matter content (and thus the dark energy content) is changed. In addition, we have introduced redshift distortions,

first through the Kaiser boost factor using $\beta = f(z; \Omega_i)$ and second introducing small scale distortions implementing a velocity dispersion of $\sigma = 500\text{km/s}$ with a Gaussian distribution function. By varying the cosmological parameters, the shape and the amplitude of the matter power spectrum are both modified, as was already discussed. Interestingly, note how in a matter dominated universe the linear power spectrum has higher amplitudes in a small range of scales than the non-linear contribution. Notice also that when translated to redshift space, the non-linearities can be suppressed by the "fingers-of-god" effect such that the resulting matter non-linear power spectrum in redshift space could naively be reproduced by linear theory. This is clearly seen in the last plot of Fig. 3.1. This situation can be also translated to the galaxy clustering, where, together with the aforementioned distortions, galaxy bias together with an implementation of a statistical model for the halo occupation distribution of galaxies can generate variations in the shape of the galaxy power spectrum in the non linear regimes.

2.4.5 Other source of distortions

The aforementioned source of distortions are based on physical effects (non-linearities, redshift distortions) or model-dependent induced distortions (geometrical). In addition, we can mention some other effects that are relevant in order to develop a detailed analysis of n -point statistics:

- Light cone effects appear in samples spanning a large redshift interval, such that the clustering properties cannot be referred to a single redshift but corresponds to a projection along the observer's past light cone of the clustering at the different redshifts where galaxies were observed (e.g. Nishioka and Yamamoto, 1999; Yamamoto and Nishioka, 1999; Suto et al., 2000).
- Photometric methods provide a large number of galaxies with measured redshifts, in comparison with spectroscopic methods. Nevertheless, photometric redshift errors are large enough to introduce strong distortions along the line of sight in the clustering pattern (just in the same direction as the FoG effect) with a dispersion $\sigma(z) \approx (1+z)\sigma_0$ (e.g. Phleps et al., 2006; Zhan et al., 2006). Typical redshift errors of $\sim 4\%$ can lead to smearing of structures almost two orders of magnitude larger than the FoG distortions. Narrow band photometry involving a large number of photometric filters are being implemented (e.g. Benitez et al., 2009) to derive precise estimates of redshifts.
- In flux-limited surveys, selection effects introduce systematic distortions in the two point statistics, mainly due to the fact that the most intrinsic luminous objects are the ones detected at the high redshifts and therefore these objects can probe the large-scale distribution of galaxies, while faint objects can only probe small scales. Given the fact that the clustering of galaxies depend on their luminosities, the presence of different luminosities in a sample translates to a tilt (toward negative slopes) in the power spectrum. This distortion pattern has been identified in galaxy surveys (Tegmark et al., 2004) and corrections can be made by measuring the luminosity bias e.g Norberg et al., 2001; Percival et al., 2004b and implementing two-point statistics estimators as a function of the luminosity. In this work, we will explore the relevance of this flux-selection induced distortions in a galaxy cluster catalogue.

2.5 Probes for cosmological parameters

2.5.1 CMB

The cosmic microwave background radiation (CMB) contains the information about the initial conditions from which structure formation took place. The photons detected by the CMB experiments such like *COBE*⁷ and *WMAP*⁸ were emitted at the redshift of matter-radiation decoupling ($z \approx 1090$) and not only they carry information concerning the seeds for structure formation, but also transport relevant information about the physical processes in galaxy clusters via the Sunyaev-Zeldovich effect⁹ (Sunyaev and Zeldovich, 1980) (hereafter SZ). CMB experiments have yielded the most precise example of a black body radiation distribution of intensity with a mean temperature of $T_{\text{CMB}} \approx 2.74\text{K}$. Despite its remarkable homogeneity on large scales, second order fluctuations in the temperature distribution have been precisely measured, showing structures with the largest amplitudes on scales of $\sim 1^\circ$. These are used as standard rules to constrain cosmological parameters via geometrical tests. The information contained in the CMB can thus provide tight constraints on parameters such as the matter content ($\Omega_{\text{mat}}, \Omega_{\text{bar}}$), the curvature Ω_{k} , the distance to the last scattering surface, the optical depth of the last scattering surface, among others (see Komatsu et al., 2010). CMB observations also suggest that on large scales the spectrum of temperature fluctuations is almost scale-invariant, in agreement with the Harrison-Zeldovich spectrum. Dark energy, playing no important role at the decoupling redshift, is poorly constrained by CMB data. The information encoded in angular power spectrum C_ℓ of the temperature anisotropies can be summarized in a set of distance priors determined by Komatsu et al. (2009). A first distance prior is given by the ‘‘acoustic scale’’ $\ell_A \equiv \pi D_A(z_*)/r_s(z_*)$, where $D_A(z)$ is the comoving diametral angular distance, $r_s(z)$ is the comoving sound horizon and z_* is the redshift at the matter-radiation decoupling. This quantity can be associated to the inverse of the typical angular scale of temperature fluctuations (Peebles, 1980). A second distance-prior is the so-called ‘‘shift parameter’’ introduced by Bond et al. (1997) as $R = \sqrt{\Omega_{\text{m}} H_0^2} D_A(z_*)/c$. Komatsu et al. (2009) provides measurements on these priors together with z_* and the respective covariance matrix to be used in combination with other experiments to obtain constraints on cosmological parameters. Note that these priors contain the information on cosmological parameters associated to the homogeneous universe, while not giving information concerning the spectral index and the amplitude of primordial fluctuations. A set of extended priors are provided by the WMAP analysis by introducing the energy content in baryons Ω_b .

The great agreement found between the Λ CDM model and the CMB observations represents the triumph of linear theory of gravitational collapse and the analysis carried out with these observational probe provides the initial conditions from where the inhomogeneous Universe evolved.

2.5.2 Galaxy clustering and BAO

The measurement of the inhomogeneous spatial distribution of galaxies from large galaxy redshift surveys is a very important probe for cosmological parameters. The success of linear theory confirmed by means of CMB measurements determined a scenario from which the

⁷<http://lambda.gsfc.nasa.gov/product/cobe/>

⁸<http://lambda.gsfc.nasa.gov/product/map/dr2/>

⁹This refers to shifts in the CMB spectrum due to the scattering of low CMB energy photons by high-energy electrons in galaxy clusters.

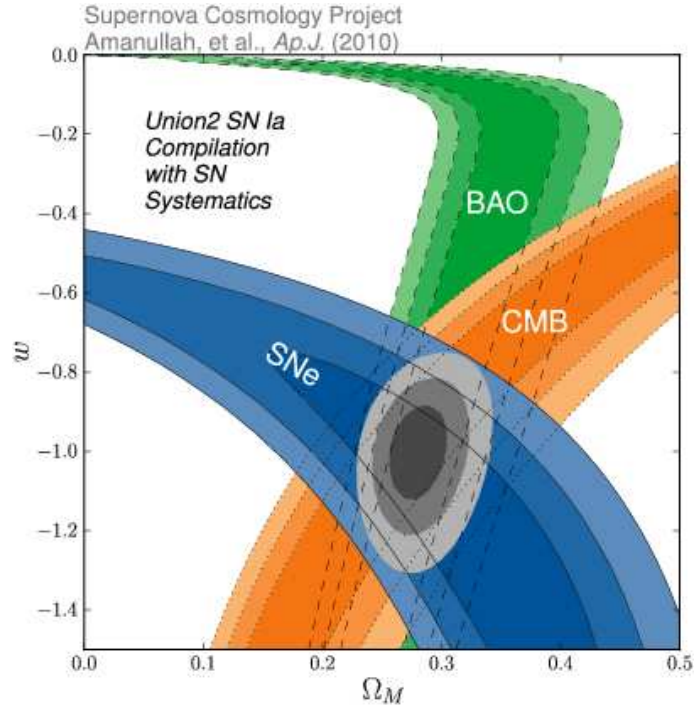


Figure 2.4: Confidence levels in the parameter space $\{w, \Omega_{\text{mat}}\}$ derived from SNe Ia, CMB and BAO. Contours correspond to the 68.3%, 95.4% and 99.7% confidence regions. The combined analysis generates tight constraints in the dark energy equation of state, compatible with a cosmological constant. Taken from Amanullah et al. (2010).

measurements of the matter and galaxy power spectrum are expected to develop in an almost scale independent fashion on large scales, followed by a decreasing power on small scales. Early measurements of the galaxy power spectrum utilized the shape and amplitude of the clustering signal to set constraints on cosmological parameter using the shape parameter, based on the detection of the turn-over $\{\Omega_{\text{mat}}, n_s\}$ (e.g. Lin et al., 1996; Tadros et al., 1998; Schuecker et al., 2001). Also, the amplitude of the power spectrum can set constraints on the parameter space $\{\Omega_{\text{mat}}, \sigma_8\}$, with a degeneracy that can be broken with the detection of the turn-over of the power spectrum. Nevertheless, non-linearities start to dominate the galaxy clustering signal and therefore usually only large scales (where the galaxy bias has been observed to be constant) are useful for unbiased constraints. These large scales are however strongly affected by cosmic variance, and therefore require either larger surveyed volumes are required or a better understanding of small scale fluctuations, or ideally both.

The detection of the BAO in the galaxy power spectrum (Percival et al., 2007a; Hütsi, 2007; Percival et al., 2010; Angulo et al., 2008) and in the correlation function (Eisenstein et al., 2005; Sánchez et al., 2008) has allowed the use such feature as a standard ruler from which cosmological information can be extracted via geometrical arguments. These acoustic signatures have only been detected in the last decade thanks to the advent of large galaxy redshift surveys such as SDSS and 2dFGRS. The use of BAO as a standard ruler implies a percent level of precision in the theoretical prediction of the position of the peaks. Some initially preferred analysis (Glazebrook and Blake 2005; Percival et al. 2007a, e.g.) isolated the acoustic wiggles by dividing the signal by a smooth power spectrum. This however lead to results which were inconsistent with SNe Ia data. Recent investigations (e.g. Sánchez et al., 2009) have shown that the modeling of the full shape of the correlation function (and thus the full shape of the power spectrum) lead to tighter and more consistent results. This is of course only applicable when a good model of non-linearities is available, which is the case of renormalized perturbation theory (Croce and Scoccimarro, 2006). The analysis of

the clustering of the luminous red galaxy sample from SDSS allows for the introduction of an additional distance prior to the set given by Komatsu et al. (2009). Sánchez et al. (2009) introduced the distance prior $G(z_m) = D_A(z_m)h(z)^{0.8}$, where $z_m = 0.35$ is the mean redshift of the galaxy sample (LRG) from SDSS implemented in the analysis of the galaxy correlation function.

2.5.3 Other probes

The current analysis of astronomical observations towards the determination of cosmological parameters is developed by combining the aforementioned probes with other highly relevant measurements, such as the peak luminosity of distant SNe Ia (e.g. Kowalski et al., 2008; Amanullah et al., 2010) and weak lensing analysis (e.g. Clowe et al., 2004). Weak lensing probes directly the distribution of matter in the Universe using the measurement of distortions in the shape of galaxies due to the foreground matter distribution, encoding the relevant cosmological information in two-point shear correlation function and power spectrum (e.g. P. G. Castro, 2005; Kitching et al., 2008; D. Munshi and van Waerbeke et al., 2008; Camera et al., 2010).

2.6 Galaxy clusters

Galaxy clusters are the largest gravitationally bound structures in the Universe. With typical masses of $10^{14}M_\odot$ ¹⁰ and radii of $\approx 2\text{Mpc}$, these objects are formed in the recent past of the cosmic history ($z \lesssim 2$)¹¹ due mainly to merging of smaller groups of galaxies, within what is known as the hierarchical structure formation scenario. Galaxy clusters are mostly composed ($\sim 87\%$ of its total mass) by a dark matter halo (i.e. non-baryonic matter) and therefore the formation and structure of galaxy clusters is closely linked to those of dark matter haloes. The size of the original region containing the matter from which the dark matter haloes are built (typical size of $\sim 10\text{Mpc}$) suggests that the content of dark matter and baryonic matter in galaxy clusters can be considered as fair representations of the content of these type of matter in the Universe (White and Frenk, 1991), providing a probe to determine cosmological parameters. Recent observations have shown however a mild gradient of the gas fraction (i.e., the ratio of mass in gas to the mass of the dark matter in the cluster) with the halo mass, which can be explained with physical inputs concerning the evolution of the baryonic component (e.g. Giodini et al., 2009).

The baryonic matter represents the second ingredient in the composition of galaxy clusters, accounting for $\sim 13\%$ of the total mass of the cluster. From this 13%, approximately 11% is embodied in the hot intra-cluster medium (ICM hereafter) with temperatures up to 10^8K . The remaining $\sim 2\%$ is represented in stars, molecular clouds, dust or galaxies in general (e.g. Allen et al., 2002; Pratt et al., 2009).

Within the structure formation paradigm, galaxy clusters of different sizes and masses are thought to be scaled versions of themselves, in what is called “self-similar” evolution. The abundance of galaxy clusters is directly related to the abundance of dark matter haloes. Therefore the galaxy cluster number-counts (e.g. Böhringer et al., 2004) and the galaxy cluster luminosity function can provide direct links between astronomical observations and the

¹⁰Up to date, the most massive galaxy cluster has been measured to be $(1.0^{+0.4}_{-0.6}) \times 10^{15}M_\odot$ with temperature $T = 7.5\text{keV}$ at a redshift $z = 1.061$ (reported by M. Brodwin at the *Galaxy Cluster conference* in Garching, 2010.)

¹¹Up to date, the most distant cluster has been found at $z \approx 1.94$ (JKCS041) (Andreon et al., 2009).

parameters of the underlying cosmological model (e.g. Böhringer et al., 2002; Vikhlinin et al., 2009b). The cluster abundance is especially sensitive to the matter content in the Universe Ω_{mat} and the r.m.s of mass fluctuations σ_8 e.g. Reiprich and Böhringer, 2002b; Schuecker et al., 2003b. Future galaxy cluster surveys will provide a large number of objects and larger (than the current ones) volumes will be covered. This will allow us to use the cluster abundance and the mass (luminosity) function of galaxy cluster as probes for dark energy models (e.g. Allen et al., 2004; Manera and Mota, 2006; Basilakos et al., 2010). Furthermore, measurements of the cluster luminosity function have not only cosmological implications, but can in principle also provide links to the properties of the intra-cluster medium (ICM hereafter), which are encoded in scaling relations (e.g. mass-X ray luminosity, mass-temperature) (e.g. Mantz et al., 2010; Vikhlinin et al., 2009a). Due to their large size, galaxy clusters have been proposed to be treated as fair samples of the baryon abundance in the Universe (e.g. White and Frenk, 1991; Allen et al., 2004). Nevertheless it has been shown that the observed gas fraction increases with the cluster mass (e.g. Pratt et al., 2009), and numerical simulations have shown that non-thermal physics in the inner regions of the hot-gas in clusters can lead to deviations from self-similar and/or introduce deviations from simple predictions of self-similar scaling relations.

Galaxy clusters are also considered as standard candles to determine cosmological distances by means of the SZ effect (Sunyaev and Zeldovich, 1980), mainly due to the fact that the spectrum of the SZ effect has universal and unique signature that do not depend on redshift of the cluster. Similarly, detailed observations of X-ray emitting cluster mass fraction can be used to set constraints on cosmological parameters due to the fact the inferred cluster radius and the gas density profile depends on the comoving diametral distance $D_A(z)$ to the redshift of the cluster z (e.g. Ettori et al., 2009).

Masses

A key quantity to study the abundance and clustering of galaxy clusters is their total mass, since it allows us to connect observations with theoretical models, as will be describe in Chapter 3. Its determination is currently achieved by three methods. These are 1) the detection of X-ray emission from the ICM, 2) galaxy dynamics in clusters, and 3) weak lensing analysis.

The study of the dynamics of its galaxy members was for many years the only probe to determine the mass of a galaxy cluster, as was originally developed by Zwicky (1933) based on optically selected clusters. Basically, the method consists in determining the membership of galaxies to a cluster and measuring their redshifts, which can be translated to velocities. Under the assumption of virial equilibrium, the mass of the cluster can be determined from these velocities (e.g. Sarazin, 1988). This method lead F. Zwicky to propose the existence of an invisible component, the dark matter, that could account for the high galaxy velocities observed.

During the last decade, the determination of the total mass in galaxy clusters has been mainly achieved from X-ray observations of the ICM thanks to the high resolution X-ray observatories such as CHANDRA and XMM NEWTON. Based on the assumption that the ICM component is in hydrostatic equilibrium within the dark matter halo, the mass enclosed at certain radius r can be determined from the set of conservations laws (mass, momentum and energy). Equilibrium is expressed through the pressure balance equation $\nabla p_{\text{gas}} = -\rho_{\text{gas}} \nabla \Phi$ where Φ is the total gravitational potential (i.e, the potential generated by the gas and the dark

matter halo). Under the assumption that the ICM satisfies the equation of state of an ideal gas, it can be shown that pressure equilibrium is translated to an estimation for the enclosed mass at a radius r (e.g. Binney and Tremaine, 1987), written as a function of the gradient of the temperature and the gas density profile:

$$M(< r) = -\frac{k_B T_{\text{gas}}(r)r}{\mu m_p} \left[\frac{d \ln \rho_{\text{gas}}(r)}{d \ln r} + \frac{d \ln T_{\text{gas}}(r)}{d \ln r} \right], \quad (2.19)$$

where k_B is the Boltzmann constant, m_p is the proton mass and μ is the mean molecular weight. Conversely, this expression is also used to determine the gas profile assuming that the total mass of the cluster is dominated by the dark matter halo, whose density profile can be calibrated from N -body simulations (see Section 3.4). In this case, assuming an isothermal distribution for the gas, it can be shown that the gas density profile is completely determined by the properties of the dark matter density profile.

Isothermal models are often adopted for the ICM (e.g. Makino et al., 1998b) leading to profiles of the form $\rho_{\text{gas}} \propto r^{-2}$. Polytopic equation of state $p_{\text{gas}} \propto \rho_{\text{gas}}^\gamma$ have been also applied to fit observations (e.g. Finoguenov et al., 2007) as suggested by numerical simulations. Approximations to the isothermal model based on the King profiles¹² (e.g. Binney and Tremaine, 1987) allow us to derive the gas density profile from the hydrostatic equilibrium equations yielding the so-called β -model (Cavaliere and Fusco-Femiano, 1976), which has been also supported by numerical simulations (e.g. Borgani et al., 2004) and mostly by the fact that its corresponding surface brightness describe very well the observed X-ray brightness profiles of galaxy clusters. The gas profile in the β -model is given by

$$\rho_{\text{gas}}(r) = \rho_0 \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-3\beta/2}, \quad (2.20)$$

where $\beta = m_p \mu \sigma_r^2 / k T_{\text{gas}}$ represents the ratio of the kinetic energy of the galaxies (with radial velocity dispersion σ_r^2) to the kinetic energy of the gas (with temperature T_{gas}). Recent observations by Plagge et al. (2010) based on SZ detected X-ray clusters suggest value of $\beta \approx 0.86$ and $r_c/R_{500} \approx 0.20$, where R_{500} is the radius comprising matter with an overdensity $\Delta = 500$ (see Section 3.4). With constant temperature, the mass enclosed at R_{500} is then defined by the parameters β , the core radius r_c and the gas temperature T_{gas} .

Contrary to X-ray studies, weak lensing analysis (e.g. Clowe et al., 2004; Mandelbaum et al., 2010) can provide direct estimates of the cluster mass without any assumption on the state of the galaxy cluster. It has been observed that the masses derived from the assumption of hydrostatic equilibrium (HSE) display a correlation with the weak lensing masses $M_{HSE} \propto M_{WL}$ with small scatter and possibly a small ($\sim 10\%$) low bias for M_{HSE} .

Scaling relations

Being the real mass of the galaxy cluster not a direct observable, it is mandatory to provide reliable links between the measurable cluster properties such as its X-ray luminosity, its temperature or the baryon fraction, with the total mass. The precise determination of the scaling relations and their intrinsic scatter must be taken into account when galaxy cluster are used as probes for cosmological models, as we discussed in Section 3.5.4.

The preferred set of the so-called mass-proxies for galaxy clusters are 1) the X-ray luminosity L_X , 2) the temperature of the hot gas T_X (e.g. Hartley et al., 2008), and 3) the mass

¹²The King profile is an approximate solution of the hydrostatic equilibrium equation for a system of collisionless particles

fraction in hot gas M_{gas} (e.g. Giodini et al., 2009). A combination of the last two quantities, $Y_X \equiv M_{\text{gas}}T_X$ has been shown to provide tighter correlations with the total mass (e.g. Vikhlinin et al., 2009a).

Thermal bremsstrahlung (e.g. Padmanabhan, 1993a) is the responsible processes of the X-ray emission from hot gas with temperatures above ~ 3 keV¹³. Basically, the gas, predominantly hydrogen and helium, falls into the deep potential wells of the dark matter haloes and is shock-heated. The gas reaches a state of (approximately) virial equilibrium with temperatures of the order of 10^7 K, capable of ionize the hydrogen and helium atoms. This, translated to luminosities of $L_X \sim 10^{44}$ erg/s. These values rank galaxy clusters as one of the brightest objects in the X-ray band.

The bolometric luminosity associated to bremsstrahlung radiation is the integral over the volume occupied by the ICM of the total emissivity $\epsilon_{\text{ff}}(T_{\text{gas}}) = \Lambda(T_{\text{gas}})\rho_{\text{gas}}^2(r)$:

$$L_{X \text{ bol}} = \int d^3r \Lambda[T(r)]\rho_{\text{gas}}^2(r), \quad (2.21)$$

For an isothermal model of the gas, it can be shown that this expressions lead to a scaling relation of the form $L_X \propto \Lambda(T)M_{\text{vir}}$, with $M_{\text{vir}} \propto r_{\text{vir}}^{-3}$. The emissivity of thermal bremsstrahlung is proportional to the square root of the temperature (e.g. Padmanabhan, 1993a), $\Lambda(T) \propto T^{1/2}$, such that the bolometric luminosity due to this process scales as $L_X \propto T^{1/2}M_{\text{vir}}$. On the other hand, assuming virial equilibrium it can be shown that the resulting temperature of the gas is related to the hosting halo mass via $T \propto M_{\text{vir}}^{2/3}$. These results lead to a scaling relation between the cluster mass and its bolometric X-ray luminosity of the form $L_{X \text{ bol}} \propto M_{\text{vir}}^{4/3}$, which is valid as long as the gas fraction is constant and do not depend on the halo mass. In the *ROSAT* energy band, not only thermal bremsstrahlung contributes to the total X-ray emissivity, but also emissions lines, leading to a temperature-independent integrated emissivity, which, under the assumption of virial equilibrium, leads to a scaling relation of the form $L_X \propto M_{\text{vir}}$.

Under the assumption that only gravity and thermal bremsstrahlung are the processes responsible for the cluster dynamics and evolution, it is expected that clusters of different masses are just re-scaled copies of themselves. This is referred to as the self-similar evolution. Concerning the mass-luminosity relation, we can write $\bar{L}_X = A(z)M^p$, where the exponent p characterizes the self-similar evolution (or deviations from it), while the function $A(z)$ depends on cosmological parameters via $A(z) = A_0E(z)^\gamma$ with $E(z) = H(z)/H_0$ (see Equation 2.1). The exponent γ in the function $A(z)$ takes the value $\gamma = 7/3$ for self-similar evolution, while $\gamma = 0$ means no evolution. The exponent γ is found to behave very close to the self-similar result (Maughan, 2007). Nevertheless, a more conservative treatment prefers to infer the mass-luminosity relation independent of theoretical model by assuming a scale relation with an amplitude varying as $A(z) = A_0(1+z)^\gamma$ (e.g. Borgani, 2001).

Recent investigations based on X-ray observations (*ROSAT* mission and the *Chandra* space observatory) have shown that scaling relations such as the mass-temperature or the luminosity-temperature are affected by the non-thermal processes taking place in the inner regions of the ICM. These processes, such as AGN activity and SN feedback contribute to the observed intrinsic dispersion about these mean relations. As an example, Pratt et al. (2009) quotes the following results for the temperature-luminosity relation for emission detected within a radius R_{500} ,

$$L_X = (6.07 \pm 0.58) T^{2.70 \pm 0.24} \quad R < R_{500}, \quad (2.22)$$

¹³Below 3 keV emission lines from heavy atoms in the ICM are responsible for the emission mechanism.

with intrinsic scatter of $\sigma_{\ln L} = 0.663 \pm 0.116$. They applied the same method on the same sample after having excised the inner part of the X-ray emission, leading to

$$L_X = (3.89 \pm 0.18) T^{2.78 \pm 0.13} \quad 0.15 R_{500} < R < R_{500}, \quad (2.23)$$

with intrinsic scatter of $\sigma_{\ln L} = 0.269 \pm 0.055$. Thus, a considerable change is observed in the normalization of the mass X-ray luminosity relation, but most important, the intrinsic scatter has been reduced almost 50% when the inner regions are omitted in the analysis. Similar conclusions are derived from the analysis of Mantz et al. (2010).

2.7 Summary

Within the last two decades we have witnessed how cosmology reached phase characterized by an unprecedented precision in the determination of the most fundamental cosmological parameters. Astronomical observations based on the detection and light-curve measurements of SN Type Ia supernovae, the power spectrum of the temperature fluctuations in CMB and the spatial distribution of galaxies have provided an accurate description of the large scale structure of the Universe. Such successful road of twenty years has attempted to answer one of the most fundamental questions of modern science, namely, what is the Universe made of? and it is still a long way to obtain a satisfactory theoretical description of the Universe we observe. With the advent of large X-ray surveys such as *eROSITA*, X-ray detected galaxy clusters will provide further information concerning the evolution of dark matter and dark energy.

In the process of understanding the Universe we observe, the two-point statistics based on the two-point correlation function and the power spectrum have received lot of attention since the first cluster catalogues (e.g. Abell, 1958; Bahcall, 1988) were published. These quantities are sufficient to quantify the statistical properties of a Gaussian random field and their dependence with the underlying cosmological model marked them as excellent probes for cosmological parameters. In the recent years, more attention has been paid to high order correlation functions, which are sensitive to non-Gaussianities (either primordial or induced by non-linear evolution).

The so-called era of precision cosmology provides increasingly smaller confidence regions of the cosmological parameters, in a race that, to date, seems to be pointing towards the complete establishment of the Λ CDM cosmological model. In an attempt to start providing answers to the fundamental questions of modern cosmology, large galaxy and galaxy cluster surveys are being carried out or will be launched soon. Their results will not only help to further shrink the confidence levels of the cosmological parameter-space, but will generate a description of how such parameters have been evolving with time. In particular, the detection of a time-evolution of the dark energy equation of state would open the possibilities for alternative theories to explain the Universe we observe.

Dark matter haloes: abundance, bias and clustering

3.1 Introduction

A basic step towards the analysis of the clustering of galaxy clusters is the understanding of the properties of their primary components, i.e, the dark matter haloes. In this chapter we introduce key concepts for the study of these properties, i.e. the abundance, biasing and clustering, and provide a quantitative description of these quantities by means of N -body simulations.

In Section 3.2 we introduce the mass function of dark matter haloes. This quantity determines the abundance of haloes of a given mass and strongly depends on the models of growth of fluctuations, and thereby is a very sensitive probe for cosmological parameters. We describe the most currently used models and the available fitting formulae calibrated with N -body simulations. In Section 3.3 we study the halo-matter bias, which provides information concerning the way dark matter haloes are distributed with respect to the underlying dark matter. We describe some of the theoretical grounds and show useful fitting formulae. The main properties of dark matter haloes, mass and density profiles, are then briefly exposed in Section 3.4. Having these quantities at hand, we focus on the clustering properties of dark matter haloes, based on the so-called *halo model* for gravitational clustering, introduced in Section 3.5. We provide a brief presentation of the formalism, applied to the clustering of dark matter, galaxies and galaxy clusters. For these last two types of objects, we describe some inputs of the model such like the halo occupation statistics for galaxies and the conditional luminosity function, necessary tools for describing the way dark matter haloes are populated by baryonic matter ¹ i.e, galaxies, hot gas etc.

In Section 3.6 we make use of N -body simulations to measure the halo abundance and the halo-mass bias. Since the modeling of the clustering properties of galaxy clusters depend on the aforementioned quantities, it is very important to compare our measurements with the available fitting formulae in order to apply the latter to a posterior theoretical modeling. In Section 3.6.2 we concentrate on the two-point statistics of dark matter haloes and measure the power spectrum and marked correlation function for a halo catalogue. We verify the

¹In the cosmological context, baryonic matter refers to non-dark matter.

detection of the BAO signal in the halo power spectrum as well as in the marked correlation function. We end with a discussion in Section 3.7.

3.2 Dark matter halo mass functions

The halo mass function $n(M, z)dM$ represents the number density of dark matter haloes with mass M in the range $(M, M + dM)$ at a given redshift z . It is a key element to understand the distribution of matter in the Universe and stands as a very useful probe for cosmological models and parameters. Within the last thirty years, theoretical and numerical efforts have been made to determine the halo mass function with high precision and accuracy. Let us here briefly describe the theoretical modelings and numerical fits of the halo mass function.

We start with the Press Schechter formalism (hereafter PS) (Press and Schechter, 1974), which yields an estimate of the abundances of dark matter halos based on the assumptions of a) Gaussian primordial fluctuations and b) bound objects as formed following the spherical collapse model (e.g. Padmanabhan, 1993b). The PS formalism considers matter overdensities smoothed by a filter function $W(\mathbf{r}; R)$ with some characteristic scale R (see Equation 2.10) that is allowed to collapse once it crosses the critical density provided by a matter-dominated model of the Universe, $\delta_c \approx 1.68$. The probability distribution $P(\delta; R)d\delta$ for the smoothed overdensity $\delta(R)$ in the range $\delta, \delta + d\delta$ to collapse given the scale R is represented by a Gaussian distribution with dispersion $\sigma(R)$ (see Equation 2.11)². The PS formalism generates a fraction of cosmic mass in collapsed objects given by

$$F(\delta > \delta_c) = \int_{\delta_c}^{\infty} P(\delta; R)d\delta = \frac{1}{2} \operatorname{erfc} \left(\sqrt{\frac{\nu}{2}} \right), \quad (3.1)$$

where the parameter ν measures the amplitude of the matter fluctuation in units of the mass dispersion (Bond et al., 1991)

$$\nu(M, z) = \frac{\delta_c^2}{\sigma^2(M, z)}. \quad (3.2)$$

The well known weak point of the PS formalism follows from Equation (3.1) in the limit when $\nu \rightarrow 0$, namely, half of the collapsed haloes are missing when the total number of collapsed overdensities is determined.

An immediate solution for the missing half of the collapse haloes in the PS formalism is to arbitrarily multiply the resulting fraction by a factor 2 (Press and Schechter, 1974). Nevertheless, more physically-motivated solutions were found within the so-called excursion-set approach (or extended PS theory). In this context, the determination of the abundance of dark matter haloes can be expressed as a stochastic problem with known solutions dating back to Chandrasekhar (1943) and put in a cosmological context by Bond et al. 1991; Heavens and Peacock 1990. In this formalism, it can be shown that the smoothed matter overdensity satisfies the evolution equation associated to a Markov process (e.g. Reif, 1965). Furthermore, assuming a top-hat filter function in Fourier space, the probability distribution $P(\delta, \sigma)$ satisfies a diffusion equation with boundary condition $P(\delta_c, \sigma) = 0$ (with $\sigma(R) \neq 0$) (Bond et al., 1991), which represents an absorbing (with constant height) barrier in the random walk of $\delta(\sigma)$ at $\delta = \delta_c$. The resulting probability distribution then consists of a Gaussian distribution

²Note that in Equation (2.11), $\sigma(M)$ is properly called mass dispersion since we have a defined $M(R)$ relation provided by the smoothing function

from which the contributions of trajectories that have reached and overcome the required threshold to collapse at some other filtering scale are subtracted. In this way, the extended PS formalism correctly accounts for the missing haloes problem (e.g. Heavens and Peacock, 1990).

In the context of the random walk with a constant barrier, a key quantity is the first-crossing distribution $f(\nu)d\nu$ (also called multiplicity function) which determines the probability that a matter perturbation crosses the density threshold δ_c by a first time when smoothed by scale in the range $\nu(R), \nu(R) + \delta\nu(R)$. This of course assumes that there is a well-defined relation between the filtering scale and the mass within the smoothing scale M such that $\nu(R)$ is defined by Equation (3.2). It can be shown that the mass function at a given z can be written in terms of the the first-crossing distribution $f(\nu) = \partial F/\partial\nu$ as

$$n(M, z)dM = \frac{\bar{\rho}_0}{M} \nu f(\nu) \frac{\partial \ln \nu}{\partial M} dM, \quad (3.3)$$

where $\bar{\rho}_0$ is the today matter energy density. In the extended PS formalism, the function $f(\nu)$ is given by:

$$f(\nu) = \sqrt{\frac{1}{2\pi\nu}} e^{-\nu/2}. \quad (3.4)$$

Numerical experiments with N -body simulations (e.g. Kauffmann et al., 1999; Jenkins et al., 2001; Warren et al., 2006; Angulo et al., 2008) showed that the measured dark matter halo mass function deviates from the predictions of the extended PS theory, which overestimate the abundance of low mass haloes and underestimates the abundance of the haloes on the high-mass tail of the mass function. Furthermore, concerning the validity of the spherical collapse model, it has been noticed that such assumption leads to over-predict the abundance of haloes by a factor ~ 1.5 (Lemson and Kauffmann, 1999) compared to numerical simulations.

An impressive improvement of the PS formalism comes from exploring the ellipsoidal collapse model (e.g. Peebles, 1980) coupling it into the excursion-set approach (Sheth and Tormen, 1999; Sheth et al., 2001), yielding a better agreement when compared with N -body simulations. A key element in this improvement comes from the fact that the density contrast that must be reached during an ellipsoidal collapse depends on the mass of the collapsing overdensity, thus leading to excursion-set approach with a moving barrier generating a first crossing distribution written as (Sheth and Tormen, 1999):

$$f(\nu) = \frac{A(p)}{\nu} (1 + (q\nu)^{-p}) \exp\left(-\frac{1}{2}q\nu\right), \quad (3.5)$$

where A is determined from $\int d\nu b(\nu)f(\nu) = 1^3$ giving ,

$$A = A(p) = \left[1 + \frac{2^{-p}\Gamma(1/2 - p)}{\sqrt{\pi}}\right]^{-1} \approx 0.332 \quad (3.6)$$

with $p = 0.3$ and $q \approx 0.75$. The extended PS formalism (spherical collapse) is recovered by setting $q = 1$ and $p = A = 1/2$.

Direct calibrations to the mass function measured from N -body simulations have been also provided in the literature. For example, Jenkins et al. (2001) fitted the first-time crossing probability with the formula

$$\nu f(\nu) = 0.315 \exp\left(-|\ln(\sqrt{\nu}/\delta_c) + 0.61|^{3.8}\right). \quad (3.7)$$

³This represents the assumption that dark matter is not biased with respect to itself

Reference	First-crossing probability $f(\nu)$	Parameters
Press and Schechter (1974)	$\sqrt{\frac{1}{2\pi\nu}}\exp(-\nu/2)$	Spherical collapse
Sheth and Tormen (1999)	$2\nu\sqrt{\frac{2}{\pi}}A\exp(-a\nu/2)(1+(\nu/a)^p)$	Elip.+Fit to N -body sim.
Jenkins et al. (2001)	$2\sqrt{\nu}A\exp(- \ln\sigma^{-1}+B ^p)$	Fit to N -body sim.
Warren et al. (2006)	$2\sqrt{\nu}A\exp(\sigma^{-a}+b)\exp(-c/\sigma^2)$	Fit to N -body sim.
Tinker (2007)	$2\sqrt{\nu}A((\sigma/b)^{-a}+1)\exp(-c/\sigma^2)$	Fit to N -body sim.
Pillepich et al. (2010)	$2\sqrt{\nu}(a\sigma^{-a}+b)\exp(-c/\sigma^2)$	Fit to N -body sim.

Table 3.1: Examples of fitting formulae for the multiplicity function $f(\nu)$ from the literature.

This parameterization is built to fit mass function of dark matter haloes identified with a friend-of-friend algorithm (FoF) in the the Hubble Volume Simulation (Evrard et al., 2002). Tinker et al. (2008) generated an improved fit for the halo mass function written in terms of masses defined by spherical overdensities, shown in table 3.1. It is important thus to have a precise definition of what the mass of a halo is in order to model the halo abundances.

It has been of great interest to explore the universality of the halo mass function (Pillepich et al., 2010; Tinker et al., 2008). When written in terms of the parameter ν , the multiplicity function acquires a universal form, i.e, it does not depend on redshift, the power spectrum or the cosmological parameters, since these dependences are encapsulated in the parameter ν . The behavior of the mass function in the presence of primordial non-Gaussianities has been explored with N -body simulations by Pillepich et al. (2010).

Indirect measurements of the halo mass function are obtained from the observation of galaxy clusters (Reiprich and Böhringer, 2002b). The determination of the halo mass function is therefore affected by the relations linking astronomical observables (e.g. X-ray luminosity, temperature) and the real mass of the cluster. Numerical simulations have shown that baryonic physics causes deviations of the halo mass function ($\sim 30\%$) when compared to pure cold dark matter haloes (Stanek et al., 2006).

3.3 The dark matter-halo bias

Theoretically and numerically, it is now well established that dark matter haloes are biased tracers of the underlying dark matter distribution. That is, at a given redshift z , the fluctuations of the halo number density δ_h can be written (in a simplified picture) as a bias factor $b(\mathbf{r}, z)$ multiplied by the fluctuations of the underlying matter density δ_{mat} (see Equation 2.5):

$$\delta_h(\mathbf{r}, z) = b(\mathbf{r}, z)\delta_{\text{mat}}(\mathbf{r}, z). \quad (3.8)$$

The determination of the halo-mass bias is closely related to the description of the halo abundance. As described in Section 3.2, two theoretical approaches have been designed to describe the abundance and distribution of haloes with respect to the dark matter, namely, the excursion set approach (Bond et al., 1991) and the theory of peaks (Bardeen et al., 1986). In the context of the clustering of peaks, the central ideal is to imagine the overdensity field δ as composed of a short δ_s and a large wavelength component δ_l (Efsthathiou et al., 1988) such that the long wavelength mode modulates the number counts of the short wavelength modes

(peaks). The density threshold required for a peak δ_l to collapse is then reduced (if $\delta_s > 0$) by the large wave mode via $\delta_l = \delta_c - \delta_s$. In other words, the background reduces (or enhances) the threshold and then it is more (or less) probable to find more collapsed objects where δ_l is large (small or ≤ 0) than in regions with low (high) background overdensities (Kaiser, 1984).

Within the excursion-set approach, a key step to determine the halo-dark matter bias is the conditional mass function (Bond et al., 1991). The idea consist in imagining the universe divided in large regions characterized by a comoving scale R_0 and overdensity δ_0 at some time t_0 (characterized by a redshift z_0), with mass M_0 and volume $V = \bar{\rho}/M_0$ e.g. Mo and White, 1996; Sheth and Tormen, 1999. At some time $t_1 > t_0$ (characterized by a redshift $z_1 < z_0$), a certain number of haloes of mass M will collapse within this region⁴. Let us denote by $N(M_1, z_1|M_0, V, z_0)dM_1$ the number of haloes with masses in the interval $M_1, M_1 + dM_1$ that collapsed at redshift z_1 while being inside of an uncollapsed region which at redshift z_0 had mass M_0 and overdensity δ_0 . The fraction of dark matter embedded in such dark matter haloes, denoted by $F(M_1, z_1|M_0, V, z_0)dM_1$, is expressed with the same functional form of the cross-time probability of the extended PS formalism but with the replacements $\sigma^2 \rightarrow \sigma_1^2 - \sigma_0^2$ and $\delta_c \rightarrow \delta_1 - \delta_0$ where $\sigma_i^2 = \sigma^2(R_i)$ and $\delta_i = \delta_1(\delta_c, z)$ ($i = 0, 1$). The conditional mass function is written as

$$N(M, z_1|M_0, V, z_0)dM_1 = \frac{M_0}{M_1} F(M_1, z_1|M_0, V, z_0) \frac{\partial \ln(\sigma_1^2 - \sigma_0^2)}{\partial M_1} dM_1. \quad (3.9)$$

The halo matter bias is then explored by comparing the abundance of haloes in these cells with the overall abundance, given by the mass function $n(M, z)$ (Mo and White, 1996). This is

$$\delta_h(M, z_1|M_0, V, z_0) \equiv \frac{N(M, z_1|M_0, V, z_0)}{Vn(M_1, z)} - 1, \quad (3.10)$$

such that in the large cell limit $R_0 \rightarrow \infty$, $\sigma_0 \ll \sigma_1$ and $\delta_0 \ll \delta_1$, it can be shown that Equation (3.10) leads to $\delta_h(M, z_1|M_0, V, z_0) \rightarrow b_{hm}^L \delta_0$ where b_{hm} is the Lagrangian halo matter bias⁵,

$$b_{hm}^L(\nu_1) \equiv \frac{\nu_1 - 1}{\delta_1}, \quad (3.11)$$

which only depends on the mass of the collapsed haloes and the redshift at which those haloes are observed, since $\nu = \nu(M, z)$. Given that $\sigma(M, z)$ is a monotonically decreasing function of the mass, Equation (3.11) implies that there exist a characteristic halo mass M_\star such that haloes with $M < M_\star$ ($M > M_\star$) are anti-biased (biased) with respect to the regions with characteristic overdensity δ_0 .

As long as the universe evolves, the Lagrangian regions can amass sufficient material and can decoupled from the Hubble flow, collapse and form bound objects. In that case, the required conditional mass function can be written exactly in the same way as in Equation (3.9) but supplying the relations between the original scale R_0 and the radius R of the collapsed region, as well as a relation $\delta_0 = \delta_0(\delta)$, where now δ represents the overdensity of the collapsed region. These relations are provided by non-linear collapse models (e.g. Mo and White, 1996; Sánchez-Conde et al., 2007). The resulting bias, referred as to Eulerian bias, is defined by $\delta_{hh} = b_{hm}^E \delta$ and given by Mo and White (1996) as

$$b_{hm}^E(\nu) = 1 + b_{hm}^L(\nu). \quad (3.12)$$

⁴This can be thought by a two-barrier problem in the context of the extended PS formalism as stated by Bond et al. (1991)

⁵By Lagrangian bias we refer to the bias with respect to a region that is still coupled to the Hubble expansion and thus its characteristic scale R_0 is measured in comoving coordinates

This means that halo fluctuations have an unbiased component with respect to the collapsed region (the term $1+$) while being enhanced due to the bias of the collapsed region with respect to the uncollapsed background (the term $b_{hm}^L(\nu)$). An important lesson to learn from this description is that the resulting bias depends on the shape of the halo mass function, which means that in order to precisely establish the large scale halo-matter bias, a precise halo-mass function must be derived or measured (e.g. Sheth and Tormen, 1999; Manera and Gaztanaga, 2009).

Improvements

In the context of the ellipsoidal collapse model, the resulting mass-dependent threshold (moving barrier) generates a halo mass function which is in better agreement with those measured from numerical simulations (Sheth and Tormen, 1999; Sheth et al., 2001; Kauffmann et al., 1999). By means of the peak-background split (Bond et al., 1991), the conditional halo mass function can be written such that the associated large scale halo-matter (Eulerian) bias reads (Sheth et al., 2001)

$$b(\nu) = 1 + \frac{q\nu - 1}{\delta_c(z)} + \frac{2p\delta_c(z)^{-1}}{1 + (q\nu)^p}. \quad (3.13)$$

It is worth testing whether this prescription (along with the corresponding halo mass function) describes the measurements from N -body simulations and real observations. It is noticeable that initially the mass function of Sheth and Tormen (1999) was proposed as a fit (in the context of the peak-background split) to measurements from numerical simulations, and a refining of the excursion-set method by Sheth et al. (2001) derived from the ellipsoidal collapse generated an almost similar mass function, with differences in the values of the parameters involved, due mainly to the intrinsic features of the N -body simulations implemented.

On intermediate scales, non-linear evolution induces the presence of higher order statistics which introduces complexity in the description of the halo-mass bias. A way to account for non-linearities was proposed by Fry and Gaztanaga (1993) through a perturbative analysis by assuming that the halo overdensity is a function of the underlying matter overdensity, $\delta_{hh} = F[\delta_m]$. As long as the matter overdensity evolves in the linear regime, the halo overdensity can be written as a Taylor expansion $\delta_h(\mathbf{x}, M, z) = \sum_{k=0} b_k(M, z)\delta_m^k(\mathbf{x}, z)/k!$, where the term b_0 is constrained by the condition $\langle \delta_h \rangle = 0$. The remaining coefficients must satisfy

$$\frac{1}{\bar{\rho}} \int dM n(M) b_i(M) = \begin{cases} 1 & i = 1 \\ 0 & i > 1, \end{cases}$$

which means that the mass of the universe is contained in dark matter haloes. In the large-cell limit and using the mass function of Equation (3.5), the coefficients b_2, b_3, b_4 can be determined (e.g. Cooray and Sheth (2002)), while the coefficient b_1 is just given by Equation (3.13). These coefficients can also be constrained by the measurements of higher order statistics (e.g. Manera and Gaztanaga, 2009).

Together with a non-linear biasing scheme, stochastic bias (Dekel and Lahav, 1998; Taruya and Suto, 2000) has also been proposed in the literature by introducing a bias conditional distribution $P(\delta_h|\delta_m)$ such that the mean halo overdensity can be written as $\delta_h = \int d\delta'_h P(\delta'_h|\delta_m)\delta'_h$. Stochasticity in the bias of dark matter haloes can arise due to the different formation redshifts of haloes with the same mass, as described by Taruya and Suto (2000); Manera and Gaztanaga (2009).

Fitting formulae

Scale dependent halo-mass bias has been observed on small scales using N -body simulations through the measurements of the halo-halo and matter-matter correlation function. A numerical fit has been provided by Tinker et al. (2005), parameterizing the halo-mass bias as $b(r, M, z) = B(r, z)b(M, z)$ with a scale-dependency given by

$$B(r, z) = \frac{(1 + 1.17\xi_{\text{mat}}(r, z))^{1.49}}{(1 + 0.69\xi_{\text{mat}}(r, z))^{2.09}}. \quad (3.14)$$

This has been found to give a good description of scale dependent bias in configuration space. However it must be kept in mind that it is only a parameterization and not a physical modeling of the non-linear bias. One might be thus concerned when using this expression to develop clustering analysis, for in principle this does not tell us much about the physics on small scale that give rise to non-linearities or a scale dependent bias. Nevertheless, since it is parameterized in terms of the dark matter correlation function (fully non-linear), we can consider that all the physics entering in the determination of $\xi_{\text{mat}}(r)$ are properly taken into account in the analysis. Also, this parameterization is not valid for any set of cosmological parameters, since it has been determined in a small region of the (cosmological) parameter space. Nevertheless the chosen region in parameter space contains the most probable values according to the latest constraints (e.g. Sánchez et al., 2009). The function $B(r)$ is an increasing function of the scale r up to $r \approx 4 \text{ Mpc } h^{-1}$ with a mild maximum at $r = 6 \text{ Mpc } h^{-1}$ displaying afterwards asymptotic behavior towards $B(r, z) \approx 1$. This behavior implies that on scales below $2 \text{ Mpc } h^{-1}$, the low mass haloes are anti-biased with respect to the matter distribution. This might be due to the fact that these low mass haloes are indeed substructures of larger haloes resolved in the high resolution simulations used by Tinker et al. (2005).

Using high resolution N -body simulations, a recent parameterization of the halo-mass bias in terms of six parameters is given by Tinker et al. (2010)

$$b(M, z) = 1 - A \frac{v^a}{v^a + \delta_c^a} + Bv^b + Cv^c, \quad (3.15)$$

where the set of parameters $\{A, a, c\}$ depend on the density contrast Δ . For $\Delta = 200$, we obtain $A = 1.04$, $a = 0.132$, $B = 0.183$, $b = 1.5$, $C = 0.262$ and $c = 2.4$. Below we will compare the numerical fits provided in this section with direct measurements from N -body simulations.

3.4 Properties of dark matter haloes

The mass of a halo

The spherical collapse model defines a dark matter halo as a conglomerate of dark matter particles whose density overcame a critical threshold and then collapses inside a virial radius equal to half the radius it had at the time it decoupled from the Hubble expansion⁶. Following Busha et al. (2005), three regions can be distinguished in a dark matter halo: 1) the inner structure with null radial velocities, 2) an intermediate phase where in-fall or out-fall of mass is present and 3) the exterior shells where still the Hubble flow dominates the dynamics of dark matter. In the final virial state, the dark matter halo is characterized by a density

⁶This is valid in a matter-dominated Universe. With dark energy, the relation $r_{\text{max}}/r_{\text{vir}}$ might depend on the equation of state (e.g. Wang, 2000).

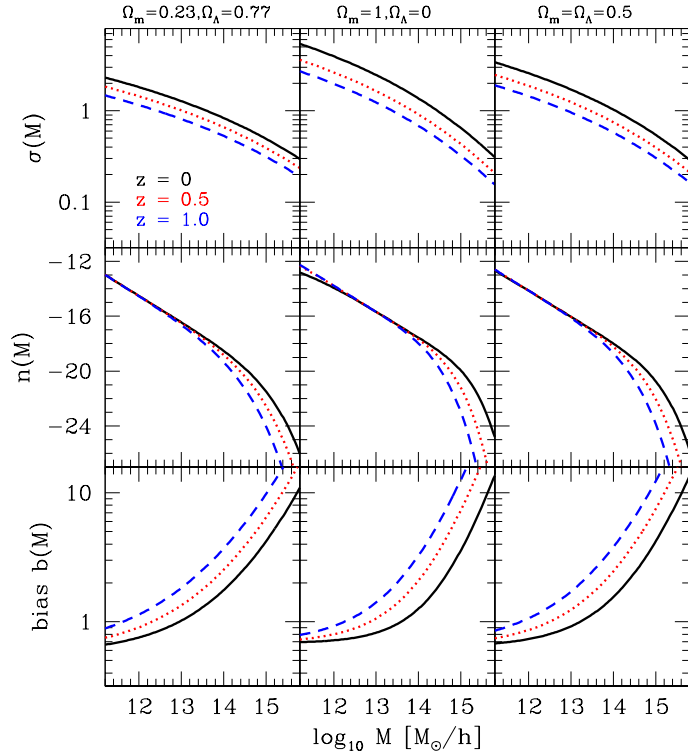


Figure 3.1: Mass variance (top row), halo mass function (middle row) and mass bias (bottom row) for three different redshifts and three different sets of cosmological parameters Ω_m, Ω_Λ maintaining the other parameters fixed to our fiducial cosmology. The Jenkins et al. (2001) halo mass function and the Sheth and Tormen (1999) fitting formulae have been used to show the halo mass function and the halo-mass bias respectively.

contrast $\Delta_{\text{vir}} \equiv \bar{\rho}/\rho_{\text{crit}}$, where $\bar{\rho}$ is the mean density of the virialized halo and ρ_{crit} is the critical density of the universe. In a matter dominated universe this corresponds to $\Delta_{\text{vir}} \approx 178$ (e.g. Padmanabhan, 1993b). For a Λ CDM cosmology, the density contrast at virialization can be approximately written (with $\Omega_{\text{rad}} = 0$) as (Bryan and Norman, 1998)

$$\Delta_{\text{vir}}(z) = 18\pi^2 + 82\Omega_{\text{mat}}(z) - 39\Omega_{\text{mat}}(z)^2, \quad (3.16)$$

with 1% percent precision with respect to the exact solution (e.g. Eke et al., 1996).

In numerical simulations the definition of the halo mass depends strongly on the method used to define dark matter haloes (e.g. White 2001, 2002; Lukic et al. 2010). Spherical overdensities algorithms (SO) (Lacey and Cole, 1993) identify dark matter haloes by finding spherical regions with a characteristic value of Δ , being $\Delta = 200, 500, 1000$ some common values (e.g. Tinker et al., 2010). As an example, in X-ray studies (see Chapter 4), values of $\Delta = 500$ or higher are preferred since it can be associated to more compact and relaxed structures. On the other hand, Friend-of-Friend (FoF) halo finders (Efsthathiou et al., 1985) are based on percolation algorithms defining dark matter particles to belong to a FoF group if they are separated a by a distance smaller or equivalent to a linking length of any other particle in the group. The linking length is usually taken to be a 20 per cent of the mean inter-particle separation in the simulation. The resulting FoF haloes have overdensities of $\Delta \approx 80$ (e.g. Lukic et al., 2010), which implies that FoF finders selects dark matter haloes with masses that are in good agreement (with a few per cent scatter in the most massive haloes) with the masses derived from virial equilibrium. Nevertheless, the FoF algorithms give an overestimation of the halo masses when compared to the masses derived from SO algorithms, especially in the low mass haloes, where resolution effects are important. Tinker et al. (2008) explores the distribution of the ratio $r_M \equiv M_\Delta/M_{\text{FoF}}$ finding that the distribution of these values is not symmetric with respect to $r_M = 1$, displaying instead broader wings at $r_M < 1$. The broadening

(and the asymmetry) is observed to increase with the redshift. Warren et al. (2006) use high-resolution N -body simulations with different resolutions applying these two halo-finders to create a dark matter halo catalogue, providing a phenomenological correction for the FoF masses

$$M_{\text{FoF}}^c = M_{\text{FoF}}(1 - N_p^{-0.6}). \quad (3.17)$$

This correction has been shown to work fairly well when compared to high resolution simulations (e.g. Crocce et al., 2010). Given the different nature of the halo-finder algorithms, there is not a direct link between FoF masses and SO masses. White (2001) showed that these are correlated with some scatter, and Lukic et al. (2010) showed that, under the assumption that dark matter haloes follow the NFW profile (see below), a mapping between these two masses can be determined once the concentration in the density profile is given.

Density profiles

Gravitational collapse of dark matter predicts density profiles following a power law (Bertschinger, 1985) $\rho \propto r^\alpha$ with a slope varying in the range $-3 \leq \alpha < -1$ ($\alpha = -2$ corresponds to the isothermal sphere), depending on whether there is a singularity in the center of the collapsing object. Theoretical predictions based on the Jeans equations (e.g. Hansen, 2004) also allows for power law solutions under certain considerations. Nevertheless the complexity of gravitational collapse imposes serious difficulties to predict a density profile for dark matter haloes.

The structure of the phase-space of dark matter particles has been recently analyzed by White and Vogelsberger (2009); Vogelsberger and White (2009). The shape of the dark matter distribution has been calibrated with N -body simulations. The limiting resolution of simulations are still shadowing the behavior of dark matter density profiles close to the center of dark matter haloes, though. Early N -body simulations suggested a cusp-like density profile, parameterized as (Navarro et al., 1996)

$$\rho(r) = \rho_s \left(\frac{r}{r_s}\right)^{-\alpha} \left(1 + \frac{r}{r_s}\right)^{\alpha-\beta}, \quad (3.18)$$

The NFW profile (Navarro, Frenk & White, 1996) is identified with the set ($\alpha = 1, \beta = 3$), while the Hernquist model (see Binney and Tremaine 1987) is reached with ($\alpha = 1, \beta = 4$). These families of profiles have two parameters ρ_s and r_s which can be calibrated with numerical simulations and expressed in terms of virialization conditions inspired by the top-hat collapse model, namely, $\bar{\rho} = \Delta_{\text{vir}}\rho_{\text{crit}}$, where $\bar{\rho}$ represents the mean density of the dark matter halo. The parameter r_s marks the radius where the slope of the density profile equals the slope of the density in an isothermal sphere, $\gamma \equiv d \ln \rho / d \ln r = -2$. The enclosed mass is analytically integrable for the NFW and Hernquist profiles.

A key parameter characterizing the matter distribution in dark matter haloes is the concentration parameter defined as $c \equiv r_{\text{vir}}/r_s$. N -body simulations have shown correlation between the concentration parameter and the dark matter halo mass with an intrinsic scatter. This scatter might arise because the differences in formation time and the merger history of the halo. N -body simulations have shown a power law dependence of the concentration parameters with the halo mass given by (Cooray and Sheth, 2002; Maccio et al., 2008) $\bar{c}(M, z) = 9(1+z)^{-1} (M/M_\star)^{-0.13}$ where M_\star corresponds to the mass scale at which the mass dispersion is $\sigma(M, z) = 1$ for a given redshift z . The distribution of concentrations around this

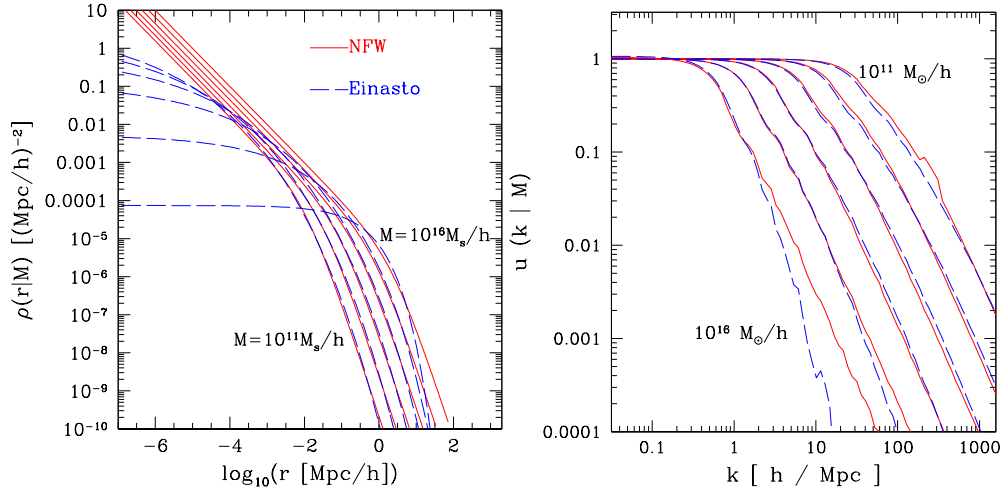


Figure 3.2: NFW and Einasto density profiles (left) and their Fourier transform for different halo masses (right).

mean relation is shown to be well described by a log-normal distribution with a dispersion of $\sigma \ln c \approx 0.2$.

Recent numerical simulations (e.g. Navarro et al., 2008) show that the density profile of dark matter haloes can be fairly well parameterized by a r -dependent slope profile. One example is the profile proposed by (Einasto, 1965), written as

$$\rho(r) = \rho_e \exp \left[-\frac{2}{\alpha} \left(\left(\frac{r}{r_e} \right)^\alpha - 1 \right) \right]. \quad (3.19)$$

and originally introduced to model the distribution of stars in the Milky Way galaxy. The Einasto profile has also a concentration parameter defined as $c = r_{\text{vir}}/r_e$ with a similar mass dependence as that of the NFW profile (see Duffy et al. (2008)).

Assuming spherically symmetric dark matter haloes, the Fourier transform of the density profile is

$$u(k; M) = \frac{4\pi}{M} \int_0^\infty j_0(kr) r^2 \rho(r; M) dr \quad (3.20)$$

with analytical solutions for the NFW profile (Cooray and Sheth, 2002). The comparison of these two density profiles is shown in Fig. 3.2 where we also show the corresponding Fourier transform. The Einasto profile develops a constant density towards the halo's core and a faster drop close and beyond the virial radius, compared to the NFW profile. Their Fourier transforms do not display relevant differences. Only for the most massive haloes we can see non-negligible differences present on intermediate-scales $k \sim 1 h \text{Mpc}^{-1}$. By convolving the density profiles with a log-normal distribution, per cent differences at small radii are observed for dispersions up to 30% around the mean concentration.

3.5 The halo model for gravitational clustering

3.5.1 Introduction

The Halo model constitutes one of the most currently used analytical tools to describe the clustering of matter, galaxies and galaxy clusters. In the recent years, a large amount of

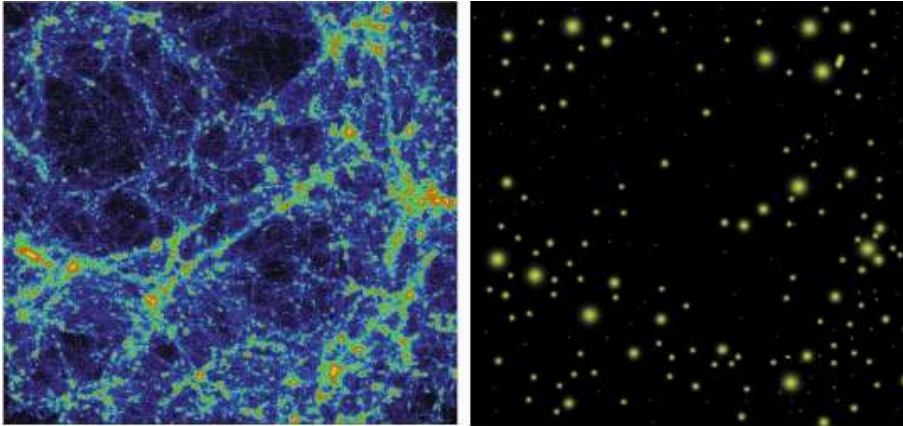


Figure 3.3: Left: complex structures like filaments, voids, super-clusters found in numerical simulations. Right: The halo model assumes that the matter content of the Universe is condensed in elementary blocks, the dark matter haloes. Figure taken from Cooray and Sheth (2002).

efforts have been made to understand and develop methods for the analysis of large scale structure (e.g. Cooray and Sheth, 2002; Hütsi and Lahav, 2008; Berlind and Weinberg, 2002; Skibba and Sheth, 2009; Wake et al., 2008; Zheng et al., 2005; Neyrink et al., 2005; Rozo et al., 2004). In particular this has been used to describe the clustering of galaxies as a function of intrinsic properties as luminosities, color (e.g. Cooray, 2006) and the dependence of the clustering on environment effects (Sheth, 2005). The assumptions made within the halo model can be described in the following:

- All the mass of the Universe is contained in dark matter haloes. These objects are assumed to be evenly distributed throughout space. This of course represents a simplified picture of the distribution of matter in the Universe. For instance, N -body simulations and observational data show that on large scales the matter distribution displays a filamentary structure with large voids (see Fig. 3.3). On the other hand, the hierarchical picture of structure formation predicts that haloes are biased tracers of the matter distribution.
- Dark matter haloes represent regions where the density perturbations were above a critical threshold δ_c such that they have collapsed and reached virial equilibrium in a roughly spherically symmetric volume with smooth density profiles. N -body simulations have shown that dark matter haloes are not completely spherically symmetric but have a triaxial configuration (e.g. Jing and Suto). Moreover, their density profile is far from being smooth because of the formation of substructures (e.g. Moore et al., 1998).
- Haloes are assumed to be small with respect to the separation between them, but are considered to be large enough to contain galaxies within them. This condition is not always fulfilled. For instance in the case of interacting clusters and galaxies, the corresponding dark matter haloes overlap. In this case haloes cannot be treated as a point-like distribution and the effects of their finite size must be taken into account.

Within the context of galaxy clustering, the basic idea is to assume that the galaxy distribution can be separated in two different regimes: large-scale regime, ruled by the behavior

of the distribution of haloes in the universe, and the small-scale regime, ruled by the distribution of galaxies within haloes. Therefore the fundamental inputs for the model are the distribution of galaxies within dark matter haloes, the density profile of dark matter haloes and their distribution on large scales.

The halo model is, in some sense, a non-dynamical approach to the understanding of the galaxy and matter distribution, because it assumes that there exist already collapsed structures. The model does not follow the dynamics of matter in the stages previous to collapse and virialization. Also, the halo model does not take into account the history of halo mergers. The latest improvements in the halo model consist in the introduction of sub-structure by introducing the statistics of sub-haloes (e.g. Sheth and Jain, 2003; Giocoli et al., 2010). Also, triaxial dark matter haloes have been considered by Smith et al. (2006).

3.5.2 Dark matter power spectrum

We can express the matter density at a position \mathbf{x} as a contribution of the matter density of each halo as

$$\rho(\mathbf{x}) = \sum_i \rho(\mathbf{x} - \mathbf{x}_i; M_i) = \int dM' d^3x' \rho(\mathbf{x} - \mathbf{x}'; M') N(\mathbf{x}; M'), \quad (3.21)$$

where the sum is done over all haloes. Here $N(\mathbf{x}; M) \equiv \sum_i \delta(M - M_i) \delta^3(\mathbf{x} - \mathbf{x}_i)$ is the number density of haloes with mass M centered at a position \mathbf{x} . Note that the equation above is implicitly assuming that the density profile of dark matter haloes possess an universal form parameterized by the mass of the halo. The halo mass function $n(M)$ is determined by the ensemble average of the spatial number density $n(M) \equiv \langle N(\mathbf{x}; M) \rangle$. Similarly, the mean density is the average of the total density $\rho(\mathbf{x})$, $\bar{\rho} = \int dM M n(M)$. With this we can write the correlation function, written in terms of the density field as $\xi(\mathbf{x} - \mathbf{y}) + 1 = \frac{1}{\bar{\rho}^2} \langle \rho(\mathbf{x}) \rho(\mathbf{y}) \rangle$. The statistics is now being focused on the quantities $N(\mathbf{x}; M)$, namely,

$$\langle N(\mathbf{x}; M) N(\mathbf{y}; M') \rangle = n(M) \delta(M - M') \delta^3(\mathbf{x} - \mathbf{y}) + n(M) n(M') (1 + \xi_{hh}(\mathbf{x} - \mathbf{y}; M, M')). \quad (3.22)$$

The first term in this expression is associated with shot-noise or auto-correlation. The second term contains the contribution from the random probability ($n(M)n(M')$) and the clustering embodied in the correlation function $\xi_{hh}(\mathbf{x} - \mathbf{y}; M, M')$, measuring the probability of having two haloes centered at \mathbf{x} and \mathbf{y} with masses M and M' respectively.

The resulting two-point correlation function can be split in two main components. Accordingly, the corresponding power spectrum is also separated as $P_{\text{mat}}(k) = P_{1h}(k) + P_{2h}(k)$, where the first term describes the clustering of dark matter particles within a single halo,

$$P_{1h}(k, z) = \int dM \frac{M^2 n(M, z)}{\bar{\rho}^2} |u_{\text{dm}}(k, z; M)|^2, \quad (3.23)$$

with $u(k, z; M)$ being the Fourier transform of the density profile of a halo of mass M at a redshift z . The power spectrum associated with pairs in different haloes reads as

$$P_{2h}(k, z) = \int dM \frac{M n(M)}{\bar{\rho}} \int dM' \frac{M' n(M', z)}{\bar{\rho}} u_{\text{dm}}(k, z; M) u_{\text{dm}}(k, z; M') P_{hh}(k, z; M, M'), \quad (3.24)$$

where $P_{hh}(k; M, M')$ is the power spectrum of pairs of dark matter haloes with masses M and M' . For the matter power spectrum on large scales, a linear halo-matter biasing scheme is assumed, such that at a given redshift,

$$P_{hh}(k, z; M, M') = b(M, z) b(M', z) P_{\text{lin}}(k, z), \quad (3.25)$$

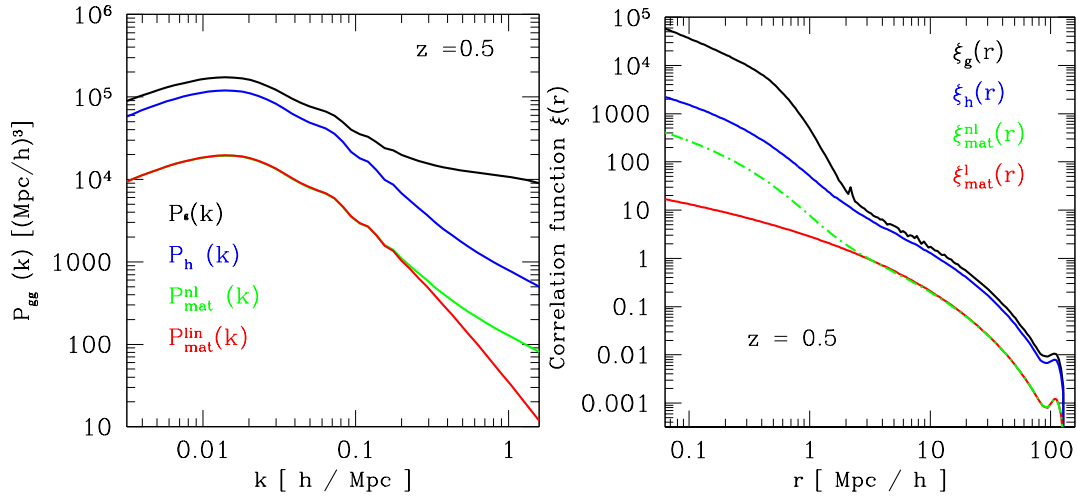


Figure 3.4: Halo model power spectrum (left) and correlation function (right). The matter and galaxy power spectrum. Note that on large scales, the contribution from the small scales is non negligible.

where $P_{\text{lin}}(k, z)$ is the linear matter power spectrum. This allows us to separate the contributions from different haloes and write $P_{2h}(k) = b_{\text{eff}}^2(k)P_{\text{lin}}(k)$ with an effective bias defined

$$b_{\text{eff}}(k, z) \equiv \int dM \frac{Mn(M, z)}{\bar{\rho}(z)} b(M, z) u_{\text{dm}}(k, z; M), \quad (3.26)$$

On large scales we have $u(k; M) \rightarrow 1$ and the effective bias becomes scale-independent. Nevertheless, this also imposes a problem within the model, because the 1-halo term acquires a constant value that becomes dominant on very large scales introducing non-physical large-scale power. The removal of such excess of power cannot simply be done by subtracting this value from the resulting power in the same way at all scales. This unpleasant situation is treated with compensated profiles (e.g. Cooray and Sheth, 2002).

The halo model thus concentrates its efforts in modeling the non-linear clustering of dark matter with the 1-halo term, mainly through the behavior of the Fourier transform of the density profile. Nevertheless N -body simulations and recent theoretical modeling of non-linearities (e.g. Crocce and Scoccimarro, 2008) have shown that this model fails to reproduce the observed power spectrum in the non-linear regime. This can be seen from the set of equations defining the halo model, where the typical non-linear scale given by the 1 halo term is the virial radius of the density profile, while containing no information on the typical BAOs scale $\sim 100 \text{ Mpc } h^{-1}$. Despite these two problematic issues within the halo model, it has still plenty of applications on galaxy clustering analysis, as we discuss in the following section.

3.5.3 Halo model for gravitational clustering

Galaxy clustering is widely explored in the framework of the halo model (e.g. Wake et al., 2008; White et al., 2007a; Blake et al., 2008; Ross et al., 2010), which is a suitable theoretical base to study clustering as a function of galaxy properties, and to investigate galaxy merger histories (e.g. White et al., 2007b) and the way galaxies populate dark matter haloes (e.g. Berlind and Weinberg, 2002; Wake et al., 2008; Blake et al., 2008). In what concerns the intrinsic properties of galaxies, it has been found (Baldry et al., 2005) that galaxies posses

a color bi-modality: red sequence galaxies are usually spheroidal, high mass and non-star forming galaxies, while blue-sequence galaxies are low-mass, star forming galaxies dominated by discs (Longair, 2005). Regarding the way galaxies populate dark matter haloes, a trend exists in observations: central galaxies, approximately located in the center of mass of haloes, tend to be red galaxies ⁷. Satellite galaxies orbit around the center following a certain distribution, and the satellite population includes both red and blue galaxies (e.g. Tinker and Wetzel, 2009).

In this section we will briefly review the applications of the halo model for galaxy clustering. Details are given in Appendix B. Within the context of the CDM paradigm, the galaxy number density has to be written as a sum over all haloes with galaxies, similar to Equation (3.21). Furthermore, within each halo, there will be a probability for having a central or a satellite galaxy with certain luminosity L , depending on the halo mass M . This information is encoded in the conditional luminosity function $\Phi(L|M)$ (CLF hereafter) (e.g. Yang et al., 2003; van den Bosch et al., 2003), defined such that $\Phi(L|M)dL$ represents the probability of assigning a luminosity L in a range $L, L + dL$ to a galaxy hosted by a dark matter halo of mass M . Based in the different properties between central and satellite galaxies, the CLF is expected to display different properties for each of these type of objects. If the galaxy distribution follows a density profile $u_{\text{gal}}(\mathbf{r}; M)$, the galaxy number density can be written as

$$n_{\text{gal}}(\mathbf{x}, z; L) = \sum_i \left[\Phi_c(L|M_i, z) \delta^3(\mathbf{x} - \mathbf{x}_i) + \Phi_s(L|M_i, z) u_{\text{gal}}(\mathbf{x} - \mathbf{x}_i; M_i, z) \right] \quad (3.27)$$

Having an expression for the galaxy number density, we can determine the galaxy-galaxy correlation function for pairs of galaxies separated by a distance r at redshift z having luminosity L :

$$\xi_{\text{gal}}(r, z; L) + 1 = \frac{\langle n_{\text{gal}}(\mathbf{x}; z; L) n_{\text{gal}}(\mathbf{x} + \mathbf{r}; z; L) \rangle}{\Phi^2(L, z)}, \quad (3.28)$$

where $\Phi(L, z)$ is the galaxy luminosity function at a given redshift. As explained before, the galaxy power spectrum will be split as $P_{\text{gal}} = P_{\text{gal}}^{1h} + P_{\text{gal}}^{2h}$. Following the notation in Cooray (2006) we defined the following quantities

$$I_c(L, M, z) = \frac{\Phi_c(L|M, z)}{\Phi(L, z)}, \quad I_s(k, L, M, z) = \frac{\Phi_s^A(L|M, z) u_{\text{gal}}(k; M, z)}{\Phi(L, z)}, \quad (3.29)$$

such that the 1-halo contribution to the luminosity-dependent galaxy power spectrum reads as

$$P_{\text{gal}}^{1h}(k, z; L) = \int dM n(M, z) [2I_c(L, M, z) + I_s(k, L, M, z)] I_s(k, L, M, z). \quad (3.30)$$

Under the assumption of a scale independent halo-mass bias, the 2-halo term is written as:

$$P_{\text{gal}}^{2h}(k, z; L) = b_{\text{eff}}(k, z; L)^2 P_{\text{mat}}(k, z), \quad (3.31)$$

where the effective bias is given by:

$$b_{\text{eff}}(k, z; L) = \int dM n(M, z) b(M, z) [I_c(L, M, z) + I_s(k, L, M, z)]. \quad (3.32)$$

⁷This mildly depends on the halo mass, and do not applies for low-mass haloes. Also, this does not implies that all central galaxies are red.

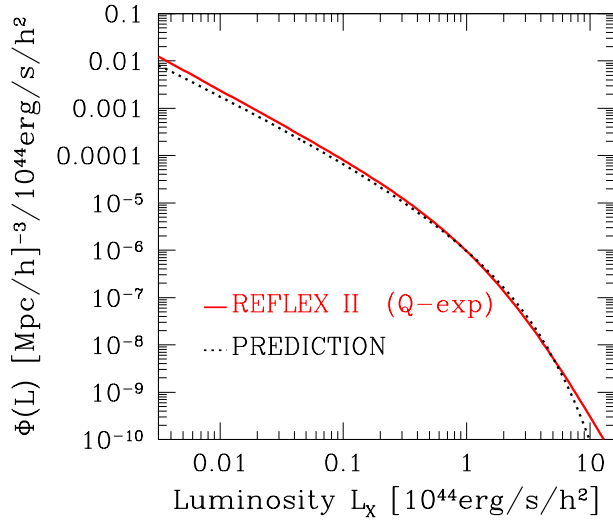


Figure 3.5: Luminosity function predicted from Equation (3.36) using the halo mass function of Jenkins et al. (2001) (black dashed line). The best parameterization from the REFLEX II luminosity function is shown by the solid red line. The underlying mass luminosity relation is given by Equation (4.11).

and $P_{\text{mat}}(k, z)$ is the full non-linear matter power spectrum. Note that the scale-dependence in this term arises through the Fourier transform of the density profile followed by galaxies. If we assume that galaxies follow a density profile such that on large scales it behaves as the NFW or Einasto profile, then on these scales the effective bias tends to a constant value (see Fig. 3.2), given by

$$b_{\text{eff}}(k \rightarrow 0, z; L) \rightarrow \frac{1}{\Phi(L, z)} \int dM n(M, z) b(M, z) \Phi(L, z|M). \quad (3.33)$$

In this limit the galaxy clustering behaves as a scaled version of the underlying matter clustering pattern. Note also that this is the effective bias we would have obtained in case we only take into account the clustering of central galaxies, haloes or galaxy clusters. On the other hand, in the small-scale limit where $u_{\text{gal}}(k; M) \rightarrow 0$ the effective bias reaches also a constant value. As we discussed in the previous section, the transition from large scales to small scales is then ruled through the Fourier transform of the galaxy density profile.

The way galaxies populate dark matter haloes is a key element in the clustering analysis built from the halo model. There is not yet a formalism that allow us to derive from first principles expressions that tell us the mean number of galaxies of a given type living in dark matter haloes of a given mass. We are therefore forced to make educated guesses on what we might expect to be the population of galaxies in dark matter in the context of the hierarchical scenario of structure formation. In Appendix B we describe some of the most popular parameterizations.

3.5.4 Clustering of galaxy clusters

In this section we will concentrate on massive galaxy clusters and explain useful relations that will be applied when exploring the clustering of X-ray detected galaxy clusters. In this context we consider the intra-cluster medium⁸, responsible for the X-ray emission, as a central object emitting with a given luminosity L_X .

The direct observables concerning X-ray detected galaxy clusters are the X-ray luminosity and temperature of the intra-cluster medium. We describe the clustering in terms of the X-ray luminosity, although it can be also described in terms of its temperature. The X-ray

⁸Hot plasma in hydrostatic equilibrium with the dark matter halo. Virial temperatures of $T \sim 3\text{keV}$

emission from galaxy clusters corresponds mostly from bremsstrahlung radiation emitted by the (almost fully ionized) hot gas. Assuming that massive galaxy clusters have only one peak in the X-ray emission (see Chapter 4), the number density of X-ray emitting regions with luminosity L embedded in a dark matter halo is

$$n_{\text{cl}}(\mathbf{x}, z; L) = \sum_i \Phi_{\text{cl}}(L|M_i, z) \delta^3(\mathbf{x} - \mathbf{x}_i), \quad (3.34)$$

and the corresponding cluster power spectrum, containing only the 2-halo component, reads $P_{\text{cl}}(k, z; L) = b_{\text{eff}}^2(z, L) P_{\text{mat}}(k, z)$, where the effective bias is given by Equation (3.32):

$$b_{\text{eff}}(k, z; L) = \frac{1}{\Phi(L, z)} \int dM n(M, z) b(M, z) \Phi_{\text{cl}}(L|M; z), \quad (3.35)$$

where $\Phi_{\text{cl}}(L, z)$ is the X-ray cluster luminosity function, written in terms of the halo mass function $n(M)$ as (see Appendix B):

$$\Phi(L, z) = \int \Phi_{\text{cl}}(L|M, z) n(M, z) dM. \quad (3.36)$$

This expression provides the direct link between the observable and measurable luminosity function and the underlying halo mass function. In the context of galaxy clusters, the CLF represents the probability of assigning a luminosity L in the X-ray band to a halo of mass M . Observational data suggest (e.g. Stanek et al., 2006; Vikhlinin et al., 2009a) that a log-normal distribution (see Equation B.16) is a good description of the conditional probability. This of course involves a mass-luminosity relation (or mass temperature when the temperature function is determined) and the determination of its intrinsic scatter. Provided that we know the mass of the clusters from independent measurements, Equation (3.36) provides a tool to extract cosmological information from the cluster abundance (e.g. Pratt et al., 2009; Mantz et al., 2010). For negligible intrinsic scatter, the luminosity function becomes a direct measurement of the halo mass function, modulated by the underlying mass-luminosity relation, $\bar{L}(M)$:

$$\lim_{\sigma_{\ln L} \rightarrow 0} \Phi_{\text{cl}}(L, z) = n[M(L), z] \left| \frac{d \ln \bar{L}(M)}{dM} \right|_{M=M(L)}^{-1}$$

Returning to the clustering of galaxy clusters, concerning the clustering of objects with a luminosity greater than some value L , the correlation function is written as

$$\xi_{\text{cl}}(r, z; > L) + 1 = \frac{\langle n_{\text{cl}}(\mathbf{x}; z; > L) n_{\text{cl}}(\mathbf{x} + \mathbf{r}; z; > L) \rangle}{\bar{n}_{\text{cl}}^2(> L, z)}, \quad (3.37)$$

where $n_{\text{cl}}(\mathbf{x}; z; > L)$ is the number density of clusters with luminosities greater than L and $\bar{n}_{\text{cl}}(> L, z)$ is the mean number density of clusters with the same situation. These are obtained as:

$$n_{\text{cl}}(\mathbf{x}; z; > L) = \int_L^\infty n_{\text{cl}}(\mathbf{x}, z; L) dL, \quad \bar{n}_{\text{cl}}(> L, z) = \int_L^\infty \Phi_{\text{cl}}(L, z) dL,$$

such that the effective bias reads as

$$b_{\text{eff}}(k, z; > L) = \frac{1}{\bar{n}(L, z)} \int n(M, z) b(M, z) G(M, L) dM. \quad (3.38)$$

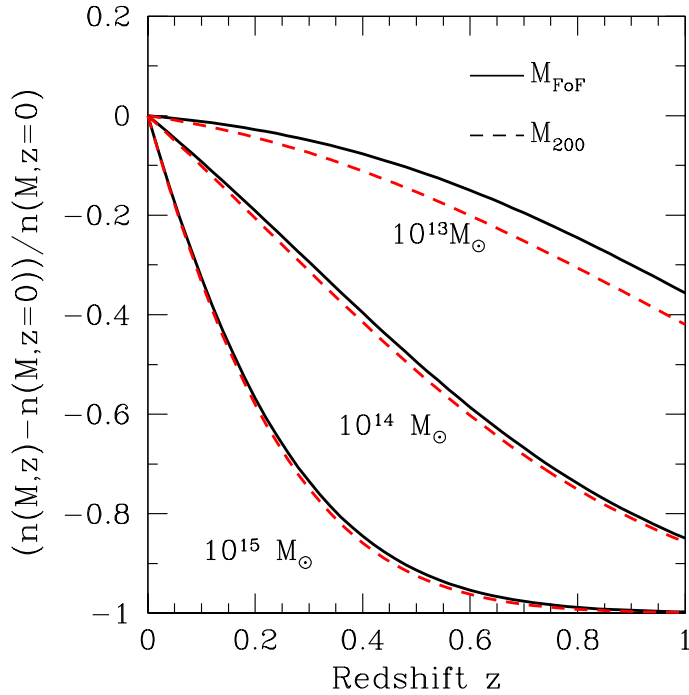


Figure 3.6: Halo abundance as a function of redshift evaluated at three different halo masses, compared to the respective value at redshift $z = 0$. Solid line represents the Jenkins et al. (2001) fitting formula. The dashed line is the Tinker (2007) fitting formula with $\Delta = 200$.

Where the function $G(M, L)$ arises from integrating the conditional luminosity function. Assuming a CLF with log-normal intrinsic scatter $\sigma_{\ln L}$ around the mean $\bar{L}(M)$ (Stanek et al., 2006, e.g.), the function $G(L, M)$ can be analytically written as

$$G(M, L) \equiv \int_L^\infty \Phi_{\text{cl}}(L|M; z) dL = \sqrt{\frac{\pi}{2}} \sigma_{\ln L} \text{Erfc} \left[\frac{\ln L - \bar{L}(M)}{2 \sqrt{\sigma_{\ln L}^2}} \right], \quad (3.39)$$

where Erfc is the complementary error function. The modeling of the clustering properties of X-ray galaxy clusters demands the knowledge of the underlying mass-luminosity relation and its intrinsic scatter. Moreover, observational effects (such as errors in the measurements of masses and luminosities) must be also considered. The formulations described in this section will be used to explain the clustering the the REFLEX II galaxy cluster survey in Chapters 4 and 5.

Light- cone effects

When modeling the measurements of the X-ray luminosity function, it is important to recall that observations are done through a past-light cone. Within it, the redshift evolution of the halo mass function can lead to significant distortions in the estimation of the luminosity function in comparison to what might be obtained if we evaluate the mass function at the effective redshift of the survey. Assuming no evolution in the mass-luminosity relation, the expected luminosity function when measured through a past-light cone can be written as in Equation (3.36) with an effective mass function $n_{\text{eff}}(M)$ given by

$$n_{\text{eff}}(M) = \frac{1}{V} \int_{z_{\text{min}}}^{z_{\text{max}}} dz \frac{dV}{dz} n(M, z). \quad (3.40)$$

where $z_{\text{min, max}}$ are the minimum and maximum redshifts of the survey and V is the volume covered by this redshift interval (Suto et al., 2000; van den Bosch et al., 2007). Light cone

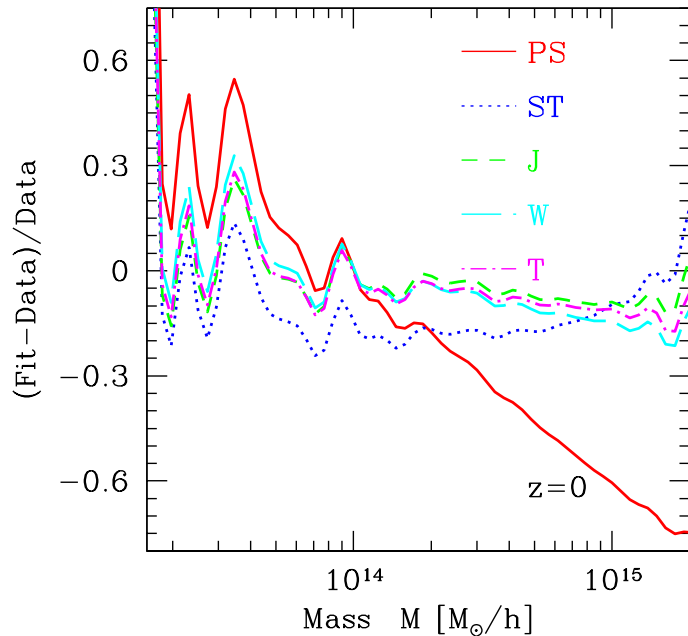


Figure 3.7: Ratio between the measured halo mass function at redshift $z = 0$ from the L-BASICC simulation with the different models and fitting functions from Press-Schechter Press and Schechter (1974), Sheth and Tormen (1999) (ST), Jenkins et al. (2001) (J), Warren et al. (2006) (W) and Tinker (2007) (T).

effects might be important if we are spanning broad redshift intervals or if we are interested in epochs when the mass function changes rapidly with redshift. In Fig. 3.6 we show the halo abundance as a function of redshift for four fixed values of halo mass, compared to the respective abundance at redshift $z = 0$. For low redshift surveys, $z \leq 0.1$ the mass function changes $\sim 10\%$ with respect to its value at $z = 0$. Reaching redshift $z = 0.2$, the mass function will display deviations of $\sim 5\%$ from its value at $z = 0$ for masses of $10^{12}M_\odot$ up to 50% deviations for haloes with masses of $10^{15}M_\odot$. Therefore it might be important to account for this effect when modeling the luminosity function of galaxy cluster catalogues, such like REFLEX II (see Chapter 4). Fig. 3.6 also shows the importance of the definition of the mass of a halo. In particular, note that the mass function in the low mass-limit ($\sim 10^{13}M_\odot$) can display differences of the order of 10% (in all the redshift range shown) when FoF and SO parameterizations are implemented.

3.6 Halo bias and two-point statistics from N -body simulations

In this section we present a brief analysis of the clustering measured from N -body simulations. We concentrate on the power spectrum and compare with the predictions of the halo model. Providing a dark matter halo catalogue, in Section 3.6.1 we measure the halo mass function. In Section 3.13 we show the behavior of the halo mass bias measured from the halo power spectrum.

3.6.1 The L-BASICC II N -body simulations

We have used $N_m = 50$ realizations from the Low resolution Baryon Acoustic Simulation at ICC (L-BASICC II) N -body simulations kindly provided by Raúl Angulo. We have used these simulations to construct a set of mock catalogues for the REFLEX II galaxy cluster survey (see chapter 4). Thus it is useful to consider some properties of this simulation such like the halo abundance, the halo-mass bias and the clustering. The latter has been analyzed

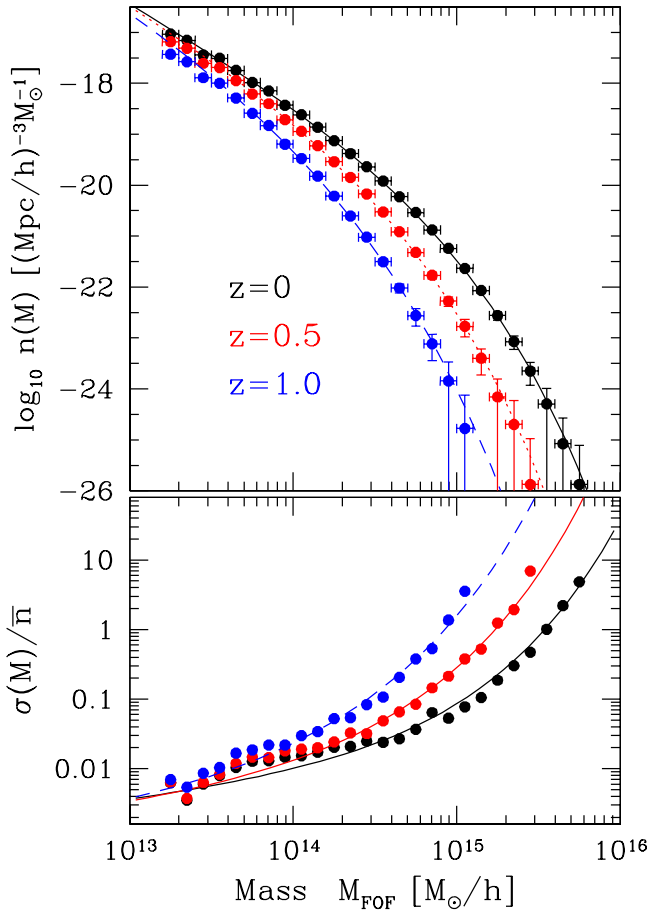


Figure 3.8: Top: L-BASICC II halo mass function (points with error bars) for the three outputs of the L-BASICC II simulation. On top of the measurements is shown the prediction based on the fit of Jenkins et al. (2001) for the corresponding redshift, following Equation (3.44). Bottom: measured variance (points) from the 50 realizations at each redshift compared with the theoretical prediction from Equation (3.46) (lines).

and discussed in real configuration space (correlation function) and in Fourier space (power spectrum) (Angulo et al., 2008; Sánchez et al., 2008).

These simulations have outputs at redshifts $z = 0$, $z = 0.5$ and $z = 1.0$ and consist of 448^3 dark matter particles. Dark matter haloes are identified via a FoF algorithm with a Plummer softening length of $50 \text{ kpc } h^{-1}$ and a linking length of $b = 0.2$ times the mean inter-particle separation. The halo mass resolution is $M = 1.75 \times 10^{13} M_{\odot} h^{-1}$ (corresponding to ten particles, and each particle with mass $m = \bar{\rho}(L/448)^3$) in a box with side $L = 1.34 \text{ Gpc } h^{-1}$, yielding $\approx 4.6 \times 10^5$ identified dark matter haloes in each realization. The low resolution allow us only to probe the high mass tail of the halo mass function. The cosmological parameters used in the simulation consist of a flat Λ CDM model with a matter energy-density parameter $\Omega_{\text{mat}} = 0.237$, baryon energy-density $\Omega_{\text{ba}} = 0.046$, dimensionless Hubble parameter $h = 0.73$, dark matter equation of state $w = -1$, the rms of mass fluctuations $\sigma_8 = 0.773$ and the scalar spectral index $n_s = 0.997$. These values are close to the latest constraints on cosmological parameters (Spergel et al., 2007; Sánchez et al., 2009; Percival et al., 2010).

The L-BASICC II mass function

The L-BASICC II halo mass function has been determined by counting the number of haloes ΔN in a mass bin ΔM in each realization. We compare the result of the mean $\Delta N/V\Delta M$ (where V is the volume of each simulation box) with the theoretical prediction of mass function

averaged in the corresponding bin,

$$\hat{n}(M_i, z) = \frac{1}{\Delta M_i} \int_{\Delta M_i} dM n(M, z). \quad (3.41)$$

This theoretical prediction is evaluated at a mass

$$M_i = \frac{1}{\bar{n}} \int_{\Delta M_i} dM M n(M, z), \quad (3.42)$$

The mean number density of haloes in the i -th mass bin $\bar{n}(M_i)$

$$\bar{n}_i = \int_{\Delta M_i} dM n(M, z). \quad (3.43)$$

In Fig. 3.7 we show the ratio of the measured halo mass function of the L-BASICC II simulation to some models and fits provided in the literature. This figure shows that the Press-Schechter mass function overestimates high mass haloes and underestimates the abundance of low-mass haloes. Note that this ratio decreases at low masses because of limited resolution. Therefore we can only describe main features on masses $\geq 10^{14} M_\odot/h$. In that range, deviations of $\lesssim 5\%$ are found in the Jenkins et al. (2001)(J), Warren et al. (2006)(W) and Tinker et al. (2008)(T) fitting formulae. The mass function of Sheth and Tormen (1999) acquires deviations of $\approx 40\%$. Figure 3.8 shows the measurements of the three outputs of the L-BASICC II simulations compared with Jenkins et al. (2001) mass function averaged in the same mass bins of the measurements. The points show the mean value for the $N_s = 50$ realizations,

$$\bar{n}(M_i) = \frac{1}{N_m} \sum_{j=1}^{N_m} \hat{n}_j(M_i), \quad (3.44)$$

where the error bars represent the 1σ variance of the ensemble given by:

$$\frac{\sigma^2(M_i)}{\bar{n}^2} = \frac{1}{N_m - 1} \sum_{j=1}^{N_m} \frac{(\bar{n}(M_i) - \hat{n}_j(M_i))^2}{\bar{n}(M_i)^2}. \quad (3.45)$$

We find that the Jenkins et al. (2001) fitting formula describes the measured halo mass function fairly well. Small discrepancies are found on both tails of the mass function. In the low mass end, resolution effects lead to an underestimation of the halo abundance. Being the fitting formulae of Jenkins et al. (2001) calibrated with N -body simulations with higher resolution, it has been proven to fairly well describe the halo mass function on larger dynamical ranges as the one explored by the L-BASICC II simulations. On the other hand, on the high mass end small discrepancies (below 1σ) can be observed, especially at redshift $z = 1$.

The expected variance in the halo mass function can be determined from the estimation of the bias treatment in Mo and White (1996). This is given as $\sigma_h^2(M, z) = b^2(M, z) \tilde{\sigma}_m^2(M)$ where $\tilde{\sigma}_m$ is the rms of the matter fluctuation given by Equation (2.11) evaluated at a scale $R = (3V/4\pi)^{1/3}$ where V is the volume of the simulation. Since we measure the mass function in bins of mass, the comparison must be done with averaged quantities in those bins. In addition, the discrete nature of the sample adds a shot noise contribution. Assuming a Poisson shot-noise, uncorrelated with the matter overdensity, this reads

$$\frac{\sigma_h^2(M_i, z)}{\bar{n}_i^2} = \hat{b}^2(M_i, z) \tilde{\sigma}_m^2(z) + \frac{1}{\bar{n}_i V}, \quad (3.46)$$

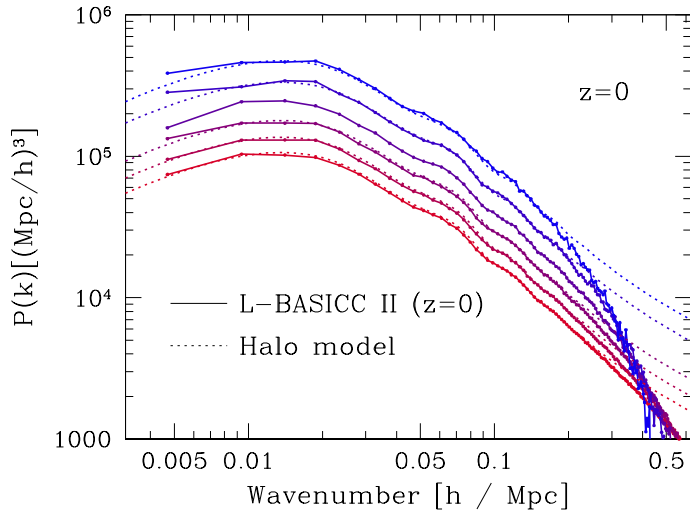


Figure 3.9: Mean halo power spectrum at $z = 0$ in real space for five different bins on halo mass: from lower to higher amplitudes ($13.7 \leq \log M \leq 13.9$, $13.9 \leq \log M \leq 14.1$, $14.1 \leq \log M \leq 14.3$, $14.3 \leq \log M \leq 14.5$, $14.5 \leq \log M \leq 14.7$). The dotted lines represent the halo model prediction with a scale independent bias.

where the effective bias reads

$$\hat{b}(M_i) = \frac{1}{\bar{n}_i} \int_{\Delta M_i} dM n(M, z) b(M, z). \quad (3.47)$$

The bottom panel of Fig.3.8 shows the comparison between the variance determined from the ensemble average of Equation (3.45) and the prediction from Equation (3.46), where a top-hat window function has been used, with the assumption of a linear halo-mass bias. This plot shows that Equation (3.46) can reproduce the ensemble averaged variance in the simulation with a very good agreement, especially at the high-mass end. A systematic increment (compared with the prediction) in the variance measured from the ensemble is detected on masses $M \lesssim 10^{14} M_\odot/h$. This deviation is stronger at redshift $z = 0$.

Note that the Jenkins et al. (2001) mass function has been calibrated with a numerical simulation (Evrard et al., 2002) that have the same characteristics of the L-BASICC II (except for the redshift). Therefore it is not surprising to find good agreement between the their fitting formula and the measurements we have shown here from the L-BASICC II simulation. The Tinker et al. (2008) mass function provides also a good fit for the mass range of the L-BASICC II. When required, we use the Jenkins et al. (2001) mass function.

3.6.2 Halo power spectrum

We have measured the halo power spectrum from the L-BASICC II simulations. In Appendix C we will discuss the power spectrum estimator in some detail. We divided the simulation box in 512^3 cells and assigned to each cell the fluctuation (Feldman et al., 1994)

$$F_i \equiv \frac{1}{N} (n_i^h - \alpha n_i^r), \quad (3.48)$$

where n_h and n_r^i denotes the number of haloes and random-distributed objects in the i -th cell and $\alpha = N_c/N_r$ is the ratio between the number of haloes and the number of random-distributed objects in each realization at the given cuts in mass. N is the normalization given by $N^2 = N_c^2/V$. The mass assignment was done using a triangular shape cloud mass assignment scheme (Hockney and Eastwood, 1981), corrected for aliasing effects. We implemented the FFTw algorithm (Frigo and Johnson, 2008) followed by an spherical averaging of the power spectrum $|F(\mathbf{k})|^2$. We finally subtracted the shot noise, given in this case by $S = \alpha(1 + \alpha)N_r V/N_c^2$

. Given the number of cells 512^3 , the fundamental mode is $\Delta k = 4.68 \times 10^{-3} h \text{Mpc}^{-1}$ and the Nyquist frequency is $k_{\text{Nyq}} = 1.198 h/\text{Mpc}$. The implementation of the estimator in the current case is simple, because no selection function is required. We have measured the power spectrum of sub-samples created by splitting the simulations in bins of mass for the three redshift outputs of the simulation, both in real and redshift space. The translation to redshift space is based on the distant-observer approximation, where we chose one axis as the line of sight and shift the coordinate of the haloes along this axis by $x \rightarrow x \rightarrow v_x/H_0$ while leaving the other two components unchanged.

In Fig. 3.9 the power spectrum in real space at redshift $z = 0$ is shown and compared with the halo model prediction. The prediction is based on the Jenkins et al. (2001) mass function and the corresponding Sheth and Tormen (1999) scale independent bias (which is a good description as we show below). We have implemented the non-linear power spectrum from the HALOFIT algorithm of Smith et al. (2003). As expected we observe that more massive objects are more clustered and thus more biased with respect to the low mass-bin. Perhaps the most interesting feature of the measurements is the suppression in the halo power spectrum due to halo exclusion effects. This shows how the clustering of the most massive haloes deviates from the expected non-linear theory at smaller wavenumbers (larger scales) compared to that of low mass haloes.

Fig. 3.10 shows the measurements of the power spectrum for the three redshift outputs of the L-BASICC II simulations and three bins of mass, determined in real and redshift space. Measurements are represented by the mean of the ensemble of 50 realizations. The shaded regions correspond to the 1σ variance drawn from the ensemble (see below). The bottom panels show the ratio of the halo power spectrum to the dark matter power spectrum. This measurements display the expected results from the halo-clustering. For instance, when we increase the redshift, the halo population is more biased with respect to the underlying dark matter. The bias is increased by the fact that the dark matter is less clustered at high redshift. In the bottom panels, each power spectrum has been divided by the corresponding dark matter power spectrum in real space (kindly provided by Francesco Montesano); these panels show that constant bias can be a good approximation on scales $0.02 < k (h \text{Mpc}^{-1}) < 0.1$ for $z = 0$ and especially in the low mass bins. For higher redshifts the measurements show a scale dependen in the explored range of wavenumbers. This is more prominent when we explore the clustering between high mass haloes. The scale dependent bias observed at $z = 1.0$ might be explained as a result of two effects. One one hand, the increasing bias indicates how the dark matter power spectrum decreases faster than the halo power spectrum. This means that non-linear effects are affecting dark matter haloes more strongly than the underlying dark matter. The trend is more evident for the high mass bin. On the other hand, while we approach to smaller scales, the exclusion effect starts to dominate the shape of the halo power spectrum and causes a rapid decrease of the amplitude of the power spectrum. The wavenumber where the ratio shown in Fig 3.10 changes its slope depends on the mass-bin. Low mass bins are less affected by exclusion than high mass haloes. At lower redshift, non-linearities affect both the dark matter and halo power spectrum in such a way that their ratio is constant up to the scale where the exclusion starts to play a role.

Covariance matrix

The covariance matrix of the power spectrum distribution is defined by

$$C(\mathbf{k}, \mathbf{k}') = \langle (\hat{P}(\mathbf{k}) - P(\mathbf{k}))(\hat{P}(\mathbf{k}') - P(\mathbf{k}')) \rangle,$$

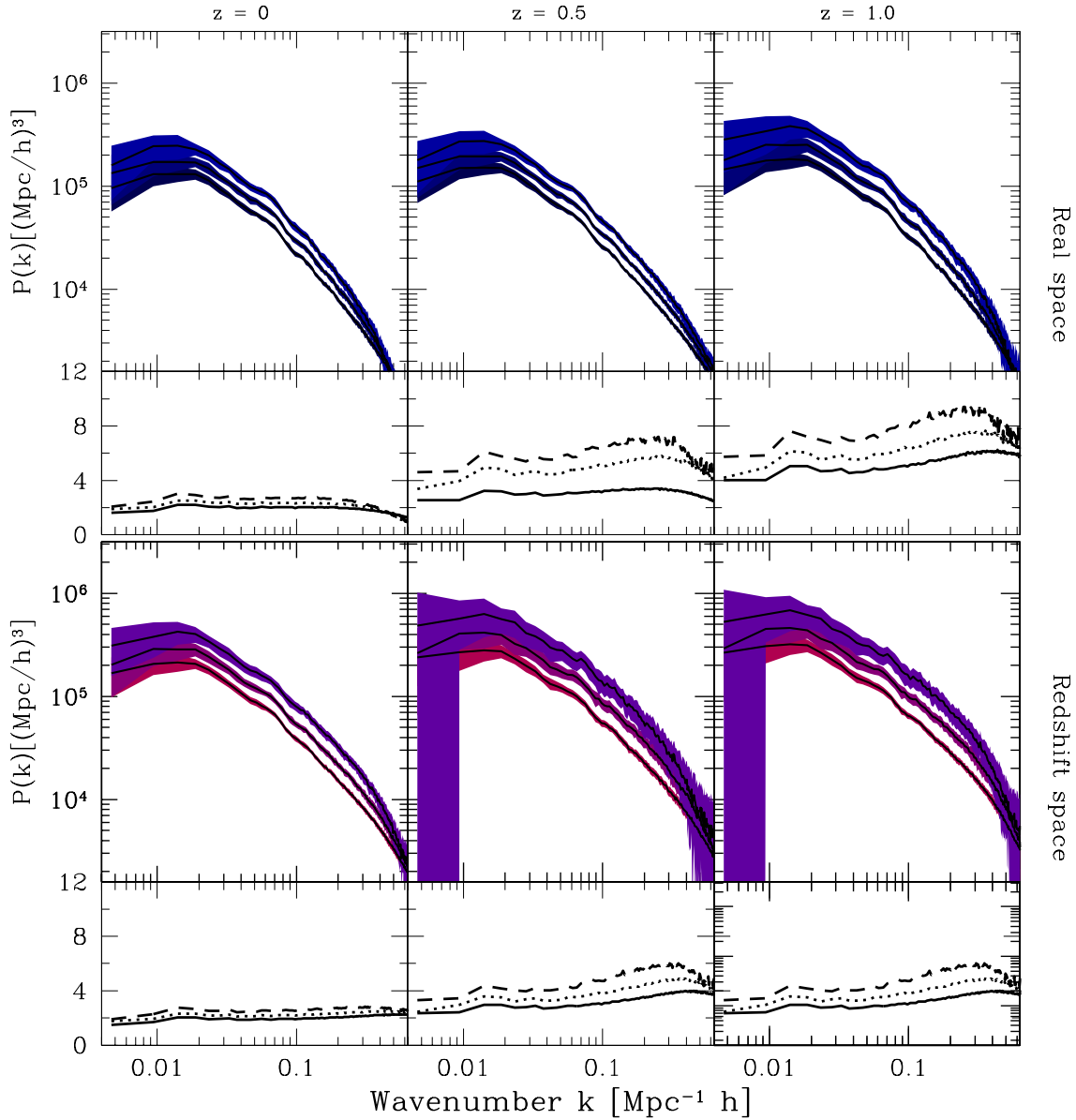


Figure 3.10: Mean halo power spectrum (solid lines) with 1σ variance drawn from the 50 L-BASICC II realizations (shaded area) for the three output redshifts and three bins of mass, ($13.7 \leq \log M \leq 13.9$, $13.9 \leq \log M \leq 14.1$, $14.1 \leq \log M \leq 14.3$), both in real and redshift space. The middle panels show the ratio of each power spectrum to the underlying matter power spectrum at the corresponding redshift.

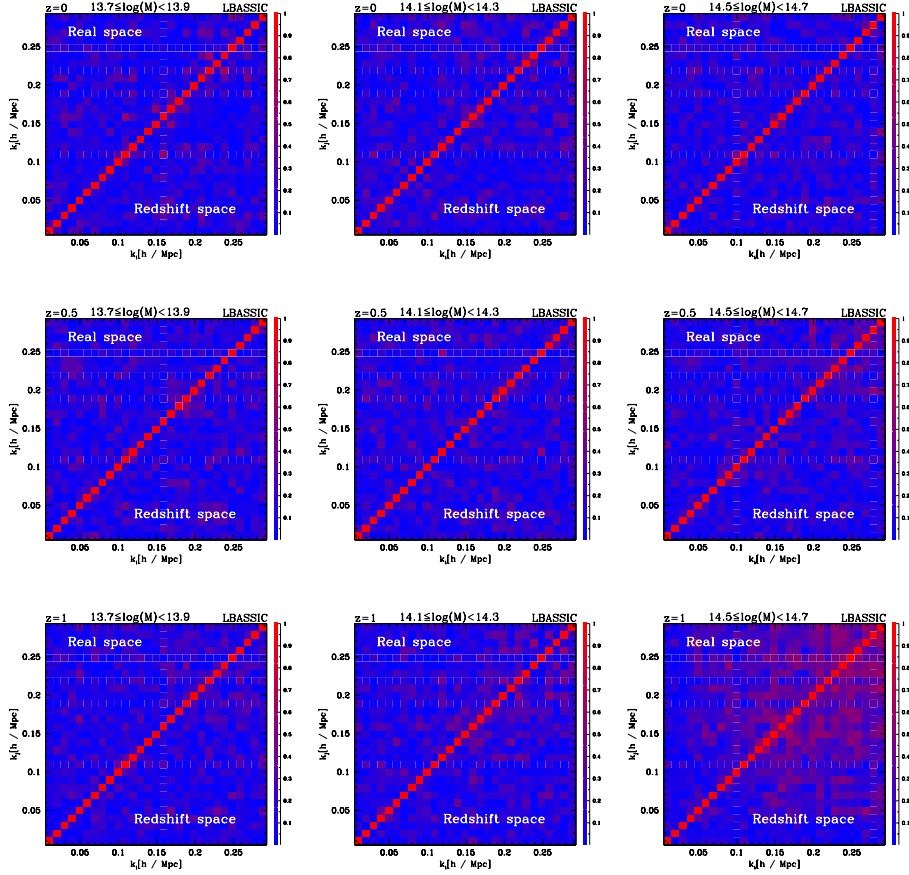


Figure 3.11: Correlation coefficients of the L-BASSIC II halo power spectrum for three different bins in mass and for the three output redshifts of the simulation, shown in real space and redshift space. As we increase the mass, non-diagonal elements are enhanced due to non-linear evolution, since the corresponding haloes are more clustered. This trend is also present if we fix the mass bin and increase the redshift, for haloes are more clustered at increased the redshift. Non-zero off-diagonal elements become important on small scales, where non-linearities are dominating where the signal is dominated by the halo exclusion.

where $\hat{P}(\mathbf{k})$ denotes the measured power spectrum and $P(\mathbf{k})$ is either the mean spectrum from the simulations or a theoretical expectation. This quantity is a key step towards obtaining constraints on cosmological parameters. Its determination represents a non-trivial task both from theoretical and numerical perspectives. From an analytical point of view, the covariance matrix can be decomposed into a Gaussian component, which is proportional to the measured power spectrum, and on off-diagonal contribution associated with the tri-spectrum. In the linear regime this off-diagonal contribution vanishes, while on intermediate and small scales the non-linear evolution of density fluctuations leads to coupling of Fourier modes, introducing non-Gaussianities and therefore a non-vanishing bi- and tri-spectrum (Matarrese et al., 1997; Verde and Heavens, 2001; Smith, 2008). Hence, to precisely determine the covariance matrix, it is necessary to have a model for these high-order statistics, together with redshift distortions, Fourier transform effects like aliasing (Cui et al., 2008), survey window function effects, beat-coupling effects (e.g. Rimes and Hamilton, 2006; Takayashi et al., 2008) and correlations introduced by bin-averages (Meiksin and White, 1999). From the numerical side,

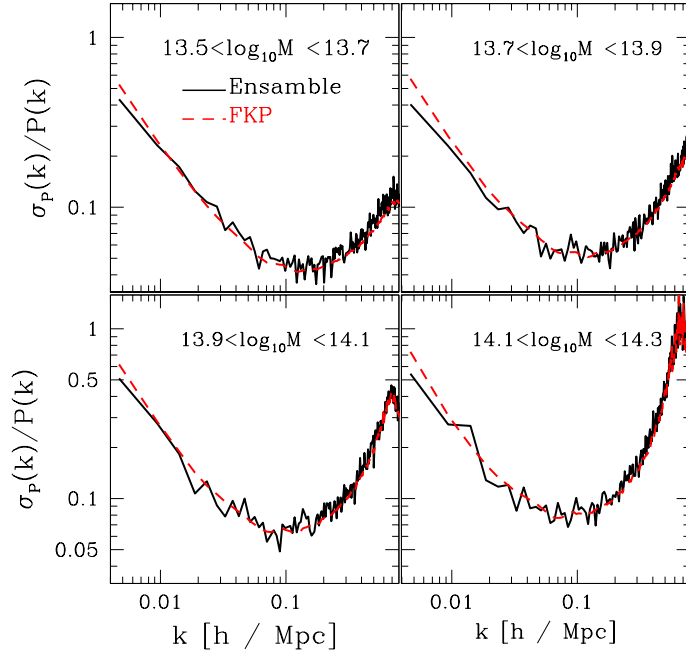


Figure 3.12: The fractional error in the power spectrum for dark matter haloes. The solid black line represents the variance drawn from the fifty realizations, while the green line represents the prediction based on Equation (C.16). The shot-noise and therefore the minimum of the fractional error increase with increasing the halo masses.

recent experiments with numerical simulations (Takayashi et al., 2008) have shown that large numbers of realizations are required for a precise determination of the covariance matrix and therefore precise determination of cosmological parameters: a 10% precision in the covariance matrix is achieved when 200 realizations are used. Following the analysis of Takayashi et al. (2008), 20% errors in the estimation of the L-BASICC II covariance matrix are expected.

Assuming that the Fourier modes have a Gaussian distribution around the mean power spectrum, after spherically averaging in shells with width greater than the correlation length and assuming that $\hat{P}(k)$ is constant at the same scale $\Delta k \ll k_i$, the bin-averaged covariance matrix is given (Feldman et al., 1994)

$$\hat{C}_{ij} = \frac{2\hat{P}(k_i)^2}{V_k V} \left(\hat{P}(k_i) + \frac{1}{\bar{n}} \right) \delta_{ij}^K, \quad (3.49)$$

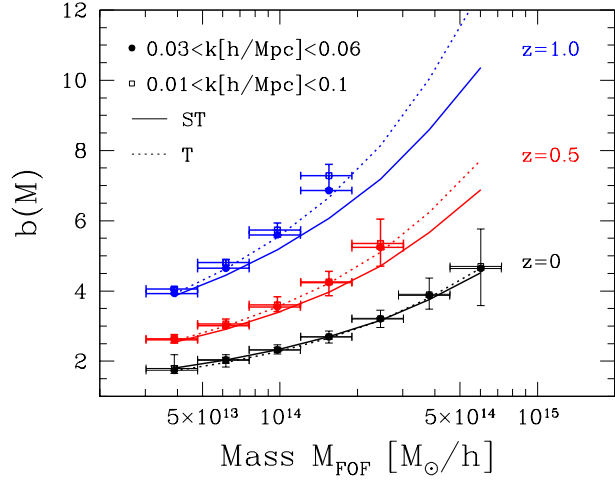
where $V_k \approx 4\pi^2 k_i^2 \Delta k / (2\pi)^3$ is the volume of the spherical shell of width Δk and V is the volume of the simulation. Equation (5.7) is mainly composed of two terms: the Poisson contribution, which is proportional to the measured spectrum $\hat{P}(k_i)$ and the shot-noise contribution $1/\bar{n}$. On large scales, where the signal dominates over the shot noise contribution, we get $V_{\text{eff}}(k) \sim V$, the volume of the sample.

This expression is only valid for a constant selection function, i.e., for a volume limited sample (see Appendix C). In Fig 3.11 we show the measurement of the correlation matrix r_{ij} , defined in terms of the measured bin-averaged covariance matrix \hat{C}_{ij} as

$$\hat{r}_{ij} = \frac{\hat{C}_{ij}}{\sqrt{\hat{C}_{ii}\hat{C}_{jj}}} \quad \hat{C}_{ij} = \frac{1}{N_m - 1} \sum_{k=1}^{N_m} (\hat{P}_i^{(k)} - \bar{P}_i) (\hat{P}_j^{(k)} - \bar{P}_j)$$

We show some examples of the three redshift outputs. This shows how the Fourier modes measured from the high-mass bins are more correlated than the low mass bins, especially on very small scales. At the lowest redshift and a fixed mass-bin, the off-diagonal elements of

Figure 3.13: Measurements of the halo-mass bias for the three output redshifts of the L-BASICC II simulations. We have also shown the prediction from the bias of Sheth and Tormen (1999) and the bias parameterization of Tinker et al. (2010). The latter, being designed for spherical-overdensity masses -as a function of Δ (see Equation 3.15)-, has been implemented with $\Delta = 200$ obtaining a good agreement.



the correlations measured in real space are almost identical to the measurements in redshift space. The most drastic change in the behavior of the off-diagonal elements takes place at the higher mass bin for the redshift output $z = 1$.

Equation (5.7) also provides an expression for the bin-averaged variance in the power spectrum, $\sigma^2(k_i) = \hat{C}_{ii}$. In Fig. 3.12 we have shown the fractional error σ/P for the L-BASICC II dark matter haloes in four bins of mass. The prediction for the variance taken from Equation (5.7) is in good agreement with the variance drawn from the measurements of power spectrum in the L-BASICC II simulation. This agreement has been also observed in configuration space through the correlation function analysis (Sánchez et al., 2008).

Halo-mass bias

Using the measurements of the halo power spectrum shown in Fig. 3.10, we have measured the halo-mass bias defined as

$$\hat{b}^2(z) = \frac{P_{\text{halo}}(k, z)}{P_{\text{mat}}(k, z)}. \quad (3.50)$$

By direct inspection, we have confirmed that the ratio defined in Equation (3.50) is compatible with a scale-independent bias within the range of wavenumbers $0.02 < k \text{ (} h \text{Mpc}^{-1}\text{)} < 0.1$. The measurements are shown in Fig 3.13, where we display the bias measured in the three redshift outputs. Furthermore, we have made measured in two ranges of wavenumbers, $(0.03 < k < 0.06)$ and $0.01 < k < 0.1$. In both cases the measurements were done by minimizing the χ^2 :

$$\chi^2 = \sum_{i,j} (\hat{b}^{(i)} - b^{\text{th}}) C_{ij}^{-1} (\hat{b}^{(j)} - b^{\text{th}}),$$

with $\hat{b}^{(i)} = \hat{b}^{(i)}(k, \bar{M}_i)$ and $b^{\text{th}} = b^{\text{th}}(\bar{M}_i)$ is constant, by assumption. The the covariance matrix drawn from the $N_m = 50$ realizations,

$$C_{ij} = \frac{1}{N_m - 1} \sum_{k=1}^{N_m} (\hat{b}_i^{(k)} - \bar{b}_i) (\hat{b}_j^{(k)} - \bar{b}_j),$$

where

$$\bar{b}_i = \frac{1}{N_m} \sum_j \hat{b}^{(j)}(k_i, \bar{M}_i)$$

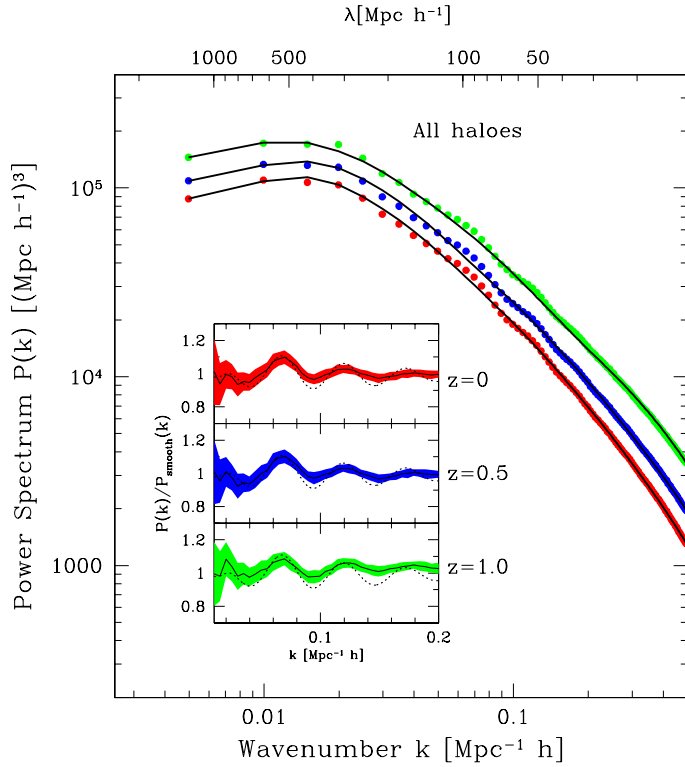


Figure 3.14: Halo power spectrum and its smooth version for different redshifts : $z = 0$ (red), $z = 0.5$ (blue) and $z = 1.0$ (green). In the inner panel the ratio of the measured power to its smooth version is shown.

is the mean bias. The best fitting values are determined with a Markov Chain Monte Carlo (MCMC) algorithm. Fig. 3.13 also contains the prediction for the halo-mass bias derived from the Sheth and Tormen (1999, ST) and the Tinker et al. (2010, T) bias schemes, using Equation (3.47). We have observed that a good agreement in the three redshift outputs is obtained when the predictions of the T fitting formulae are implemented with $\Delta = \Delta_{\text{vir}}$ given by Equation (3.16). The ST bias is a good description of the measurements only at redshift $z = 0$ for all mass bins, and tend to be a good description at higher redshifts only for the low mass haloes. The fitting formulae of Tinker et al. (2010) gives a good description of the measurements for the three redshifts shown. The ST formulae tend to underestimate the bias especially at high redshifts, where explicit dependence on the critical overdensity value must taken into account (for these plots we have set δ_c constant). Note that the agreement between our measurements and the fitting formulae of Tinker (2007) is not to be expected, since the fitting formulae of Equation (3.15) is designed for masses defined with the SO halo-finder, while the masses of the L-BASICC simulations are FoF masses.

BAOs signature

The presence of BAOs in the L-BASICC II simulation has been already explored in Fourier space by Angulo et al. (2008) and configuration space by Sánchez et al. (2008). The motivation to study this feature from the simulations lies in the fact that it is possible to explore the physical effects than induces distortions in the acoustic signal and therefore provide a more precise framework in order to model BAO. This allows one to use the BAO signature as a standard ruler from which cosmological parameters can be constraint in galaxy redshift surveys.

The use of BAOs as stander ruler dates few years back, after the discovery of the acoustic

peak in the correlation function of the SDSS LRG sample by Eisenstein et al. (2005). The detection of the BAOs in the power spectrum measured from the same sample was confirmed few years later by Percival et al. 2007b; Hütsi 2007. One technique to extract cosmological information from the acoustic wiggles in the power spectrum consists of dividing the measurement of power by an smoothed version of the same measurement $P_{\text{smooth}}(k)$ (Percival et al., 2007a, 2010), such that the effects large-scale gradients due to scale dependent bias or non-linearities are cancelled out. The model for the resulting signal most, however, take into account that the BAOs signal is further damped with respect to the linear prediction due to non-linearities. Numerical simulations (e.g. Seo et al., 2008) have shown that the damping in the BAOs in a spherically averaged power spectrum can be described by a Gaussian kernel $G(k) = \exp(-k^2\sigma^2/2)$ with $\sigma \approx 10\text{Mpc}h^{-1}$, such that the spectrum divided by its smoothed version $R(k) = P(k)/P_{\text{smooth}}(k)$ can be modeled as $R(k; \alpha, \sigma) = 1 + G(\alpha k; \sigma)(R_{\text{lin}}(\alpha k) - 1)$ where $R_{\text{lin}}(k)$ is the ratio of the linear power spectrum to its smoothed version and α is the stretch factor defined in Equation (2.17). The behavior of the set of parameters $\{\alpha, \sigma\}$ has been explored by Angulo et al. (2008) and Montesano et al. (2010).

For illustration we have here created a smoothed version of halo power spectrum measured in three mass bins. This is obtained by doing a cubic spline of the measurement in 10 logarithmic bins between $k = 0.0049h\text{Mpc}^{-1}$ and $k = 0.49h\text{Mpc}^{-1}$. The result is shown in Fig. 3.14 where we show the BAO signal for the three redshifts of the simulation and all the haloes in the simulation. The dotted lines in the inner boxes represent the curve $R(k; \sigma)$ obtained using the formulae of Eisenstein and Hu (1998) using the same cosmological parameters of the simulation (i.e., setting the stretch parameter $\alpha = 1$, see Equation 2.17). The comoving volume of the L-BASICC II simulation $2.41(\text{Mpc}h^{-1})^3$ allows for a statistically significant detection of the BAOs and currently ongoing galaxy redshift surveys such as BOSS⁹ aim for probing large volumes by measuring spectroscopic redshift of LRG and generate a precise detection of the acoustic oscillations.

3.6.3 Marked statistics with haloes

Marked correlation functions have been introduced into the LSS analysis (e.g. Sheth, 2005) as an attempt to explore the dependence of the galaxy clustering on intrinsic properties or on their environment, in order to contribute to the understanding of the process of galaxy formation and galaxy clustering. The basic idea behind the mark correlation function is to measure the clustering of a galaxy property, -the mark-. Marked-correlation functions have been used by White and Padmanabhan (2009) to break the usual degeneracies between the parameters σ_8 and Ω_{mat} . Skibba and Sheth (2009) measured the marked correlation function for galaxies in SDSS using the galaxy luminosity as a mark and modeled this statistics in the context of galaxy HOD and halo model (Cooray and Sheth, 2002). Sheth et al. (2005) uses the colour of the SDSS galaxies as a mark and Skibba et al. (2006) extended the analysis to the luminosity weighted mark correlation functions. These measurements concentrate on small scales, $r \lesssim 30\text{Mpc}h^{-1}$, where the clustering of marks such as the galaxy luminosity or colour have been detected. However no attempt to explore the behavior of the marked correlation function on the scales where the acoustic peak might be detected has been made to date, as far as I am aware of. We therefore measured the Marked correlation function from the L-BASICC II simulations on scales $60 < r < 150\text{Mpc}h^{-1}$ and try to detect the BAOs signal, as will be shown below.

⁹<http://www.sdss3.org/cosmology.php>

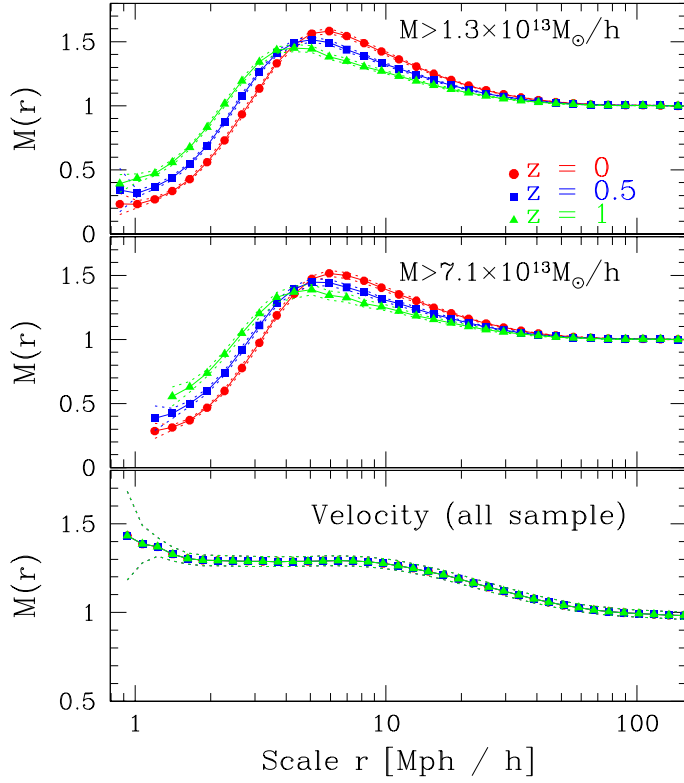


Figure 3.15: Marked correlation function using the halo mark mass as a mark. Two different samples are taken with different minimum masses, shown at the three outputs of the simulation. In the third panel we have used the velocity of the haloes has a mark

The second order mark correlation function is defined as the ratio of the mark-weighted correlation function to the (unmarked) correlation function $\xi(r)$. Following (Skibba and Sheth, 2009; Sheth, 2005), given the positions \mathbf{x} of galaxies or clusters in the sample, this is written as

$$M(r) = \frac{\sum_{i,j} m(\mathbf{x})m(\mathbf{y})\mathcal{I}(r - |\mathbf{x} - \mathbf{y}|)}{\bar{m}^2 \sum_{i,j} \mathcal{I}(r - |\mathbf{x} - \mathbf{y}|)} \equiv \frac{1 + W(r)}{1 + \xi(r)}, \quad (3.51)$$

where the sum is done over pairs separated by scales in the range $r - \frac{1}{2}\Delta < |\mathbf{r}_i - \mathbf{r}_j| < r + \frac{1}{2}\Delta$, \bar{m} is the mean mark in the sample and the function $\mathcal{I}(x) = 0$ for $x \neq 0$. This sum $\sum_{i,j} \mathcal{I}(r - |\mathbf{x} - \mathbf{y}|)$ denotes the typical number of pairs separated by a distance r , which is proportional to $\bar{\rho}(1 + \xi(r))$, where $\xi(r)$ is the correlation function. Given this, the quantity $W(r)$ is interpreted as the correlation function between pairs that we weighted with the property $m(\mathbf{x})$. One advantage of this statistic is the fact that, contrary to the measurements of correlation function or power spectrum, there is not need to construct a random catalogue. This because we can estimate the marked-correlation function using the quantity $DD(r)/RR(r)$ ¹⁰ to estimate $1 + \xi(r)$ (e.g. Hamilton, 1993) and the analogous quantity $WW(r)/RR(r)$ for $1 + W(r)$, thus canceling the random contribution (Skibba and Sheth, 2009). This of course is not valid if we implement more precise (and un-biased) estimators for the correlation function, where the combination of pairs of random-distributed objects and pairs form the real sample cannot be separated in the way aforementioned.

Returning to Equation (3.51), we measure the correlation $M(r)$ via

$$\hat{M}(r) = \frac{\sum_{i,j} m_i m_j}{\bar{m}^2 n(r)}, \quad (3.52)$$

¹⁰This the number of pairs separated by a distance r in the data sample divided by the number of pairs separated by the same scale in the random catalogue

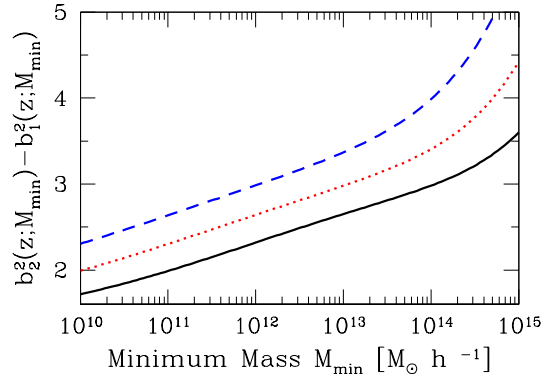


Figure 3.16: Bias factors in the marked-correlation function $b_2^2 - b_1^2$ as a function of the minimum mass and for the three redshift outputs of the L-BASICC II simulation. Solid line ($z = 0$), dotted line ($z = 0.5$) and dashed line ($z = 1.0$).

where $n(r)$ is the number of halo pairs in the interval $r - \frac{1}{2}\Delta r \leq r_{ij} < r + \frac{1}{2}\Delta$. In Fig. 3.15 the marked correlation function is shown for the haloes in the L-BASICC II simulation. In all plots we have used the variance in the mark drawn from the ensemble of realizations in order to set error-bars on the measurements. In the first two panels of that figure we have used the halo mass as the mark; the trend for the three redshifts and for the two mass cuts is similar: on large scales the marked correlation approaches to $M(r) \rightarrow 1$. This means that there is no correlation between the mark (the values of the mass) in pairs separated by scales greater than $r \approx 50 \text{Mpc } h^{-1}$. As long as we approach to smaller scales, the correlation grows reaching a maximum at a scale between $r \approx 5 - 6 \text{Mpc } h^{-1}$, mildly changing for the three redshifts shown. This increase is interpreted as that pair separated by these scales are predominantly dominated by massive haloes (i.e, haloes with masses above the mean of the sample). The maximum of the marked-correlation function can roughly tell us about the diameter of the most massive haloes. This explains the fact that the maximum in $M(r)$ does not change by imposing a minimum halo mass. Note however that the maximum moves with redshift: this is a consequence of the fact that the most massive haloes are collapsing at later times.

On scales smaller than $\approx 0.9 - 2.5 \text{Mpc } h^{-1}$ the marked correlation is $M(r) < 1$, meaning that haloes separated by these scales are likely to have masses below the mean mass of the sample. This decreasing behavior of the marked-correlation function is associated with the exclusion effect. This can be seen by comparing the marked-correlations on small scales for the two cuts in mass shown in Figure 3.15 the higher mass cut ($7.1 \times 10^5 M_\odot h^{-1}$) does not take into account haloes with virial radii of the order of $1.3 \text{Mpc } h^{-1}$. In the third panel we have implemented the peculiar velocity of the center of mass of the dark matter haloes as a mark. A smooth transition from the asymptotic value $M(r) = 1$ to a constant correlation $M(r) \approx 1.45$ is observed in the range $(1 \lesssim r \lesssim 10) \text{Mpc } h^{-1}$. This transition marks the scales where the motion of haloes due to peculiar velocities (associated with matter overdensities) are relevant, i.e, on scales $\geq 100 \text{Mpc}$, gravitational collapse cannot induce correlation between peculiar velocities.

Regarding the shape of the marked correlation function, we expect that features inherent to gravitational processes as the acoustic peak, while being present in the correlation function, will also be present in the weighted correlation function such that their ratio will still contain its signal but, perhaps, with an smaller amplitude as found in the correlation function. Assuming that the two point correlation function for haloes of different masses can be separated as $\xi_{\text{hh}}(r; M, M') = b(M)b(M')\xi_{\text{mat}}(r)$ (i.e, on linear scales), we can write $W(r) = b_2^2 \xi_{\text{mat}}(r)$ (see Appendix B) together with $\xi(r) = b_1^2 \xi_{\text{mat}}(r)$ such that the marked correlation function for haloes with masses greater than M_{min} and mark $\phi(M)$ (a function of the halo mass) can be

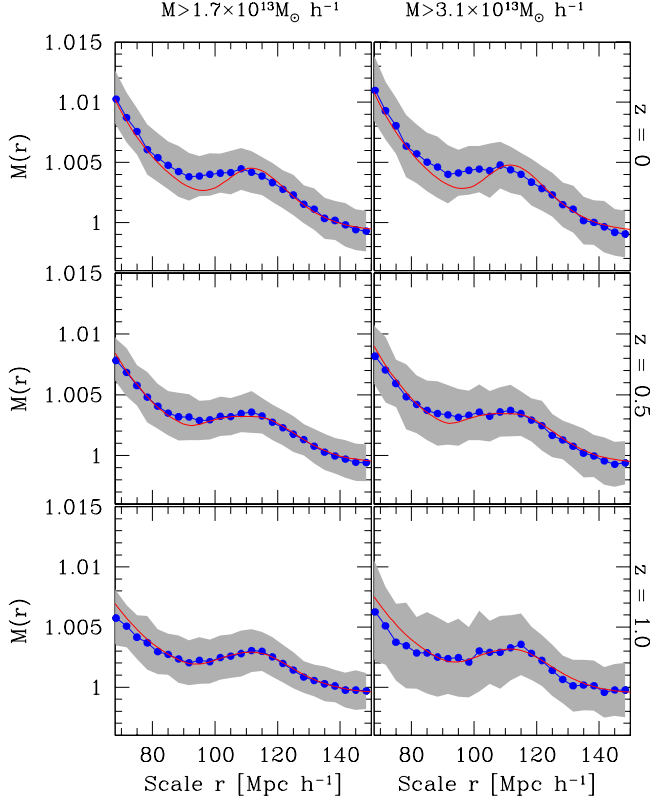


Figure 3.17: Marked correlation function on the scales of the acoustic peak, where we have used the halo mass as the mark. The three outputs of the simulations are used for two different minimum mass. The points represent the mean of the 50 simulations, and the grey shaded area represents their 1σ variance. The solid lines shows the prediction from Equation (3.53).

written as

$$M(r, z) = 1 + \left[b_2^2(z; M_{\min}) - b_1^2(z; M_{\min}) \right] \frac{\xi_{\text{mat}}(r, z)}{1 + b_1^2(z; M_{\min}) \xi_{\text{mat}}(r, z)}, \quad (3.53)$$

where b_1 and b_2 effective biases are given by

$$b_1(z; M_{\min}) = \frac{\int_{M_{\min}} dM n(M) b(M)}{\int_{M_{\min}} dM n(M)} \quad b_2(z; M_{\min}) = \frac{\int_{M_{\min}} dM \phi(M) n(M) b(M)}{\int_{M_{\min}} dM \phi(M) n(M)}. \quad (3.54)$$

Therefore, on these scales, the modeling of the marked correlation function reduces to the modeling of the matter correlation function. Since the the scale independent behavior of the halo-matter bias has been shown to be a good description on large scales, Equation (3.53) can be considered to provide a good theoretical framework to model the acoustic peak in the marked correlation function. Note that the amplitude of the acoustic peak in this statistics can be related to the differences between the two terms $b_{1,2}$. In Fig. 3.16 we show the behavior of the difference $b_2^2 - b_1^2$ as a function of the minimum mass and for the three redshift outputs of the L-BASICC II simulation.

In Fig. 3.17 we show the marked correlation function using the halo mass as the mark, $\phi(M) = M$, for two minimum masses and the three outputs of the simulation. The presence of the acoustic peak in the marked correlation function is readily confirmed, with an amplitude almost negligible compared to the maximum of the correlation found on small scales. It is nevertheless statistically significant as can be concluded by observing the $1 - \sigma$ variance (shaded region) drawn from the ensemble of simulations. The solid line in that figures represent the prediction from Equation (3.53) from which the non-linear matter correlation function is determined from the Fourier transform of the non-linear power spectrum taken

from the halo fit (Smith et al., 2003). The power spectrum is de-wiggled in order to mimic the damping in the acoustic scale due to non-linearities.

The detection of the acoustic peak in the marked correlation function shown in this work motivates the implementation of this statistics in large galaxy redshift surveys such as SDSS.

3.7 Summary

In this chapter we have explored some of the most relevant properties of dark matter haloes. We have focused on their abundance, clustering and bias with respect to the underlying matter distribution. We have also described the halo model for clustering in order to apply it to the galaxy clustering analysis.

Our study on the dark matter halo abundance was developed using a suit of 50 dark matter halo simulations, the L-BASICC simulations, with outputs at three different redshifts. For the purposes of the present and future work, the knowledge of the halo mass function is of great relevance in order to explore the abundance of galaxy clusters and extract cosmological implications from measurements of the cluster luminosity function (e.g. Wu et al., 2010). We have measured the halo abundance and the halo mass bias, the latter determined by measuring the measurements of the halo power spectrum. We have compared our results with the available models and fitting formulae based on N -body simulations with higher resolution, finding good agreement. As expected (due to the features of the simulations we have used) we have seen that the fitting formulae of Jenkins et al. (2001) can describe the measurements of the halo abundance up to a $\sim 10\%$ precision in the dynamical range of the L-BASICC II simulations. We observed that the parameterization of Tinker et al. (2008) also provides a good fit, though we can not strictly make comparisons due to the FoF masses of our simulations.

We have explored the halo clustering properties through measurements of the halo power spectrum and the marked correlation function. We have compared the predictions from the halo-model with the appropriate abundance and halo-mass bias with the measurements of clustering, showing good agreement. Concerning the marked correlation statistics developed on the L-BASICC simulations, it is worth to note that we have detected the BAO signature in the marked correlation function of dark matter haloes. This suggest that this feature might be also detected in large galaxy redshift surveys as the luminous red galaxy sample from SDSS. It is then interesting to explore the possible signature of BAO with marked statistics using different galaxy properties, such as their luminosity or color.

The knowledge of halo abundance and halo-mass bias is a necessary step to make predictions on the clustering properties of galaxy clusters as we will discuss in the following chapters. Furthermore, no less than 10% deviations in this quantities are accepted in currently on-going and future galaxy-cluster surveys when precise constraints on cosmological parameters are expected (e.g. Wu et al., 2010).

The REFLEX II galaxy cluster catalogue: sample and mock catalogues

4.1 Introduction

In this chapter we introduce the new ROSAT-ESO Flux Limited (REFLEX II) galaxy cluster survey, on which our large-scale structure analysis will be developed. Based on the properties of the REFLEX II sample and using N -body simulations, we construct a set of REFLEX II mock catalogues, which will be intensively used to test the statistical methods used in the clustering analysis. In Section 2.6 we briefly introduce some properties of galaxy clusters. In Section 4.2 we introduce the REFLEX II sample and show some properties of the data-set. In Section 4.3 we determine the X-ray luminosity function of the REFLEX II sample and create the survey selection function. Having built the REFLEX II selection function, we proceed to construct a set of 100 independent mock catalogues using the N -body simulations described in Chapter 3. The construction of these mock catalogues is described in Section 4.4 together with the description of the main properties of the resulting REFLEX II mock catalogues. We finish this chapter with a summary.

4.2 The REFLEX II galaxy cluster catalogue

4.2.1 REFLEX II sample

The REFLEX II is an extension of the ROSAT-ESO Flux Limited (REFLEX) REFLEX catalogue (Böhringer et al., 2004) based on RASS (ROSAT All Sky Survey) observations (Truemper, 1993), which are complemented with follow-up observations as described by Guzzo et al. (2009) yielding spectroscopic redshifts for $\approx 90\%$ of the sample. The lower limiting flux in the REFLEX II sample (1.8×10^{-12} erg s $^{-1}$ cm $^{-2}$) permits the inclusion of 464 clusters on top of the REFLEX sample, yielding a total of 911 clusters. This substantial increase in the number of X-ray galaxy clusters is translated into a better statistical analysis concerning mainly the cluster-two point statistics.

The REFLEX II is to date the largest and most statistically homogeneous galaxy cluster survey. Though it will be overcome by future X-ray surveys, the legacy of the REFLEX

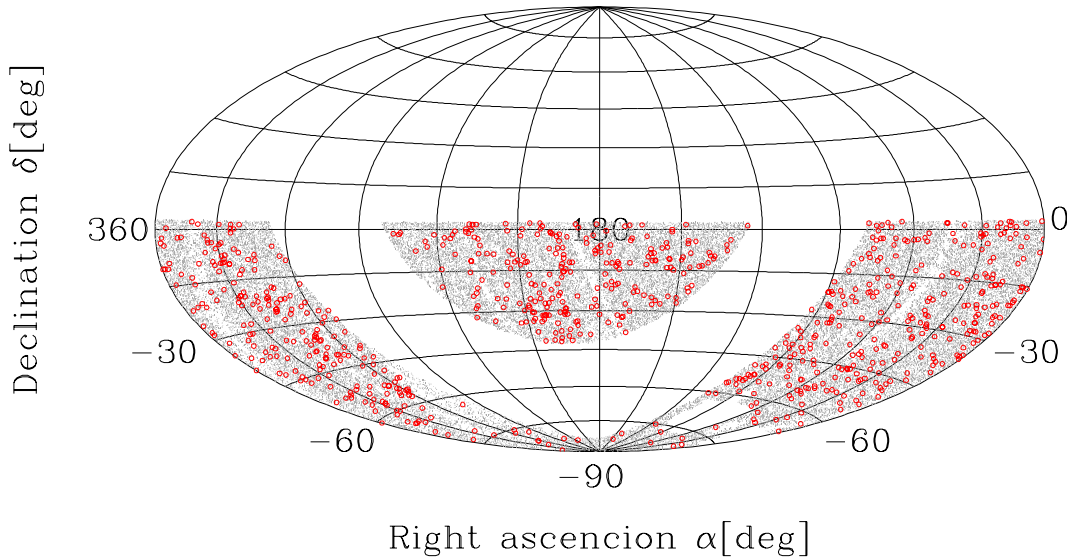


Figure 4.1: Aitoff projection of the REFLEX II galaxy clusters in equatorial coordinates. Red points represent the position of the REFLEXII clusters. Green points represents 10% of the random catalogue. The empty band corresponds to the cut (band within $|b_{II}| < 20^\circ$) of the Milky Way, while the two squared regions in the southern hemisphere are the Magellanic clouds.

catalogue and the improvements presented in this and future publications concerning the REFLEX II have set the REFLEX survey as a inevitable reference when talking about cosmology with galaxy clusters.

In Fig. 4.2 we show some properties of the REFLEX II catalogue, which we will refer to throughout this chapter. A more detailed description of the sample construction and the derivation of the cluster parameters is given in a forthcoming paper (Böhringer et al. in preparation). We just provide here a brief description of these measurements and calculations. The REFLEX II catalogue contains 911 clusters covering 13924 deg^2 (4.24 sr , 33.75% of the sky) in the southern hemisphere ($\delta < 2.5^\circ$) where the Milky Way, (band within galactic latitude $|b_{II}| < 20^\circ$), the Large (244.4 deg^2) and the Small (79.8 deg^2) Magellanic Clouds have been excised (see Fig. 4.1) to avoid contamination due to stars and regions with high X-ray absorbing neutral hydrogen column density and high extinction in the optical band. The limiting flux of the REFLEX II sample is $1.8 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2}$ and includes 464 new clusters in addition to the REFLEX sample, which had 447 clusters to a limiting flux of $3 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2}$. The detection technique is the same as that developed for the REFLEX sample (Böhringer et al., 2004; Guzzo et al., 2009), with optical identifications and spectroscopic redshifts for 760 clusters. The missing $\approx 6\%$ of the redshifts are currently being measured in follow-up observations at La Silla (Chile). We estimate that our results won't change dramatically due to this incompleteness. The sample with redshifts span luminosities in the range $4.9 \times 10^{40} < L_X < 1.96 \times 10^{45} \text{ ergs}^{-1} \text{ h}^{-2}$.

The accessible sky area defining the REFLEX II sample has been divided in $N_{\text{pix}} = 13902$ pixels each with area $\approx 1 \times 1 \text{ deg}^2$. Each pixel has equatorial coordinates (α_i, δ_i) and a limiting flux F_{lim}^i . This generates N_{pix} different selection functions as shown in Fig 4.3.

A description of the determination of the cluster luminosities can be summarized as follows:

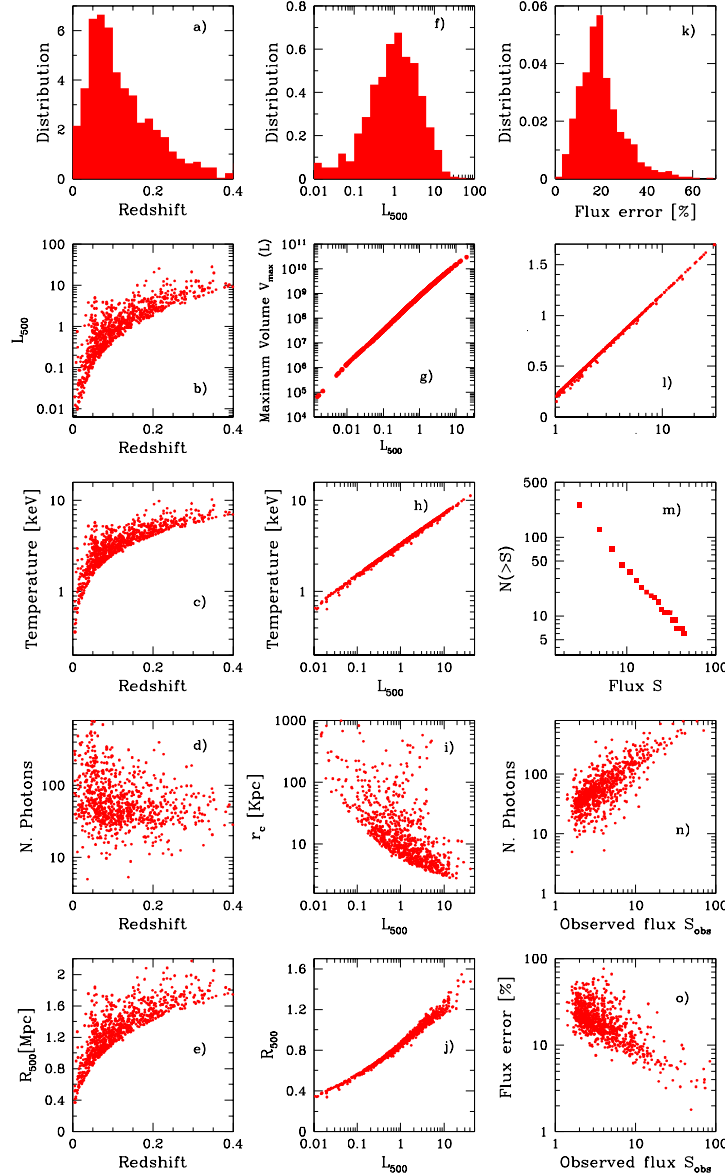


Figure 4.2: Some properties of the REFLEX II sample: a) Redshift distribution b) X-ray luminosity as a function of the redshift c) Temperature as a function of the redshift d) Number of detected photons as a function of the cluster redshift e) Radius R_{500} as a function of the redshift f) X-ray luminosity distribution g) Maximum volume as a function of the luminosity. h) Temperature vs X-ray luminosity i) Core radius for the β -model as a function of the X-ray luminosity j) Radius R_{500} vs X-ray luminosity k) Distribution of flux errors l) Observed flux vs Flux in the *ROSAT* energy band m) Number counts n) Number of photons vs flux o) Flux error vs Observed flux Properties of the REFLEX II sample.

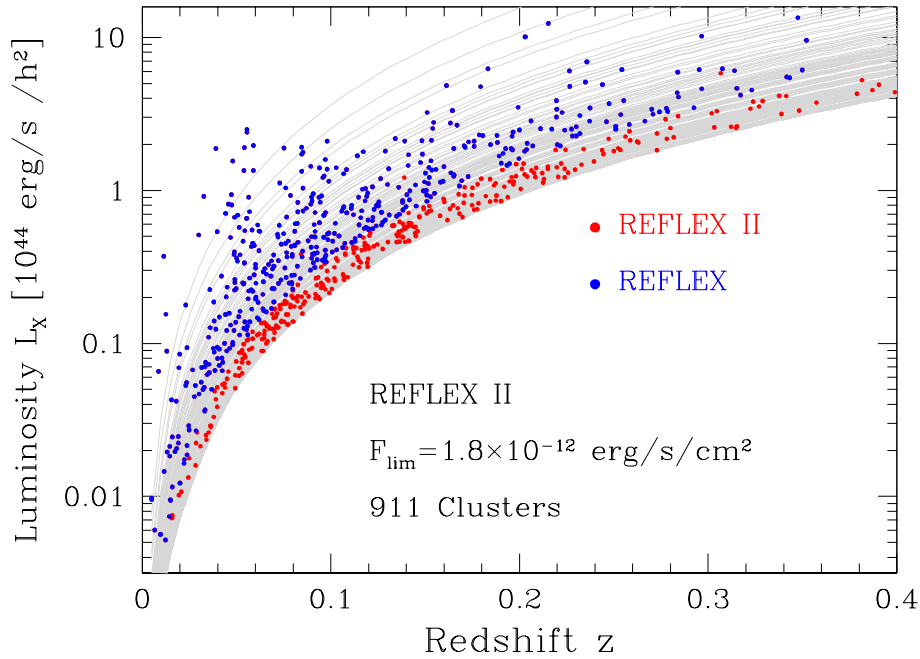


Figure 4.3: Luminosity-redshift diagram for the REFLEX (blue points) and the REFLEX II sample (red points) which includes the REFLEX clusters. The grey lines represent the limiting luminosity as a function of redshift for the N_{pix} pixels in the REFLEX II mask.

- Source counts for galaxy clusters in the RASS have been determined in the 0.5 to 2 keV *ROSAT* energy band by means of the growth curve analysis method described in Böhringer et al. (2000). The growth curve method is tailored to maximize the aperture in which the source counts are determined.
- The count rate (obtained by reference to the exposure maps of the RASS) is then converted to a nominal flux, F_n , by means of XSPEC (Arnaud, 1996) assuming a MEKAL¹ plasma model for a cluster temperature of $T_{\text{gas}} = 5$ keV, at redshift $z = 0$ a metallicity of 0.3 solar and the hydrogen column density N_H taken from the measurements of Dickey and Lockman (1993). This flux limit is imposed on the cluster sample.
- For clusters whose redshifts are known, an improved flux value, F_x , is determined by recalculating the count rate to flux conversion (see panel m) of Fig. 4.2) for a temperature estimated via the X-ray luminosity temperature relation as given by Pratt et al. (2009) (see Equation 2.23 and panel h) of Fig. 4.2) and by including the proper band corrections (analogous to the optical K-correction) for the actual cluster redshift.
- The measured flux is converted to an estimated flux within an aperture radius of r_{500} (defined with respect to the critical density of the Universe) by means of relations given in Pratt et al. (2009). The flux extrapolation (and in some cases interpolation) is achieved by assuming a cluster surface brightness following a β -model (Equation 2.20) with $\beta = 2/3$ and core radius $r_c = (1/7)r_{500}$.

¹MEKAL is a mode for the emission spectrum of hot diffuse gas which takes into account the gas temperature, the redshift of the source, the abundance of elements and the hydrogen column density in the line of sight (e.g. Mewe et al., 1985).

Sample i	L_i^{\min}	N_i	\bar{z}_i	N_{VLS}	\bar{n}_{VLS}
1	0.015	661	0.070	32	4.318
2	0.049	625	0.080	90	1.856
3	0.154	512	0.099	150	0.6562
4	0.245	441	0.112	159	0.3905
5	0.490	306	0.136	154	0.1535
6	0.588	260	0.143	157	0.1155

Table 4.1: Minimum X-ray luminosities L_i^{\min} ($i = 1, \dots, 6$) used to define sub-samples of the REFLEX II catalogue, with a maximum redshift $z_{\max} = 0.22$. N_i denotes the number of clusters in each sub-sample and \bar{z}_i is the corresponding mean redshift. N_{VLS} is the number of REFLEX II clusters and \bar{n}_{VLS} the mean density in each sample volume-limited sample. X-ray luminosity in units $10^{44} \text{erg s}^{-1} h^{-2}$ in the energy band 0.1 – 2.4 keV.

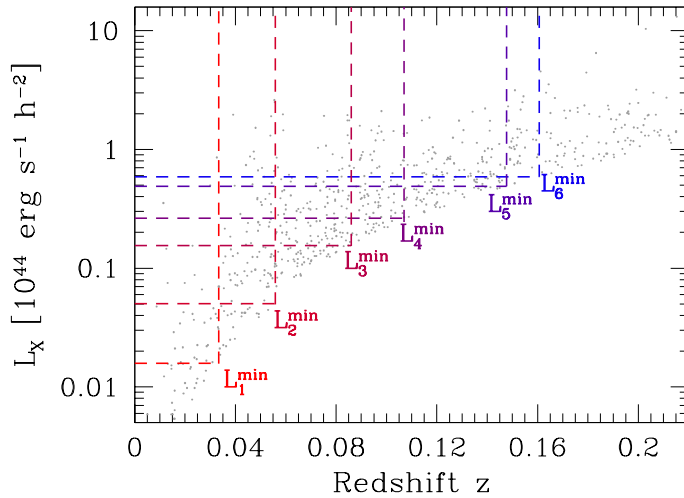


Figure 4.4: Luminosity-redshift diagram for the REFLEX II sample (points). Dashed lines schematically represents the volume-limited samples defined by the limiting luminosities of Table 4.1.

- Given the redshift of the cluster z , the cluster rest frame X-ray luminosity is calculated by means of the luminosity distance $D_L(z)$ from F_X by taking the proper band corrections for the redshift into account. The luminosity distance was determined using our fiducial cosmological model.
- For the determination of the sky-position dependent selection function, the minimum luminosity as a function of redshift and the position in the sky $L_X^{\text{lim}}(\alpha, \delta, z)$ is calculated assuming the nominal flux limit F_n and taking into account that for 5.4 per cent of the sky the exposure is too short (< 100 sec. mostly due to instrument shut down during passages of the radiation belts in South Atlantic Anomaly) to reach the nominal flux limit. The values of $L_X^{\text{lim}}(\alpha, \delta, z)$ are then derived by accounting for the proper F_X for given redshift and by performing an iterative backward engineering of the above described process.

In order to construct the REFLEX II mask, each pixel of the surveyed sky, centered on equatorial coordinates (α_i, δ_i) were assigned a limiting luminosity $L_X^{\text{lim}}(\alpha_i, \delta_i, z)$, was tabulated in the redshift range $0 \leq z \leq 0.8$. Given the minimum count rate of 20 cts and the geometrical

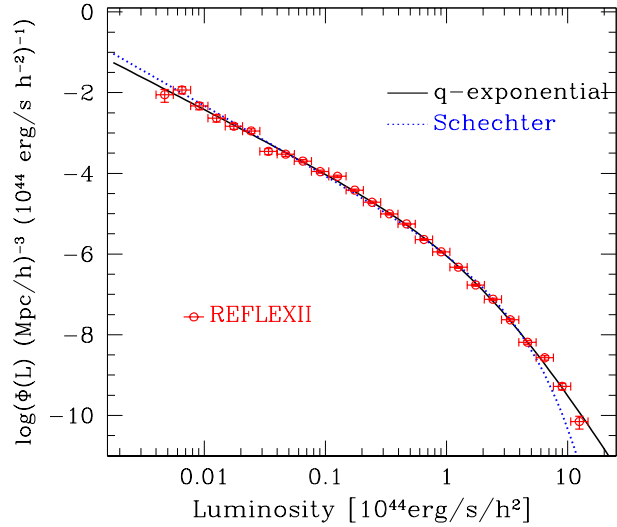


Figure 4.5: Luminosity function (red points) of the REFLEX II sample. The black solid line represents the q-exponential fit to the measurement. The red dashed line is the corresponding Schechter adjustment.

boundaries of the survey (see table 1 of Böhringer et al. 2004) we end up with a total of 787 galaxy clusters.

4.3 REFLEX II selection function

4.3.1 The REFLEX II X-ray luminosity function

We measured the REFLEX II X-ray luminosity function $\Phi(L)$ using the standard $1/V_{\max}$ estimator (Felten, 1977). This is defined as

$$\hat{\Phi}(L) = \frac{1}{\Delta L} \sum_{i=1}^{N_i} \frac{1}{V_{\max}(L_i)}, \quad (4.1)$$

where the sum is done over the clusters N_i with luminosities falling in the luminosity interval $(L - \Delta L/2, L + \Delta L/2)$, while $V_{\max}(L_i)$ is the maximum volume in which an object can be located and been detected given the minimum flux. This maximum volume is usually determined by varying the redshift of the objects in the luminosity interval such that, when comparing with the minimum luminosity at the new redshift, the object (with its luminosity fixed) can be still detected. The maximum redshift found by this means (i.e when the minimum luminosity at the new redshift equals the object's luminosity) defines the maximum volume. In our case, since the sample has different depths in the sky, the maximum volume is written as a sum over the different pixels of the REFLEX II mask:

$$V_{\max}(L_i) = \sum_{j=1}^{N_{\text{pix}}} V_{ij} \quad V_{ij} = \Delta\Omega_j \int_0^{z_{\max}^i(\alpha_j, \delta_j)} dz \frac{dV}{dz}, \quad (4.2)$$

where $z_{\max}^i(\alpha_j, \delta_j)$ is the maximum redshift that an object with luminosity L_i might have in order to be detected in the j -th pixel of the REFLEX II mask. To a good approximation, the maximum volume sampled by the REFLEX II can be characterized with a power-law $V_{\max} \propto L_X^\delta$ with slope $\delta \approx 1.38$, as is shown in the panel g) of Fig. 4.2.² The measurement in

²In a low redshift survey and assuming a flat- Λ CDM cosmological model, the maximum volume scales as $V_{\max} \propto L^{3/2}$

	α	L_\star	n_0	q	$\bar{\chi}^2$
Schechter	$-1.717^{+0.007}_{-0.003}$	$1.515^{+0.014}_{-0.015}$	$1.295^{+0.059}_{-0.055}$	1	1.65
q-exponential	$-1.536^{+0.006}_{-0.014}$	$0.628^{+0.032}_{-0.028}$	$4.082^{+0.012}_{-0.001}$	$1.313^{+0.01}_{-0.013}$	0.64

Table 4.2: Best fit parameters for the Schechter and q -extended Schechter parameterization of the XLF. Luminosity in units of $10^{44} \text{erg s}^{-1} h^{-2}$. Number density n_0 in units of $10^{-6} (\text{Mpc } h^{-1})^{-3}$. $\bar{\chi}^2$ denotes the reduced χ^2 of the fit.

finite luminosity bins is the average value of the real luminosity function in that bin,

$$\hat{\Phi}(L_i) = \frac{1}{\Delta L_i} \int_{L(i)-\frac{1}{2}\Delta L_i}^{L(i)+\frac{1}{2}\Delta L_i} \Phi(L) dL \quad (4.3)$$

We have initially parameterized the X-ray luminosity function $\Phi(L)$ by means of a Schechter form,

$$\Phi(L)dL = n_0 \left(\frac{L}{L_\star} \right)^\alpha \exp\left(-\frac{L}{L_\star}\right) d\left(\frac{L}{L_\star}\right). \quad (4.4)$$

where n_0 determines the overall normalization (abundance), the parameter α characterizes the slope of the X-ray luminosity function in the low-luminosity regime and L_\star marks the transition from power-law behavior at low luminosities to the exponential fall-off at high luminosities. In order to have a better parameterization, we have explored an extended Schechter function, written as

$$\Phi(L)dL = n_0 \left(\frac{L}{L_\star} \right)^\alpha e_q\left(\frac{L}{L_\star}\right) d\left(\frac{L}{L_\star}\right). \quad (4.5)$$

where the function $e_q(x)$ (Tsallis, 2009) is the so-called q -exponential distribution, defined as³

$$e_q(x) = \begin{cases} (1 + x(1 - q))^{1/(1-q)} & q \neq 1 \\ e^x & q = 1. \end{cases} \quad (4.6)$$

This parameterization allows for a better description of the measured luminosity function, as can be seen from Fig. 4.3. The interpretation of the set of parameters $\{\alpha, L_\star, n_0\}$ is maintained but the best-fit values are of course different from these of the Schechter parameterization. The role of the q -parameter is more relevant in the high-luminosity end of the X-ray luminosity function, where we observe that the Schechter function under-estimates the abundance of clusters. For our purposes, the behavior of the two parametric forms might lead to negligible effects in the final clustering measurement since the clusters found in the highest luminosity bins of the luminosity function will be discarded when we set a maximum redshift of $z = 0.22$ in our sample. By default we will use the q -exponential distribution.

³Four-parameters representation of the X-ray luminosity function were introduced by Bahcall (1979) by means of a two power-law function $\Phi(L) \propto (L/L_\star)^\alpha (1 + (L/L_\star)^\beta)^{-1}$.

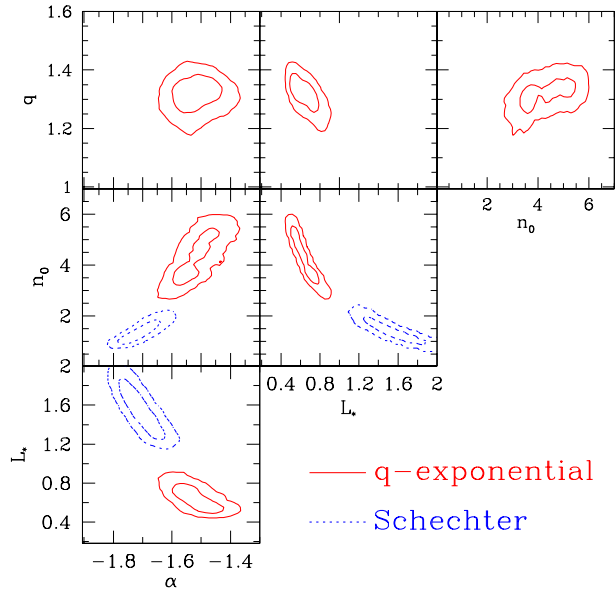


Figure 4.6: Confidence levels for the parameterization of the X-ray luminosity function using the q-exponential distribution (solid contours) and the Schechter (dashed contours). The 1σ and 2σ contours are shown. Luminosity in units of $10^{44}\text{erg s}^{-1} h^{-2}$. Number density n_0 in units of $10^{-6}(\text{Mpc } h^{-1})^{-3}$.

We have implemented a Markov Chain Monte Carlo (MCMC) algorithm (e.g. Verde, 2009; Heavens, 2009) to determine best fit for the parameters (n_0, α, L_*, q) , shown in Table. 4.2. Due to the fact that the luminosities of the REFLEX and the REFLEX II sample are determined with different way (e.g. different cosmological model, different underlying mass-temperature relations), we cannot make a direct comparison of these fits with the ones derived from the REFLEX sample (see Böhringer et al., 2002, table 1). In Fig 4.6 the constraints on the set of parameters (n_0, α, L_*, q) are shown for the q-exponential and the Schechter function. The observed confidence levels for the different combinations of parameters (n_0, α, L_*) follow the same degeneracy for both parameterizations, which is expected since the q-exponential distribution reduces to the Schechter function for $q = 1$.

4.3.2 Cosmological implications

The X-ray galaxy cluster luminosity function $\Phi(L)dL$ is defined as the number density of galaxy clusters with X-ray luminosities in the range $L_X, L_X + dL_X$. This quantity provides a link to the abundance of dark matter haloes, once a precise understanding of the underlying mass-luminosity relation and its intrinsic scatter is achieved. The luminosity function is therefore sensitive to the cosmological model (mainly, the content of matter Ω_{mat} and the normalization of the power spectrum σ_8) and also to the astrophysical processes encoded in the mass-X ray luminosity relation.

At a given redshift z , the X-ray cluster luminosity function can be understood as the result of marginalizing the conditional probability distribution $p(L_X|M, z)$ (hereafter CPD)⁴ using as a prior the dark matter halo mass function $n(M, z)$. The CPD represents the probability of assigning X-ray luminosity to the ICM given the dark matter halo mass M . This is expressed as in Equation (3.36):

$$\Phi(L, z) = \int dM p(L|M, z)n(M, z). \quad (4.7)$$

Following the analysis of Stanek et al. (2006) and Vikhlinin et al. (2009b), the CPD can be

⁴This is what we called the conditional luminosity function in Chapter 3

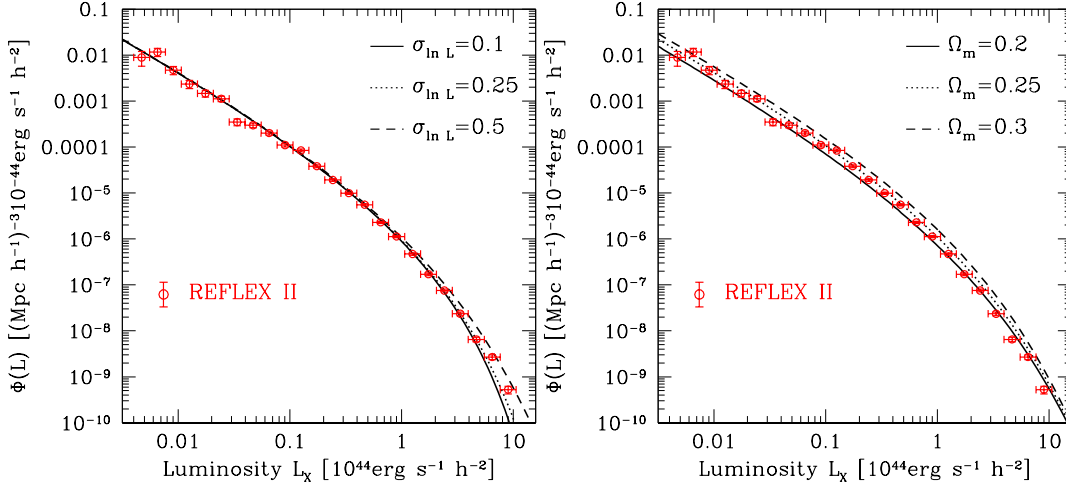


Figure 4.7: Luminosity function compared with predictions from the Λ CDM cosmological model. The left panel shows the REFLEX II measured luminosity function and three predictions for different values of Ω_m with $\sigma_{\ln L} = 0.25$. The right panel shows the same measurement with different predictions based on different values of the intrinsic scatter $\sigma_{\ln L}$ with $\Omega_m = 0.25$ and $\sigma_8 = 0.773$ fixed. We use our fiducial cosmological parameters.

represented by a log-normal distribution with an intrinsic scatter $\sigma_{\ln L}$ around the mean $M-L_X$ relation:

$$p(L|M, z)dL = \frac{1}{\sqrt{2\pi\sigma_{\ln L}^2}} \exp\left[-\frac{(\ln L - \ln \bar{L}(M, z))^2}{2\sigma_{\ln L}^2}\right] \frac{dL}{L}. \quad (4.8)$$

The resulting X-ray luminosity function depends on the shape of the halo mass function and the intrinsic scatter associated to the $M-L_X$ relation. As an example, we could assume a mass function given by a power-law $n(M, z)dM = \bar{n}M^\beta dM$ (a good approximation in the low-mass tail) and a $M-L_X$ relation of the form $\bar{L} = AM^p$. We can recover the low-luminosity tail of the luminosity function in Equation (4.5) with $\alpha = \beta/p$ and $L_* = A(2p)^{\beta/p} e^{\beta\sigma_{\ln L}^2/p}$. Therefore, the slope on the luminosity function in the low-luminosity tail is given by the slope of the mass function and the slope of the mass-luminosity relation; for a fixed p , the steeper the mass function is, the steeper the luminosity function; on the other hand, for fixed β , the smaller the exponent p is, the steeper the luminosity function becomes; in other words, a halo with a low mass will be translated into the high luminosity tail of the luminosity function, decreasing the abundance in the low-luminosity regime while increasing it in the high luminosity tail.

Note that in the limit of negligible scatter, the break in the luminosity function L_* will only depend on the exponents p and β and on the normalization of the mass-X ray luminosity relation. On the other hand, with a non-negligible intrinsic scatter (with fixed exponents β and p), the break L_* increases exponentially with $\sigma_{\ln L}$. The shape of the luminosity function is not changed, but shifted towards higher (lower) luminosities if $\sigma_{\ln L}$ increases (decreases). Therefore, at a fixed luminosity, an increment of the intrinsic scatter will represent an increment in the cluster abundance.

Due to the shape of the halo mass function, the assignment of luminosities mediated by a log-normal distribution will lead to the fact that a luminosity bin (centered at a luminosity $\bar{L}(M_0)$) will be preferentially populated by haloes with masses below M_0 which are more

abundant. This situation is more remarkable at the high-mass tail of the mass function (the exponential fall-off) and therefore the variations of the intrinsic scatter should be more sizable at the high-luminosity tail of the luminosity function. Moreover, for a fixed Ω_m , an increment in the intrinsic scatter (which approximately translates to an increment in L_*) must be accompanied by a lower value of σ_8 in order to account for the observed abundance.

In Fig. 4.7 we show comparisons of the REFLEX II X-ray luminosity function and predictions based on Equation (4.7). The left panel shows the predictions for three different values of intrinsic scatter σ_{inL} maintaining the matter energy density parameter Ω_m fixed. The right panel shows different values of Ω_m for fixed intrinsic scatter. We have used the mass-luminosity relation calibrated in Section 4.4. The figure shows that the low-luminosity tail of the luminosity function is weakly dependent on the intrinsic scatter predictions and, as was previously written, the most relevant effects of the scatter are in the high-luminosity tail. On the other hand, by fixing the scatter we observe the expected increment in the amplitude as long as we increase Ω_m . Note that by increasing the intrinsic scatter σ_{inL} or Ω_m , we can generate approximately the same net effect, i.e., an increment in the cluster abundance. The analysis of X-ray detected clusters from the RASS sample by Reiprich and Böhringer (2002b) show a degeneracy in the plane $\sigma_8 - \Omega_m$ determined from the X-ray luminosity function, approximated written as $\sigma_8 = 0.43\Omega_m^{-0.38}$.

The measurements of the REFLEX II X-ray luminosity function we have shown in this Section will be used as complement to the statistical analysis of the large scale structure in order to extract cosmological information. This will be explored in a forthcoming paper (Balaguera-Antolinez et al., in preparation).

4.3.3 Selection function

In order to explore the behavior of the clustering strength as a function of X-ray luminosity, we split the REFLEX II sample in six sub-samples characterized by a minimum luminosity L_i^{min} . These are listed in Table 4.1, together their corresponding number of clusters N and mean redshift \bar{z} . For a given luminosity cut L_i^{min} , the expected number density at a position \mathbf{r} , characterized by angular coordinates (α, δ) and redshift z , is obtained via integration of the X-ray luminosity function as

$$\bar{n}(\mathbf{r}; L_i^{\text{min}}) = \int_{\hat{L}(\mathbf{r})}^{\infty} \Phi(L_X) dL_X \quad (4.9)$$

where the lower integration limit $\hat{L}(\mathbf{r})$ is given by the REFLEX II sensitivity map as

$$\hat{L}(\mathbf{r}) \equiv \text{Max}(L_i^{\text{min}}, L_X^{\text{lim}}(\alpha, \delta, z)). \quad (4.10)$$

We also constructed six volume-limited samples (hereafter VLS) using the same minimum luminosities. These VLS are schematically represented in Fig. 4.4. Table 4.1 also lists the mean redshift and the cluster number density for these sub-samples.

Finally, we created a random catalogue of $N_r = 2 \times 10^6$ objects with luminosities greater than $L_X = 1.4 \times 10^{43} \text{erg s}^{-1} h^{-2}$. In Fig. 4.1 we show the angular positions of a subset of the random sample. Note the variations in the angular distribution of the random catalogue which follow the fluctuations in the sensitivity map of the REFLEX II survey.

4.4 Construction of the REFLEX II mock catalogs

The construction of mock catalogues is a key step towards a precise analysis of the galaxies and clusters abundance and two-point statistics. Some utilities of the mock catalogues can be summarized as follows:

- To test theoretical predictions concerning the growth of structures,
- study the propagation of systematic errors within the statistical analysis and
- to determine precise covariance matrices for the one and two-point statistics.

Mock catalogues are constructed either from N -body simulations or from log-normal realizations of the density field (Coles and Jones, 1991). N -body simulations have the advantage of properly characterizing the dynamical evolution and growth of structures in the non-linear regime, provided a tool where theoretical predictions on non-linear evolution can be directly tested. Regarding the two-point statistics, a large number of realizations $\approx 10^3$ is required to obtain precise measurements on the covariance matrix (Takayashi et al., 2008). This demands a long computational time, setting a strong constraint to create mock catalogues. On the other hand, log-normal catalogues can be massively produced in a shorter lapse (e.g. Percival et al., 2004b) and can be constructed such that their clustering matches the measured clustering from the real data only on intermediate and large scales.

In this work we will use N -body simulations to create a suit of REFLEX II mock catalogues. This has the advantages that not only non-linear growth of structure can be tested, but also the mass of the halo is provided, allowing as an example, for the study of systematics effects in the calibration of scaling relations.

We have used $N_m = 50$ realizations from the *Low resolution Baryon Acoustic Simulation at ICC* (L-BASICC II) N -body simulation, (Angulo et al., 2008; Sánchez et al., 2008) described in Chapter 3, which provides information concerning the (box) coordinates, the peculiar velocity and mass of each halo. By defining an origin (i.e, an observer), we assigned a distance for each halo and implemented our fiducial cosmological model to determine the corresponding cosmological redshift (i.e, by inverting Equation 2.3). We next transformed the coordinates of the haloes to redshift space using $r \rightarrow r + \mathbf{v} \cdot \hat{\mathbf{r}}/H_0$ where \mathbf{v} is the peculiar velocity of the center of mass of the haloes. We neglected spectroscopic redshift errors for being of the order of $\sigma \sim 10^{-3}$ (Guzzo et al., 2009). In their new coordinates, haloes were assigned their final redshift, which contains the information of peculiar velocities. At this stage, we assigned a luminosity to the dark matter haloes by means of a mass-X-ray luminosity relation, which we will discuss in the following section. Finally we observed the illuminated dark matter haloes through the REFLEX II mask, selected them with the REFLEX II selection function and obtained a set of mock catalogues with the abundance and geometry in agreement with those of the REFLEX II sample.

In this way we have constructed two sets of mock catalogues. One set consists of 50 mocks covering the full volume of the REFLEX II sample (out to $z \approx 0.5$). For the second set, we have noticed that the effective volume of the REFLEX II sample reaches a plateau at redshift $z \approx 0.2$ (see Section 5.2.1). Therefore, including clusters with higher redshift will not help to improve the signal-to-noise ratio of the measurement of power spectrum. We therefore set a maximum redshift of $z = 0.22$ which allowed us to construct 100 independent mocks catalogues out of the 50 L-BASICC II realizations. Unless otherwise stated, we use the set

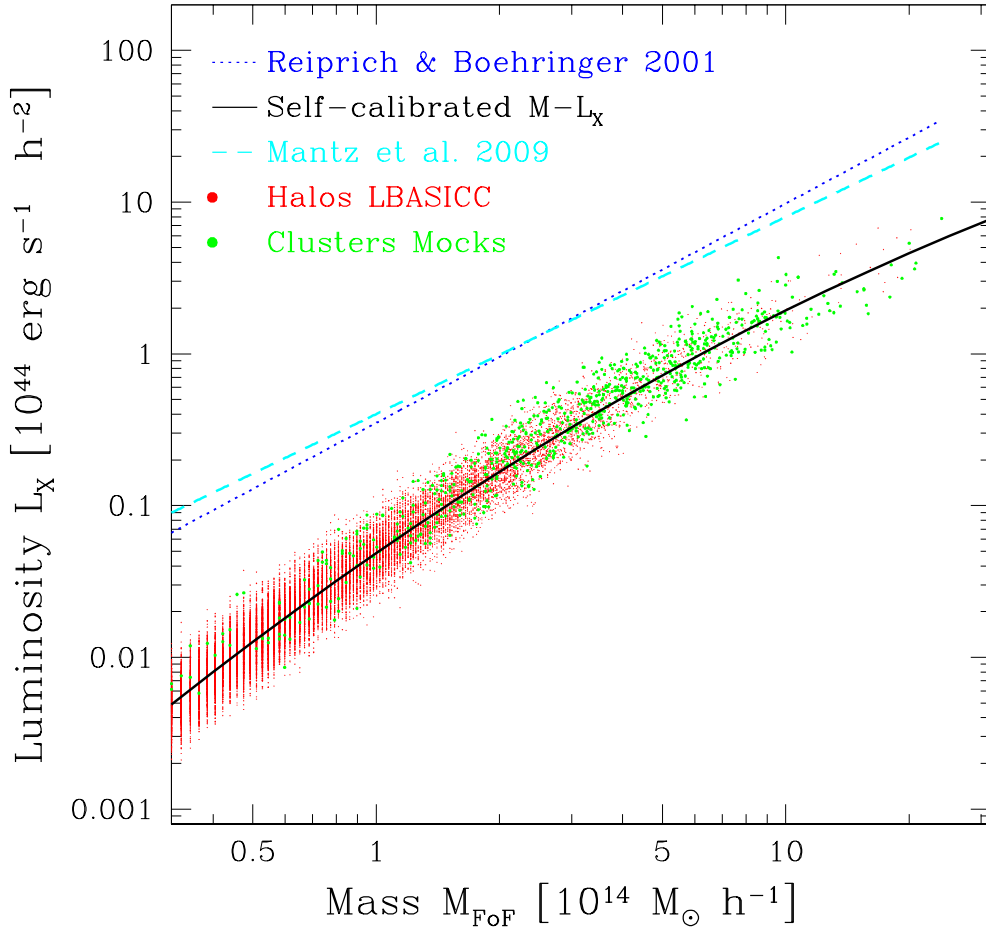


Figure 4.8: Mass-X ray luminosity relation calibrated with the REFLEX II luminosity function. Red points show the $M-L_X$ relation for one realization of the L-BASICC simulation. Green points represent one realization with the REFLEX II selection function. For comparison, the mass-luminosity relations of (Mantz et al., 2010) and (Reiprich and Böhringer, 2002a) are shown with the right slope, but arbitrary amplitude.

of 100 mocks to carry out our statistical analysis. The set of 50 mocks were only used for consistency checks.

In the analysis of the REFLEX power spectrum (Schuecker et al., 2001), the scaling relation $M-L_X$ determined by Reiprich and Böhringer (2002a) was implemented to illuminate dark matter haloes in a small suit of N -body simulations based on a Λ CDM cosmological model ($\Omega_\Lambda = 0$). No intrinsic scatter around the mean $M-L_X$, flux errors or missing flux corrections were introduced in the mocks. In Section 5.5.2 we will show that the modeling of the amplitude of the cluster power spectrum is a function of such parameters. Nevertheless, for the case of the REFLEX sample, the systematics induced by these factors are smaller than the resulting error bars in the measurements of the power spectrum.

4.4.1 Mass-X-ray luminosity relation

We have assigned X-ray luminosities to the dark matter haloes in the L-BASICC II simulations by calibrating the mass-X ray luminosity relation with the REFLEX II luminosity function. This means that we only constrained our set of mock catalogues to intrinsically follow the same

luminosity distribution observed in the real data, without imposing any constraint concerning the two point statistics. Nevertheless, since our fiducial cosmology and the cosmology used in the L-BASICC simulations are close to the latest constraints, we do not expect that the clustering observed in the REFLEX II sample differs from that observed from our mock catalogues.

Numerical tests in the early stage of this work showed that a mass-luminosity relation described by a power law with parameters derived from observations such as those quoted by Reiprich and Böhringer (e.g. 2002a); Stanek et al. (e.g. 2006); Pratt et al. (e.g. 2009); Vikhlinin et al. (e.g. 2009a) could not properly describe the observed abundance in the REFLEX II sample. Such deviation could be expected due to two facts: first, the selected clusters are not core-excised, and therefore, non-thermal physics (e.g. AGN and /or supernova feedback) can induce deviations from self-similarity; second, observations are relating the X-ray luminosity to a given mass defined, usually with an overdensity $\Delta = 500$, while our set of N -body simulations are basically FoF masses. We will return to this issue below.

We adopted a more flexible parameterization for the mass-luminosity relation implementing a power law with a mass-dependent slope. Note that technically speaking this corresponds to a relation linking Friend-of-Friend masses with X-ray luminosities in the *ROSAT* band. Writing

$$\ell = \log_{10} \left(\frac{\bar{L}_X}{10^{44} \text{erg s}^{-1} h^{-2}} \right), \quad m = \log_{10} \left(\frac{M}{10^{14} M_{\odot} h^{-1}} \right)$$

we fit a quadratic function in m

$$\ell = a + bm + cm^2. \quad (4.11)$$

The parameter a modulates the amplitude of the mass-luminosity relation; the parameter b determines the slope at a mass scale of $10^{14} M_{\odot} h^{-1}$ and the parameter c characterizes the change of slope. With this parameterization, positive values of c will generate a mass-luminosity relation that assigns higher luminosities on both extremes of the halo mass function, compared to a power-law ($c = 0$), while a negative value of c will imply that the low-mass haloes will be assigned an X-ray luminosity lower than the self-similar value; in the same way, a high mass-halo will also be assigned a luminosity below the self-similar mass-luminosity relation.

The parameters $\{a, b, c\}$ are then calibrated such that the resulting X-ray luminosity function of the mock catalogues follows the parameterization of the REFLEX II luminosity function, given by Equation (4.5). In principle, a rigorous way of calibrating the set of parameters $\{a, b, c\}$ from the REFLEX II abundance should be as follows (method 1)

- Build 100 mocks and assign redshifts to the dark matter haloes
- Given an initial set of values $\{a, b, c\}$, assign luminosities to each of the 100 realizations with intrinsic scatter $\sigma_{\ln L}$ (from this point we start referring to the dark matter haloes as clusters)
- Observe clusters in each realization with the REFLEX II mask.
- Given the cluster redshift and its assigned luminosity, assign a flux estimate and thereby assign an error to the luminosity $\sigma(L) = \delta L/L$.
- Measure the luminosity function of the 100 realizations and determine the mean luminosity function of the ensemble $\bar{\Phi}(L)$.

- Determine the χ^2 of the mean luminosity function with the best fit parameterization of the REFLEX II luminosity function by means of a χ^2 analysis:

$$\chi^2 = \sum_i (\hat{\Phi}(L_i) - \Phi^{\text{pred}}(L_i)) C_{ij}^{-1} (\hat{\Phi}(L_i) - \Phi^{\text{pred}}(L_i)),$$

where $\Phi^{\text{pred}}(L_i)$ is obtained by integrating the parameterization from Equation (4.5) as shown in Equation (4.3). The covariance matrix is determined from the 100 measurements of luminosity function as

$$C_{ij} = \frac{1}{N_m - 1} \sum_{m=1}^{N_m} (\hat{\Phi}_m(L_i) - \bar{\Phi}(L_i)) (\hat{\Phi}_m(L_j) - \bar{\Phi}(L_j)). \quad (4.12)$$

By applying a MCMC technique, we would have needed to repeat this process hundred thousand times in order to let the MCMC to converge and properly sample the posterior probability distribution of the parameters $\{a, b, c\}$. This of course requires long computational time. We have therefore tried a simplified strategy (method 2), described as follows:

- For a given set of $\{a, b, c\}$ and the fixed dispersion $\sigma_{\ln L}$, assign luminosities to the dark matter haloes in the 50 realizations.
- Assign a constant error in the luminosity $\sigma = \delta L/L$.
- Measure the luminosity distribution in the simulation boxes and determine the mean X-ray luminosity distribution from the ensemble.
- Compare the mean luminosity distribution with the best fit parameterization of the REFLEX II by determining the corresponding χ^2 from the simulation boxes.

Note that the main difference with the first method is that we do not assign flux-errors but constant luminosity errors. This procedure can speed up the calibration of the mass-luminosity relation. Nevertheless it is still slow given the number of haloes per realization and the number of steps required in the MCMC analysis.

A third method derived from the second method, (method 2.a) consists in the individual calibration of the 50 realizations and the determination of the best fit parameters from the distribution of each set of the 50 best-fit values. Finally, the most simplified method is to calibrate the mass-luminosity relation using only one simulation box, for which we might take the most representative member of the ensemble of 50 realizations.

For the final results, we have adopted the last method. We have qualitatively tested the differences between the methods 1 and 2.b in the following way:

- Using the second method (constant luminosity errors), we calibrated a set of parameters $\{a_0, b_0, c_0\}$.
- We then re-assigned luminosities to the haloes in the simulations with these parameters but with no luminosity errors.
- We then observed these haloes through the REFLEX II mask and selection function. The clusters were assigned an observed luminosity and thereafter the corresponding flux. With the flux, a flux error was determined and its luminosity was re-defined according to that error.

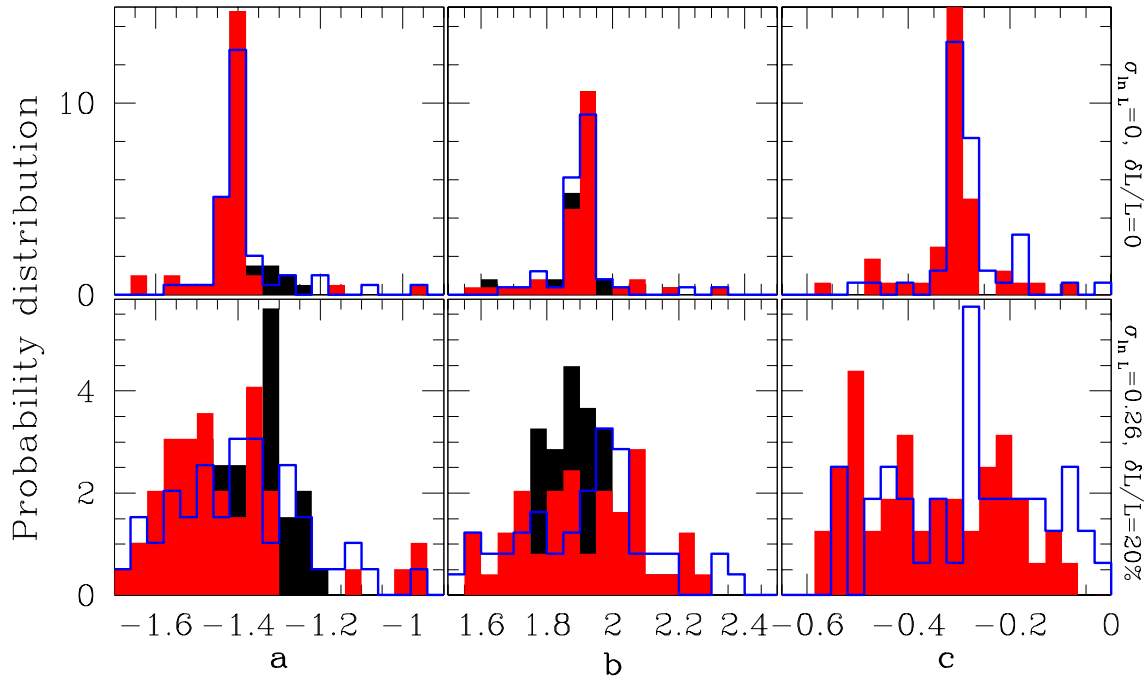


Figure 4.9: Probability distribution of the parameters $\{a, b, c\}$ of the mass-luminosity relation in the 50 realizations. The two rows show respectively two sets of values $(\sigma_{\ln L}, \sigma)$. The shaded distributions corresponds to the calibration done with the FoF halo masses, for the mass-dependent slope (red) and the power-law (black). The distribution shown with a solid line represents the calibration done with the masses corrected using Equation (3.17).

- The resulting luminosity functions arising from the previous step and that generated by the set $\{a_0, b_0, c_0\}$ display differences of the order of $\leq 3\%$. We are therefore confident that the calibration scheme given by method 2 can produce a realistic ensemble of REFLEX II mock catalogues.

In summary, we have implemented an intrinsic scatter in the natural logarithm of the X-ray luminosity as $\sigma_{\ln L} = 0.26$. This value is measured by Stanek et al. (2006) under the assumption of a flat cosmological model with $\Omega_{\text{mat}} = 0.24$ (close to our fiducial cosmology) together with the characterization of the mass-luminosity relation, using a sub-sample of the RASS sample. Although this measured scatter has been also determined in accordance to a high value of σ_8 , it is not expected to introduce considerable effects in our calibration procedure. Continuing, a 20% constant measurement error in the luminosity is assigned, being this value a good overall description of the measured luminosity errors in the REFLEX II sample, as shown in panel o) of Fig. 4.2. The best fit values drawn from the method 2.b are $a = -1.3164$, $b = 1.8769$ and $c = -0.2955$. We show the resulting mass-luminosity relation in Fig. 4.8, where we have also shown the results from Reiprich and Böhringer (2002a) and Mantz et al. (2010) regarding the slope in the mass-luminosity relation (with arbitrary normalization).

In Fig. 4.9 we present the distribution of the set of parameters $\{a, b, c\}$ defining the mass-luminosity relation in the 50 realizations. We have determined this distribution for two cases regarding the scatter and luminosity errors: $(\sigma_{\ln L}, \sigma) = (0, 0), (0.24, 0.2)$. In the absence of intrinsic scattering or luminosity errors, the distributions are well peaked around certain values $(-1.45, 1.96, -0.34)$. The observed preference in the ensemble is almost erased when

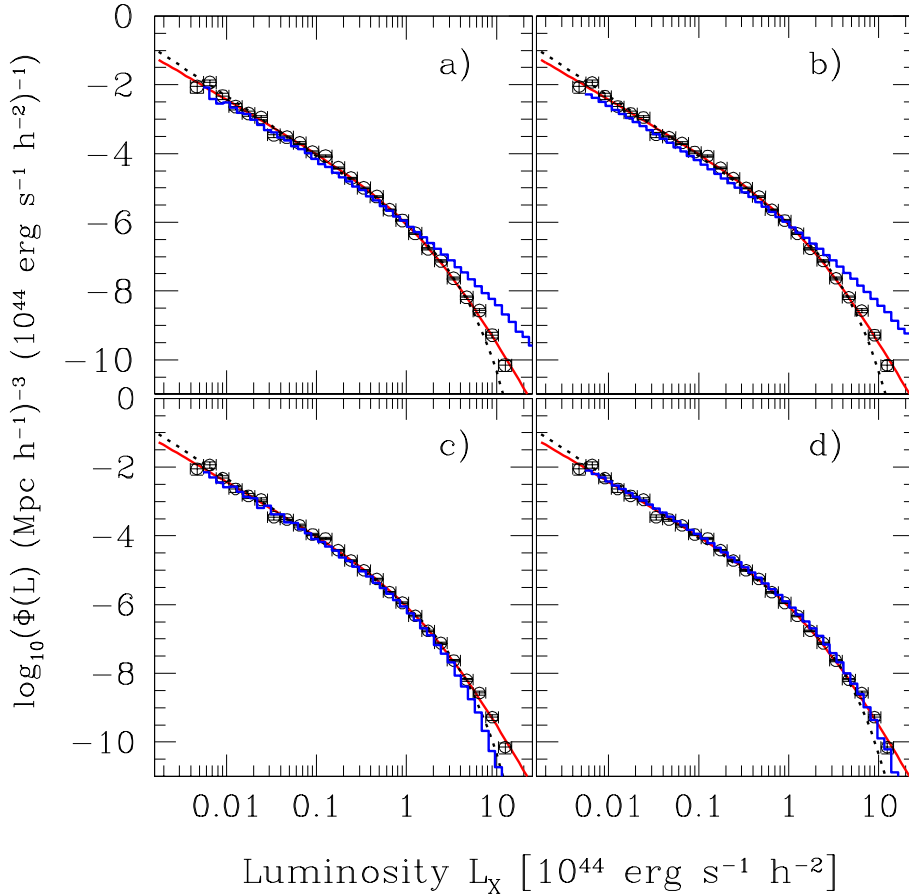


Figure 4.10: Mean luminosity distribution of the illuminated dark matter haloes in the L-BASICC II simulation (histogram) compared with the measurements of the REFLEX II luminosity function (open circles with error bars) and the best-fit from the q-exponential (solid line) and Schechter function (dotted line). Each panel shows different parameters ($a, b, c, \sigma_{\text{ln}}, \sigma$) of the underlying mass-luminosity relation: a) $(-1.45, 2.11, 0, 0, 0)$ (power-law with no-scatter), b) $(-1.61, 2.16, 0, 0.26, 0.2)$ (power-law with scatter), c) $(-1.45, 1.96, -0.31, 0, 0)$ (running-index without scatter) and d) $(-1.31, 1.87, 0.26, 0.2)$ (running-index with scatter).

we apply a intrinsic scatter. The parameter c is less sensitive to the intrinsic scatter, while the distribution of the parameters a, b are strongly modified with no clear preference. The distribution represented by a solid line in Fig.4.9 show the results of calibrating the mass-luminosity relation using the corrections of Warren et al. (2006) (see Equation 3.17). Without intrinsic scatter, the behavior of the mass-luminosity relation changes little if we correct for the systematics in the FoF masses. This conclusion can be also applied to the bottom panel, where the differences arising are mainly due to the intrinsic scatter. The resulting luminosity distribution arising from the set of parameters described by Fig. 4.9 are shown in Fig 4.10 compared to the measurements of the REFLEX II luminosity function. As shows in this figure, a mass-luminosity relation described by a power-law relation underestimates the abundance of the most luminous galaxy clusters.

In Fig. 4.11 we compare the REFLEX II luminosity function and the mean luminosity function determined from the mock catalogues. The middle panel shows the relative difference between the REFLEX II X-ray luminosity function and the mean luminosity function derived

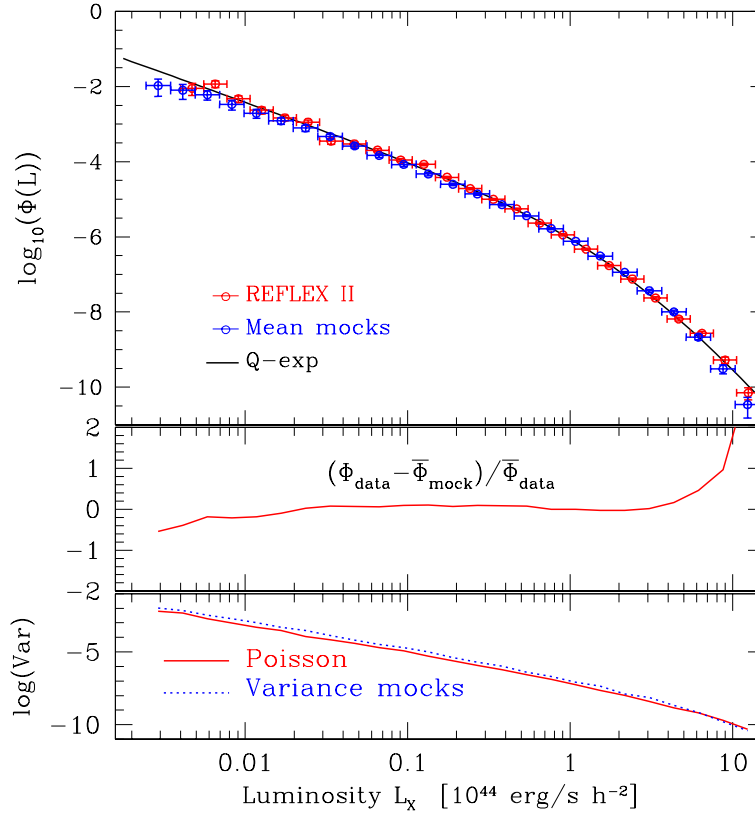


Figure 4.11: REFLEX II luminosity function compared with the mean luminosity function of the mocks catalogues obtained by assigning luminosities with Equation (4.11). The y- axis denotes the decimal logarithm of the luminosity function in units of $(\text{Mpc } h^{-1})^{-3} (10^{44} \text{ erg s}^{-1} h^{-2})^{-1}$. The middle panel shows the ratio of measured luminosity function to the mean luminosity function determined from the ensemble of 50 REFLEX II Like mock catalogues. The bottom panel shows the variance in the luminosity function: the Poisson variance agrees very well with the variance derived from the ensemble of mock catalogues.

from the ensemble of mock catalogues. This shows the differences of $\leq 3\%$ are achieved within the range of luminosities where we will explore the clustering. The third panel in that figure shows how well the Poisson-derived variance describe the variance from the ensemble of the mocks catalogues.

4.4.2 Systematics

The trend observed by the self-calibrated mass luminosity relation is in agreement with observations and N -body hydro-simulations (Puchwein et al., 2008; Stanek et al., 2010), which show deviations from the a power-law (self-similar) behavior on the low masses ($10^{13} M_{\odot} h^{-1}$) and luminosity ($10^{43} \text{ erg s}^{-1} h^{-2}$) ranges. These deviations can be associated to non-thermal processes like cooling flows, AGN or supernova feedback, as has been claimed from observations (Stanek et al., 2010). Nevertheless we do not attempt to extract physical conclusions from this calibration for a number of reasons:

- We are not assigning bolometric luminosities but K-corrected luminosities in the *ROSAT* energy band [0.1 – 2.4] keV.

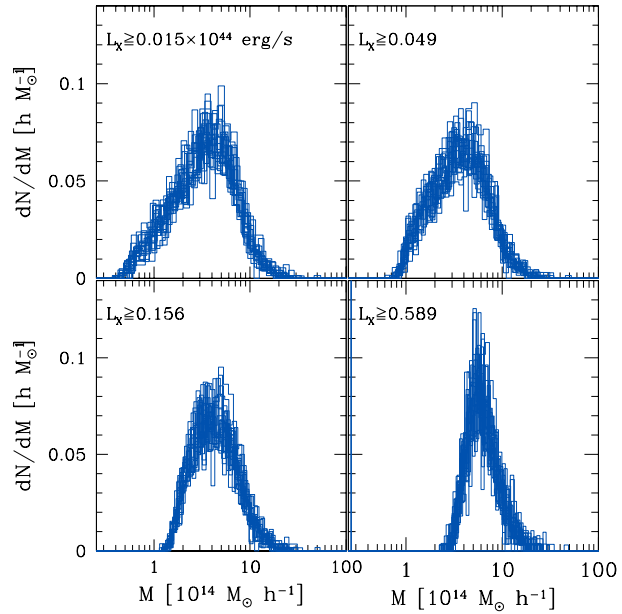


Figure 4.12: Distribution of halo masses for the REFLEX-II like mock catalogues. Within our cuts in luminosity, the REFLEX II mask selects haloes with masses in the range $3 \times 10^{14} M_{\odot} h^{-1} \leq M \leq 5 \times 10^{15} M_{\odot} h^{-1}$.

- We are using redshift $z = 0$ simulations to model clustering up to $z \approx 0.5$ instead of using a simulation at the mean redshift of the survey or a light cone. Nevertheless, the mean redshift of the survey $\bar{z} \approx 0.09$ is too small to admit a significant evolution in the clustering pattern of galaxy clusters. Furthermore, we have observed no evolution in the X-ray cluster luminosity function simply by observing that the abundance of clusters in different bins of redshift is well described with the same parameterization.
- The FoF halo-finder algorithms overestimates the halo mass due to systematics introduced by the finite number of particles. Warren et al. (2006) showed a way to correct for this systematics, which only depends on the FoF mass. This is shown by Equation (3.17). For the lowest luminosity cut shown in Table 4.1, the distribution of minimum halo masses selected by Equation (4.11) in the mock catalogues peaks at $M \approx 5 \times 10^{13} M_{\odot} h^{-1}$, which, following the correction from Equation (4.11) corresponds to an offset of ~ 40 per cent with respect to unbiased mass estimations. Nevertheless, we do not attempt to correct the L-BASICC II FoF masses, for any correction would lead to a new set of parameters (a, b, c) which will still reproduce, by construction, the observed X-ray luminosity function.
- There is not a clear one-to-one relation between FoF masses and the spherical overdensity masses (e.g. Lukic et al., 2010), which are usually implemented in the calibration of scaling relations in galaxy clusters (e.g. Pratt et al., 2009).

This last point lead us to an extension of the present work, based in the same analysis presented here using another mass definition. It has been pointed out in the literature (e.g. Crocce et al., 2010; Pillepich et al., 2010; Wu et al., 2010) how critical the definition of the halo mass is in order to compare different halo mass functions, as was also depicted in Fig. 3.6. Having a different mass definition such as the ones given by spherical overdensity finders can provide more direct links to the observed (measured) masses. It is then relevant to test whether the trend observed in the inferred mass-luminosity relation (i.e., its deviation from a self-similar behavior) is due to the underlying physical effects encoded in the cluster

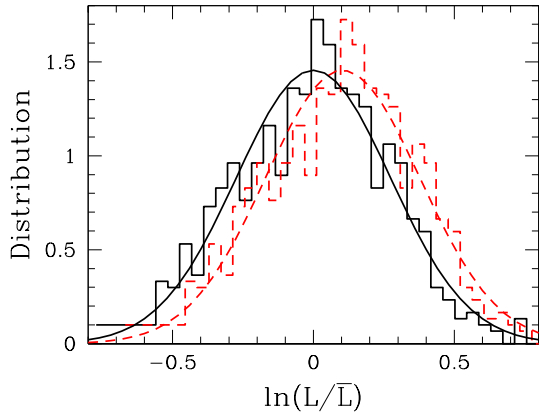


Figure 4.13: Distribution of luminosities around the mean mass luminosity relation in the one mock catalogue. The continuous solid line represents a log-normal distributions with zero mean. The dotted line is a log-normal distribution with mean $\delta\sigma_{\ln L}^2$.

luminosity function or is just an artifact of the definition of mass.

4.4.3 Properties of the mock catalogues

Some basic properties of the REFLEX II mock catalogues can be summarized as follows:

- The resulting mass distribution of the mock catalogues is shown in Fig. 4.12 for four values of minimum luminosity. For a minimum X-ray luminosity of $L_X^{\min} = 1.5 \times 10^{42} \text{erg s}^{-1} h^{-2}$, the REFLEX II selection function allowed for the detection of dark matter haloes with masses in the range $3 \times 10^{14} \leq M/M_\odot h^{-1} \leq 5 \times 10^{15}$.
- Malmquist bias is a selection effect inherent to flux limited samples. It arises due to the fact that the most intrinsically bright clusters are those detected at the largest distances. Therefore, at a given cluster mass (high mass especially), a flux limited sample is biased towards the detection of the brightest objects and therefore the resulting mean mass-luminosity relation $\langle \ln L \rangle_{FLS}$ is dominated by clusters with luminosities above the true underlying mass luminosity relation (e.g. Stanek et al., 2006; Vikhlinin et al., 2009a; Pratt et al., 2009). If no evolution in the mass-luminosity relation is assumed, Malmquist bias can be translated into overestimated slopes and underestimated amplitudes of the mass-luminosity relation (e.g. Reiprich and Böhringer, 2002a). Following Vikhlinin et al. 2009a and Pratt et al. (2009), at a given mass M the mean observed mass-luminosity relation $\langle L(M) \rangle$ differs from the underlying mass-luminosity relation $\bar{L}(M)$ by a bias factor⁵

$$\text{Bias}(L, M) \equiv \langle \ln L_X \rangle - \ln \bar{L}_X = \frac{\int d \ln L (\ln L_X - \ln \bar{L}_X(M)) p(\ln L|M) V(L)}{\int d \ln L p(L|M) V(L)} \quad (4.13)$$

where $V(L)$ is the volume probed by a cluster as a function of its luminosity, such that for a constant $V(L)$ we simply reduce to $\langle \ln L_X \rangle = \ln \bar{L}_X$. Assuming the log-normal distribution as in Equation (4.8) and a maximum volume as a power law $V(L) \propto L^\delta$, it can be shown that Equation (4.13) reduces to a mass-independent bias factor $\text{Bias}(L, M) = \delta\sigma_{\ln L}^2$. In Fig. 4.13 we show the distribution of luminosities around the mean mass-luminosity relation given by Equation (4.11). The continuous solid line represents a log-normal distribution centered at $\ln L/\bar{L} = 0$. The dashed line represents a log-normal distribution centered at $\ln(L/\bar{L}) = \delta\sigma_{\ln L}^2 \approx 0.09$, for which we have used $\sigma_{\ln L} = 0.26$ and $\delta = 1.38$ (see

⁵This shouldn't be confused with the luminosity bias appearing in the clustering analysis.

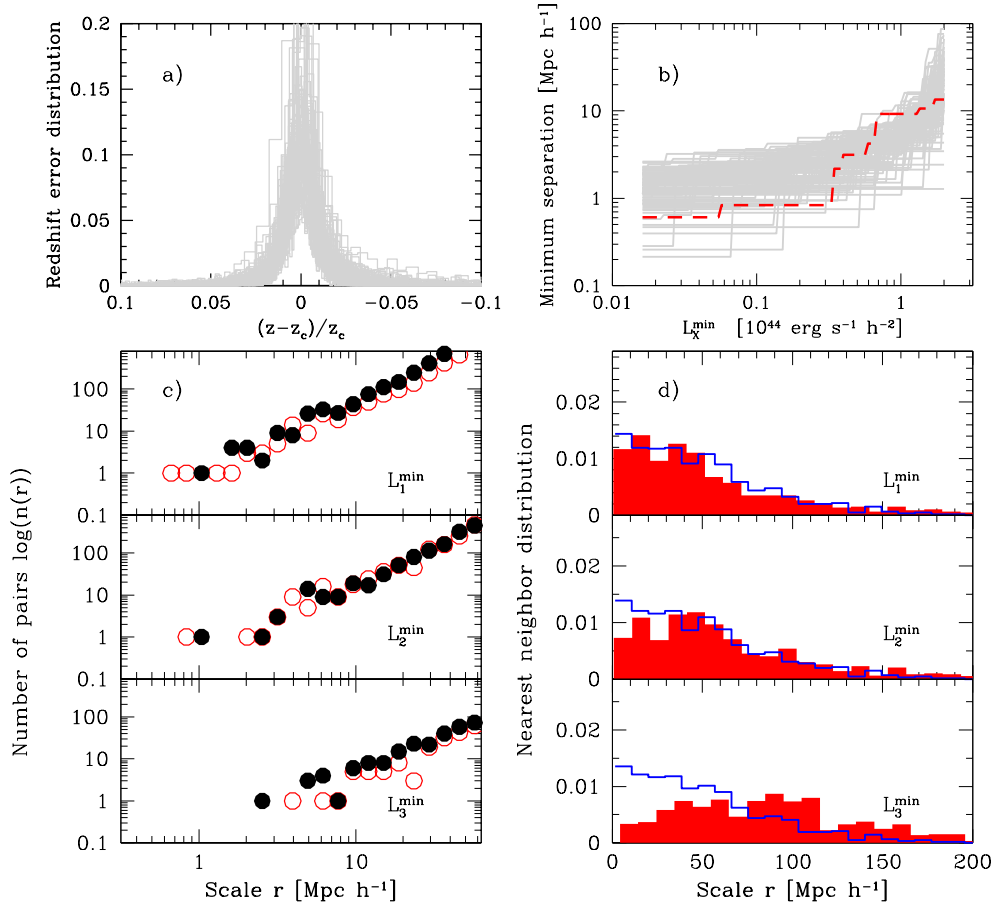


Figure 4.14: Some properties of the mock catalogues: a) redshift error distribution associated to the difference between the cosmological redshift z_c and the redshift induced by peculiar velocities z ; b) dependence of the minimum separation of clusters with the X-ray luminosity. The dashed line shows the result from the REFLEX II sample; c) number of pairs as a function of the separation for one mock (filled circles) and the REFLEX II sample (open circles); d) nearest neighbor distribution for the REFLEX II (filled histogram) and one mock catalogues (solid line).

Fig 4.2). With this values we can properly correct for the this luminosity bias, as can be seen from Fig. 4.13.

- Panel a) of Fig. 4.14 shows the redshift error distribution arising from the difference between the cosmological redshift and the redshift with the effects of peculiar velocities included, for all the mock catalogues. Panel b) of the same figure shows the behavior of the separation to the closest neighbor as a function of the minimum luminosity. Panel c) shows the number of pairs as a function of the pair separation for the REFLEX II sample and one mock catalogue. The exclusion effect (see Appendix B) can be seen as a lack of pairs on small scales in the mock catalogue when compared with the REFLEX II sample. It is more noticeable when we increase the minimum luminosity, as expected. Panel d) shows the nearest neighbor distribution for the same two data sets.
- Tables 4.3 and 4.4 show the final information contained in each mock catalogue and the random catalogue respectively. Most of this tabulated quantities are used in the

Column	Symbol	Name	Units	Definition/Comments
1	α	Right ascension	Deg	
2	δ	Declination	Deg	
3	$r(z)$	Comoving distance	Mpc h^{-1}	L-BASICC II cosmology
4	z	Total redshift		L-BASICC II cosmology
5	z_{cos}	Cosmological redshift		L-BASICC II cosmology
6	$\log_{10}L_X$	Corrected X-ray Luminosity	$10^{44}\text{erg s}^{-1} h^{-2}$	
7	$\log_{10}L_X$	Observed X-ray Luminosity	$10^{44}\text{erg s}^{-1} h^{-2}$	
8	M	Mass	$10^{14}M_{\odot} h^{-1}$	FoF L-BASICC II
9	N_{pix}	Number of pixel		
10	V_{max}	Maximum Volume	$(\text{Mpc } h^{-1})^3$	fiducial cosmology
11	$r(z)$	Comoving distance	Mpc h^{-1}	fiducial cosmology
12	$b(L_X)$	Luminosity bias		L-BASICC II
13 – 18	\bar{n}^j	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{L_{j+1}} \Phi(L)dL$ $j = 1, 6$
19 – 25	\bar{n}^i	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{\infty} \Phi(L)dL$ $i = 1, 7$
26 – 31	\bar{n}_b^j	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{L_{j+1}} \Phi(L)b^2(L)dL$ $j = 1, 6$
32 – 38	\bar{n}_b^i	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{\infty} \Phi(L)b^2(L)dL$ $i = 1, 7$

Table 4.3: Information tabulated for the mock catalogues. The total redshift z represents the cosmological redshift plus the contributions from peculiar velocities. The luminosity bias $b(L)$ assigned to each halo corresponds the fit derived from measurements of power spectrum of the illuminated haloes in the L-BASICC simulations (see Chapter 5). The selection functions tabulated in columns 13 to 38 are implemented in the clustering analysis.

clustering analysis of Chapter 5.

4.5 Summary

In this chapter we have described the REFLEX II galaxy cluster sample. This new catalogue contains 911 X-ray detected clusters down to a limiting flux of $F_{\text{lim}} = 1.8 \times 10^{-12} \text{erg s}^{-1} \text{cm}^{-2}$. Its higher sensitivity provides a better statistics (compared to the original REFLEX sample) which will allow us to explore in some detail interesting properties of the clustering of galaxy clusters. The sky-coverage and detection strategy of the REFLEX II is nearly the same ⁶ as in the REFLEX sample. In order to properly test the REFLEX II selection function, we have measured the REFLEX II X-ray luminosity function. The implementation of these measurements to extract cosmological information will be presented in a forthcoming paper (Böhringer et al., in preparation).

The second main goal of this chapter was to describe the construction of the REFLEX II mock catalogues, which are to be used in the large scale structure analysis to provide measurements of the covariance matrix for the different methods applied on the data. Based on N -body simulations, we built a suit of 100 mock catalogues which are constraint to intrinsically follow the best-fit parameterization of the REFLEX II X-ray luminosity function

⁶REFLEX II has a slightly modified sky-coverage

Column	Symbol	Name	Units	Definition/Comments
1	α	Right ascension	Deg	
2	δ	Declination	Deg	
3	$r(z)$	Comoving distance	Mpc/h	L-BASICC II cosmology
4	z	Redshift		L-BASICC cosmology
5	$\log_{10} L_X$	X-ray Luminosity	$10^{44} \text{erg s}^{-1} h^{-2}$	
6	N_{pix}	Number of pixel		
7	$b(L_X)$	Luminosity bias		L-BASICC II
8 – 13	\bar{n}^j	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{L_{j+1}} \Phi(L) dL$ $j = 1, 6$
14 – 20	\bar{n}^i	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{\infty} \Phi(L) dL$ $i = 1, 7$
21 – 26	\bar{n}_b^j	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{L_{j+1}} \Phi(L) b^2(L) dL$ $j = 1, 6$
27 – 33	\bar{n}_b^i	Selection function	$(\text{Mpc } h^{-1})^{-3}$	$\int_{L_i}^{\infty} \Phi(L) b^2(L) dL$ $i = 1, 7$

Table 4.4: Description of the synthetic catalogue. The description is the same as in Table 4.3.

together with the REFLEX II selection strategy and sky coverage. By demanding the mock catalogues to follow the REFLEX II X-ray luminosity function, we have calibrated a mass-X-ray luminosity relation for our mock clusters. Being this one of the potentials of the X-ray luminosity function (once the cosmology is known), we did not attempt to compare this calibration with measurements of scaling relations Pratt et al. (e.g 2009), since our procedure required intrinsic scatter and luminosity errors that are usually simultaneously determined with the parameters of the mass-X ray luminosity relation itself. These parameters, when changed, lead to another set of parameters characterizing our mass-X ray luminosity relation, which is still reproduced by construction the REFLEX II luminosity function. Nevertheless the knowledge of an underlying mass-X-ray luminosity relation, its intrinsic scatter and the luminosity errors for the mock catalogues make of them a valuable tool to explore systematic effects concerning the determination of the mass-X ray luminosity relation or the constraints on cosmological parameters.

Power spectrum analysis of the REFLEX II sample

5.1 Introduction

In this chapter we present the analysis of the power spectrum of the new REFLEX II catalogue, described in Chapter 4 (Balaguera-Antolinez et al., 2011).

Besides the advantages provided by a larger cluster sample, the power spectrum analysis presented here represents an improvement over that of Schuecker et al. (2001) in a number of ways. In particular, our analysis is complemented with a set of N -body simulations, the L-BASICC II (Angulo et al., 2008; Sánchez et al., 2008), from which we constructed a suit of 100 REFLEX II mock catalogues. These catalogues were calibrated to reproduce the measured REFLEX II X-ray luminosity function. Selection criteria of the REFLEX II sample were applied in their construction, yielding a large suit of mocks that can be used to analyze the statistical methods applied to the data.

This chapter is organized as follows. In Section 5.2 we describe the power spectrum estimator and show the measurements of the REFLEX II window function and the covariance matrix. In Section 5.3 we make a parenthesis and explore the clustering properties from the L-BASICC- simulations and the mock catalogues as a function of the X-ray luminosity. This will allow us to test models for the the amplitude of the REFLEX II power spectrum. In Section 5.4 we explore the sensitivity of the REFLEX II sample to distortions induced by flux-selection effects. In Section 5.4.1 we model the shape of the power spectrum. The model of the shape and the amplitude is applied to the REFLEX II sample in Section 5.5. In Section 5.6 we briefly describe the measurements of the cluster correlation function. We end with our conclusions in Section 5.7.

5.2 The measurement of $P(k)$

We have measured the power spectrum of the REFLEX II sample containing clusters with luminosities $L_X \geq L_1^{\min}$, which represents 760 objects in the redshift interval covered by the REFLEX II survey. We implemented the standard minimum variance weighting power spectrum estimator of Feldman et al. 1994 (hereafter FKP), which defines a weighted density

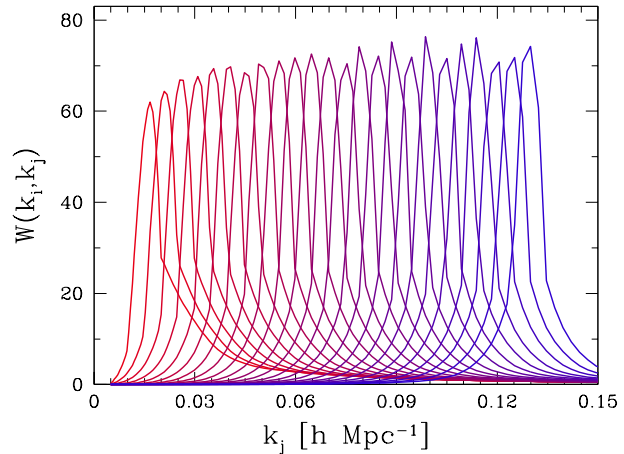


Figure 5.1: Example of the REFLEX II shell averaged window (matrix) function W_{ij} for different modes k_i determined for the sub-sample with limiting luminosity L_2^{\min} . Each curve is normalized to $\int dk_j W(k_i, k_j) = 1$.

fluctuation

$$F(\mathbf{r}) = w(\mathbf{r}) (n_c(\mathbf{r}) - \alpha n_r(\mathbf{r})), \quad (5.1)$$

where $n_c(\mathbf{r})$ ($n_r(\mathbf{r})$) is the number density of clusters in the real (random) catalogue for a given luminosity cut (where we dropped the L^{\min} dependence to avoid clutter). The parameter α , given by

$$\alpha = \frac{\int w(\mathbf{r}) n_c(\mathbf{r}) d^3 r}{\int w(\mathbf{r}) n_r(\mathbf{r}) d^3 r}, \quad (5.2)$$

forces the fluctuation to have zero mean, $\int F(\mathbf{r}) d^3 r = 0$. The optimal normalized weights $w(\mathbf{r})$ are given by Feldman et al. (1994) as

$$w(\mathbf{r}) = \left(\frac{1}{1 + \bar{n}(\mathbf{r}) P_{\text{est}}} \right) \left(\int \frac{\bar{n}(\mathbf{r}')}{1 + \bar{n}(\mathbf{r}') P_{\text{est}}} d^3 r' \right)^{-1}, \quad (5.3)$$

where P_{est} is an estimate of the power spectrum to be measured, for which we have chosen $P_{\text{est}} = 2 \times 10^4 (\text{Mpc } h^{-1})^3$. We used the FFTw algorithm (Frigo and Johnson, 2008) embedding the REFLEX II volume in a cube divided in $N_{\text{grid}} = 512^3$ cells and implemented a triangular shaped cloud mass assignment (Hockney and Eastwood, 1981) correcting afterwards for aliasing effects. The length of the sides of the cube are determined by $L_{\text{box}} = 2r(z_{\text{max}})$, with $z_{\text{max}} = 0.22$ which corresponds to a box size of $1.25 \text{ Gpc } h^{-1}$ for our fiducial cosmology. The fundamental mode is $\Delta k = 2\pi/L_{\text{box}} = 0.0049 h \text{ Mpc}^{-1}$. The Nyquist frequency for this box is $k_{\text{Nyq}} = 1.27 h \text{ Mpc}^{-1}$, and we can ignore aliasing effects on wavenumbers smaller than $k \approx 0.7 h \text{ Mpc}^{-1}$. We subtract the shot noise and average in spherical shells to obtain the bin-averaged power spectrum $\hat{P}(k)$ (see Appendix C). This measurement is the convolution of the underlying cluster power spectrum with $|W(\mathbf{k})|^2$, the square of the Fourier transform of the REFLEX II window function given by

$$W(\mathbf{k}) = \int \bar{n}(\mathbf{r}) w(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d^3 r. \quad (5.4)$$

We follow the procedure of Cole et al. (2005) to construct the window function in matrix form W_{ij} , by using a Gauss-Legendre integration scheme (Press et al., 2002), described in in Appendix C. The measured power spectrum can be written as a matrix multiplication (see

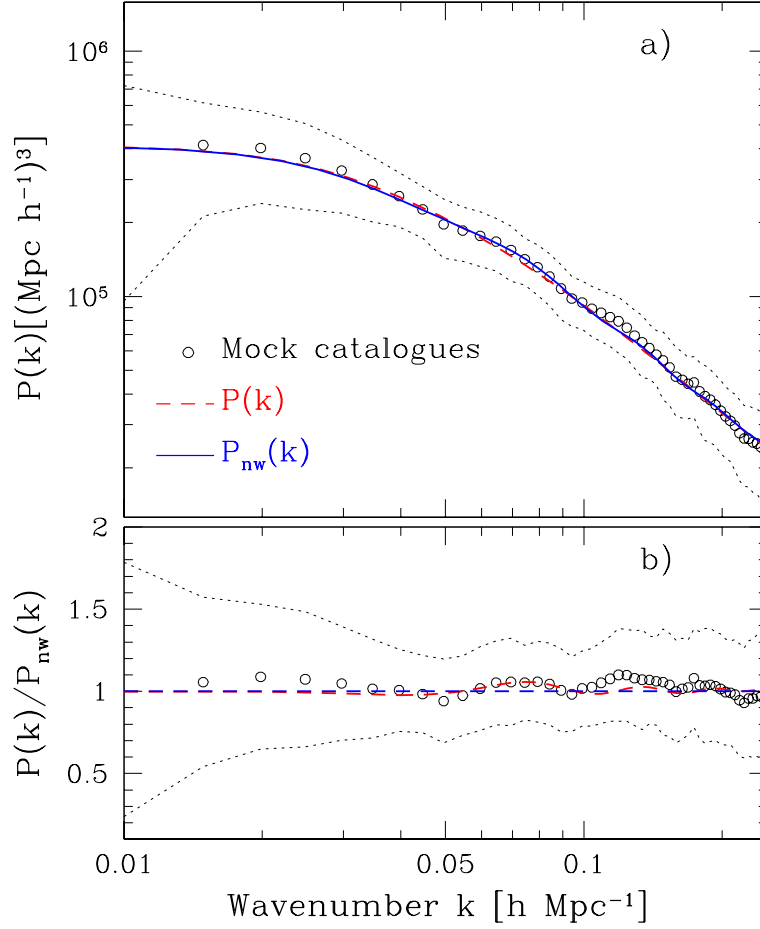


Figure 5.2: a) Mean power spectrum from the mock catalogues (open points) with its 1σ variance (dotted lines) for the sub-sample characterized with the limiting luminosity L_2^{min} . The dashed line represents a non-linear matter power spectrum $P(k)$ with BAO and convolved with the REFLEX II window function. The solid line $P_{\text{nw}}(k)$ is the same theoretical prediction from which the BAO have been excluded, also convolved with the window function. b) Ratio of the mean mock power spectrum (open points), and the power spectrum $P(k)$ to the power spectrum $P_{\text{nw}}(k)$. Dotted lines denotes the 1σ variance.

Equation C.19)

$$\hat{P}(k_i) = \sum_j W_{ij} P(k_j) - \beta W_{i0}, \quad (5.5)$$

where $P(k_j)$ is the underlying power spectrum and the term βW_{i0} accounts for the integral constraint (Percival et al., 2007a; Reid et al., 2009) such that $P(0) = 0$. As an example, Fig. 5.1 shows some elements of the window matrix of the sample defined by the minimum luminosity L_2^{min} (panel a) and the volume limited sample defined by the same luminosity cut (panel b). As expected, large scale modes receive contributions from intermediate and even small ($k \geq 0.3h \text{ Mpc}^{-1}$) scales.

We used the window matrix to assess the possibility of the detection of the signal from the BAO in the measured REFLEX II power spectrum. The signature of BAO in the dark matter halo distribution of the L-BASICC II simulations has been analyzed both in the spatial two-point correlation function (Sánchez et al., 2008) as well as in the power spectrum (Angulo

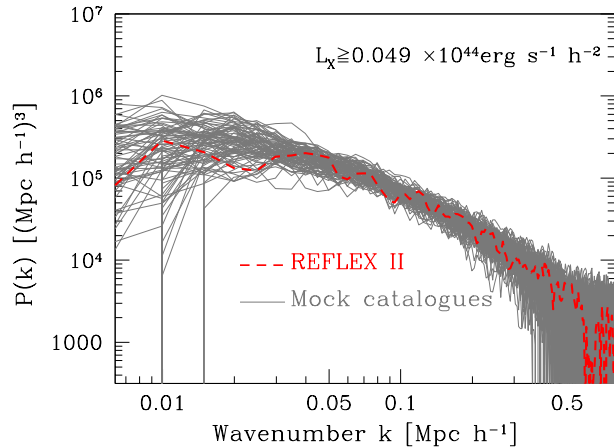


Figure 5.3: Individual power spectra from the mock catalogues compared with the measurements from the REFLEX II sample.

et al., 2008).

We used the fitting formulae of Eisenstein and Hu (1998) to compute the matter power spectrum for the cosmological model of the L-BASICC II simulations, including a non-linear correction computed with HALOFIT (Smith et al., 2003). We also computed a model with the same broad-band shape but without any baryonic oscillations, $P_{\text{nw}}(k)$. Panel a) of Fig. 5.2 shows the comparison of these theoretical models (solid and dashed lines), convolved with the REFLEX II window function, with the mean power spectrum from the mock catalogues (open circles). It can be seen that the difference between these two models is much smaller than the variance in $P(k)$ that corresponds to the REFLEX II volume, which can be determined from the ensemble of mock catalogues (dotted lines, see Section 5.2.1). This can be more clearly seen in panel b) of the same figure, which shows the ratio of these power spectra to $P_{\text{nw}}(k)$. The convolution with the window function washes out the acoustic oscillations in the power spectrum to a level where they can not be distinguished from a model without BAO. Furthermore, we computed the χ^2 of these two models (analytically marginalizing over the amplitude as described in Lewis and Bridle, 2002) and found a difference of less than ~ 3 per cent between them. We thus conclude that, due to the survey volume, no statistically significant signal of BAO can be detected in the REFLEX II power spectrum.

In Fig. 5.3 we show individual power spectra for the 100 mock samples compared to the measurement done on the REFLEX II sample (to be analyzed in Section 5.5).

5.2.1 Covariance matrix

We used our ensemble of 100 mock catalogues to obtain an estimate of the bin-averaged covariance matrix $\hat{C}(k_i, k_j)$ of the REFLEX II power spectrum by

$$\hat{C}(k_i, k_j) = \frac{1}{N_m - 1} \sum_{n=1}^{N_m} (\hat{P}_i^n - \bar{P}_i) (\hat{P}_j^n - \bar{P}_j), \quad (5.6)$$

where $\hat{P}_i^n = \hat{P}^n(k_i)$ is the measured power spectrum in the n -th mock catalogue in the bin centered at k_i and \bar{P}_i is the mean power spectrum from the ensemble of mocks at the same bin. As an example, panel a) of Fig. 5.4 shows examples of the correlation coefficients r_{ij} defined from the covariance matrix via $\hat{r}_{ij} = \hat{C}_{ij} / (\hat{C}_{ii} \hat{C}_{jj})^{1/2}$ for L_3^{min} (upper triangular part) and L_6^{min} (lower triangular part). For comparison, panel c shows the correlation matrix inferred from the clustering of the illuminated halos in the L-BASICC II simulation for the same

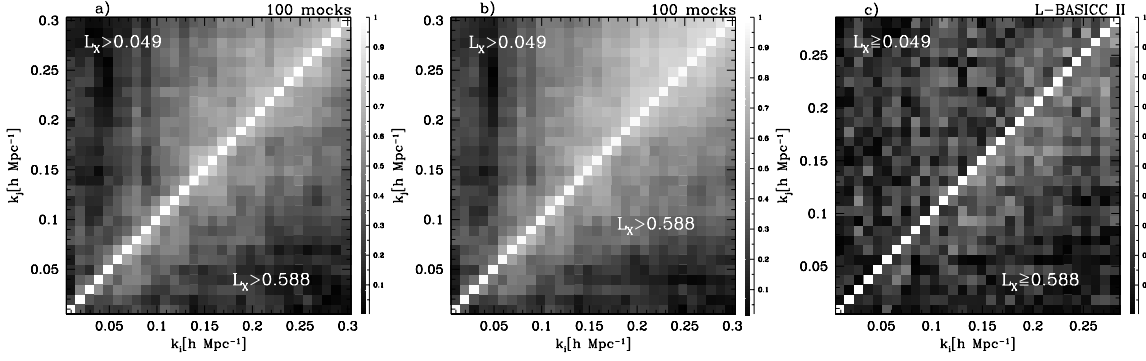


Figure 5.4: Comparison of the correlation matrix $|r(k_i, k_j)|$ of the power spectra measured with a) the FKP estimator and b) the PVP estimator (see Section 5.4) for two luminosity-cuts. c) Correlation matrix determined from the illuminated dark matter halos of the L-BASICC II simulation with same luminosity-cuts.

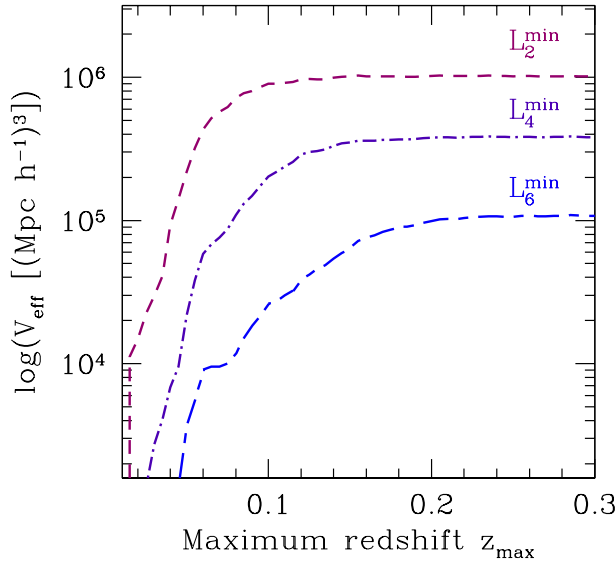


Figure 5.5: Effective volume probed by the REFLEX II sample (see Equation C.15) as a function of the maximum redshift for three different cuts in luminosity.

luminosity cuts. The covariance matrix of the mocks contains important off-diagonal terms which arise from the mode coupling induced by the window function. All the statistical analyzes performed in this work are based on the covariance matrix defined by Equation (5.6).

Feldman et al. (1994) derived an approximated expression for the variance of the spherically averaged power spectrum under the assumption that the Fourier modes are Gaussian-distributed. This is given by

$$\frac{\sigma^2(k)}{P(k)^2} = \frac{2}{V_k V_{\text{eff}}(k)}, \quad (5.7)$$

where $V_k \approx 4\pi k^2 \delta k / (2\pi)^3$ is the volume of a spherical shell of width δk and $V_{\text{eff}}(k)$ is the effective (coherence) volume probed by the survey at a scale k , defined by Tegmark (1997) as

$$V_{\text{eff}}(k) = \int \left(\frac{\bar{n}(\mathbf{r}) P(k)}{1 + P(k) \bar{n}(\mathbf{r})} \right)^2 d^3 r. \quad (5.8)$$

Equation (5.7) assumes that the power spectrum $P(k)$ is smooth on scales δk and applies for wavenumbers $k_i \gg \delta k$. Fig. 5.5 shows the effective volume for different luminosity cuts as a

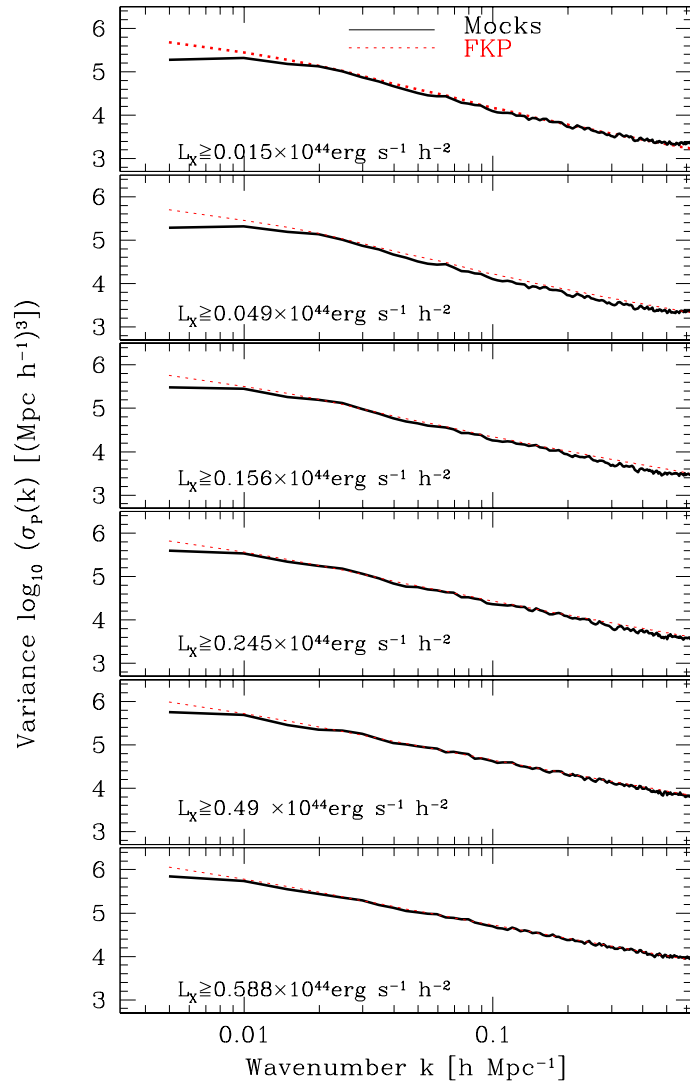


Figure 5.6: Bin-averaged variance of the REFLEX II power spectrum for our six sub-samples defined in Table 4.1. Solid line represents the variance determined from the set of mock catalogues following Equation (5.6). The dotted line shows the prediction from the Equation (5.7).

function of the maximum redshift of the sample. As pointed out in Section 4.4, the effective volume gained by including objects with redshifts $z \geq 0.22$ is very small and only leads to an increase of the shot-noise.

The effective volume of the SDSS-LRG sample is about one order of magnitude above our sample (Eisenstein et al., 2005). At scales of $k = 0.15h/\text{Mpc}$ with $P(k) = 2 \times 10^4$, the SDSS LRG DR7 sample probes a volume of $0.4 \times 10^9 (\text{Mpc}/h)^3$ (Percival et al., 2010), more than one order of magnitude compared to the maximum REFLEX II value.

Fig. 5.6 shows a comparison of the theoretical variance computed using Equation (5.7) with that derived from the diagonal elements of the covariance matrix of the mocks (Equation 5.6) for the luminosity cuts defined in Section 4.3. The theoretical predictions were computed using a linear theory power spectrum with an amplitude rescaled to match that of the REFLEX II measurements for the correspondent luminosity cut. This comparison shows a very good agreement between the results obtained from the ensemble of mocks and the theoretical prediction for all values of L^{\min} . Equation (5.7) thus provides a good estimate of the error bars in the measured power spectra.

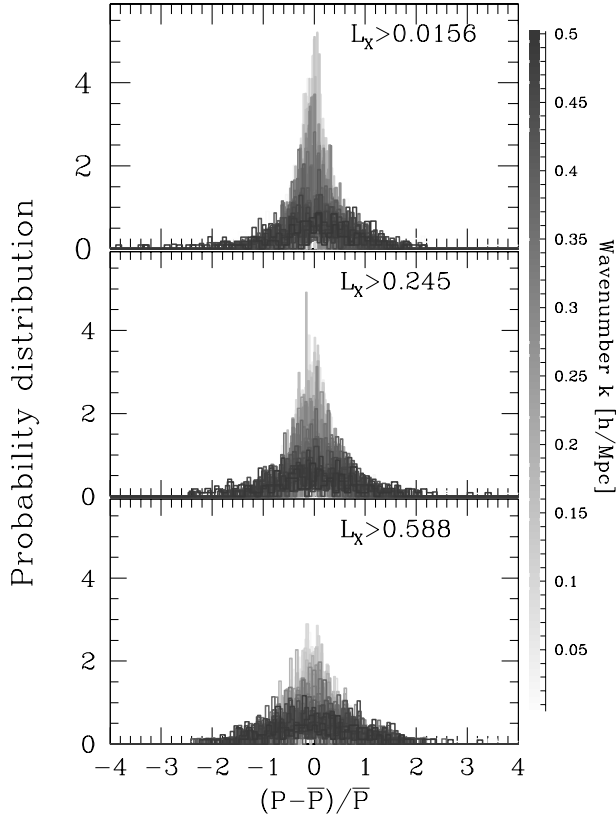


Figure 5.7: Probability distribution of the 100 mock power spectrum with respect to their mean $\bar{P}(k_i)$ for 100 Fourier modes (depicted by the grey scale) and three cuts in luminosity. The lines represents a Gaussian with zero mean and variance given by the fractional error in the mean power spectrum of the mocks at the corresponding scale k .

On Gaussianity

The distribution of the Fourier modes in the ensemble of mock catalogues is shown in Fig.5.7 for three cuts in luminosity. A quick judgment of this plot would suggest that the distribution of Fourier modes is roughly approximated by a Gaussian distribution. On small scales we observe how these distributions acquire broad dispersions due to non-Gaussianities induced by the window function. Also, the broadening of the distribution is more evident as long as we increase the luminosity cut. A more carefully observations of the distributions reveal that they are not symmetric and there is a mild tendency for these distributions to be broader towards positive values of the variable shown in the x -axis. In order to quantify these deviations we determined the skewness $S(k)$ and the kurtosis excess $K(k)$ of the distributions. These quantities vanish if Fourier modes are Gaussian-distributed. The skewness and kurtosis excess are respectively determined with

$$S(k_i) = (N_m - 1)^{1/2} \frac{\sum_{n=1}^{N_m} (\hat{P}_i^n - \bar{\hat{P}}_i)^3}{\left[\sum_{n=1}^{N_m} (\hat{P}_i^n - \bar{\hat{P}}_i)^2 \right]^{3/2}}, \quad K(k_i) = (N_m - 1) \frac{\sum_{n=1}^{N_m} (\hat{P}_i^n - \bar{\hat{P}}_i)^4}{\left[\sum_{n=1}^{N_m} (\hat{P}_i^n - \bar{\hat{P}}_i)^2 \right]^2} - 3. \quad (5.9)$$

These are shown in Fig.(5.8) for the same three luminosity cuts. We have shown with a solid line the prediction for $S(k) = \sqrt{8}N_k^{-1/2}$ and $K(k) = 12/N_k$, where N_k is the number of modes in the spherical shell of width δk . These are valid in the limit when the Fourier modes follow a χ^2 distribution (Takayashi et al., 2008). Figure 5.8 confirms our observations from Fig.5.7 in the sense that the distribution of Fourier modes displays a positive skewness. This is slightly more noticeable in the low luminosity cuts. The behavior of the kurtosis excess is less evident, but there is a mild tendency towards positive values, especially in the higher luminosity cut

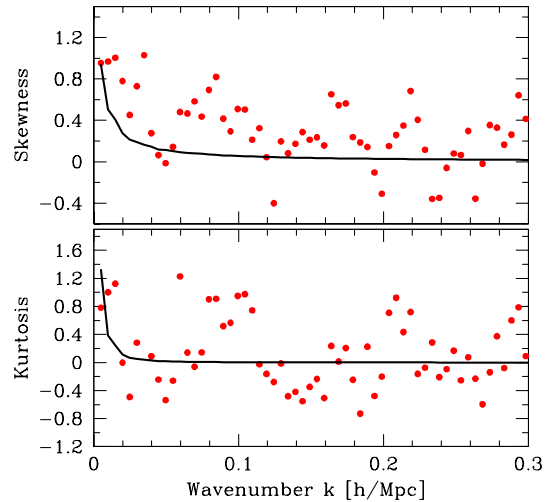


Figure 5.8: Skewness and Kurtosis evaluated from Equation (5.9) using the power spectra measured from the REFLEX II mock catalogues.

shown in the figure.

Due to the relative low number of realizations we have implemented, it is not clear where the presented non-Gaussian features are real. When setting constraints for cosmological parameters using our measurements of power spectrum, we will assume that Gaussianity is a valid approximation.

5.3 Understanding the REFLEX II power spectrum

In this section we use the L-BASICC II simulations to explore the behavior of the luminosity bias using the results described in Chapter 4. The main goal of this section is to show how, having the correct mass-luminosity relation and estimates of intrinsic scatters and luminosity errors, we can model the amplitude of the X-ray galaxy cluster power spectrum when measured as a function of the X-ray luminosity.

5.3.1 Luminosity bias

As was discussed in Chapter 3, under the assumption that on intermediate and large scales the underlying halo-mass bias is scale independent, at a given redshift z we can write the power spectrum of galaxy clusters with a given luminosity L as:

$$P_{\text{cl}}(k, z; L) = b^2(L, z)P_{\text{mat}}(k, z), \quad (5.10)$$

where $P_{\text{mat}}(k, z)$ is the matter power spectrum. The luminosity bias is written in terms of the halo mass function $n(M, z)$ and the underlying halo mass bias $b(M, z)$ as

$$b(L, z) = \frac{\int dM n(M, z) b(M, z) p(L|M, z)}{\int dM n(M, z) p(L|M, z)}, \quad (5.11)$$

where $p(L|M)$ is the probability distribution of assigning luminosity L to a dark matter halo of mass M with scatter $\sigma_{\text{ln}L}$ ¹. In order account for the luminosity errors (that we have introduced in the construction of the mocks) we assume a constant fraction $\sigma = \delta L_X / L_X$ (see Section 4.4)

¹What we called $\Phi(L|M)$ in Chapter 4

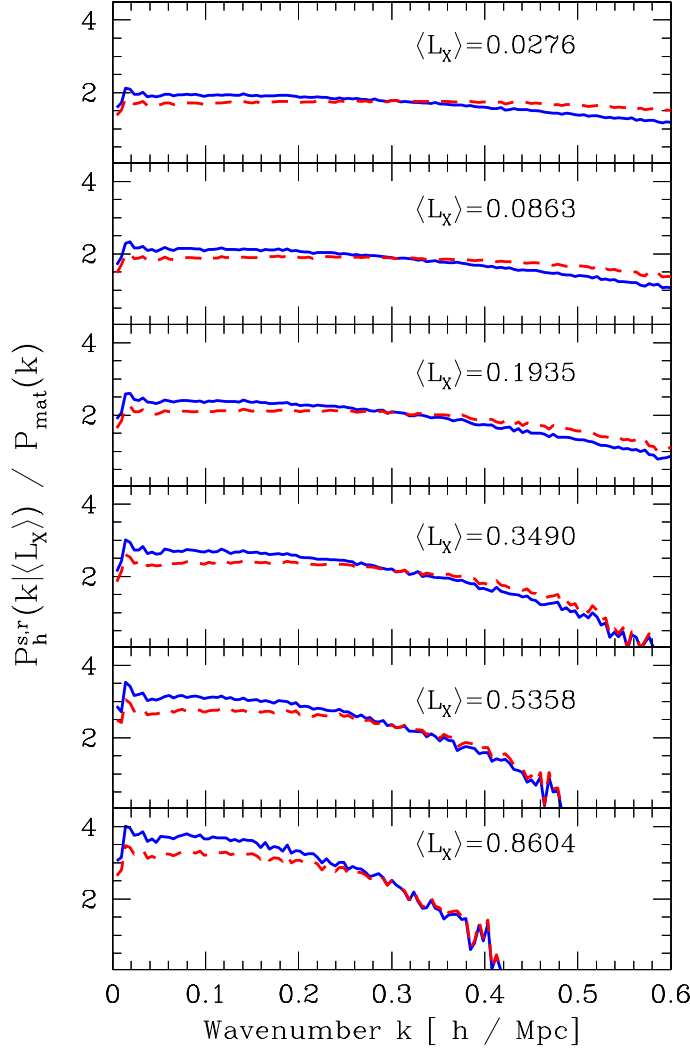


Figure 5.9: Behavior of the ratio halo power spectrum to matter power spectrum for different bins of X-ray luminosity, for which the mean X-ray luminosity is shown in each panel. Solid (dashed) lines represent the ratio in real (redshift) space. Luminosities are given in units of $10^{44} \text{erg s}^{-1} h^{-2}$

and convolve the probability distribution $p(L_X|M)$ with a second log-normal distribution of width σ , such that these two scatters add in quadrature, that is, the effective scatter is

$$\tilde{\sigma}^2 = \sigma_{\text{inL}}^2 + \sigma^2. \quad (5.12)$$

We have assigned luminosities to the dark matter halos in the L-BASICC II simulations, as discussed in Section 4.4, and measured the power spectrum $\hat{P}(k)$ in bins of luminosity. In Fig. 5.9 we show the ratio of the halo power spectrum (in real and redshift space) to the underlying matter power spectrum measured from the same simulation. The ratio is shown as a function of the wavenumber for different bins in luminosity. The mean luminosity in each luminosity bin of width ΔL is shown in the different panels, and its determined with the X-ray luminosity function

$$\langle L_X \rangle_i = \frac{\int_{L_i - \Delta L/2}^{L_i + \Delta L/2} \Phi(L) L dL}{\int_{L_i - \Delta L/2}^{L_i + \Delta L/2} \Phi(L) dL}.$$

With solid lines we show the ratio obtained using the halo power spectrum in real-space. The dashed line shows the corresponding result in redshift-space. A nearly scale-independent

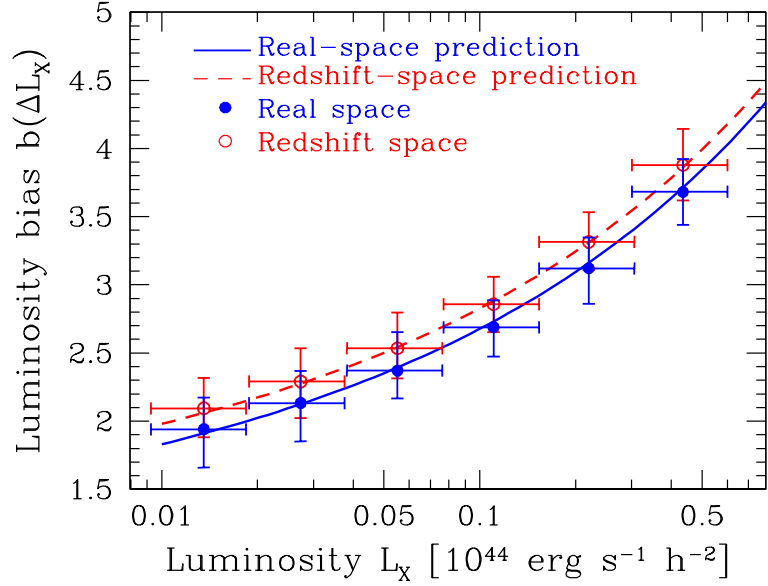


Figure 5.10: Measured luminosity bias from the L-BASICC II simulation in real and redshift space.

bias $\hat{b}_s(\bar{L}_X)$ is observed in the range $0.04 \leq k[\text{Mpc } h^{-1}] \leq 0.1$, both in real and redshift-space. Large scales ($k \geq 0.02h \text{ Mpc}^{-1}$) are slightly dominated by sampling variance (see Fig. 3.9), while small scales are mainly dominated by the halo-exclusion (see Appendix B). Given the direct link between the halo mass and the X-ray luminosity, the halo-exclusion depends on X-ray luminosity of the haloes. The low-luminosity haloes display a constant bias in wider ranges of wavenumbers than that of the high-luminosity haloes. From Fig. 5.9 it can be also seen that halo-exclusion is stronger in real-space than in redshift-space. As explained in Chapter 3, this is due to the fact that on cluster scales, redshift distortions are mainly caused by peculiar velocities such that when viewed in redshift-space, galaxy clusters appear closer, allowing us to probe scales even smaller than the true exclusion-scale.²

Given the halo power spectrum in different bins of luminosity, we measured the ratio

$$\hat{b}_{r,s}(\bar{L}_X) \equiv \sqrt{\frac{\hat{P}_h^{r,s}(k, \bar{L}_X)}{\hat{P}_{\text{mat}}(k)}}. \quad (5.13)$$

for each luminosity bin. We obtain the best estimate for this ratio and compared these results with the prediction from Equation (5.11). Since we are assuming the distant-observer approximation and the large-scale limit, we relate the bias in redshift-space \hat{b}_s to its real-space counterpart \hat{b}_r by means of the Kaiser boost factor (see Equation 2.14):

$$b_s^2(z, L) = b_r^2(z, L)S(L, z) \quad (5.14)$$

where

$$S(L, z) \equiv 1 + \frac{2}{3} \frac{f(z)}{b_r(z, L)} + \frac{1}{5} \left(\frac{f(z)}{b_r(z, L)} \right)^2. \quad (5.15)$$

where $f(z)$ is the growth index defined in Equation (2.15). The comparison between the measured bias and the theoretical prediction of Equation (5.14) is shown in Fig. 5.10. This agreement depends fundamentally on how well we know the halo mass function and the underlying halo-mass bias of the L-BASICC II simulations, together with the underlying

²For a sample of haloes with masses M and radius $R \propto M^{1/3}$, the smaller available scale to be probe by two-point statistics is $\sim 2R$.

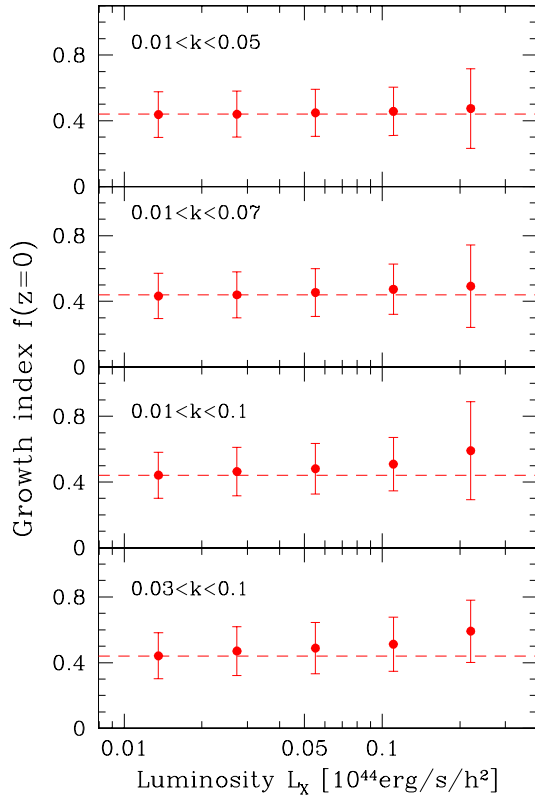


Figure 5.11: Recovered growth index from the illuminated haloes in the L-BASICC II simulation at $z = 0$ assuming the Kaiser boost in the different wavenumbers (in units of $h \text{ Mpc}^{-1}$) as a function of the X-ray luminosity.

mass-luminosity relation. Following the results from Chapter 3, we have used fitting formulae of Jenkins et al. (2001) and Tinker (2007) for the halo mass function and the halo-matter bias respectively.

In forthcoming analysis we will require a measure of the absolute luminosity bias from the L-BASICC II simulations. Although the predictions from Equation (5.11) have provided a good description of the measurements, we prefer to fit the results luminosity bias with three parameters (μ, ν, λ) in the form

$$b_r(L_X) = \mu + \nu L_X^\lambda. \quad (5.16)$$

which is linked to our real-space measurements through the X-ray luminosity function $\Phi(L_X)$ via

$$\hat{b}_r(\bar{L}_X) = \frac{\int_{\Delta L} \Phi(L_X) b_r(L_X) dL_X}{\int_{\Delta L} \Phi(L_X) dL_X}. \quad (5.17)$$

In Table 5.1 we show the fitting values of the set of parameters (μ, ν, λ), where we show the fits obtained in four ranges of wavenumber, both in real and redshift space. The parameters determined from redshift space are derived using Equation (5.14) and the bias given by Equation (5.17). The parameters determined from the two spaces are the same within 2σ of their errors, which implies that on the scales where the bias was measured, the Kaiser approximation is valid.

A simple way of testing the Kaiser boost factor in a given range of wavenumbers and luminosities (or masses) consists in recover the growth factor $f(z)$ using the measurements of the bias. The result from this simple exercise is shown in Fig. 5.10. The data points are

$\Delta k (h \text{ Mpc}^{-1})$	μ	ν	λ	Recovered $\hat{f}(z=0)$
Real space:				
$0.01 \leq k \leq 0.05$	1.26 ± 0.02	$3.20^{+0.04}_{-0.05}$	0.36 ± 0.01	0.451
$0.01 \leq k \leq 0.07$	$1.25^{+0.04}_{-0.05}$	3.24 ± 0.04	0.36 ± 0.01	0.458
$0.01 \leq k \leq 0.1$	$1.30^{+0.05}_{-0.03}$	3.22 ± 0.02	0.38 ± 0.01	0.497
$0.03 \leq k \leq 0.1$	$1.27^{+0.05}_{-0.03}$	$3.18^{+0.01}_{-0.02}$	0.36 ± 0.01	0.500
Redshift space:				
$0.01 \leq k \leq 0.05$	$1.26^{+0.02}_{-0.06}$	3.24 ± 0.04	0.37 ± 0.01	
$0.01 \leq k \leq 0.07$	1.19 ± 0.04	3.29 ± 0.02	0.35 ± 0.01	
$0.01 \leq k \leq 0.1$	$1.33^{+0.03}_{-0.05}$	3.36 ± 0.04	0.40 ± 0.01	
$0.03 \leq k \leq 0.1$	$1.29^{+0.03}_{-0.05}$	3.32 ± 0.04	0.38 ± 0.01	

Table 5.1: Numerical values of the parameters (μ, ν, λ) characterizing the absolute luminosity bias drawn from the L-BASICC simulations with illuminated haloes following the REFLEX II luminosity function. The measurements are done in four ranges of wavenumber, both in real and redshift space. The fitting in redshift space uses the Kaiser boots. The fourth column of the first panel shows the average (of the five luminosity bins) of the measured growth factor assuming the Kaiser approximation. The fiducial cosmology of the simulation yields $f(z=0) = 0.44$, following Equation (2.15).

obtained by using the fitting to the luminosity bias obtained in real-space and using Equation (5.14) to solve for the growth factor. The recovered growth factor is shown in Table 5.1. We observed that the bias measured in the range of wavenumbers $0.01 < k (h \text{ Mpc}^{-1}) < 0.05$ provides unbiased estimates of $f(z=0)$ even in the high luminosity bins. On wider ranges of wavenumbers, only the low-luminosity bins can provide an unbiased estimate of the growth index.

5.3.2 Effective bias

The results of Equations (5.14) and (5.17) can also be used to estimate the effective bias of a power spectrum measurement in the REFLEX II catalogue for a given luminosity cut L^{\min} (by effective bias we mean the bias of objects with luminosities above certain value). In order to achieve this, we first write the cluster power spectrum in redshift-space at a given redshift z in terms of the linear dark-matter power spectrum at $z=0$ as

$$P_{\text{cl}}^s(k, z; > L^{\min}) = b^s(> L^{\min}, z)^2 P_{\text{mat}}(k, z=0), \quad (5.18)$$

where

$$b^s(z, > L^{\min})^2 = b(> L^{\min}, z)^2 S(z, > L^{\min}) g^2(z). \quad (5.19)$$

where $g(z)$ denotes the growth factor arising from the matter power spectrum at a redshift z (see Section 2.3.2). In this expression we take into account that the minimum luminosity included in the sample varies with z following the REFLEX II selection function. Accordingly, we determine the bias $b(z, > L^{\min})$ following Equation (5.14) with the redshift dependence given by the REFLEX II sensitivity map:

$$b(z, > L^{\min}) = \frac{1}{\bar{n}(z, > L^{\min})} \int_{L(z)}^{\infty} \Phi(L) b(L) dL, \quad (5.20)$$

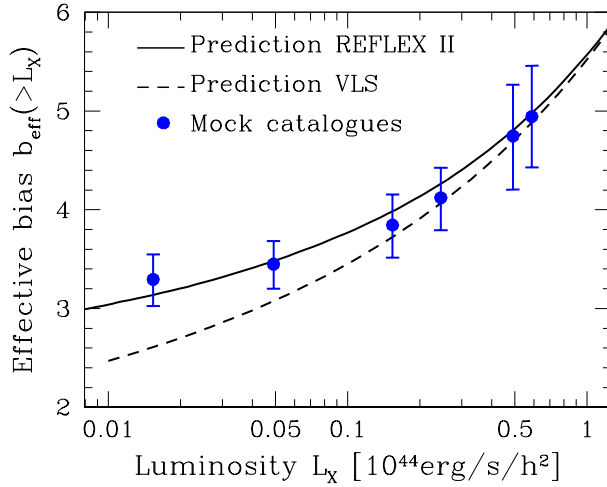


Figure 5.12: Effective luminosity bias measured from the REFLEX II mock catalogues (filled squares with error bars). The solid line shows the prediction from Equation (5.21) with the fiducial values $\sigma_{\ln L} = 0.26$ and constant luminosity error of 20 per cent. The dashed line is the prediction for a VLS.

where the mean number density $\bar{n}(z, > L^{\min})$ and the lower integration limit $\hat{L}(z)$ are given by Equations (4.9) and (4.10) respectively. Note that $\hat{L}(z)$ should not only depend on the redshift but also on the angular position, according to the REFLEX II sensitivity map. However, in order to give an estimate of the effective bias at a fixed redshift we made an average of the values of $L^{\min}(\mathbf{r})$ of all the N_{pix} pixels within the REFLEX II mask. Individual pixels displayed small differences compared to the average in the final result.

The effective bias of the full sample will then be given by the average of the bias factors of Equation (5.20) over the observed volume as (e.g. Suto et al., 2000; Moscardini et al., 2000)

$$b_{\text{eff}}(> L^{\min})^2 = \frac{\int_z [\bar{n}(z, > L^{\min}) b^s(> L^{\min}, z)]^2 \frac{dV}{dz} dz}{\int_z \bar{n}(z, > L^{\min})^2 \frac{dV}{dz} dz}, \quad (5.21)$$

where the integrals are evaluated in the redshift interval of the sample with a volume element $dV/dz = r(z)^2/H(z)$ according to our fiducial cosmology. This therefore allows us to write a prediction for the observed power spectrum from a sub-sample characterized by the luminosity cut L^{\min} as

$$P_{\text{cl}}(k; > L^{\min}) = b_{\text{eff}}(> L^{\min})^2 P_{\text{mat}}(k, z = 0). \quad (5.22)$$

The solid line in Fig. 5.12 shows the prediction for the effective bias of the REFLEX II catalogue as a function of the minimum luminosity L^{\min} computed using Equation (5.16) (setting $D(z) = 1$ since, by construction, the mock catalogues assume no redshift evolution). For comparison, the dashed line in Fig. 5.12 represents the equivalent prediction for a volume limited sample. Equation (5.21) gives an excellent description of the direct measurements obtained from the mock catalogues (shown by the filled circles). Then, this model provides a means to extract the important cosmological information contained in the amplitude of the measured REFLEX II power spectrum.

5.4 Distortions induced in a flux limited sample

Due to the flux-limited nature of the REFLEX II catalogue, large scales might be probed predominantly the most luminous clusters, with a higher clustering amplitude. This would artificially increase the measured power spectrum on large scales, introducing a scale-dependent

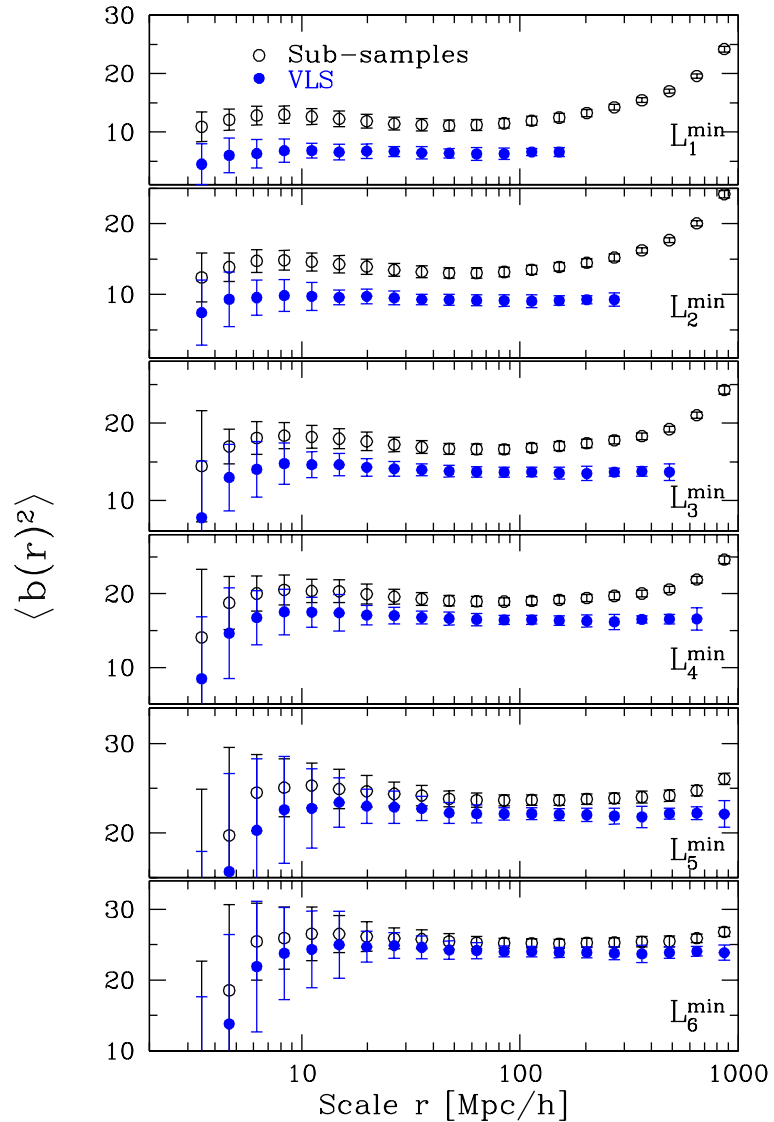


Figure 5.13: Mean squared luminosity bias for pairs separated by a scale r in the REFLEX II mock catalogues, for the sub-sample L_2^{\min} (filled circles) and its corresponding VLS (open circles).

distortion with respect to the volume limited case. Schuecker et al. (2001) analyzed this problem in detail for the REFLEX sample and concluded that no significant effect can be detected for scales $r < 150 \text{ Mpc } h^{-1}$. In this section we perform a similar analysis on the REFLEX II sample. As the REFLEX II sample spans a wider range of luminosities, and covers a larger volume than that used by Schuecker et al. (2001), it is necessary to test whether this systematic effect can affect our measurements.

Fig. 5.13 shows the number of pairs with separation r weighted by the individual luminosity bias factors of each pair member (computed using Equation 5.16), yielding the average squared bias factors

$$\langle b^2(r) \rangle = \frac{1}{n(r)} \sum_{i,j} b(L_i)b(L_j), \quad (5.23)$$

where the sum is done over pairs separated by scales in the range $r - \frac{1}{2}\Delta < |r_i - r_j| < r + \frac{1}{2}\Delta$ and $n(r)$ is the number of clusters in the same interval. The open circles show the results obtained from the sample with minimum luminosity L_2^{\min} (see Table 4.1), while the filled

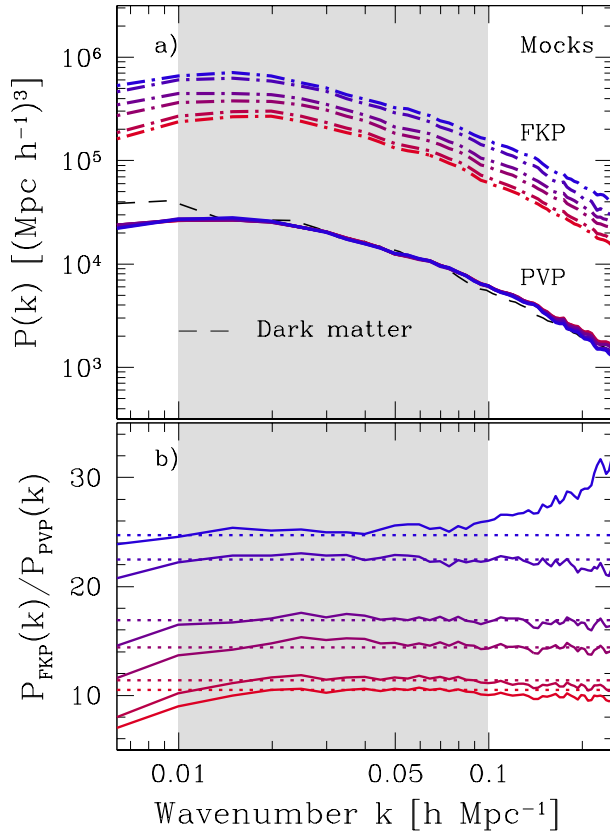


Figure 5.14: a) Comparison of the resulting mean power spectrum from the mock sample using the PVP estimator (solid lines) and the FKP result (dashed lines) for our six cuts in luminosity defined in Table 4.1. The dashed line represents the measurement of the dark matter power spectrum of the L-BASICC II simulations. The shaded region shows the scales where the absolute bias is fitted ($0.01 \leq k \text{ (h Mpc}^{-1}) \leq 0.1$) b) Ratio of FKP power spectra to the PVP measurements for the six luminosity cuts.

circles correspond to the equivalent measurement from the VLS determined by the same minimum luminosity. The error bars are drawn from the variance of the ensemble. Two prominent features can be observed from Fig. 5.13. On one hand, there is an increase in the mean bias on scales $r \approx 10 \text{ Mpc } h^{-1}$. This is understood as to show that pairs of clusters separated by these scales have at least one cluster with luminosity bias higher than the mean value of the sample. This is therefore a consequence of gravitational clustering. On scales smaller than $r \sim 9 \text{ Mpc } h^{-1}$ the mean bias decreases as a consequence of the halo-exclusion (e.g. Porciani et al., 1998). On the other hand, on scales $r \gtrsim 150 \text{ Mpc } h^{-1}$, a systematic increase of the average squared bias factor exists in the sub-sample compared with the volume-limited case. This is a direct consequence of the flux-limited nature of the survey. Naively, the scale where this flux-selection distortions is relevant would correspond to $k = 2\pi/r = 0.04 \text{ h Mpc}^{-1}$, suggesting that for wavenumbers smaller than this limit no scale-dependent distortion affects the measurements from the REFLEX II catalogue. Comparing with other sub-samples, we observe that this distortion is damped in the high luminosity cuts, which almost behave like a volume-limited sample (see Fig. 4.4).

To analyze this issue in more detail we used the FKP estimator as implemented by Percival et al. (2004b, hereafter PVP). This is also a minimal variance weighting estimator which takes into account the luminosity bias to obtain an estimate of the power spectrum free of the distortions induced by the flux-limited selection of the sample. Panel a) in Fig. 5.14 shows a comparison of the mean power spectra obtained by the FKP (dot-dashed lines) and PVP (solid lines) algorithms for the different luminosity cuts defined in Table 4.1 in our ensemble of mock catalogues. In the PVP method, we weighted each object by the inverse of its luminosity bias, computed using a fit to the redshift-space results shown in Fig. 5.11.

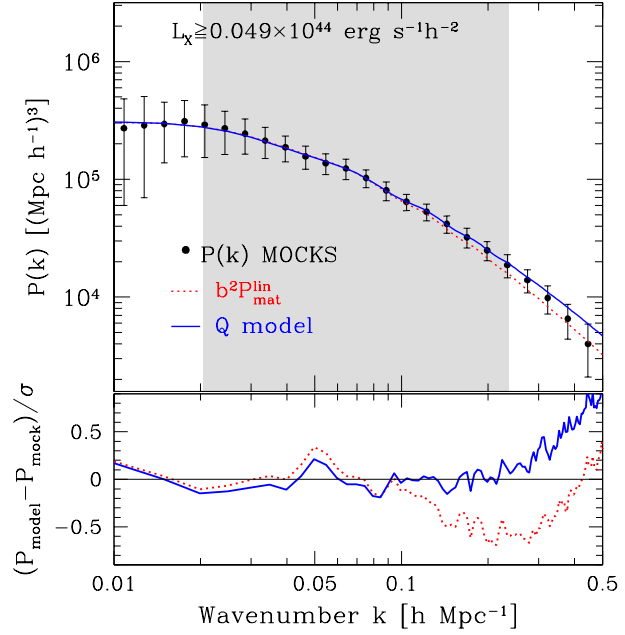


Figure 5.15: Q -model description of the mock power spectrum. The shaded region shows the scale where the Q -model was used to fit the mean power spectrum from the mock catalogues. The bottom panel show the ratio of the difference between the Q -model and the measurements to the variance of the mocks.

The shaded area in Fig. 5.14 represents the range of scales used to measure these bias factors ($0.01 \leq k/(h\text{Mpc}^{-1}) \leq 0.1$). This produces a power spectrum normalized as that of the dark matter distribution (shown by the dashed line in Fig. 5.14). Panel b) shows the ratios between the mean power spectra obtained using the FKP and PVP estimators for each luminosity cut. In the absence of a scale-dependent distortion, these ratios should correspond to the bias factors $b_{\text{eff}}^2(> L_{\text{min}})$, shown by the dotted lines. These ratios show no clear signature of a scale-dependent distortion for $k \leq 0.02 h\text{Mpc}^{-1}$. Contrary to what might be expected from this systematic effect, Fig. 5.14 shows a weak indication of a decrease in power on large scales for the lower luminosity cuts, which is smaller than the variance of the measurements. From this analysis we conclude that no significant distortion is introduced in the shape of the power spectra estimated with the FKP method. Note however that this statement applies exclusively to the REFLEX II catalogue, as this effect has been detected in galaxy surveys (e.g. Tegmark et al., 2004; Percival et al., 2007a).

Panel b) of Fig. 5.4 shows the correlation matrix inferred from the ensemble of mock catalogues of the power spectra for two luminosity cuts obtained using the PVP estimator. A comparison with panel a) shows that the PVP estimator induces a stronger covariance between the power spectrum measurements in different bins. For this reason, together with the lack of systematic distortions in the FKP measurements, we chose to use the FKP estimator to analyze the data from the REFLEX II catalogue.

5.4.1 Modeling the shape of $P(k)$

In this section we use our ensemble of mock catalogues to test a model of the shape of the REFLEX II power spectrum. We focus on the clusters with luminosities greater than L_2^{min} . The filled circles in Fig. 5.15 show the mean redshift-space power spectrum of the mocks for this luminosity cut with error-bars determined from the variance of the ensemble. It can be clearly seen that this measurement exhibits an excess of power at small scales with respect to the predictions from linear perturbation theory, shown by the dashed line. This is due to

the combined effect of non-linear evolution and redshift-space distortions.

In recent years the distortions in the shape of the power spectrum produced by these effects have been intensively studied using large N -body simulations and recent advances in perturbation theory (e.g. Smith et al., 2007; Crocce and Scoccimarro, 2006; Angulo et al., 2008; Sánchez et al., 2008; Montesano et al., 2010). These analysis have produced accurate descriptions of these distortions to the level of accuracy demanded by forthcoming surveys, which will probe volumes much larger than present day catalogues. Due to the moderate volume probed by the REFLEX II catalogue, percent-level accuracies in the treatment of these effects are not required.

We now test whether or not the Q -model of Cole et al. (2005) (modified as in Sánchez et al., 2008) can provide a good description of the non-linearities observed in the power spectra of our mock catalogues. In this model the shape of the cluster power spectrum is given by

$$P_{\text{cl}}(k, > L) = b_{\text{eff}}(> L)^2 \left(\frac{1 + Qk^2}{1 + Ak + Bk^2} \right) P_{\text{mat}}^{\text{lin}}(k), \quad (5.24)$$

where $P_{\text{mat}}^{\text{lin}}(k)$ is the linear theory matter power spectrum. We follow Cole et al. (2005) and fix the value of $A = 1.4$ as obtained from the analysis of N -body simulations, while Q and B are left as free parameters whose values will depend on the limiting luminosity of the sample. We assumed all the cosmological parameters to be known and fitted for Q and B , analytically marginalizing over the amplitude (as described in Lewis and Bridle, 2002). From this analysis we obtain the values $Q = 24.9 \pm 1.1$ and $B = 12.0 \pm 2.1$, corresponding to the sub-sample defined by L_2^{min} . The best fit model obtained this way is shown by the solid line in Fig. 5.15. It can be clearly seen that the model of Equation (5.24) gives an accurate description of the shape of the mean power spectrum from our ensemble of mock catalogues. This can be also seen in panel b) of the same figure, where we show the ratio between the difference of the mean mock power spectrum and the best fit-model to the variance from the ensemble. The parameters B and Q fitting the power spectrum of the sub-sample L_2^{min} follow a degeneracy that can be described approximately by $B(Q) = 0.805Q - 8.15$. This degeneracy is maintained if the amplitude of the model is fixed according to Equation (5.21). We can thus use this degeneracy to reduce the number of degrees of freedom when constraining cosmological parameters using the measured power spectrum. The best fit value of Q increases with the limiting luminosity of the sample, varying from $Q = 20.7 \pm 0.9$ for L_1^{min} to $Q = 44.9 \pm 2.3$ for L_6^{min} . The general trend in the degeneracy $B(Q)$ is maintained for different luminosity cuts. In Section 5.5.3 we compare the predictions of this model with the measurement of the REFLEX II power spectrum.

5.5 Analysis of the REFLEX II power spectrum

5.5.1 Measurements

The measured power spectrum for the REFLEX II sample with limiting luminosity L_1^{min} is shown by the filled points in Fig. 5.16 with error bars drawn from the FKP method (see Equation 5.7). The solid line represents a Λ CDM linear power spectrum, convolved with the window function of the survey. This theoretical prediction was computed using the fitting formulae of Eisenstein and Hu (1998), with amplitude rescaled to match that of the REFLEX II measurement. This simple exercise shows that the shape of the REFLEX II power spectrum is consistent with the predictions of the Λ CDM cosmological model.

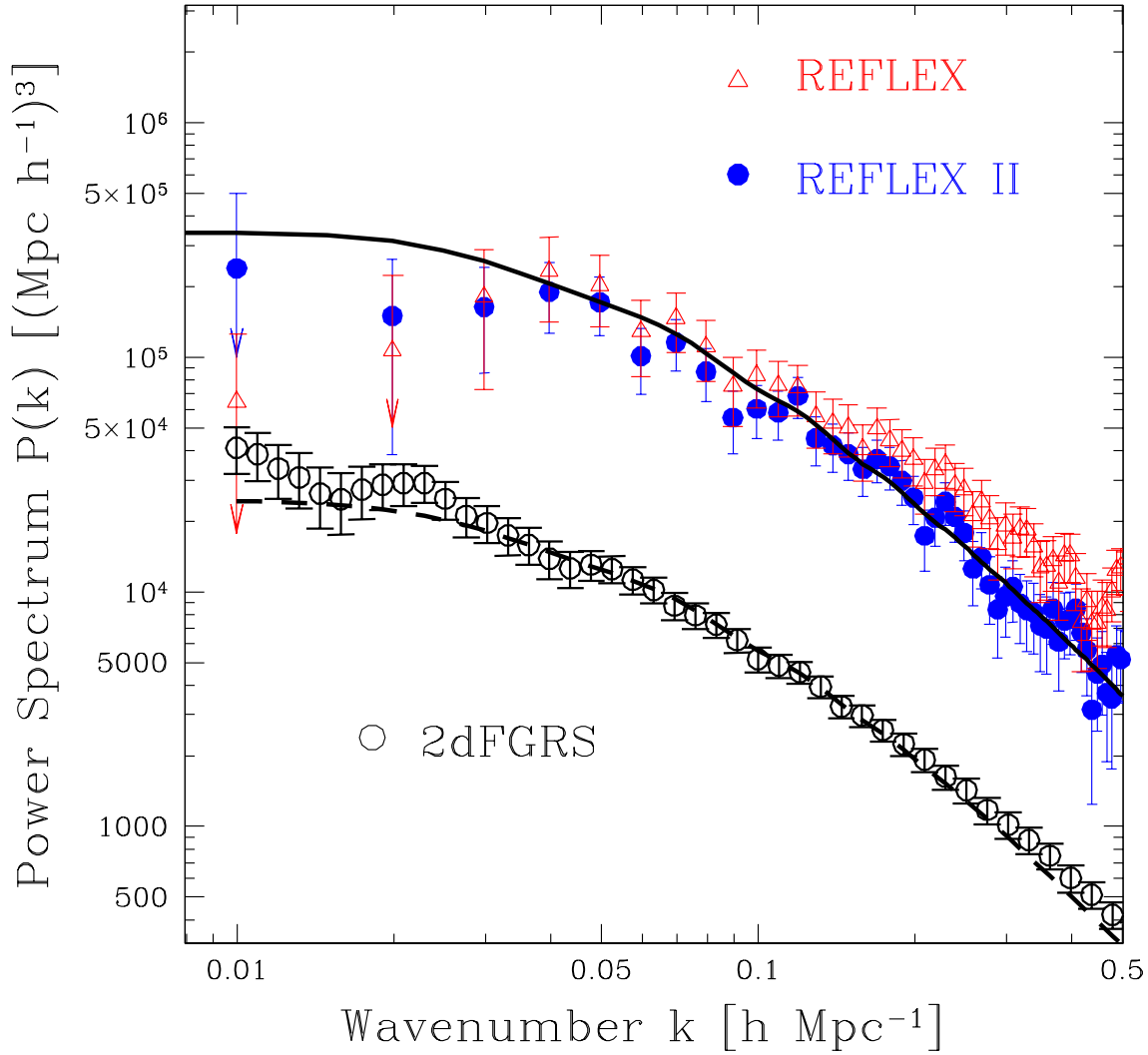


Figure 5.16: REFLEX II power spectrum (filled circles with error bars) for clusters with luminosities $L_X > L_1^{\min}$. The REFLEX power spectrum is shown by the open triangles. These error bars for these two measurements are taken from Equation (5.7). For comparison we also show the measured power spectrum from the 2dFGRS taken from Cole et al. (2005) (empty circles). The dashed line represents the Λ CDM power spectrum convolved with the REFLEX II window function. Error-bars exceeding the range of the plot are represented by arrows.

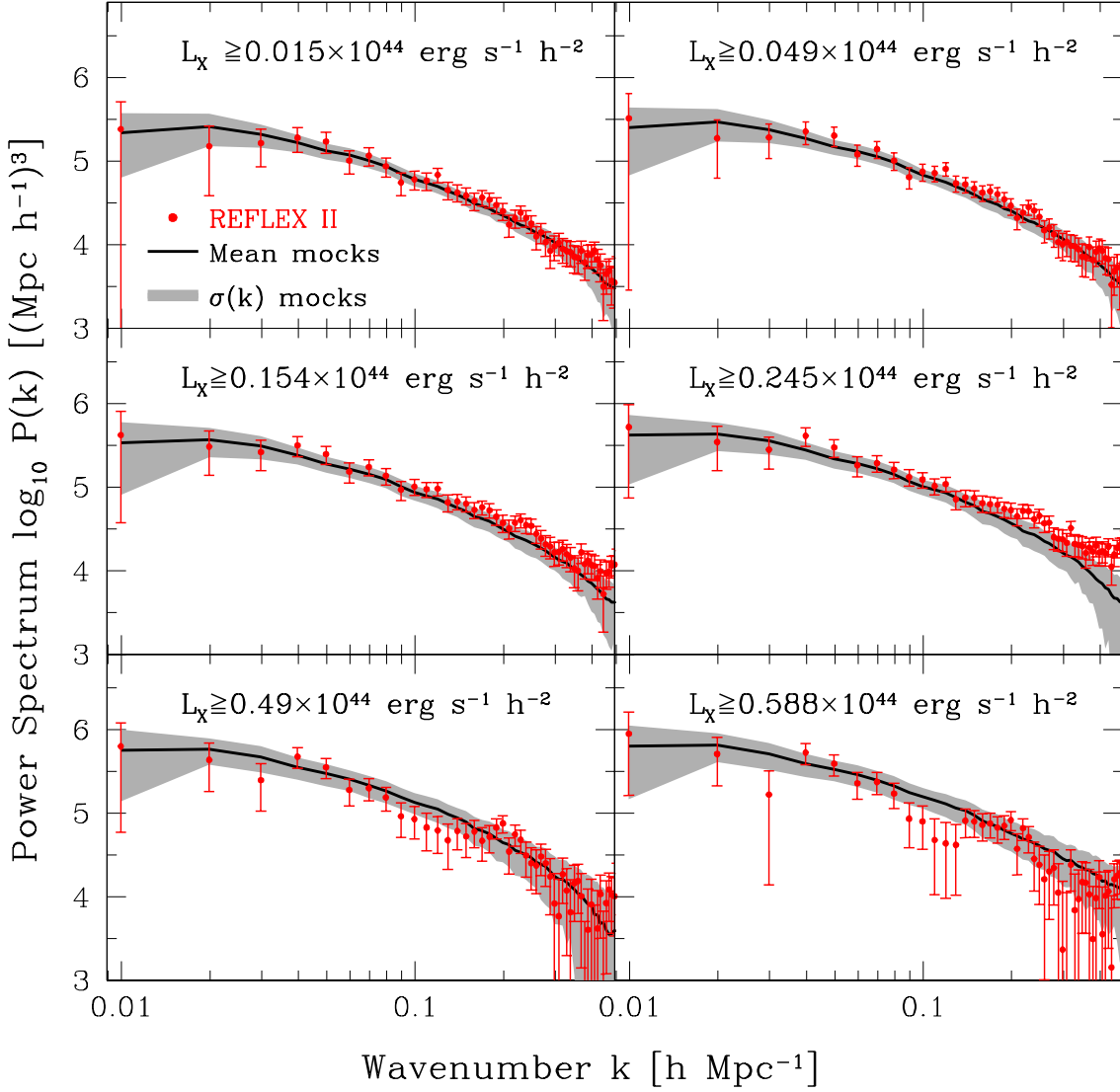


Figure 5.17: Measured REFLEX II power spectrum for the different sub-samples defined in Table 4.1. FFTw is done in a box of side $L_{\text{box}} = 1263.8 \text{ Mpc } h^{-1}$ and $P_{\text{est}} = 2 \times 10^4 (\text{Mpc } h^{-1})^3$. The fundamental mode is $\delta k = 2\pi/L_{\text{box}} = 0.0049 \text{ h Mpc}^{-1}$ and the spherically averaged is done on shells with width $\delta k = 2\Delta k$. The Nyquist frequency is $k_N = 1.27 \text{ h Mpc}^{-1}$. Points represent the REFLEX II measurements with error bars drawn from Equation (5.7). The shaded region represents the 1σ variance determined from the mocks catalogues. The solid line represents the mean mock power spectrum.

Fig. 5.16 also shows a new estimation of the power spectrum of the original REFLEX sample (open triangles). The REFLEX power spectrum has a higher amplitude, as expected from the higher flux-limit of this sample, and its shape is in good agreement with that of the REFLEX II measurement. The larger volume probed by the new catalogue reduces the impact of cosmic variance on large scales, where the REFLEX II power spectrum exhibits a higher amplitude than the measurement in the original REFLEX sample. Fig. 5.16 also shows the galaxy power spectrum measured from the 2dFGRS (Cole et al., 2005). The dashed line represents the same Λ CDM power spectrum described above convolved with the 2dFGRS window function. This shows that, once their respective window functions have been taken into account, the large-scale ($k < 0.1 h\text{Mpc}^{-1}$) shape of the power spectra inferred from the REFLEX II and the 2dFGRS are in good agreement and can be described with the same cosmological model. At smaller scales, redshift-space distortions and non-linear evolution produce deviations in the shapes of these power spectra.

Fig. 5.17 shows a comparison of the measured REFLEX II power spectra (points with error bars) for the six cuts in luminosity described in Section 4.3, and the corresponding mean power spectra from the mock catalogues (solid lines), with their corresponding 1σ variance (shaded regions). The error bars of the REFLEX II power spectrum correspond to the theoretical prediction of the FKP method (see section 5.2.1). We observe that the spectra measured in the mocks are compatible within 1σ with the REFLEX II clustering up to $k \approx 0.3 h\text{Mpc}^{-1}$ for all luminosity cuts. Notice that the mocks were only calibrated to follow the X-ray luminosity function of the REFLEX II sample. This agreement allows us to use the covariance matrices inferred from the ensemble of mock catalogues when analyzing the REFLEX II measurements.

On scales $0.05 < k < 0.2$ we can fit the measured power spectrum with a power law $P \propto k^n$. For the mean of the mock catalogs determined from the sub-sample L_2^{min} we obtained $n = -1.22^{+0.16}_{-0.15}$. The spectral slope in the volume limited samples with the same limiting luminosity is $n = -1.11^{+0.21}_{-0.19}$ (in both cases, error bars are determined with the FKP method, see section 5.2.1). The steeper slope in the sub-samples is a signature of the selection effects, i.e, the presence of different luminosities in the sample. Nevertheless, these values are compatible within the 1σ error bars, which suggest that such effect might not significantly affect the measurements obtained from a flux-limited sample, as we have discussed in Section 5.4.

On small scales the power spectra inferred from the mock catalogues are affected by the halo exclusion effect. Dark matter halos in the simulations have been counted as separate entities when they did not overlap with their radii of r_{FoF} . On the other hand, in the REFLEX II catalogue clusters have been treated as distinct if their X-ray emission does not significantly overlap. Due to the short exposures in the RASS, the outer boundary of the X-ray emission (in two dimensional images, which is significantly smaller than the aperture radius determined from one-dimensional profiles) is smaller than the radii of r_{500} . This produces differences between the REFLEX II power spectra and the results from the mock catalogues on scales $k > 0.2 h\text{Mpc}^{-1}$.

Two noticeable features are observed in the REFLEX II power spectrum. The first one appears as a bump at scale $k \approx 0.5 h/\text{Mpc}$ in the spectra measured for the six luminosity cuts. The strength of the feature increases with the limiting luminosity. Given the Nyquist frequency of the measurement, we expect that aliasing associated to the mass assignment scheme occurs on wavenumbers $k \geq 0.7 h/\text{Mpc}$. Therefore this feature cannot be associated to aliasing effects. This is confirmed by the fact that this feature is not present in power spectrum from the mock catalogues. This signal has been also detected in the measurements

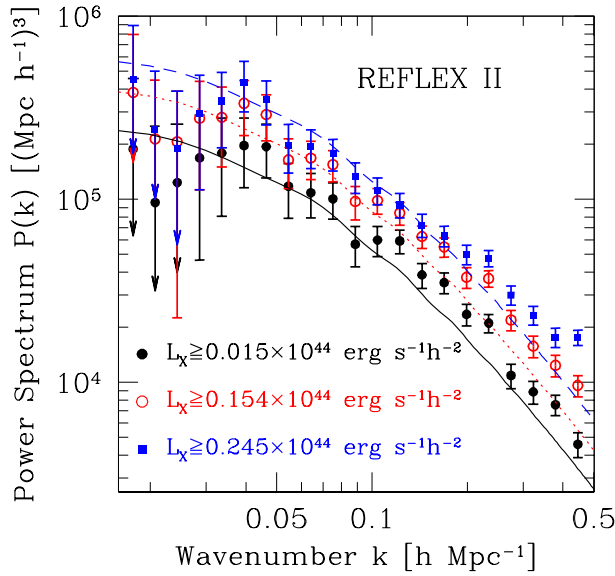


Figure 5.18: REFLEX II power spectrum for three different minimum luminosities. Lines represent a linear-perturbation theory power spectrum with our fiducial cosmology and amplitude given by Equation (5.21). Error bars correspond to the variance drawn from the mock catalogues. Arrows denote the error bars exceeding the range of the plot.

with different values of P_{est} . On these scales there is however a strong mode coupling, as can be seen from the covariance matrix (Fig. 5.4); thus a small increment of power on this scales is transferred to the neighbor modes generating such a feature. We have measured the REFLEX II spectra with different maximum redshifts and in luminosity bins: we still found this feature. We cannot associate this excess of power to a particular feature in configuration space, as can be seen from the behavior of the distribution of pairs shown in Fig. 4.14. The second feature is visible in the higher luminosity cuts as a suppression in the REFLEX II power spectrum on two different scales: the first occurs at $k \approx 0.03h/\text{Mpc}$ and the second at $k \approx 0.1h/\text{Mpc}$. The first of these features could be also interpreted as a peak on scales $\approx 0.05h/\text{Mpc}$, in agreement with the observations from early measurements of power spectrum of optical galaxy clusters (e.g. Einasto et al., 1997). Note that the error bars of our measurements are such that these are consistent with the mean of the power spectrum measured from the mean of the ensemble of REFLEX II mocks. Note however that in the highest luminosity cut, the Fourier modes on these two scales have an amplitude which is almost 2σ away from the amplitude of the mean power spectrum of the mocks. Nevertheless, being these scales dominated by cosmic variance, we do not draw physical conclusions from these features.

Regarding the redshift incompleteness of the REFLEX II catalogue (around 10 per cent), we verified that our results are not substantially modified when the power spectrum is measured after randomly subtracting up to 20 per cent of the total number of clusters.

5.5.2 Amplitude of the REFLEX II $P(k)$

Fig. 5.18 shows the measurements of the REFLEX II power spectra for three of the subsamples defined in Table 4.1. The increase in the amplitude with increasing minimum luminosity can be clearly seen, showing the signature of luminosity bias.

In Section 5.3.1 we showed that the measurements of the effective bias of the REFLEX II mock catalogues are well described by the predictions of Equation (5.21). In this section we confront this prediction with the power spectra measured from the REFLEX II sample. In order to avoid using the underlying dark matter power spectrum, we test Equation (5.21) by means of the a relative luminosity bias $r(L_X)$, defined as the ratio between the power spectrum

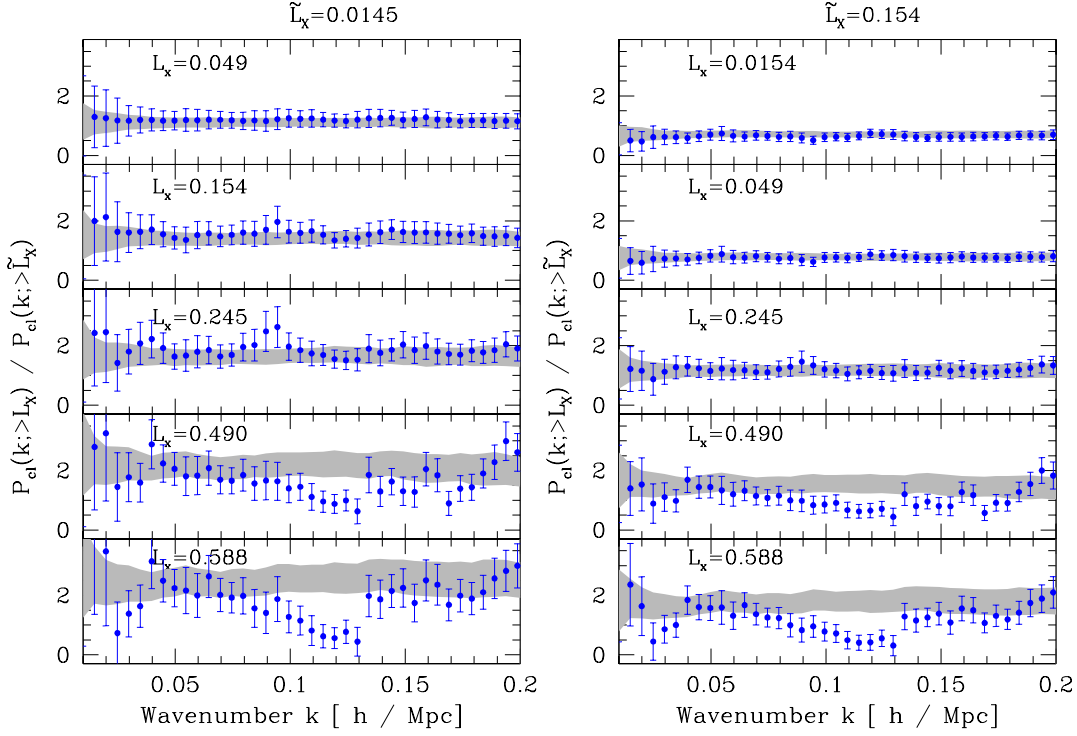


Figure 5.19: Relative luminosity bias in the REFLEX II sample using as reference luminosities the values L_1^{\min} and L_3^{\min} . The clustering of the REFLEX II is compatible, on large scales, with a constant relative bias. The higher luminosity cuts display scale-dependent features that can be associated to statistical noise due to the low sampling power of these sub-samples.

of the sub-sample defined by a minimum luminosity L_X to that of clusters with luminosities greater than a reference value \tilde{L}_X :

$$r(L_X) \equiv \frac{b_{\text{eff}}(> L_X)}{b_{\text{eff}}(> \tilde{L}_X)}. \quad (5.25)$$

The behavior of this ratio as a function of the wave number is shown in Fig. 5.19. This figure shows how the constant relative bias is a good approximation up to scales $k \approx 0.15 h \text{Mpc}^{-1}$. The results are shown in Fig. 5.20. The open squares show the measurements from the REFLEX II data and the filled points correspond to the respective measurement from the mock catalogues. The solid line shows the prediction from Equation (5.21), while the dashed line is the prediction of the effective bias for a VLS. The prediction from Equation (5.21) provides a good description of the REFLEX II measurements in the low luminosity cuts. As can be seen in Fig. 5.17, the spectra measured from the last two luminosity cuts are not consistent with a constant bias, even on large scales. This might explain the points corresponding to these luminosity cuts lie below the theoretical prediction. These results confirm the validity of Equation (5.21) to model the amplitude of the REFLEX II power spectrum.

5.5.3 Shape of the REFLEX II $P(k)$

We used the Q -model to analyze the shape of the REFLEX II power spectrum for the sub-sample defined by L_2^{\min} , following the same procedure as in Section 5.4.1. The best-fit of the

Wavenumber k	$P_{\text{REFLEXII}}(k)$	$\sigma_{\text{REFLEXII}}(k)$	$\bar{P}_{\text{mock}}(k)$	$\sigma_{\text{mo}}(k)$
0.03522	220204	114433	202063	69608.4
0.07045	117872	42863.1	108441	22370.4
0.10568	63247.4	22339	59637.3	11170.7
0.14091	40167.8	14366.5	40924.4	7123.94
0.17614	32901	10828.1	27600.2	5884.77
0.21136	13764.3	7381.84	21819.3	5355.84
0.24659	21537.7	6931.1	16010.6	4151.87
0.28182	5421.56	4901.49	12447.5	2942.32
0.31705	11267.4	4760.07	8986.62	2546.83
0.35228	6660.7	4000.63	7121.6	2891.89
0.38751	6964.62	3657.29	5865.72	2295.65
0.42273	6224.22	3313.47	3679.41	2573.55
0.45796	5980.59	3043.57	3334.65	2211.65
0.49319	6121.52	2833.01	2820.51	2140.65
0.52842	4111.92	2557.12	1854.4	2150.58

Table 5.2: Numerical values for the power spectrum of the REFLEX II flux limited sample with $L_X > 0.049 \times 10^{44} \text{erg}/sh^{-2}$ and $z_{\text{max}} = 0.22$. The variance obtained from Equation (5.7) $\sigma_{\text{RE}}(k)$ are shown. The corresponding mean value of the mocks power spectrum $P_{\text{mocks}}(k)$ is shown with the variance from the 100 mock catalogues $\sigma_{\text{mock}}(k)$. The power spectrum and the variance are in units $[(\text{Mpc}/h)^3]$. The wavenumber is in units $[\text{h}/\text{Mpc}]$. We show the linear spaced Fourier modes each eight modes.

Q -model is shown by the solid line in the upper panel of Fig. 5.21. For this measurement we find $Q = 24.7 \pm 1.5$ and $B = 8.6 \pm 1.1$, with a degeneracy described by $B(Q) = 0.72Q - 9.25$. This degeneracy is maintained when fixing the amplitude of the model power spectrum according to Equation (5.21). As in Fig. 5.15, the bottom panel shows the ratio of the difference between the model and the measurements to the variance determined from the mock catalogues. This shows that a model including a correction for non-linearities provided by Equation (5.24) gives a better description of the shape of the REFLEX II power spectrum than the predictions from linear perturbation theory.

5.6 Correlation function

The clustering of the REFLEX II galaxy clusters has been also explored through the two-point correlation function. These measurements will be analyzed in a forthcoming paper (Sánchez et al., in preparation). The two-point correlation function has been defined in Equation (2.6) and corresponds to the Fourier transform of the power spectrum (see Equation 2.9). By definition, these two-point statistics encode the same information concerning the inhomogeneous spatial distribution of clusters. Therefore similar constraints on cosmological parameters would be obtained from them. Nevertheless, the selection strategy affect these these measurements in different forms, which might lead to different results regarding the constraints on cosmological parameters. As we have explained before, in Fourier space the resulting measurement of power spectrum is the convolution of the underlying cluster power spectrum with the survey window function. In configuration space, the window function acts as a multiplicative (scale

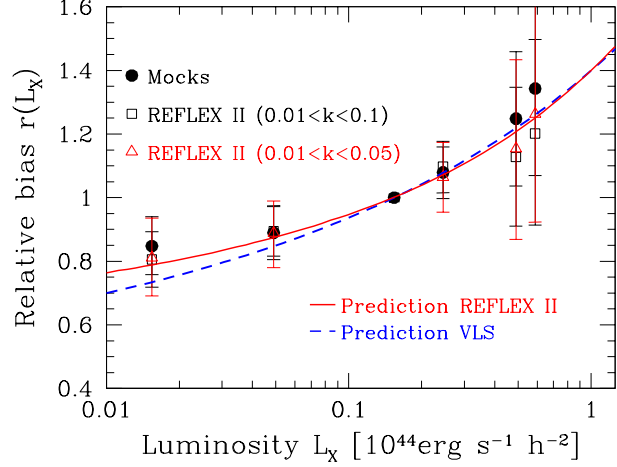


Figure 5.20: Relative luminosity bias measured from the REFLEX II sample (open squares) and the mock catalogues (filled points). The reference luminosity is $\tilde{L} = L_3^{\min}$. The dashed line is the prediction from Equation (5.25) for a VLS, while the dotted line is the prediction from Equation (5.21) use in Equation (5.25).

dependent) factor and thus is simpler to decouple the effects of the window function in the correlation function than in the power spectrum. This is one of the reasons why it might be preferred to develop clustering analysis in configuration space instead of working in Fourier space. However, the time required to measure a correlation functions scales quadratically with the number of objects in the real sample and the synthetic catalogue, while the power spectrum can be measured relatively fast using the FFTw algorithms.

We implemented the correlation function estimator designed by Landy (2002), where the correlation function in a bin of scales centered at r_i is determined from a set of n clusters and m objects in the synthetic catalogue as

$$\hat{\xi}(r_i) = \frac{\frac{2}{n-1}\langle DD \rangle - \frac{2}{m}\langle DR \rangle + \frac{2}{m-1}\langle RR \rangle}{\frac{2}{m-1}\langle RR \rangle}. \quad (5.26)$$

The quantities $\langle DD \rangle$, $\langle RR \rangle$, $\langle RD \rangle$ represent the number of data-data, random-random and data-random pairs respectively separated by a distance r_i in the interval $r_i - \Delta/2, r_i + \Delta/2$. The multiplicative factors account for the number of independent pairs. The bin-averaged covariance matrix of the correlation function is determined from the ensemble of mock catalogues as

$$\hat{C}_{ij} = \frac{1}{N} \sum_{k=1}^N (\hat{\xi}_k(r_i) - \bar{\xi}(r_i))(\hat{\xi}_k(r_j) - \bar{\xi}(r_j)), \quad (5.27)$$

where $\bar{\xi}(r_i)$ is the mean correlation function of the ensemble. In Fig. 5.22 we show the measured correlation function for the six sub-samples defined in Table 4.1. On scales $10 < r/(\text{Mpc } h^{-1}) < 60$, the correlation function is well described by power law $\xi(s) = (s/s_0)^{-\gamma}$. In Table 5.3 we show some the best fit parameters derived from the measurements of the correlation function from the mock catalogues for the six sub-samples (Sánchez et al., in preparation). In agreement with the trend observed in the power spectrum, the correlation length increases with the luminosity cut ranging from $r_0 = 18.7 \text{ Mpc } h^{-1}$ for the L_1^{\min} sub-sample to $r_0 = 28 \text{ Mpc } h^{-1}$ for the highest luminosity cut. The slope of the power-law also increases with the minimum luminosity, accordingly to the fact that most luminous clusters sample larger scales. Figure 5.22 shows again the halo-exclusion effect as a lack of clustering power on small scales in the mock samples compared to the REFLEX signal.

Together with the analysis of the correlation function in redshift space (the monopole contribution), the two-dimensional correlation function $\xi(\sigma, \pi)$ has been also measured. Such

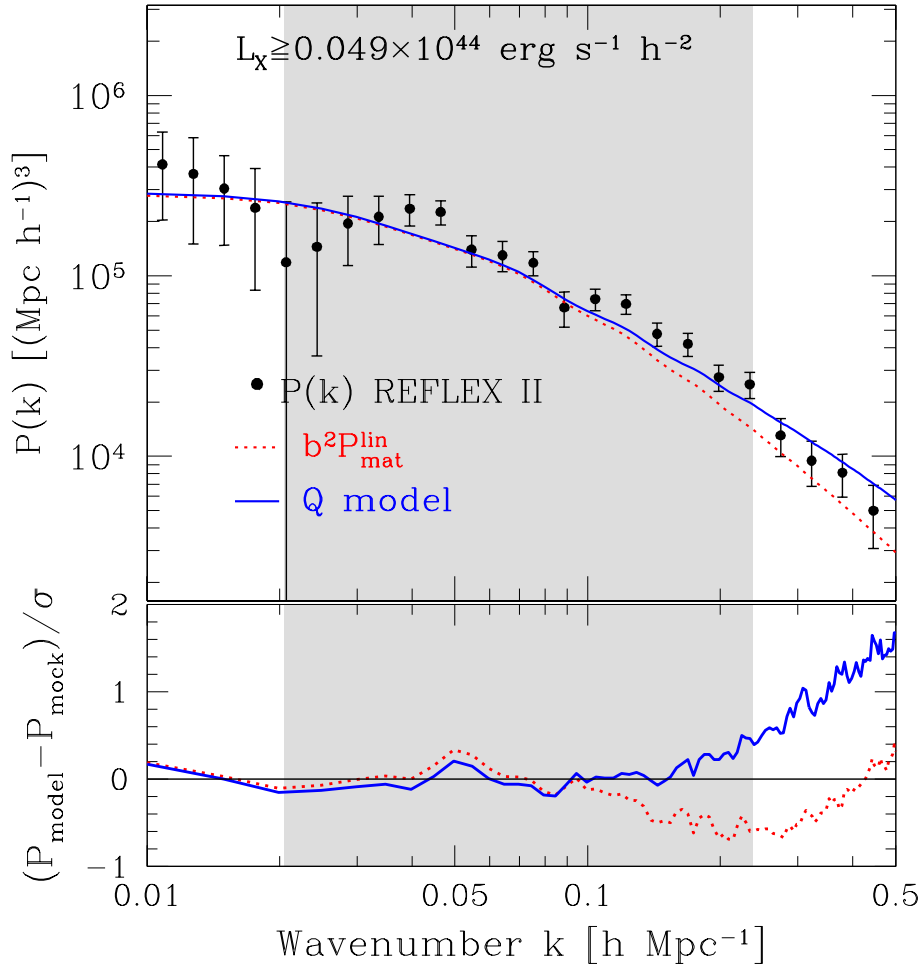


Figure 5.21: Best-fitting Q -model for the REFLEX II mock power spectrum (points with error bars). See Fig. 5.15 for description.

measurement will be implemented to explore redshift distortions (Guzzo et al., in preparation). Following the work of Sánchez et al. (2005), the projected cluster-galaxy cross-correlation function has been also measured using the 2MASS galaxy survey. These measurements will be analyzed in a forthcoming paper (Sánchez et al., in preparation).

5.7 Summary

In this chapter we presented the measurement and analysis of the power spectrum from the new REFLEX II catalogue which is an extension of the original REFLEX sample (Böhringer et al., 2004) to a lower limiting flux ($1.8 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2}$). The new sample contains 911 X-ray detected galaxy clusters of which 860 have measured redshifts in the range $0 \leq z \leq 0.6$ and X-ray luminosities in the range $4.9 \times 10^{40} \leq L_X / (\text{erg s}^{-1} \text{ h}^{-2}) \leq 1.96 \times 10^{45}$. The total flux and X-ray luminosities are estimated using the up-to-date scaling relations based on the REXCESS Survey (Böhringer and et al., 2007; Pratt et al., 2009).

The new sample allowed us to perform a detailed study of the full shape and amplitude of the power spectrum of X-ray detected galaxy clusters. We complemented this analysis by

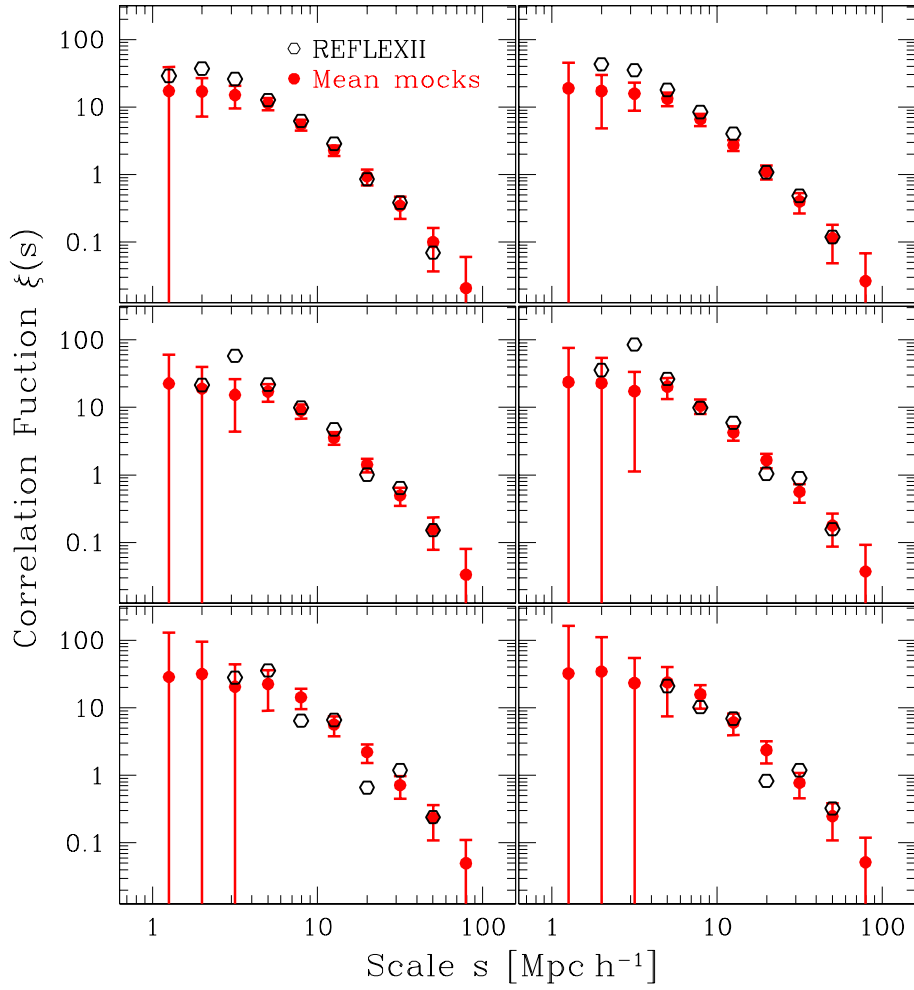


Figure 5.22: Redshift space correlation function of the REFLEX II sample for the six sub-samples defined in Table 4.1 (Sánchez et al., in preparation). Error bars are determined from the variance of the ensemble of mock catalogues.

Sample	Amplitude s_0 [$\text{Mpc}h^{-1}$]	Slope γ
1	18.68 ± 0.10	2.16 ± 0.02
2	20.15 ± 0.16	2.17 ± 0.02
3	22.78 ± 0.12	2.20 ± 0.01
4	24.27 ± 0.33	2.24 ± 0.02
5	27.33 ± 0.29	2.31 ± 0.03
6	28.06 ± 0.13	2.31 ± 0.03

Table 5.3: Fits of the power-law function describing the redshift correlation function of the REFLEX II mock catalogues for the different sub-samples defined in Table 4.1.

using a set of 100 independent mock catalogues constructed to match the selection function of the REFLEX II survey. The clustering properties of these mock catalogues are in good agreement with those measured in the REFLEX II sample. Thus, this ensemble provides a reliable tool to test the statistical methods applied to the data. In particular, we used the mock catalogues to test a model for the luminosity dependence of bias, to construct covariance matrices of the the REFLEX II power spectrum and to analyze the possible systematic effects that might affect this measurement.

Due to the flux-limited selection of the REFLEX II survey, the clustering pattern of galaxy clusters might be affected by scale-dependent distortion, as has been observed in galaxy surveys (e.g. Tegmark et al., 2004; Percival et al., 2007a). Using the mock catalogues, we have shown that these distortions might affect the clustering in configuration space (i.e., when measured with the cluster-correlation function) on scales $r \geq 150 \text{Mpc}h^{-1}$, which would naively correspond to scales $k \leq 0.04h\text{Mpc}^{-1}$ in Fourier space. In order to test the impact of this flux-selection effect on the final measurements of power spectrum, we implemented the luminosity dependent estimator of (Percival et al., 2004b), which is designed to correct for this distortion. We observed that the shape of the power spectrum measured by means of the FKP estimator does not show significant distortions compared to the results from the PVP estimator. This implies that the flux-selection of the REFLEX II sample does not introduce a significant systematic effect in the measurement of the power spectrum of this catalogue.

The shape of the mean power spectrum from our ensemble of mock catalogues is in good agreement with the measured power spectrum from the REFLEX II sample, and is statistically distinguishable from the linear perturbation theory predictions on intermediate scales. This implies a clear signature of non-linear evolution in the X-ray cluster spatial distribution. Nevertheless, given the precision of the measurements of power spectrum in the REFLEX II sample, it is sufficient to model these distortions with moderate accuracy using the Q -model of Cole et al. (2005). We find that this prescription provides a good description of the measurements from the mock catalogues on intermediate scales ($0.02 \leq k/(h\text{Mpc}^{-1}) \leq 0.25$). This model can also be used to describe the shape of the measured REFLEX II power spectrum, providing a valuable tool to extract the cosmological information contained in the shape of this statistic. The next generation of X-ray galaxy clusters surveys, such as *eROSITA*³ and *WFXT*⁴, will provide measurements of the two-point statistics of the cluster population with higher accuracy than present-day samples, for which a more detailed modelling of non-linearities will be required (e.g. Crocce and Scoccimarro, 2006; Montesano et al., 2010).

³<http://www.mpe.mpg.de/heg/www/Projects/EROSITA/main.html>

⁴<http://wfxt.pha.jhu.edu/>

Our measurements of the REFLEX II power spectrum are compatible with the prediction of the Λ CDM cosmological model and shows good agreement with the previous results from the REFLEX sample (Schuecker et al., 2001), saved the expected differences due to the lower limiting flux of the REFLEX II sample. We showed that our measurements cannot provide a statistically significant detection of BAO, which is mainly due to the moderate volume probed by the survey (compared to the volume probed by current galaxy redshift surveys). We found that the power spectra measured from the REFLEX II sample and the mock catalogues were compatible with a scale-independent effective bias in the range of wavenumbers $0.01 \leq k/(h \text{Mpc}^{-1}) \leq 0.1$, and that a simple theoretical prediction, based on the halo-mass bias, the halo mass function and the mass-luminosity relation, was able to describe these measurements. This, together with the modeling of the shape of the power spectrum given by the Q -model, provides a link to the cosmological models and allows our measurements to reach their full constraining power.

Conclusions

Galaxy clusters represent the highest peaks of the underlying matter density fluctuations. Therefore, the characterization of their spatial distribution is a key element towards the determination of the matter distribution in the Universe and in turns, towards an understanding of the cosmological model.

In order to properly interpret the observed clustering properties of galaxy clusters, we first need to understand the large-scale properties of their primordial component, namely, the dark matter haloes. To achieve this, in Chapter 3 we have used a set of low-resolution N -body simulations to determine the halo abundance, the halo power spectrum and the halo bias. A critical issue is the definition of the halo mass, its corresponding halo mass function and how might this be related to astronomical observations concerning galaxy clusters (e.g. luminosity, abundance).

In Chapter 4 we introduced the REFLEX II sample and described the survey-selection function. We measured the REFLEX II luminosity function and discussed some cosmological implications. As an improvement with respect to the REFLEX analysis by Schuecker et al. (2001), a suit of 100 REFLEX II mock catalogues has been provided. The mock catalogues were constructed to reproduce the observed luminosity function by calibrating the underlying mass-X ray luminosity relation. We have shown that the observed REFLEX II luminosity function cannot be reproduced by a mass-X ray luminosity relation represented by simple power law. This can in principle lead to physical interpretations (e.g., deviations from self-similar evolution), once the underlying halo mass function is known for the real clusters. Being that not the case, we did not make any attempt to interpret the resulting mass-X ray luminosity relation. The set of mock catalogues contain relevant information concerning cluster properties such as masses, X-ray luminosities, luminosity errors, redshifts etc, and therefore are an excellent tool, not only for the large-scale structure analysis, but also for the analysis of systematics in determination of cluster masses from the observed luminosities and abundances.

In Chapter 5 we presented the measurement and analysis of the power spectrum of the REFLEX II catalogue. The observed clustering strength is in good agreement with the previous measurements from the REFLEX sample by Schuecker et al. (2001). Given a larger number of clusters in the REFLEX II (due to a lower flux limit), we could perform detailed analysis on the clustering properties of X-ray clusters, such as the luminosity dependence of the bias. We observed that the clustering properties of the mock catalogues are in good

agreement with the clustering measured in the REFLEX II, thus providing reliable estimates of the covariance matrix for the power spectrum. We showed that the REFLEX II sample cannot provide a statistically significant detection of BAO, mainly due to the moderate volume probed by the survey. We observed the expected increase in the power spectrum amplitude with increasing minimum luminosity. These measurements were in good agreement with theoretical predictions, once key parameters as the dispersion in the mass-luminosity relation and luminosity errors were provided.

Using the mock catalogues we determined the relative luminosity bias and observed how the clustering signal of the REFLEX II is likely to be consistent with an scaled version of the matter power spectrum on large scales $0.05 \leq k[h/\text{Mpc}] \leq 0.15$.

We observed that our measurements of power spectrum lack of significant distortions induced by flux-selection effects. This is mainly due to the range of X-ray luminosities covered by the REFLEX II sample, for which the resulting luminosity bias changes smoothly with the luminosity. Although the quantitative conclusions we have drawn here are only valid to the REFLEX sample, it is important to note that the machinery used to explore such effects can be applied to future flux-limited galaxy cluster sample.

The new REFLEX II catalogue allowed us to detect for the first time a signature of non-linear evolution in the galaxy-cluster population. We modeled the amplitude and the shape of the cluster power spectrum using and tested these models with our measurements of power from the mock catalogues. Given the fact that our mocks follow the same REFLEX II selection function and that the observed clustering is compatible with that of the REFLEX II sample, we successfully applied the model to reproduce the measurements from the REFLEX II sample.

The REFLEX II is to date the largest X-ray galaxy cluster catalogue. The measurement of the cluster power spectrum and the precision achieved in this work represents the state-of-the-art concerning the large scale structure of the universe as traced with X-ray clusters. The measurements of the cluster power spectrum (and correlation function) are currently being used to extract information concerning the parameters of the concordance cosmological model, mainly, the amount of dark matter Ω_m , the dark energy equation of state w , the spectral index of primordial fluctuations and the Hubble parameter. The results will be shown and discussed in a forthcoming paper. These constraints will be determined using also the information from the power spectrum of temperature fluctuations from *WMAP*. Furthermore, the analysis will comprise the information contained in the measured X-ray luminosity function. As an example, simple constraints derived with the distance priors from *WMAP* (e.g. Komatsu et al., 2010) and only using the information of the measured shape of the power spectrum shows that the parameter degeneracy observed in *WMAP* is hardly broken with the REFLEX II data (e.g. Sánchez et al., 2009), yielding constraints in the dark energy equation of state of $w = -0.86^{+0.50}_{-0.66}$ and the content of dark matter of $\Omega_m = 0.34^{+0.15}_{-0.18}$ which approximately corresponds to the same precision and exactitude as the CMB constraints from *WMAP*. Being the model for the luminosity bias that we have described in Chapter 5 in agreement with the measurements, smaller error bars and more exact values can be achieved by incorporating the information of the amplitude to our analysis.

In the near future it is planned to extend the galaxy cluster catalogue by the characterization of the RASS observations in the northern hemisphere (NORAS). At its best, the combination REFLEX+NORAS will help to increase the surveyed volume in a factor of 2. Together with the increase in the number of clusters, it will provide more precise measurements of cluster power spectrum. Being galaxy clusters rare objects, the volume they probe

will always be overcome by that of galaxy redshift surveys. Despite this fact, it would be interesting to characterize the clustering properties of a larger sample, for which some forecasts are currently being determined.

The legacy of the ROSAT mission, embodied in our context in the REFLEX catalogue, has left after two decades, valuable information concerning the large scale structure of the Universe. The completion of an all sky catalogue (REFLEX+NORAS) and the subsequent analysis of its cosmological implications will be the final product, placing us at the portals of a new era in X-ray astronomy. This new age is embodied in the *eROSITA* mission, to be launched in 2012. The volume probed by this mission will overcome the volume surveyed by the *ROSAT* mission and will certainly provide not only results on large scale structure comparable in precision with those derived from galaxy surveys, but also a better understanding of the intrinsic properties of galaxy clusters.

Spherical collapse model

The spherical collapse model describes the non-linear evolution of matter density perturbations. Let us consider a spherical region of radius r filled with dark matter, with an overdensity $\delta(r)$ with respect to the underlying dark matter distribution. With time, this region can amass enough mass such that it reaches a critical density contrast and collapse to form bound structures. In general, the collapse of large scale structures can be divided in three main stages:

- Expansion of the region contained the matter with the Hubble flow.
- The turn-around, when the system acquires a critical density, decouples from the Hubble flow and starts to collapse
- Virialization, corresponding to the final stage where the system reaches equilibrium via violent relaxation processes. From this point, the properties of the collapsed object will evolve independently of the background (the so called "stable clustering hypothesis").

We will consider this spherical region as composed of thin shells, each with radius r and containing a mass \mathcal{M} . If the scale of the perturbation is smaller than the Hubble horizon, we can use the Newtonian limit, where the equation of motion for the spherical shell of radius r can be written as:

$$\frac{\ddot{r}}{r} = \frac{\ddot{a}}{a} - \frac{\delta M}{r^3}. \quad (\text{A.1})$$

where δM corresponds to the amount of dark matter mass due to the excess of density in the cloud with respect of the background:

$$\delta M(r) = \int_0^r 4\pi s^2 \delta(s) ds.$$

We now consider the acceleration equation, which arises from the Einstein-Hilbert field equations with the FRW line element (e.g. Carroll, 2003). Taking into account all forms of matter-energy encoded in the acceleration equation:¹

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi [\rho_{\text{cdm}}(t) + \rho_x(t)\eta_x(t)], \quad \eta_x(t) \equiv 1 + 3\frac{p_x(t)}{\rho_x(t)}.$$

¹Note: one must recall the hidden c - factors: the r.h.s. term in the acceleration equation is written as $\sim \rho + c^{-2}p$, and the equation of state is $p = c^2\omega\rho$, and hence the c -factors cancel out!

Let us consider the simplest case in which we only take into account cold dark matter with $\rho_b \rightarrow \rho_{\text{cdm}}$. We then have

$$\ddot{r} = -\frac{\delta M}{r^2} - \frac{4}{3}\pi\rho_b(t)r = -\frac{\delta M}{r^2} - \frac{M_b}{r^2} \equiv -\frac{\mathcal{M}(r,t)}{r^2}. \quad (\text{A.2})$$

where $\mathcal{M}(r,t) = \delta M(r) + M_b(t)$ is the contribution from the background M_b and the collapsed matter in the same radius. With $M_b(t,r) = \frac{4}{3}\pi r^3 \rho_b(t)$ and $\rho_b(t) \sim a(t)^{-3}$. We then write the total mass as $\mathcal{M}(r,t) = M_b(t)(1 + \bar{\delta}(r,t))$, where $\bar{\delta}$ denotes the volume average of the overdensity δ . One next assumes that different shells labeled at some initial time $t = t_i$ as 1, 2, 3, ... with radius $r_1 < r_2 < r_3 < \dots$ does not cross each other during the expansion or the collapse. This implies that the mass \mathcal{M} contained within each shell of radius r is constant:

$$\mathcal{M}(r_i, t_i) = \mathcal{M} = \frac{4}{3}\pi r_i^3 \rho_b(t_i)(1 + \bar{\delta}_i) = \text{constant}, \quad (\text{A.3})$$

with $\bar{\delta}_i = \bar{\delta}(r_i, t_i)$. Once the systems decouples from the Hubble expansion, it is assumed that the dynamics of the falling material can be described with the equations of motion that describe dynamics motion for a test object under the gravitational influence of a central object with mass \mathcal{M} . One can therefore determine analytical solutions and derive first integral of motion. For instance, given some initial value t_i and by using Hubble's law (neglecting initial peculiar motions), the first integral of motion, (which for a test particle is its energy) we can written as

$$\mathcal{E} = -\frac{1}{2}H_i^2 \dot{r}_i^2 [\Omega_i(1 + \bar{\delta}_i) - 1], \quad (\text{A.4})$$

where $\Omega_i = \Omega_b(t_i)$ is the density parameter of the background at the initial time t_i . The condition to be fulfilled in order to allow collapse is then $\Omega_i(1 + \bar{\delta}_i) > 1$. In other words, collapse may be allowed as long as the total density is greater than the critical density at the time t_i . Equation (A.4) allows us to write the relevant parameters of this model in terms of initial conditions. At turn-around we have $\dot{r} = 0$ and the system acquires its maximum radius r_{max} :

$$r_{\text{max}} = \left[\frac{\Omega_i(1 + \bar{\delta}_i)}{\Omega_i(1 + \bar{\delta}_i) - 1} \right] r_i. \quad (\text{A.5})$$

The solution for the equation of motion $\ddot{r} = -\mathcal{M}/r^2$ can be written in through the parametric angle θ as

$$r = \tilde{A}(1 - \cos \theta), \quad t + T = \tilde{B}(\theta - \sin \theta), \quad (\text{A.6})$$

where T is some initial time which we can neglect. The coefficients \tilde{A} and \tilde{B} must satisfy $\tilde{A}^3 = \mathcal{M}\tilde{B}^2$ in order to satisfy the equation of motion. In this solution we see that the time grows while the radius reaches a maximum values and decreases. The evolution of the collapsed region can be described in terms of the parametric angle θ as follows

Expansion	$\theta = 0$	$r = 0$	$t = 0$
Turn-around	$\theta = \pi$	$r = r_{\text{max}} = 2\tilde{A}$	$t = t_{\text{ta}} = \pi\tilde{B} - T$
Collapsed	$\theta = 2\pi$	$r = 0$	$t = t_{\text{coll}} = 2\pi\tilde{B} - T$

Assuming a flat- matter dominated universe such that $a(t) \propto t^{2/3}$, the mean density of the collapsed configuration can be written on terms of the parameter θ as (Padmanabhan, 1993b)

$$\bar{\delta}(\theta) + 1 = \frac{9(\theta - \sin \theta)^2}{2(1 - \cos \theta)^3}. \quad (\text{A.7})$$

Note that at the virialization time $t_{\text{vir}} = 2\pi$ we get $\delta \rightarrow \infty$. Nevertheless, collapsing structures reached a finite size, which can be determined assuming that the final configuration is in virial equilibrium. It is interesting to compare the evolution in the nonlinear regime with the linear approximation. In the linear regime, the density contrast can be written as

$$\bar{\delta}_L(t) = \frac{3}{5} \left(\frac{3}{4} \right)^{2/3} (\theta - \sin \theta)^{-2/3}.$$

As the over-density evolves, one has the following behaviors:

- $\theta = \pi/2$: $\bar{\delta}_L = 0.341$ and $\bar{\delta} = 0.466$. Both density contrast lie in the linear regime. Nevertheless, an overdensity evolving in the non-linear regime is almost 40% denser than the corresponding linear overdensity.
- $\theta = 2\pi/3$: $\bar{\delta}_L = 0.568$ and $\bar{\delta} = 1.01$. This value of θ then marks the transition between the linear and the non-linear behavior.
- $\theta = \pi$: $\bar{\delta}_L = 1.06$ and $\bar{\delta} = 4.55$. The over-density acquires its maximum radius at $r_{\text{max}} = 2\tilde{A}$ with $\bar{\rho}(t_{\text{max}}) \approx 5.6\rho_b(t_{\text{max}})$. This happens at a redshift given by $(1 + z_{\text{max}}) = 0.56(1 + z_i)\bar{\delta}_i$. We can relate the time for maximum radius and the time for onset of non-linearity as $(1 + z_{\text{max}}) = 0.53(1 + z_{\text{nl}})$,
- $\theta = 3\pi/2$. This is the second time the radius acquires half of the maximum radius (the first time was at $\theta = \pi/2$ during the expansion). This corresponds to the radius at virial equilibrium as we will see below. Linear theory predicts $\bar{\delta}_L = 1.48$, while the non-linear model predicts $\bar{\delta} = 146.8$.
- $\theta = 2\pi$: This is the time needed for the cloud to collapse to a singularity at $r = 0$. The linear regime yields $\bar{\delta}_L = 1.686$. The non-linear evolution gives $\bar{\delta} \rightarrow \infty$, but we can derive a finite number from the results derived above using the fact that the *collapse* radius is finite. We then assume that the time needed to reach virial equilibrium is given by $t_{\text{vir}} = t_{\text{col}} = 2\pi\tilde{B}$.

For an spherical cloud with total mass M and radius r_{vir} , the virial theorem reads

$$2\mathcal{K}_{\text{vir}} = -U_{\text{grav}}(r) = \frac{3M^2}{5r_{\text{vir}}},$$

where \mathcal{K} is the kinetic energy and U_{grav} is the gravitational potential energy. Energy conservation at the maximum radius and at the virial radius implies finally $r_{\text{max}} = 2r_{\text{vir}}$, such that

$$\delta_{\text{vir}} + 1 = \frac{\bar{\rho}(t_{\text{vir}})}{\rho_b(t_{\text{vir}})} = 8 \frac{\bar{\rho}(t_{\text{max}})}{\rho_b(t_{\text{vir}})} = 8(5.6) \frac{\rho_b(t_{\text{max}})}{\rho_b(t_{\text{vir}})} = 44.8 \frac{(1 + z_{\text{max}})^3}{(1 + z_{\text{vir}})^3}.$$

and we obtain the density contrast at virialization to be $\bar{\delta}_{\text{vir}} = 18\pi^2 \approx 177.65$. Virialization takes place at a redshift given by $(1 + z_{\text{vir}}) = 0.63(1 + z_{\text{max}})$ and the mean density of the a virialized structure can be written as

$$\rho_{\text{vir}} = 9.39t_{\text{vir}}^{-2} = 21.24H_0^2(1 + z_{\text{vir}})^3 \approx 1.6 \times 10^{-33}(1 + z_{\text{vir}})^3 \text{ gr cm}^{-3}.$$

This implies that collapsing clouds in early times are denser than collapsing clouds forming at the present time. In other words, dark matter halos around individual galaxies are formed

first than the dark matter haloes of galaxy clusters. The radius of the virialized configurations can be written as

$$r_{\text{vir}} = 215(1 + z_{\text{vir}})^{-1} \left(\frac{M}{10^{12} M_{\odot} h^{-1}} \right)^{1/3} \text{ kpc } h^{-1} \quad (\text{A.8})$$

A small (galactic) dark matter halo with $M \sim 3 \times 10^{11} M_{\odot}$ and $R \approx 50$ kpc then is expected, in this simple approximation, to have formed at $z \approx 2.6$.

Scaling realtions: mass-temperature

After the time of decoupling, the free electrons were able to recombine and form hydrogen, helium, and (in a small fraction) heavier elements. The evolution of matter and baryon overdensities then show that baryons fall in the deep dark matter potential wells. As long as the baryonic component falls, the total kinetic energy becomes a contribution of random and ordered motions (associated to the dispersion or pressure tensor). Hydrogen and helium are shock-heated and ionized, emitting radiation in the X-ray band. In the hierarchical structure formation scenario, these objects will merge and form larger structures composed with hot gas, galaxies and the dark matter halo. In such scenario, the assumption of hydrostatic equilibrium allows us to relate the thermal energy associated to the gas with the gravitational potential energy represented by the motion of galaxies in the cluster:

$$\mathcal{K} = \frac{1}{2} M \langle v^2 \rangle = \frac{3}{2} \beta \left(\frac{M}{\mu m_p} \right) k_B T_{\text{gas}},$$

where β is the ratio of the kinetic energy to the temperature and μ is the mean molecular weight of the gas, arising from the contribution of free electrons μ_e and ions μ_I as $\mu = \mu_e \mu_I / (\mu_I + \mu_e)$ and m_p is the mass of the proton. Assuming that the gas is fully ionized and ignoring the contribution from heavy elements, $\mu = 4/(3+5X)$ where X is the relative abundance of hydrogen: for $X = 0.75$, $\mu = 0.59$. Under the assumption of virial equilibrium, the kinetic energy of the system (i.e., of the gas for dark matter is cold and pressureless) equals twice the potential energy of the system, such that we can relate the temperature of the gas with the gravitational potential $T \sim \Phi$. Taking into account the effects of a background (see Equation A.1) composed of dark matter and dark energy represented by a cosmological constant, the virial theorem can be written as (e.g. Wang, 2000; Balaguera-Antolinez et al., 2007)

$$\frac{3}{2} \frac{M}{\mu} k_B T_{\text{gas}} = \frac{1}{2} U_{\text{grav}} - \frac{2}{3} \pi G_N \mathcal{I} \Omega_{\Lambda} \rho_c(z) \quad (\text{A.9})$$

where \mathcal{I} is the moment of inertia. This expressions can be evaluated once the density profile of the gas and dark matter are known. Neglecting the contribution from the hot-gas and assuming an homogeneous dark matter halo, we can use Equation (A.8) for the radius at virial equilibrium and solve for the temperature $T \propto M^{2/3}$. Neglecting any contribution from the cosmological constant, this becomes

$$T_{\text{gas}} = 2.82 \times 10^5 (1 + z_{\text{vir}}) \left(\frac{M}{10^{12} M_{\odot} h^{-1}} \right)^{2/3} \text{ K}. \quad (\text{A.10})$$

This expression is only valid for a matter dominated universe and for spherical and homogeneous configurations. Though incorrect, Nevertheless it provides a good estimate for the temperature of intra-cluster medium in galaxy clusters. For a dark matter halo with $M = 10^{14} M_{\odot} h^{-1}$ collapsing at $z_{\text{vir}} \approx 1$ we obtain $T_{\text{gas}} \approx 10^7$ K.

X-ray luminosity

The intra-cluster medium in galaxy cluster has been identified with a hot-diffuse gas with temperatures of the order of 10^7 K. The main contribution to the total radiated energy from this gas comes in form of bremsstrahlung radiation, which is a free-free transition taking place in a fully ionized gas. For lower temperatures, other radiation mechanism such like recombination (free-bound) and two-photon decay of metastable energy levels are also contributing to the total radiated power.

The emissivity $\epsilon(\nu)$ is defined as the amount of energy radiated per unit time, per volume per frequency:

$$\epsilon_\nu \equiv \frac{dL}{dVd\nu}. \quad (\text{A.11})$$

The radiation emitted after a free-free collision a hot plasma of temperature T_{gas} is characterized by an emissivity (e.g. Sarazin, 1988)

$$\epsilon_{ff}(\nu, T_{\text{gas}}) = \frac{32\pi e^6}{3m_e c^3} \left(\frac{2\pi}{3m_e k_B} \right)^{1/2} Z^2 n_i n_e g_{ff}(Z, T_{\text{gas}}, \nu) T_{\text{gas}}^{-1/2} e^{-h\nu/k_B T_{\text{gas}}}, \quad (\text{A.12})$$

where m_e is the electron mass, Z is the ion charge, n_i and n_e are the ion and electron number density respectively. g_{ff} is the so-called Gaunt factor, accounting for quantum-mechanical corrections. For the ROSAT energy band $h\nu \approx 2\text{keV} \approx k_B T_{\text{gas}}$, this correction is of order $g_{ff} \approx 1$. If we assume that the gas is fully ionized and mainly composed of Hydrogen and Helium, then the mass density of the gas can be written as $\rho_{\text{gas}} = \mu(1+f)m_p n_e$ where f is the ratio of the number density of ions to electrons in the gas. Using again $X = 0.75$, this fraction is $f = 0.8$ and $\rho_{\text{gas}} = 1.06n_e m_p$. Therefore we can write the emissivity as a function of ρ_{gas}^2 ,

$$\epsilon_\nu(\nu, T_{\text{gas}}) \equiv \rho_{\text{gas}}^2 \Lambda(T_{\text{gas}}, \nu),$$

where the function Λ can be defined from Equation (A.12).

The X-ray luminosity emitted at a frequency ν from a hot-gas with temperature T_{gas} in a volume V is written as $L_\nu = \int_V dV \epsilon_{ff}(\nu, T)$. From this, luminosity emitted in a energy range characterized by the frequencies ν_1 and ν_2 is

$$L_{[\nu_1, \nu_2]} = \int_V dV \rho_{\text{gas}}^2(r) \int_{\nu_1}^{\nu_2} d\nu \Lambda(\nu, T_{\text{gas}}).$$

Assuming a constant gas temperature (isothermal gas), we obtain

$$L_{[\nu_1, \nu_2]} = \int dV n_e(r)^2 \int_0^\infty d\nu \Lambda(\nu, T_{\text{gas}}) = AF(T_{\text{gas}}, \nu_1, \nu_2) \langle \rho_{\text{gas}}^2 \rangle V, \quad (\text{A.13})$$

where A is a constant that can be read from Equation (A.12) and $\langle \rangle$ denotes volume average (for the β -model described in Chapter 4, using $\beta = 2/3$ and $r_{\text{vir}}/r_0 = 7$ we have $\langle \rho_{\text{gas}}^2 \rangle = 0.123\rho_0^2$). The function $F(T, \nu_1, \nu_2) \equiv \int_{\nu_1}^{\nu_2} d\nu \Lambda(T, \nu)$ is characterized by the energy interval and the temperature of the gas. For bremsstrahlung emission, integrating over all frequencies we have $F(T, \nu_1, \nu_2) = F(T) = T^{1/2}$ such that the bolometric luminosity reads $L_X = AT_{\text{gas}}^{1/2} V \langle \rho_{\text{gas}}^2 \rangle$. Assuming spherical symmetry and $\langle \rho_{\text{gas}}^2 \rangle \approx \langle \rho_{\text{gas}} \rangle^2$, we have

$$L_{X \text{ bol}} \propto T_{\text{gas}}^{1/2} r_{\text{vir}}^3 \langle \rho_{\text{gas}}^2 \rangle \sim T_{\text{gas}}^{1/2} M (f_{\text{gas}}^2 M^2 r_{\text{vir}}^{-3}) \propto f_{\text{gas}}^2 T^{1/2} M \propto f_{\text{gas}}^2 M^{4/3},$$

where $f_{\text{gas}} = M_{\text{gas}}/M$ and M is the total mass of the cluster. In the last step we have used the virial condition $T_{\text{gas}} \propto M^{2/3}$. Therefore, if the fraction of baryons scales with the total mass as $f_{\text{gas}} \propto M^\alpha$, the X-ray luminosity should scale with the cluster mass as $L_{X \text{ bol}} \propto M^{\frac{4}{3}+2\alpha}$.

When observed in a energy band, the mass-luminosity relation is expected to scale with a different slope. For the ROSAT energy band $[0.1, 2.4]\text{keV}$, integration of the emissivity given by Equation (A.12) yields $L_{X[0.1-2.4]\text{keV}} \propto M^{1+2\alpha}$.

The halo model for gravitational clustering

IN this appendix we show some explicit expressions for the two-point statistics for galaxies in the context of the halo model. Let us consider two type of galaxies accordingly to their location within the dark matter haloes, namely, central and satellite galaxies. Also, let us assume that the galaxy population can be split in two types, namely, red and blue galaxies. We refer to this properties with the letters A and B. Note that these properties do not imply a preferred location of the galaxies within the dark matter haloes, i.e, in principle both types can have a representative galaxy being central or satellite. This yields the following scenario for the population of dark matter haloes i) haloes with one galaxy type A, ii) haloes with one galaxy type B, iii) haloes with more than one galaxy and a central type A, iv) haloes with more than one galaxy and a central type B, vi) haloes without central galaxies. By assumption, when only one galaxy is found in a dark matter halo, it is assumed that it is located in the center of mass of the halo. We also assume that if a dark matter halo have one or more galaxies, it has a central galaxy.

Correlation function

Let us consider an intrinsic property of galaxies which can be related with the mass of the hosting dark matter halo (i.e, luminosity, color, velocity dispersion). Assume that the distributions of galaxies can be described in a six-dimensional phase-space density characterized by a distribution function $\mathcal{F}(\mathbf{x}, \mathbf{v}, z; L|M)$, such that the number density of galaxies in a volume element d^3x centered at \mathbf{x} at redshift z with a luminosities in the interval $L, L+dL$ hosted by dark matter halos with masses in the mass range $M, M+dM$ can be determined by marginalizing with respect to the galaxy velocities:

$$n_{\text{gal}}(\mathbf{x}, z|L)d^3x dL dM = \int d^3v \mathcal{F}(\mathbf{x}, \mathbf{v}, z; L|M) d^3x dL dM. \quad (\text{B.1})$$

We now assume that the total distribution function can be written as a linear combination of the distribution function associated to each dark matter halo:

$$\mathcal{F}(\mathbf{x}, \mathbf{v}, z; L|M)dM = \sum_i \delta(M - M_i) \mathcal{F}_i(\mathbf{x}, \mathbf{v}, z|L, M_i) dM,$$

where the sum is done over the halos. Furthermore, let us assume that in each halo the distribution function is separable and can written as a product of a normalized velocity

distribution $\mathcal{P}(\mathbf{v}|M_i, z)$ ($\int d^3v \mathcal{P}(\mathbf{v}|M_i, z) = 1$) and a generalized density profile. If the central galaxies are in the center of mass of the dark matter halo, then the density profile they follow is just $\delta^3(\mathbf{x} - \mathbf{x}_i)$ and their velocity distribution is $\mathcal{P}_c(\mathbf{v} - \mathbf{v}_i|M_i, z) = \delta^3(\mathbf{v} - \mathbf{v}_i)$ where \mathbf{v}_i is the velocity of the center of mass of the dark matter halo. On the other hand, satellites galaxies follow a density profile $u_{\text{gal}}(\mathbf{x} - \mathbf{x}_i|M, z)$. Each density profile is weighted by the conditional probability for a galaxy to have luminosity L in the range $L, L + dL$ given the fact that it is hosted by a halo of mass M . This corresponds to the *conditional luminosity function* (CLF) $\Phi(L|M, z)dL$, which is also split for centrals and satellites (or early and late type galaxies). In symbols,

$$\mathcal{F}(\mathbf{x}, \mathbf{v}, z; L, M)dLdM = \sum_i \delta(M - M_i) \left[\Phi_c(L|M_i, z)\delta^3(\mathbf{v} - \mathbf{v}_i)\delta^3(\mathbf{x} - \mathbf{x}_i) \right. \\ \left. + \Phi_s(L|M_i, z)\mathcal{P}_s(\mathbf{v} - \mathbf{v}_i|M_i, z)u_{\text{gal}}(\mathbf{x} - \mathbf{x}_i|M_i, z) \right] dL dM.$$

The total velocity distribution function of type A galaxies (centrals and satellites) is obtained by marginalizing the distribution function over the spatial coordinates. Accordingly, the total mean density of galaxies with a luminosity L can then be written as

$$n_{\text{gal}}(\mathbf{x}, z; L) = n_{\text{gal}}^A(\mathbf{x}, z; L) + n_{\text{gal}}^B(\mathbf{x}, z; L) \\ = \sum_i \left[\Phi_c^A(L|M_i, z)\delta^3(\mathbf{x} - \mathbf{x}_i) + \Phi_s^A(L|M_i, z)u_{\text{gal}}^A(\mathbf{x} - \mathbf{x}_i; M_i, z) \right] \\ + \sum_i \left[\Phi_c^B(L|M_i, z)\delta^3(\mathbf{x} - \mathbf{x}_i) + \Phi_s^B(L|M_i, z)u_{\text{gal}}^B(\mathbf{x} - \mathbf{x}_i; M_i, z) \right],$$

where the sum is done over all dark matter haloes with galaxies. Using Equation (3.21), we can write the galaxy number density in terms of the number density of dark matter haloes at a position \mathbf{x} with mass M , $N(\mathbf{x}; M)$, as

$$n_{\text{gal}}(\mathbf{x}, z; L) = \int d^3x' dM \left[\left(\Phi_c^A(L|M, z) + \Phi_c^B(L|M, z) \right) \delta^3(\mathbf{x}' - \mathbf{x}) \right. \\ \left. + \Phi_s^A(L|M, z)u_{\text{gal}}^A(\mathbf{x}' - \mathbf{x}; M, z) + \Phi_s^B(L|M, z)u_{\text{gal}}^B(\mathbf{x}' - \mathbf{x}; M, z) \right] N(\mathbf{x}'; M), \quad (\text{B.2})$$

such that applying Equation (3.22) we can determine the galaxy correlation function (setting $\xi_{\text{gal}}^{1h} \equiv \xi_{\text{gal}}^{1h}(\mathbf{x} - \mathbf{y}; L, L')$)

$$1 + \xi_{\text{gal}} = \frac{\langle n_{\text{gal}}(\mathbf{x}, z; L')n_{\text{gal}}(\mathbf{y}, z; L) \rangle}{\bar{n}_{\text{gal}}(z; L)\bar{n}_{\text{gal}}(z; L')} = (1 + \xi_{\text{gal}}^{1h}) + (1 + \xi_{\text{gal}}^{2h}), \quad (\text{B.3})$$

where $\bar{n}_{\text{gal}}(z; L) = \Phi(L, z)$ is the luminosity function. It can be shown that the 1-halo contribution to the correlation function is

$$1 + \xi_{\text{gal}}^{1h} = \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) \delta^3(\mathbf{x} - \mathbf{y}) \\ + \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_s^A(L'|M) + \Phi_s^B(L'|M) \right) u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M) \\ + \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M) \\ + \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \left(\Phi_s^A(L'|M) + \Phi_s^B(L'|M) \right) \\ \times \int d^3x' u_{\text{gal}}(\mathbf{x}' - \mathbf{x}; M) u_{\text{gal}}(\mathbf{x}' - \mathbf{y}; M), \quad (\text{B.4})$$

where we have assumed that type A and type-B galaxies follow the same density profile $u_A(\mathbf{x} - \mathbf{y}; M) = u_B(\mathbf{x} - \mathbf{y}; M) = u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M)$ and spherical symmetry via $u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M) = u_{\text{gal}}(\mathbf{y} - \mathbf{x}; M)$.

The first term corresponds to the self-correlation associated to the central galaxy. The second and the third terms represent the clustering associated to pairs central-satellite of the two types within the same halo; here the scale dependence of the correlation function is in this case ruled directly by the shape of the density profile followed by galaxies B . The last term in Equation (B.4) is the associated to the pairs formed by satellite galaxies of the two types within the same halo; this involves the convolution of the density profile of followed by the two types of galaxies.

On the other hand, on large scales the contributions from the 2-halo term reads as (the third term of Equation (3.22) yields a 1 that cancels the 1 preceding the ξ_{gal}^{2h} in Equation (B.3))

$$\begin{aligned}
\xi_{\text{gal}}^{2h} &= \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) \xi_{hh}(\mathbf{x} - \mathbf{y}; M, M') \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_s^A(L'|M') + \Phi_s^B(L'|M') \right) \\
&\times \int d^3 y' \xi_{hh}(\mathbf{x} - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M') \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L'|M') + \Phi_c^B(L'|M') \right) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \\
&\times \int d^3 y' \xi_{hh}(\mathbf{x} - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M) \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \left(\Phi_s^A(L'|M') + \Phi_s^B(L'|M') \right) \\
&\times \int d^3 x' d^3 y' \xi_{hh}(\mathbf{x}' - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M) u_{\text{gal}}(\mathbf{y}' - \mathbf{y}'; M')
\end{aligned} \tag{B.5}$$

The first term of this expression corresponds to the clustering between central galaxies of the two types in different haloes. The second and the third and third encodes the clustering between central and satellite galaxies of the two types, while the fourth term gives the clustering between satellites of the two types in different haloes. Equations (B.4) and (B.5) thus provide the halo-model description of the galaxy clustering in terms of a galaxy property, the luminosity, and the conditional luminosity function. From these expressions we can then derive the correlation function between central or satellite galaxies of different types (colour) and luminosities.

Power spectrum

The determination of the galaxy power spectrum is strait-forward, once the correlation function has been written. We only need to use properly the Parseval's theorem. The 1-h contribution to the galaxy power spectrum can be written as

$$\begin{aligned}
(2\pi)^3 \delta^3(\mathbf{k}) + P_{\text{gal}}^{1h}(k) &= \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) \tag{B.6} \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_s^A(L'|M) + \Phi_s^B(L'|M) \right) u_{\text{gal}}(k; M) \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) u_{\text{gal}}(k; M) \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM n(M) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \left(\Phi_s^A(L'|M) + \Phi_s^B(L'|M) \right) |u_{\text{gal}}(k; M)|^2.
\end{aligned}$$

The (first) constant term appearing on the right-hand side of Equation (B.6) arises from the auto-correlation, and therefore is not taken into account for practical purposes. Similarly, the term appearing on the left-hand side is only non-negligible at $k = 0$, which is in practice very difficult to measure. The 2-h contribution to the power spectrum is given by

$$\begin{aligned}
P^{2h}(k) &= \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_c^A(L'|M) + \Phi_c^B(L'|M) \right) P_{hh}(k; M, M') \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L|M) + \Phi_c^B(L|M) \right) \left(\Phi_s^A(L'|M') + \Phi_s^B(L'|M') \right) \\
&\times P_{hh}(k; M, M') u_{\text{gal}}(k; M') \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_c^A(L'|M') + \Phi_c^B(L'|M') \right) \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \\
&\times P_{hh}(k; M, M') u_{\text{gal}}(k; M) \\
&+ \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) \left(\Phi_s^A(L'|M') + \Phi_s^B(L'|M') \right) \\
&\times P_{hh}(k; M, M') u_{\text{gal}}(k; M) u_{\text{gal}}(k; M')
\end{aligned}$$

It is usual to assume that on the typical scales associated to the separation between dark matter haloes, the halo-halo power spectrum can be decomposed as $P_{hh}(k; M, M') = b(M)b(M')P_{\text{mat}}(k)$ where $b(M)$ is the halo-mass bias and $P_{\text{mat}}(k)$ is the matter power spectrum. Under this assumption, the 2-h contribution to the galaxy power spectrum is simplified to (recovering the notation in full glory)

$$P^{2h}(k, z; L, L') = b_{\text{eff}}(k, z; L) b_{\text{eff}}(k, z; L') P_{\text{mat}}(k, z) \quad (\text{B.7})$$

where the effective bias is given by

$$b_{\text{eff}}(k, z; L) \equiv \frac{1}{\bar{n}_{\text{gal}}(L)\bar{n}_{\text{gal}}(L')} \int dM dM' n(M)n(M') b(M)b(M') \Phi_{\text{eff}}(L|M, k), \quad (\text{B.8})$$

where we defined

$$\Phi_{\text{eff}}(L|M, k) \equiv \Phi_c^A(L|M) + \Phi_c^B(L|M) + \left(\Phi_s^A(L|M) + \Phi_s^B(L|M) \right) u_{\text{gal}}(k; M). \quad (\text{B.9})$$

On large scales, the Fourier transform of the density profile goes asymptotically to 1 leading to a scale-independent bias in Fourier space.

Conditional luminosity function

The conditional probability distribution $\Phi_i(L|M, z)$ determines the probability of assigning a luminosity L to a galaxy that resides in a dark matter halo with mass M . The CLF is a key link between astrophysical and cosmological scenarios, since it can depend on the local physics in dark matter haloes or clusters and can also be used to construct the luminosity function of galaxies $\Phi(L)$ which of course contains enormous information on cosmological parameters through the underlying halo mass function $n(M)$.

At a given redshift, the joint probability for a galaxy to have a luminosity L and being hosted by a dark matter halo of mass M can be written as $P(L, M) = \Phi(L|M)n(M) = \Phi(M|L)\Phi(L)$ from which we obtain

$$\Phi(M|L, z) = \frac{\Phi(L|M, z)n(M, z)}{\Phi(L, z)}, \quad (\text{B.10})$$

such that the luminosity function of galaxies can be written as a marginalization over the masses of dark matter haloes:

$$\Phi(L, z) = \int dM \Phi(L|M, z) n(M, z).$$

Equation (B.10) represents the Bayes' theorem. In a Bayesian language, the luminosity function is the marginalized probability and the halo mass function is the prior distribution function. $\Phi(M|L, z)$ represents the posterior probability distribution while the conditional luminosity function represents the likelihood function. Note that in practice, the luminosities of galaxies in cluster is the observable and therefore the conditional probability $\Phi(M|L, z)$ is the key element to determine the mass distribution of haloes from observations. The Bayes theorem provides an improvement of the prior, the halo mass function, once the galaxy luminosities have been measured.

The average luminosity of galaxies in halo of mass M at a given redshift z is

$$\bar{L}(M, z) = \int dL \Phi(L|M, z) L.$$

such that for a given population (satellite or centrals), this expression generates a $M_{\text{halo}} - L_{c,s}$ relation.

Inspired by the fact that the luminosity function of optically detected galaxy clusters as well as all (i.e. red and blue) galaxies (e.g. Norberg et al., 2001) can be well represented by a Schechter function, the CLF is usually parameterized in the same way (e.g Yang et al., 2003):

$$\Phi(L|M) dL = \Phi^*(M) \left(\frac{L}{L_*(M)} \right)^{\alpha(M)+1} \exp\left(-\frac{L}{L_*(M)}\right) d \ln L,$$

where the three Schechter parameters $\{\Phi^*, \alpha, L_*\}$ are function of the halo mass. There is not a physical motivated model for the mass dependence of these parameters (neither of the Schechter parameterization). This forces the introduction -together with an educated guess for the functional form- of free parameters that need to be constrained via measurements of the luminosity function or clustering. This free-parameter scenario is then translated to the HOD analysis, where different models have been used in the last few years and we will show their impact on the final clustering within the HM in a future section.

Some relevant quantities can be derived from the CLF. The mean number of objects with luminosities greater than L hosted by a halo with mass greater then M is given by

$$N_{s,c}(z; > M; > L) = \int_M^\infty \int_L^\infty dM' dL' \Phi_{s,c}(L'|M', z).$$

The mean number of galaxies in the with luminosities $> L$ is

$$\bar{n}_{\text{gal}}(z|>L) = \int dM n(M, z) N_{\text{gal}}(M, z|>L),$$

where the expected number of galaxies in a halo of mass M is

$$N_{\text{gal}}(M, z|>L) = N_c^{\text{gal}}(M, z|>L) + N_s^{\text{gal}}(M, z|>L).$$

Similarly, if we are interested in the halo distribution, then we can set $\langle N_B|M \rangle = 1$ and $u_B(\mathbf{x} - \mathbf{x}'|M) = \delta^3(\mathbf{x} - \mathbf{x}')$; we obtain

$$n_h(\mathbf{x}, z|>M_*) = \int_{M_*}^\infty dM N(\mathbf{x}'|M, z).$$

We then see that we can determine all the possible combinations for the two-point statistics of different objects: galaxies, halos (clusters) and matter.

Marked statistics

A more general treatment of the two-point statistics can be reached by the so-called marked correlation functions (e.g. Sheth, 2005). The idea is, instead of exploring the clustering of galaxies using some property as a label (as we have just shown in the previous section), we explore the clustering of that property itself. We do this by weighting each galaxy by a physical property such as its luminosity, colour, mass, etc. Let us thus assume that we weight each galaxy by a property $w(M)$ which we can relate to the mass of its hosting halo. Defining $p(w|M)dw$ as the probability distribution of a galaxy to possess a property w in the interval $w, w + dw$ being hosted by a dark matter halo of mass M , the expected value of the weight given the mass M is just

$$\langle w|M \rangle \equiv w(M) = \int dw w p(w|M)$$

and the mean mark of a sample is

$$\bar{w} = \int dM n(M) \langle w|M \rangle \quad (\text{B.11})$$

The conditional probability distribution function $p(w|M)$ is just the CLF in case our mark is the luminosity of the galaxy. We weight each galaxy by the factor $W(M) \equiv w(M)/\bar{w}$, that is, the expected mean value of w given the mass of the halo where the galaxy resides. The weighted galaxy mean number density of galaxies can be written as

$$n_{\text{gal}}^w(\mathbf{x}, z) = \int d^3 x' dM \left[(W_c^A(M) + W_c^B(M)) \delta^3(\mathbf{x}' - \mathbf{x}) + W_s^A(M) u_{\text{gal}}^A(\mathbf{x}' - \mathbf{x}; M, z) + W_c^B(M) u_{\text{gal}}^B(\mathbf{x}' - \mathbf{x}; z; M) \right] N(\mathbf{x}'; M),$$

and the two-point galaxy w -marked correlation function for galaxies with luminosities L and L' reads as

$$1 + W_{\text{gal}} = \langle n_{\text{gal}}^w(\mathbf{x}, z) n_{\text{gal}}^w(\mathbf{y}, z) \rangle = (1 + W_{\text{gal}}^{1h}) + (1 + W_{\text{gal}}^{2h}), \quad (\text{B.12})$$

For $w(M) = 1$, we obtain $\bar{w} = \bar{n}$ and we reduce to the standard (unmarked) correlation function. The 1-h and 2-h contributions to the marked correlation function can be read from Equations (B.4) and (B.5) (omitting the self-correlation terms and setting $W_{c,s}(M) = W_{c,s}^A(M) + W_{c,s}^B(M)$):

$$\begin{aligned} 1 + W_{\text{gal}}^{1h} &= \int dM n(M) \left[W_c(M) + W_s(M) u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M) \right] + \int dM n(M) W_c(M) W_s(M) u_{\text{gal}}(\mathbf{x} - \mathbf{y}; M) \\ &+ \int dM n(M) W_c(M) W_s(M) \int d^3 x' u_{\text{gal}}(\mathbf{x}' - \mathbf{x}; M) u_{\text{gal}}(\mathbf{x}' - \mathbf{y}; M), \end{aligned}$$

and

$$\begin{aligned} W_{\text{gal}}^{2h} &= \int dM dM' n(M) n(M') W_c(M) W_c(M') \xi_{hh}(\mathbf{x} - \mathbf{y}; M, M') \\ &+ \int dM dM' n(M) n(M') W_c(M) W_s(M') \int d^3 y' \xi_{hh}(\mathbf{x} - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M') \\ &+ \int dM dM' n(M) n(M') W_c(M') W_s(M) \int d^3 y' \xi_{hh}(\mathbf{x} - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M) \\ &+ \int dM dM' n(M) n(M') W_s(M) W_s(M') \int d^3 x' d^3 y' \xi_{hh}(\mathbf{x}' - \mathbf{y}'; M, M') u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M) u_{\text{gal}}(\mathbf{y} - \mathbf{y}'; M') \end{aligned}$$

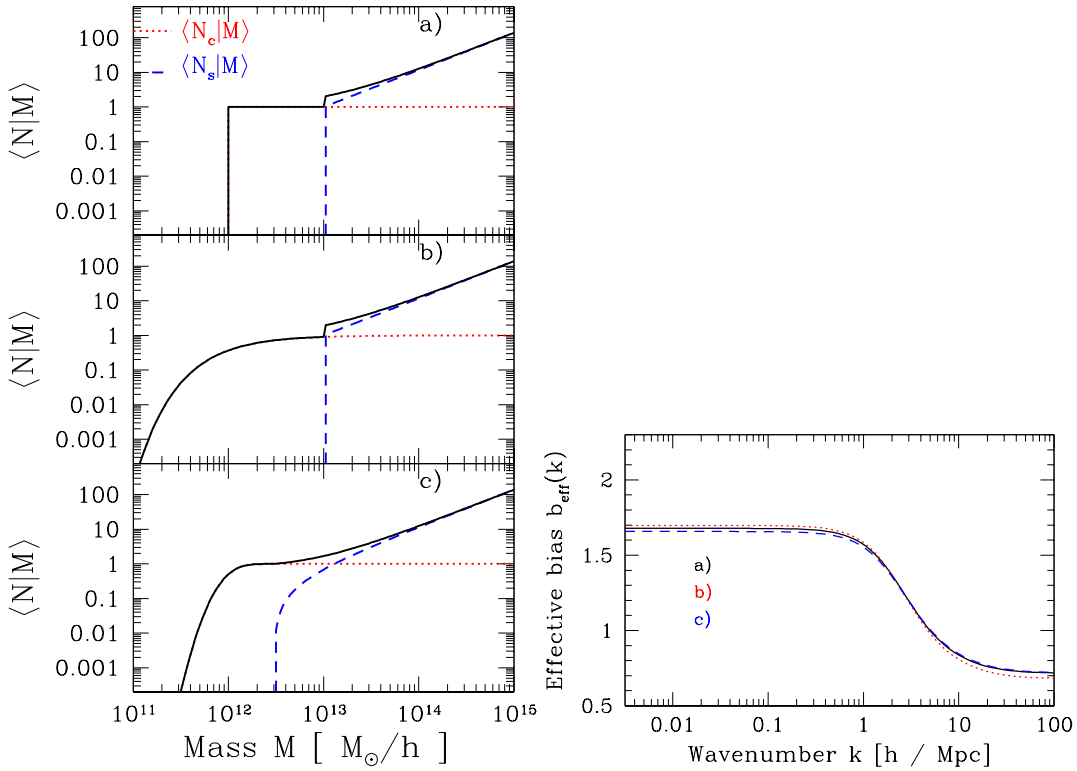


Figure B.1: Different parameterizations of the HOD following the models commented in the text. In this plot $\alpha = 1.07$, $M_1 = 10^{13} M_\odot/h$, $M_2 = 10^{12} M_\odot/h$, $\sigma_{\ln M} = 0.46$. The right panel shows the effective galaxy bias at $z = 0$ for the three HOD of the left panel, displaying small variations and even behaving very similar on scales where the scale dependency is acting.

The example developed in Chapter (3) utilized the halo mass to determine the marked correlation functions in dark matter haloes. In this case, we are only interested in the 2-h contribution. Furthermore, only the first line in the previous equation contributes to that example, since it is the one associated to the correlations between central objects. Under the typical assumption that the halo-halo correlation function can be written as $\xi_{hh}(r; M, M') = b(M)b(M')\xi_{\text{mat}}(r)$, the marked correlation function for haloes reads as $W_h(r) = b_2^2 \xi_{\text{mat}}(r)$ where

$$b_2 = b_2(z, M_{\min}) \equiv \int_{M_{\min}} dM n(M, z) b(M, z) W(M, z). \quad (\text{B.13})$$

Halo occupation distribution

The reasonable assumptions we can make in order to model the HOD for central and satellite galaxies can be described as follows:

- The more massive the dark matter halo the more galaxies it contains.
- High mass haloes are more likely to have a central galaxy. Therefore the fraction of haloes with central galaxies should behave as $N_c(M) \rightarrow 1$ as long as $M \rightarrow \infty$.
- There can exist halos which low masses such that that it potential wells weren't deep enough to amass baryonic gas to the center and form galaxies. Thus there exist a critical

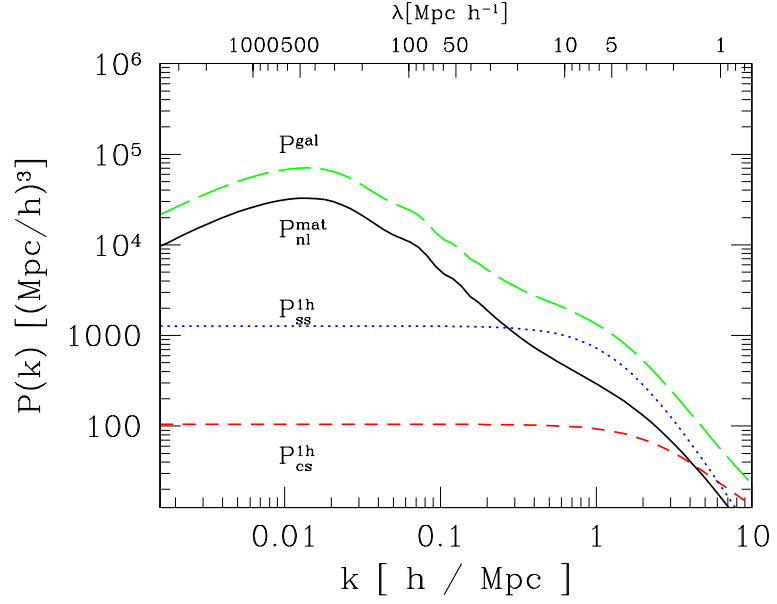


Figure B.2: Galaxy power spectrum from the halo model. The contributions from pairs satellite-central and satellite-satellite in the 1-h term are shown.

mass M_1 below which a halo cannot host a central galaxy. It means that the expected number of galaxies (both satellites and centrals) should behave as $\langle N|M \rangle \rightarrow 0$ for $M \rightarrow M_1$.

- When dark matter haloes exceed the value M_1 , then the hosted galaxy is a central galaxy. When the halo reaches a mass $M_2 > M_1$, it is allowed to host satellite galaxies. That is, $\langle N_s|M \rangle = 0$ for $M < M_2$. In this case the number of galaxies (satellites and centrals) in a halo of mass M is written as

$$N_c(M)(1 + N_c(M)) \equiv \langle N_c|M \rangle + \langle N_s|M \rangle,$$

This represents the number of galaxies in a halo (the '1' counting for the central galaxy plus satellites) weighted by the fraction of haloes with centrals.

These assumptions can be encoded in a simple set of parameterizations. For instance, the top-hat HOD parameterization, widely used (e.g Berind and Weinberg 2002) reads for central galaxies

$$\langle N_c|M \rangle_a = \begin{cases} 0 & M < M_1 \\ 1 & M \geq M_1 \end{cases},$$

while for satellite galaxies it is assumed that they follow a Poisson distribution with a mean given as a power law above some critical mass:

$$\langle N_s|M \rangle_{(a)} = \begin{cases} 0 & M < M_2 \\ (M/M_2)^\alpha & M \geq M_2 \end{cases}.$$

With three free parameters, this parameterization provides a very simplistic model of galaxy occupation number. In reality the process of galaxy formation, the process of virialization of haloes and its dependence with redshift is complicated enough that a more detailed modeling is required. For instance, to allow a slow transition between haloes containing no galaxies and galaxy-populated haloes, (Wake et al., 2008) one can model the fraction of halos with

central galaxies as an exponential function (we label it as model a) with four free parameters

$$\langle N_c | M \rangle_{(b)} = \exp \left[-\frac{M_1}{M} \right], \quad (\text{B.14})$$

while the fraction of satellites is still given by the power law. The parameters M_1, M_2 and α are fitted through the measurements of the galaxy correlation functions e.g. Wake et al., 2008.

Processes concerning baryon physics (such as galaxy formation) might introduce scatter in the values of halo masses at which a dark matter halo is allowed to contain a given number of galaxies, generating a soft transition between the empty and occupied haloes. In order to account for this transition, a four free parameters HOD has been introduced (Zheng et al., 2005; White et al., 2007a): centrals galaxies are modeled via

$$\langle N_c | M \rangle_{(c)} = \frac{1}{2} \operatorname{erfc} \left(\frac{1}{\sigma_{\ln M}} \ln \left[\frac{M_{\min}}{M} \right] \right), \quad (\text{B.15})$$

where $\operatorname{erfc}(x)$ is the complementary error function. Here M_{\min} is the minimum mass that an halo might have in order to have a central galaxy and the parameter $\sigma_{\log M}$ controls the transition from empty haloes to one-galaxy haloes. Note that this distribution can be viewed as the result of assuming a conditional luminosity function (see Section 3.5.4) for central galaxies given by a log-log distribution with a mean central galaxy luminosity-halo mass relation $L_c(M)$ and an intrinsic scatter $\sigma_{\ln L}$ about the mean e.g. Skibba and Sheth, 2009

$$\Phi(L_c | M) dL_c = \frac{1}{\sqrt{2\pi}\sigma_{\ln L_c}} \exp \left[-\frac{(\ln L_c - \ln \bar{L}_c(M))^2}{2\sigma_{\ln L_c}^2} \right] \frac{dL_c}{L_c}. \quad (\text{B.16})$$

Within the same parameterization, satellite galaxies are still modeled with a power law but with an independent cut-off at αM_{\min} :

$$\langle N_s | M \rangle_{(c)} = \left(\frac{M - \alpha M_{\min}}{M_1} \right)^\alpha, \quad (\text{B.17})$$

such that halos are only allowed to host satellites for mass es greater than αM_{\min} . These models are displayed in fig.B.1.

Halo model in redshift space

In order to model the redshift space power spectrum in the framework of the halo model, let us remember that there are two physical sources of redshift distortions: on large scales, the bulk motion of halos towards regions of high density induces an enhance in the signal of the two point statistics along the line of sight, lead by the mapping $s_z = z - u_z$, wherein this chapter we use $u_z = -v_z/H$. On small scales, the random motion of galaxies in virialized halos induce a suppression of power due to the large galaxy velocity dispersions, in the so-called fingers-of-god effect (fog hereafter): this information will be contained in the halo-distribution function. Following Equation (B.1) and a naive version of the streaming model, we can argue that the number density of galaxies in redshift space can be written as the convolution of the real space number density with a streaming function. Since the bulk motions are the one defining the translation to redshift space coordinates, we just have

$$n_{\text{gal}}^s(\mathbf{s}) = \int d^3x n_{\text{gal}}(\mathbf{x}) \delta_D^3(\mathbf{s} - (\mathbf{x} - \mathbf{v})).$$

In order to save notation, let us not consider any particular property of galaxies and let us concentrate in a single galaxy type. In terms of the distribution function via Equation (B.1), and assuming again that each halo distribution function can be separable, this can be written as

$$n_{\text{gal}}^s(\mathbf{s}) = \sum_i \int d\mathbf{z} d\mathbf{x}_\perp \mathcal{P}(v_\perp, z - s_z | M_i) \left(N_c^i \delta_D^2(\mathbf{x}_\perp - \mathbf{x}_\perp^i) \delta_D(z - z_i) + N_s^i u_{\text{gal}}(\mathbf{s}_\perp - \mathbf{s}_\perp^i, z - z_i | M_i) \right), \quad (\text{B.18})$$

where we name for simplicity $N_{c,s} = N_{c,s}(M, z)$. We can further assume that the galaxy velocity distribution function in dark matter haloes is isotropic such that we can separate it in three one-dimensional velocity distribution functions

$$\mathcal{P}(\mathbf{v} | M_i) = \mathcal{P}_{1D}(v_x | M) \mathcal{P}_{1D}(v_y | M) \mathcal{P}_{1D}(v_z | M),$$

each satisfying $\int dv_i \mathcal{P}_{1D}(v_i | M) = 1$, such that

$$n_{\text{gal}}^s(\mathbf{s}_\perp, s_z) = \sum_i \left[N_c^i \mathcal{P}_{1D}(s_z - z_i | M_i) \delta_D^2(\mathbf{s}_\perp - \mathbf{s}_\perp^i) + N_s^i \int d\mathbf{z} \mathcal{P}_{1D}(z - s_z | M_i) u_{\text{gal}}(\mathbf{s}_\perp - \mathbf{s}_\perp^i, z - z_i | M_i) \right].$$

The contribution to the number density in redshift space is the resulting of the distribution of the central galaxies, which, being in the center of dark matter haloes have a velocity \mathbf{v} representing the halo peculiar velocity (this implies that we indirectly assume that the central galaxy is in the center of mass of the halo, to which its peculiar velocity is referred to), together with the distribution of the satellite galaxies, which in redshift space follow a density profile represented by the convolution of the matter density profile followed by galaxies and the velocity distribution function of each halo. This can be rewritten in a compact form by introducing some Dirac's delta functions (and keeping in mind that the distortion acts along the z axis such that $\mathbf{x}_\perp = \mathbf{s}_\perp$):

$$n_{\text{gal}}^s(\mathbf{s}) = \sum_i \left[N_c^i \mathcal{P}(\mathbf{s} - \mathbf{x}_i | M_i) + N_s^i \int d^3x' \mathcal{P}(\mathbf{x}' - \mathbf{s} | M_i) u_{\text{gal}}(\mathbf{x}' - \mathbf{x}_i | M_i) \right].$$

The net step is try to write $n_{\text{gal}}^s(\mathbf{s})$ as in terms of $N(\mathbf{x} | M, z)$. This last equation allow us to say that in redshift space the power spectrum will contain the same dependence found before on the Fourier transform on the density profile, with the change $u(k_z | M) \rightarrow \tilde{u}(k\mu | M) = u(k | M) \mathcal{P}(k\mu | M)$ where $\mu = \hat{z} \cdot \hat{\mathbf{k}}$ is the projection of the wave number perturbation \mathbf{k} along the line of sight. Therefore, in redshift space the 1-halo contribution to the power spectrum is written as

$$P_{\text{gal}}^{1h,s}(k, \mu) = \frac{1}{\bar{n}_A^2} \int dM n(M) \left[2 \langle N_{\text{gal}}^c | M \rangle + \langle N_{\text{gal}}^s | M \rangle \tilde{u}_{\text{gal}}(k, \mu | M) \right] \langle N_A^s | M \rangle \tilde{u}_A(k, \mu | M), \quad (\text{B.19})$$

For the 2-halo term, we will assume that the halo-halo power spectrum can be written following the small angle approximation:

$$P_{hh}^s(k, \mu | M, M') = \left(1 + \beta(M)\mu^2 \right) \left(1 + \beta(M')\mu^2 \right) b(M)b(M') P_{\text{mat}}(k), \quad (\text{B.20})$$

The full $2h$ term in redshift space reads as

$$P_{\text{gal}}^{2h,s}(k, \mu) = \frac{1}{\bar{n}_g^2} \int dM dM' n(M)n(M') \hat{N}(k, M, M') b(M)b(M') \mathcal{P}(k\mu | M) \mathcal{P}(k\mu | M') \left(1 + \beta(M)\mu^2 \right) \left(1 + \beta(M')\mu^2 \right) P_{\text{mat}}(k)$$

where we have defined

$$\hat{N}(k, M, M') \equiv (\langle N_A^c | M \rangle + \langle N_A^s | M \rangle u_A(k | M')) (\langle N_A^c | M' \rangle + \langle N_A^s | M' \rangle u_A(k | M')).$$

Let us now explore the behavior of the power spectrum in the limit of large and intermediate scales. It is important to note that on large scales, the main distortion comes from the Kaiser effect, while on small scales, fog distortions are the dominant ones.

Multipole expansion in the halo model

On large scales, where the $2h$ term dominates we can neglect fog-like distortions and develop a Legendre expansion of the redshift power spectrum and use this to estimate the peculiar velocity parameter β , the multipole expansion

$$P_\ell^s(k) = (2\ell + 1) \int_0^1 d\mu \mathcal{P}_\ell(\mu) P_{gg}^{2hs}(k, \mu),$$

lead to the usual contributions from the $\ell = 0, 2, 4$ -moments out of which the real space power spectrum can be recovered. One obtains the usual expressions,

$$P_0^s(k) = \left(1 + \frac{2}{3}\langle\beta\rangle + \frac{1}{5}\langle\beta\rangle^2\right) P_{gg}^{2h}(k), \quad P_2^s(k) = \left(\frac{4}{3}\langle\beta\rangle + \frac{4}{7}\langle\beta\rangle^2\right) P_{gg}^{2h}(k), \quad P_4^s(k) = \frac{8}{35}\langle\beta\rangle^2 P_{gg}^{2h}(k),$$

where the effective β - parameter is given by

$$\langle\beta\rangle = \frac{\int dM n(M) (\langle N_c | M \rangle + u_g(k|M) \langle N_s | M \rangle) \beta(M)}{\int dM n(M) (\langle N_c | M \rangle + u_g(k|M) \langle N_s | M \rangle)} \approx \frac{1}{\bar{n}} \int dM n(M) \langle N | M \rangle \beta(M),$$

and $\beta(M) = f(z)/b(M)$. The last approximation is valid on large scales where $u(k|M) \approx 1$. With this, the isotropic (monopole) redshift power spectrum reads as

$$P_{gg}^{2hs}(k) = \left(1 + \frac{2}{3}\langle\beta\rangle + \frac{1}{5}\langle\beta\rangle^2\right) P_{gg}^{2h}(k). \quad (\text{B.21})$$

Including small-scale distortions

When we include the fog correction via a Gaussian or a Lorentzian velocity distribution, the angle average of the 1-halo term power spectrum will contain integrals of the form

$$F_0(k|M) \equiv \int_{-1}^1 d\mu \mathcal{P}(k, \mu|M), \quad F_1(k|M) \equiv \int_{-1}^1 d\mu |\mathcal{P}(k, \mu|M)|^2$$

which are in principle simple to evaluate for a Gaussian and Lorentzian velocity distributions. The general form of the $1-h$ term is then

$$P_{\text{gal}}^{1hs}(k) = \frac{1}{\bar{n}_A^2} \int dM n(M) \left[2 \langle N_{\text{gal}}^c | M \rangle F_0(k|M) + \langle N_{\text{gal}}^s | M \rangle u_{\text{gal}}(k|M) F_1(k|M) \right] \langle N_A^s | M \rangle u_{\text{gal}}(k, \mu|M), \quad (\text{B.22})$$

Similarly, the $2h$ term will contain integrals of the form

$$F_2(k|M, M') \equiv \int_{-1}^1 d\mu \mathcal{P}(k, \mu|M') \mathcal{P}(k, \mu|M) (1 + \beta(M)\mu^2)(1 + \beta(M')\mu^2).$$

In general we can decompose these integrals F_i as

$$F_i(k|M, M') = A_i(k|M, M') + \frac{1}{3} B_i(k|M, M') (\beta(M) + \beta(M')) + \frac{1}{5} C_i(k|M, M') \beta(M) \beta(M'),$$

where $A_0 = F_0(k|M)$, $B_0 = C_0 = 0$, $A_1 = F_1(k|M)$, $B_1 = C_1 = 0$ and

$$\begin{aligned} A_2(k|M, M') &= \int_{-1}^1 d\mu \mathcal{P}(k, \mu|M) \mathcal{P}(k, \mu|M') \\ B_2(k|M, M') &= \int_{-1}^1 d\mu \mu^2 \mathcal{P}(k, \mu|M) \mathcal{P}(k, \mu|M') \\ C_2(k|M, M') &= \int_{-1}^1 d\mu \mu^4 \mathcal{P}(k, \mu|M) \mathcal{P}(k, \mu|M') \end{aligned}$$

For the 2-h contribution we can rearrange terms and write

$$P_{gg}^{2hs}(k) = \left(\mathcal{F}_{gg}(k) + \frac{2}{3}\mathcal{F}_{gv}(k) + \frac{1}{5}\mathcal{F}_{vv}(k) \right) P_{\text{mat}}(k). \quad (\text{B.23})$$

where the functions $\mathcal{F}(k)$ are defined as

$$\begin{aligned} \mathcal{F}_{gg}(k) &\equiv \frac{1}{\bar{n}_g^2} \int dM dM' n(M) n(M') b(M) b(M') A_2(k|M, M') \hat{N}(k, M, M') \\ \mathcal{F}_{gv}(k) &\equiv \frac{1}{\bar{n}_g^2} \int dM dM' n(M) n(M') b(M) b(M') B_2(k|M, M') \hat{N}(k, M, M') \frac{1}{2} (\beta(M) + \beta(M')) \\ \mathcal{F}_{vv}(k) &\equiv \frac{1}{\bar{n}_g^2} \int dM dM' n(M) n(M') b(M) b(M') C_2(k|M, M') \hat{N}(k, M, M') \beta(M) \beta(M') \end{aligned} \quad (\text{B.24})$$

The set of Equations (B.22) and (B.23) represent the galaxy power spectrum in redshift space.

Halo exclusion

Exclusion effect arises due to the fact that dark matter haloes defined via fog algorithms are not allowed to overlap (e.g Porciani et al., 1998; Smith et al., 2007). In other words when exploring halos with virial radii R_{vir} , scales below $\approx 2R_{\text{vir}}$ cannot be probed by the halo-halo clustering analysis. In order to see how this exclusion might distort the shape of the power spectrum or correlation function, we can start by analyzing the correlation function for dark matter haloes with masses above some limit M_\star . The exclusion effect is then introduced via a non-overlapping probability function $P(x = r/\tilde{R})$ where r is the scale probed by the clustering and $\tilde{R}(M, M') = R_{\text{vir}}(M) + R_{\text{vir}}(M')$ is the sum of the virial radii of two halos with masses M and M' . The exclusion not only introduces a scale dependent bias through this function $P(r)$ but also a scale dependent halo number density such that the cluster-cluster correlation function reads as

$$\xi_{\text{cl}}(r; > M_\star) = \frac{1}{\bar{n}_h^2(r)} \int_{M_\star}^{\infty} \int_{M_\star}^{\infty} dM dM' n(M) n(M') P(x = r/\tilde{R}) \xi_h(r; M, M')$$

where the scale dependent mean halo density is written as

$$\bar{n}^2(r) = \int_{M_\star}^{\infty} \int_{M_\star}^{\infty} dM dM' n(M) n(M') P(x = \tilde{R}/r) \quad (\text{B.25})$$

Therefore we cannot write anymore the halo correlation function as the product of an effective bias times the matter correlation function, even if we make the usual assumption of a scale independent halo-mass bias. For spherical halos, the non-overlapping probability is just given by a top-hat function $P(x) = 1, 0$ for $x > 1$ or $x \leq 0$ respectively. Recent analysis consider the ellipsoidal scenario, though in order to be consistent one should also consider the correct density profile for ellipsoidal haloes when using the HM as in Smith et al. (2006); for ellipsoidal haloes Tinker (2007) has found

$$P(x) = \begin{cases} 3y^2 - 2y^3, y = (x - 0.8)/0.29 & 0 \leq y \leq 1 \\ 0 & y < 0 \\ 1 & y > 0 \end{cases} .$$

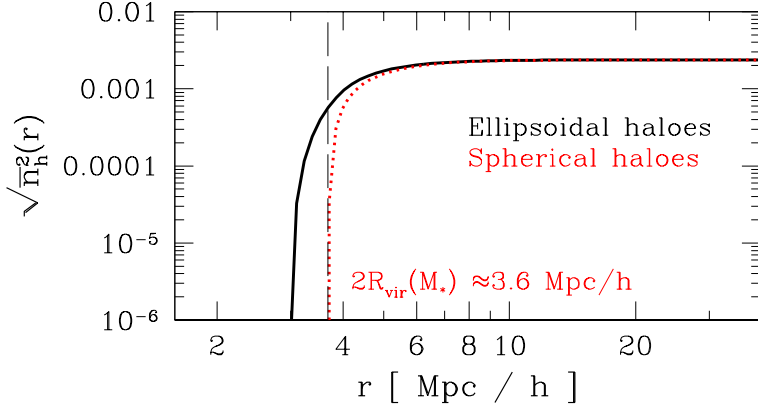


Figure B.3: Scale dependent mean density determined from Equation (B.25) for the spherical and ellipsoidal dark matter haloes for a mass limit such that its corresponding virial radius is $\approx 1.8 \text{ Mpc}/h$. The vertical line marks the value of $2R_{\text{vir}}$. One large scales the the mean density recovers its non-exclusion value.

Halo exclusion is then translated to a scale dependent bias that makes the passage from configuration space to Fourier space less straightforward. Combining both scale-dependent effects we can write the cluster power spectrum as

$$P_{\text{cl}}(k) = \int_0^\infty dk' k'^2 \hat{G}(k, k') P_{\text{mat}}(k),$$

where the kernel

$$G(k, k') = \frac{2}{\pi} \int_0^\infty dr r^2 j_0(kr) j_0(k'r) \hat{B}^2(r),$$

and

$$\hat{B}^2(r) = \frac{B^2(r)}{\bar{n}^2(r)} \int_{M_*}^\infty dM \int_{M_*}^\infty dM' n(M) n(M') b(M) b(M') P(x = r/\tilde{R}).$$

such that with $\hat{B}^2(r) = b_{\text{eff}}^2$ for $r \geq r_1 \approx 100 \text{ Mpc}/h$ and $B_{\text{ex}}^2(r) = 0$ for $r < r_0 \approx 1.5 \text{ Mpc}/h$. Hence we can split the power spectrum in two contributions: a first term which is the power spectrum we already determined without exclusion plus the exclusion effect:

$$P_{\text{cl}}(k) = b_{\text{eff}}^2 P_{\text{mat}}(k) + P_{\text{excl}}(k) \quad (\text{B.26})$$

with

$$P_{\text{excl}}(k) = \int_0^\infty dk' k'^2 \hat{G}(k, k') P_{\text{mat}}(k) \quad (\text{B.27})$$

and the kernel

$$\hat{G}(k, k') = \int_{r_0}^{r_1} dr r^2 j_0(kr) j_0(k'r) (\hat{B}^2(r) - b_{\text{eff}}^2). \quad (\text{B.28})$$

In this way, the exclusion effect introduces a cut-off in the two-point statistics when it is analyzed in configuration space, while translated to Fourier space it is embodied in a suppression of power on small scales followed by low amplitude oscillations. Before we analyze this, it is worth to mention that the introduction of the exclusion effect increases the time of calculation required for the modeling the two point statistics. A way to speed the calculations has been proposed by Tinker et al. (2005), which consists in determining the mean density

$$\bar{n}' = \int_0^{M_{\text{lim}}(r)} dM n(M)$$

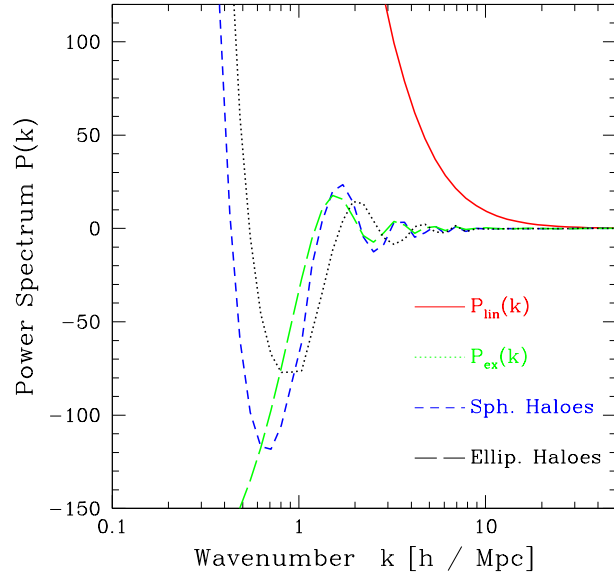


Figure B.4: Power spectrum with halo exclusion: comparison between the exclusion effect obtained from Equation (B.26) with the pure geometrical approach of Equation (B.29.)

such that at a given scale r we can solve for the limiting mass $M_{\text{lim}}(r)$ in order to match the scale dependent mean density $\bar{n}'^2 = \bar{n}^2(r)$. In this way we can therefore determine an effective bias in configuration space just by calculating

$$\hat{B}'(r) = \frac{\int_{M_\star}^{M_{\text{lim}}(r)} dM n(M) B(r) b(M)}{\int_{M_\star}^{M_{\text{lim}}(r)} dM n(M)}$$

These different ways of calculating the bias with exclusion effect are similar to a 1%, being the second the less time consuming. A comparison between these two methods can be seen in Fig. B.3.

A simple geometrical approach to the exclusion is usually applied based on the analysis of Smith et al. (2007). Knowing the minimum mass of the sample, we can write the correlation function as $\xi_{\text{cl}}(r) = \xi_{\text{cl}}^{\text{no ex.}}(r) + \xi_{\text{exc}}(r)$ where the exclusion is modeled as

$$\xi_{\text{cl}}(r) = \begin{cases} -1 & r \leq 2R_{\text{vir}}(M_\star) \equiv \bar{R} \\ 0 & r > 2R_{\text{vir}}(M_\star) \end{cases} \quad (\text{B.29})$$

then the power spectrum will have the same form as in Equation (B.26) but now the exclusion term is purely a geometrical term, corresponding to the Fourier transform of a top-hat function

$$P_{\text{cl}}(k) = -\frac{4\pi\bar{R}}{k} j_1(k\bar{R})$$

The comparison between these the approaches is shown in Fig. B.4 for the power spectrum.

The power spectrum estimator

The FKP estimator (Feldman et al., 1994) has been designed to obtain unbiased measurements of the underlying power spectrum of interest (galaxy, galaxy cluster) with minimal (optimal) variance. This is achieved. The synthetic catalog is constructed with the same selection function of the real survey and must possess a large number of objects such that it can be used as a reliable estimate of the volume of the survey. For a galaxy redshift survey, it is often used a synthetic catalogue with ~ 10 times more objects than the real one (e.g. Cole et al., 2005; Percival et al., 2007a). For galaxy cluster survey like the REFLEX II, we used a synthetic catalog with $\sim 10^4$ more objects in the synthetic catalog. The FKP estimator measures the object (galaxy or galaxy cluster) power spectrum by defining a weighted object fluctuation in the form

$$F(\mathbf{r}) \equiv \frac{w(\mathbf{r})}{N}(n_g(\mathbf{r}) - \alpha n_s(\mathbf{r})), \quad (\text{C.1})$$

where N is a normalization constant,

$$N^2 = \int d^3r \bar{n}^2(\mathbf{r}) w(\mathbf{r})^2.$$

The parameter α matches the number densities of real and synthetic catalogs and sets the zeroth of the galaxy fluctuation $\int d^3r F(\mathbf{r}) = 0$,

$$\alpha = \frac{\int d^3r w(\mathbf{r}) n_g(\mathbf{r})}{\int d^3r w(\mathbf{r}) n_r(\mathbf{r})}. \quad (\text{C.2})$$

This implies that the density field which fluctuation is $F(r)$ is characterized is *forced* to be comparable with the real mean density. In other words, when writing α in this way we are assuming that our sample is a fair representation of our Universe on large scales. This is of course not true for the volume we have access to is finite and limited. We therefore must correct for this fake information in the power spectrum estimator. Nevertheless, there is no way to know -beforehand- the true mean number density of objects given the limits in mass, or luminosity our survey is characterized with.

We then take mean ensemble averages of the product of the inhomogeneous cluster distribution $\langle F(\mathbf{r})F(\mathbf{r}') \rangle$. This implies the knowledge of three terms, namely, the cluster-cluster correlation, the cluster-random correlation and the random-random correlation. The cluster-cluster correlation can be written as

$$\langle n_{\text{cl}}(\mathbf{r})n_{\text{cl}}(\mathbf{r}') \rangle_{\text{ens}} = \bar{n}_{\text{cl}}(\mathbf{r})\delta^3(\mathbf{r} - \mathbf{r}') + \bar{n}_{\text{cl}}(\mathbf{r})\bar{n}_{\text{cl}}(\mathbf{r}') + \bar{n}_{\text{cl}}(\mathbf{r})\bar{n}_{\text{cl}}(\mathbf{r}')\xi(\mathbf{r} - \mathbf{r}').$$

Here $\xi(\mathbf{r} - \mathbf{r}')$ represents the underlying correlation function and $\bar{n}_{\text{cl}}(\mathbf{r})$ is the expected number density of clusters at a location \mathbf{r} . The cluster-random term only possesses a Poisson component (that is, the objects the two catalogues are not correlated):

$$\langle n_{\text{cl}}(\mathbf{r})n_{\text{r}}(\mathbf{r}') \rangle_{\text{ens}} = \frac{1}{\alpha} \bar{n}_{\text{cl}}(\mathbf{r})\bar{n}_{\text{cl}}(\mathbf{r}'), \quad (\text{C.3})$$

and finally the random-random correlation possesses a Poisson component and also a shot noise contribution,

$$\langle n_{\text{r}}(\mathbf{r})n_{\text{r}}(\mathbf{r}') \rangle_{\text{ens}} = \frac{1}{\alpha^2} \bar{n}_{\text{cl}}(\mathbf{r})\bar{n}_{\text{cl}}(\mathbf{r}') + \frac{1}{\alpha} \bar{n}_{\text{cl}}(\mathbf{r})\delta^3(\mathbf{r} - \mathbf{r}').$$

Collecting these terms, the correlation function in the FKP estimator can be written as

$$\langle F(\mathbf{r})F(\mathbf{r}') \rangle_{\text{ens}} = W(\mathbf{r})W(\mathbf{r}')\xi_f(\mathbf{r} - \mathbf{r}') + S(\mathbf{r} - \mathbf{r}'), \quad (\text{C.4})$$

where $W(\mathbf{r})$ is the survey window function in configuration space

$$W(\mathbf{r}) = \frac{1}{N} \bar{n}(\mathbf{r})w(\mathbf{r}),$$

and $S(\mathbf{r})$ is the shot-noise (or auto-correlation) contribution

$$S(\mathbf{r} - \mathbf{r}') = \frac{1 + \alpha}{N^2} \bar{n}(\mathbf{r})w^2(\mathbf{r})\delta^3(\mathbf{r} - \mathbf{r}').$$

Using Parseval's theorem, the Fourier transform Eq.(C.4) leads us to write the covariance between the fluctuations $F(\mathbf{r})$ in Fourier space as

$$\langle F(\mathbf{k})F^*(\mathbf{k}') \rangle_{\text{ens}} = \int \frac{d^3k''}{(2\pi)^3} P(k'')W(\mathbf{k}'' - \mathbf{k})W^*(\mathbf{k}'' - \mathbf{k}') + S(\mathbf{k} - \mathbf{k}'), \quad (\text{C.5})$$

where the Fourier transform of the shot-noise like term $S(\mathbf{r})$ is given as

$$S(\mathbf{k}) = \frac{(1 + \alpha)}{N^2} \int d^3r \bar{n}(\mathbf{r})w^2(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}. \quad (\text{C.6})$$

From Equation (C.5) we can write

$$\langle |F(\mathbf{k})|^2 - S(0) \rangle_{\text{ens}} = \int \frac{d^3k''}{(2\pi)^3} |W(\mathbf{k} - \mathbf{k}'')|^2 P_f(k''). \quad (\text{C.7})$$

That is, the statistical average of the quantity $|F(\mathbf{k})|^2 - S(0)$ represents the convolution of the underlying power spectrum with the survey window function. In practice we have only access to only one realization of our ensemble, i.e, the universe we observe. The FKP estimator is then defined as the shell average of the shot-noise subtracted quantity

$$\hat{P}(\mathbf{k}) \equiv |F(\mathbf{k})|^2 - S(0),$$

such that the measured power spectrum

$$\hat{P}(k_i) \equiv \langle \hat{P}(\mathbf{k}) \rangle_{V_{k_i}} = \frac{1}{V_{k_i}} \int_{V_{k_i}} d^3k (|F(\mathbf{k})|^2 - S(0)),$$

can be a good description of the quantity $W^2 \otimes P$ as long as the volume covered by the survey is large. In this expression $V_{k_i} = 4\pi\Delta k k_i^2$ is the volume of the spherical shell of width $\Delta k \ll k_i$. In terms of the FKP estimator $F(\mathbf{r})$ reads

$$\hat{P}(k_i) = \int_V d^3r d^3r' F(\mathbf{r})F(\mathbf{r} - \mathbf{r}') \langle e^{-i\mathbf{k}\cdot\mathbf{r}'} \rangle_{V_{k_i}} - S(0),$$

where $\langle e^{-i\mathbf{k}\cdot\mathbf{r}'} \rangle_{V_{k_i}}$ is the spherical average of the a plane wave in the volume V_{k_i} , which can be analytically written using the Raylight expansion of plane waves.

The FKP estimator generates a measurement of power spectrum $\hat{P}(k_i)$ which in principle is equal to the shell-average of $W^2 \otimes P_f$, that is

$$\hat{P}(k_i) = \int \frac{d^3 k''}{(2\pi)^3} \langle |W(\mathbf{k} - \mathbf{k}'')|^2 \rangle_{V_{k_i}} P_f(\mathbf{k}''). \quad (\text{C.8})$$

We therefore need to measure the survey window function in order to compare a theoretical model $P(k)$ with our measurements. Note that the same procedure applied to estimate the power spectrum can be translated to the measurements of the window function in Fourier space. Going back to Equation (C.1), and turning-off the number density of the real objects and using Equation (C.3) we can define an estimate for the squared of the Fourier transform of the window function $|F_W(\mathbf{k})|^2$ as

$$|F_W(\mathbf{k})|^2 = \left| \int d^3 r \bar{n}(\mathbf{r}) w(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \right|^2 - \frac{\alpha}{1+\alpha} S(0), \quad (\text{C.9})$$

which is our definition of window function with the corresponding subtraction of a shot-noise term, arising due to the fact that we use the finite number of random-distributed objects to trace the window function.

Covariance

The weighting function is chosen for the FKP estimator such that the variance of the power spectrum estimator is minimized with respect to variations in $w(\mathbf{r})$. The covariance matrix of the power spectrum is defined as

$$C(\mathbf{k}, \mathbf{k}') = \langle (\hat{P}(\mathbf{k}) - P(\mathbf{k})) (\hat{P}(\mathbf{k}') - P(\mathbf{k}')) \rangle,$$

In order to give error bars on our spherical averaged power spectrum, we need to average the covariance matrix in the same spherical shells of width Δk ; this is written as

$$\hat{C}_{ij} = \frac{1}{V_{k_i}} \int_{V_{k_i}} d^3 k \frac{1}{V_{k_j}} \int_{V_{k_j}} d^3 k' C(\mathbf{k}, \mathbf{k}').$$

Under the assumption that the fluctuations around the mean can be described with a Gaussian distribution function, Feldman et al. (1994) showed that the covariance matrix can be written as

$$C(\mathbf{k}, \mathbf{k}') = |\langle F(\mathbf{k}) F^*(\mathbf{k}') \rangle|^2. \quad (\text{C.10})$$

Thus, the bin averaged covariance matrix is given by

$$\hat{C}_{ij} = \frac{1}{V_{k_i}} \int_{V_{k_i}} d^3 k \frac{1}{V_{k_j}} \int_{V_{k_j}} d^3 k' \left| S(\mathbf{k} - \mathbf{k}') + \int \frac{d^3 k''}{(2\pi)^3} P(k'') W(\mathbf{k}'' - \mathbf{k}) W^*(\mathbf{k}'' - \mathbf{k}') \right|^2$$

In this expression $P(k)$ is still the underlying power spectrum. The expression for the covariance matrix is simplified for a volume limited sample, where

$$S(\mathbf{k} - \mathbf{k}') = \frac{(1+\alpha)}{\bar{n}V} \delta^3(\mathbf{k} - \mathbf{k}') \equiv S \delta^3(\mathbf{k} - \mathbf{k}'), \quad (\text{C.11})$$

and we recover

$$\hat{C}_{ij}^{VLS} = \frac{1}{V_{k_i}} \int_{V_{k_i}} d^3k \frac{1}{V_{k_j}} \int_{V_{k_j}} d^3k' \delta^3(\mathbf{k} - \mathbf{k}') |P(\mathbf{k}') + S|^2, \quad (\text{C.12})$$

such that if the shell average is done on shells with widths $\Delta k \gg |\mathbf{k} - \mathbf{k}'|$ and if the power spectrum $P(\mathbf{k})$ is constant on the scale Δk , the covariance matrix reduces to

$$\hat{C}_{ij}^{VLS} = \frac{1}{V_{k_i}^2} (P(k_i) + S)^2 \quad (\text{C.13})$$

Under the assumption of Gaussianity, the variance in the power spectrum reads as

$$\sigma_p^2(k_i) = \left\langle \frac{1}{V_{k_i}^2} \int_{V_{k_i}} d^3k d^3k' \delta P(\mathbf{k}) \delta P(\mathbf{k}') \right\rangle = \frac{1}{V_k^2} \int_{V_k} d^3k d^3k' \langle |F(\mathbf{k}) F(\mathbf{k}')|^2 \rangle,$$

with $\delta P = \hat{P}(k) - P(k)$. Under the assumptions that the window function is compact such that the power spectrum is constant on modes associated to scales $k \ll 1/R$, the variance takes the form

$$\sigma_p^2(k_i) = \frac{1}{V_{k_i}} \int d^3k' |P(k) Q(\mathbf{k}') + S(\mathbf{k}')|,$$

where

$$Q(\mathbf{k}) \equiv \frac{1}{N^2} \int d^3r w^2(\mathbf{r}) \bar{n}^2(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}}.$$

Using Parseval's theorem these integrals can be converted to integrals in configuration space. We reduce to

$$\sigma_p^2(k_i) = \frac{(2\pi)^3 \int d^3r \bar{n}^2(\mathbf{r}) w^4(\mathbf{r}) (1 + \bar{n}P(k_i))^2}{V_k \left[\int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r}) \right]^2},$$

such that minimizing $\sigma_p^2(k)$ with respect to the weighting function we obtain the FKP result:

$$w(\mathbf{r}) = \frac{1}{1 + \bar{n}(\mathbf{r})P(k)}. \quad (\text{C.14})$$

Usually in flux limited samples the expected number density $\bar{n}(\mathbf{r})$ is a fast decreasing function of \mathbf{r} , such that on large scales we have $\bar{n}(\mathbf{r}) \rightarrow 0$ one has $w(\mathbf{r}) \rightarrow 1$ and we weight each galaxy by an equal weight ~ 1 . On the other extreme, on small scales the number density dominates such that we weight each galaxy by $w(\mathbf{r}) \rightarrow 1/\bar{n}(\mathbf{r})$. The usual conflictive part of using the FKP estimator is the fact that the weighting function requires a previous knowledge of the power spectrum we want to measure. As a remedy it has been adopted to use a constant value as an estimation of the amplitude of the power on the scales of interest.

Coming back to the variance of the power spectrum, the bin-averaged minimum variance weighting function it reads as

$$\frac{\sigma_p^2(k_i)}{P(k_i)^2} = \frac{(2\pi)^3}{V_{k_i}} \left[\int d^3r \left(\frac{P(k_i) \bar{n}(\mathbf{r})}{1 + P(k_i) \bar{n}(\mathbf{r})} \right)^2 \right]^{-1}.$$

This expression defines the effective volume probed by a survey (Feldman et al., 1994; Tegmark, 1997)

$$V_{\text{eff}}(k_i) \equiv \int d^3r \left(\frac{P(k_i) \bar{n}(\mathbf{r})}{1 + P(k_i) \bar{n}(\mathbf{r})} \right)^2. \quad (\text{C.15})$$

and is used as a tool to characterize the statistical relevance of a survey characterized by a selection function $\bar{n}(\mathbf{r})$. On large scales, where the signal dominates over the shot noise contribution, we get $V_{\text{eff}}(k) \sim V$, the volume of the sample. On scales dominated by the shot noise, the effective volume becomes $V_{\text{eff}}(k) \sim NP(k)$, where N is the total number of objects in the sample. The variance of the power spectrum scales then as $V_{\text{eff}}^{-1/2}$,

$$\frac{\sigma_p^2(k_i)}{P(k_i)^2} = \frac{2\pi^2}{\Delta k k^2 V_{\text{eff}}(k)}.$$

For a constant selection function (a simulation or a volume limited sample) this reduces to

$$\frac{\sigma_p^2(k_i)}{P(k_i)^2} = \frac{2\pi^2}{\Delta k k_i^2} \left[1 + \frac{1}{P(k_i)\bar{n}} \right], \quad (\text{C.16})$$

which just says that, even if $P(k) = 0$ we will still find a contribution to the variance associated to the finite volume the survey is probing, the so-called cosmic variance. The covariance matrix is then written as

$$C(\mathbf{k}, \mathbf{k}') \approx |\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle|^2 = (2\pi)^6 \delta^3(\mathbf{k} - \mathbf{k}') \left(P(k) + \frac{1}{\bar{n}} \right)^2.$$

Thus the bin-averaged covariance matrix is written by

$$\hat{C}_{ij} = \frac{1}{V_{k_i} V_{k_j}} \int_{V_{k_i}} \int_{V_{k_j}} d^3k d^3k' C(\mathbf{k}, \mathbf{k}'),$$

such that assuming that the underlying power spectrum is constant within the shell of width Δk , we obtain $\hat{C}_{ij} = \delta_{ij} \sigma_p^2(k_i)$. On scales where $P\bar{n} \gg 1$, the fractional error decreases as $\sigma/P \sim k^{-1}$. On the other hand, on scales where the shot-noise dominates $P\bar{n} \ll 1$, the fractional error behaves as $\sigma/P \sim k^{n-1}$ provided that on these scales $P \sim k^{-n}$. Thus, for $n > 1$, the fractional error increases on scales dominated by the shot-noise.

Window matrix

The convolution of the underlying power spectrum and the window function can be easily determined if we assume an isotropic underlying power spectrum, i.e, if we neglect redshift distortions such that $P(\mathbf{k}) = P(|\mathbf{k}|) = P(k)$. In that case, the spherical average and the convolution commute $\hat{P} = P \otimes \langle |W|^2 \rangle_{\text{sph}}$. The ensemble average of the estimator can be determined as

$$\hat{P}(k_i) = \frac{1}{4\pi^2} \int_0^\infty dq P(q) q^2 \int_{-1}^{+1} d\mu \hat{W}^2 \left(\sqrt{k_i^2 + q^2 - 2k_i q \mu} \right). \quad (\text{C.17})$$

Using a Gauss-Legendre integration method (Press et al., 2002), this can be written as a matrix product,

$$\langle \hat{P}(k_i) \rangle = \sum_j W_{ij} P_f(k_j),$$

where the window matrix $W_{ij} \equiv W(k_i, k_j)$ is defined as

$$W_{ij} = \frac{\omega_{GL}(k_j)}{4\pi^2} k_j^2 \int_{-1}^{+1} d\mu \hat{W}^2 \left(\sqrt{k_i^2 + k_j^2 - 2k_i k_j \mu} \right). \quad (\text{C.18})$$

where $\omega_{GL}(k_j)$ are the weights given by the Gauss-Legendre integration method.

Integral constraint

Equation (C.2) defined the level from which the FKP estimator uses as reference to determine the power of fluctuations at a certain scale. The amplitude of the power spectrum is then sensitive to the parameter α , which in turn depends on the mean density of objects in the universe. Having no way of determining this information, we have to live with the uncertainty provided by assuming our sample as a fair sample of the Universe. Nevertheless, what could be the effect if we have fluctuations around this value?. Let us follow Percival et al. (2004b) and ask what would happen under a variation of the mean number density such that $\alpha \rightarrow \alpha = \alpha_0 + \beta$. The fluctuation defined in Equation (C.1) will vary as

$$F(\mathbf{k}) \rightarrow F(\mathbf{k}) = F_0(\mathbf{k}) - \beta W(\mathbf{k}),$$

where $F_0(\mathbf{k})$ is the fluctuation in Equation (C.1) for $\beta = 0$. Under the assumption that β is small compared with α_0 , the final estimate of the power spectrum reads as

$$\hat{P}(\mathbf{k}) = \langle |F(\mathbf{k})|^2 \rangle - S(0) = \langle |F_0(\mathbf{k})|^2 \rangle - \left(\frac{\beta}{\alpha_0} \right)^2 (|W(\mathbf{k})|^2 + S(0)).$$

The shot-noise contribution in the second term arises from what we subtract from the spherically averaged window function, according to Equation (C.9). We therefore see that a variation in the mean number density (i.e, a variation in α) is translated to subtract a scaled copy of the window function in Fourier space. Note that in a volume-limited sample this is just the subtraction of a term proportional to a Dirac delta function centered at the mode $\mathbf{k} = 0$. Clearly, we do not have any method to do this correction, for we do not know β . What we can do is to use the fact that, given our α_0 and the assumption that the level it sets corresponds to fluctuations with zero mean, then the power spectrum estimate should be such that on the largest scales, we recover homogeneity, $\hat{P}(\mathbf{k} = 0) = 0$.

In terms of the window matrix defined in Equation (C.18), the measured power spectrum can be written as the multiplication of the underlying power spectrum with the window matrix minus the scaled version of the window matrix. Such scaling is given by the constrain $P(0) = 0$. Therefore, given the underlying power spectrum $P(k)$, an estimate of the measured power spectrum taking into account the integral constraint can be written as

$$\langle \hat{P}(k_i) \rangle = \sum_j W_{ij} P(k_j) - \frac{W_{i0}}{\sum_j W_{0j} P(k_j)} \quad (\text{C.19})$$

Note that the correction is still cosmology dependent through, even if we fix the window matrix with our fiducial cosmology. This nevertheless have been shown to not provide biased estimations on cosmological parameters.

Implementation

In order to use a FFTw algorithm as FFTw (Frigo and Johnson, 2008), we need a use a mass-assignment scheme and assign particles to the FFT grid. Following the definition of the parameter α , we can write as a sum over the real and synthetic catalog:

$$\alpha = \frac{\sum_{i=1}^{N_c} w_i}{\sum_{j=1}^{N_r} w_j},$$

such that for a volume limited sample the parameters α reduces to the ratio of the number of galaxies to the number of randomly distributed objects. To determine the shot-noise

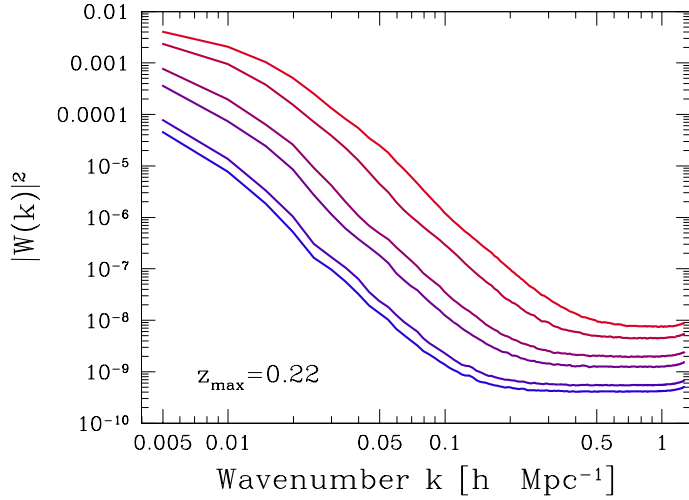


Figure C.1: Shell-averaged window function $|W(k)|^2$ for the six sub-samples defined in Table 4.1. The sub-sample L_1^{\min} is the higher curve.

contribution from Equation (C.6), we use the approximation $\bar{n}(\mathbf{r}) \approx \alpha n_r(\mathbf{r})$ in order to convert the integrals over the surveyed volume by sums over random objects: and

$$S(0) = \frac{\alpha(\alpha + 1)}{N^2} \sum_{j=1}^{N_r} w_j^2.$$

The FFT is done by determining the weight w_i of each object and assign it via a mass-assignment scheme the grid. We have chosen the triangular-shaped-cloud, although the cloud-in-cell and the nearest grid point gives similar results in the scales where we use our measurements to extract cosmological information.

Similar, the window function is obtained by sampling in a grid the quantity $\bar{n}_i w_i$, followed by the Fourier transform shell-average and shot-noise subtraction following Equation (C.9). With this we obtain the $\langle |W(\mathbf{k} - \mathbf{k}'')|^2 \rangle_{V_{k_i}}$ which is thereafter interpolated with a spline subroutine (Press et al., 2002) in order to create the window matrix. In Fig. C.1 it is shown the shell-averaged window function of the REFLEX II sample for the six sub-samples defined in Table 4.1. The amplitude of the window function varies according to the parameter α for each sub-sample. On intermediate and small scales, the window function reaches a plateau which embodies the fact that we have a finite number of random objects sampling the surveyed volume. Note that this constant value is reached at different scales for the different sub-samples.

Bibliography

- G. O. Abell. *Astrophys. J. Suppl.*, 3:211, 1958.
- G. O. Abell, H. G. Corwin, and R. P. Olowin. *Astrophys. J. Suppl.*, 70:1, 1989.
- C. Alcock and B. Paczynski. *Nature*, 281:358, 1979.
- S. Allen, R.W Schmidt, and H. Ebeling et al. *MNRAS*, 353:457–467, 2004.
- S. W. Allen, R. W. Schmidt, and A.C. Fabian. *MNRAS*, 334:L11–L15, 2002.
- R. Amanullah, C. Lidman, and D. Rubin et al. *ApJ*, 2010.
- S. Andreon, B. Maughan, G. Trinchieri, and J. Kurk. *A&A*, 507:1, 2009.
- R. Angulo, C. M. Baugh, and C. S. Frenk et al. *MNRAS*, 383:755, 2008.
- K. A. Arnaud. *Astronomical Data Analysis Software and Systems*, 101:17, 1996.
- M. Arnaud, G. W. Pratt, and R. Piffaretti et al. *arXiv:0910.1234v1 [astro-ph.CO]*, 2009.
- N. A. Bahcall. *Astrophys. J.*, 232:L83–L88, 1979.
- N. A. Bahcall. *ARAA*, 26:631, 1988.
- N. A. Bahcall and R. M. Soneira. *Astrophys. J.*, 270:20–38, 1983.
- A. Balaguera-Antolinez, D. F. Mota, and M. Nowakowski. *MNRAS*, 382:621, 2007.
- A. Balaguera-Antolinez, A. G. Sánchez, and H. Böhringer et al. *MNRAS*, In Press, 2011.
- I. K. Baldry, K. Glazebrook, and K. Budavári et al. *MNRAS*, 358:441–456, 2005.
- W. E. Ballinger, J. A. Peacock, and A. Heavens. *MNRAS*, 282:877, 1996.
- J. M. Bardeen, J. R. Bond, and N. Kaiser et al. *Astrophys. J.*, 304:15, 1986.
- S. Basilakos, M. Plionis, and J. A. S. Lima. *arXiv:1006.3418v1 [astro-ph.CO]*, page 1, 2010.
- N. Benitez, E. Gaztanaga, and R. Miquel et al. *Astrophys. J.*, 691:241–260, 2009.
- A. Berlind and D. Weinberg. *Astrophys. J.*, 575:587–616, 2002.
- R. Bernabei, P. Belli, and F. Capella et al. *Eur. Phys. J.C*, 56:333–355, 2008.
- E. Bertschinger. *Astrophys. J. supp*, 58:39–66, 1985.
- J. Binney and S. Tremaine. *Galactic Dynamics*. Princeton University Press, 1987.
- C. Blake, A. Collister, and O. Lahav. *MNRAS*, 385:1257–1269, 2008.
- H. Böhringer and P. Schuecker et al. *A&A*, 469:363, 2007.
- H. Böhringer, W. Voges, and J. P. Huchra. *Astrophys. JS*, 129:435, 2000.
- H. Böhringer, C. Collins, and L. Guzzo et al. *ApJ*, 566:93–102, 2002.
- H. Böhringer, P. Schuecker, and L. Guzzo et al. *A&A*, 425:367–383, 2004.

- J. R. Bond, S. Cole, and G. Efstathiou *et al.* *Astrophys. J.*, 379:440, 1991.
- J. R. Bond, G. Efstathiou, and M. Tegmark. *MNRAS*, 291:L33, 1997.
- S. Borgani. *Astrophys. J.*, 561:13, 2001.
- S. Borgani, G. Murante, and V. Springer *et al.* *MNRAS*, 348:1078–1096, 2004.
- H. A. Borges and S. Carneiro. *Gen.Rel.Gravit*, 38:1385, 2005.
- G. L. Bryan and M. Norman. *Astrophys. J.*, 80:495, 1998.
- T. Buchert. *Gen.Rel.Gravit*, 33:1381, 2001.
- M. Busha, A. Evrard, and F.C. Adams *et al.* *MNRAS*, 363:L11–L15, 2005.
- S. Camera, T. Kitching, and A. F. Heavens *et al.* *arXiv:1002.4704v1 [astro-ph.CO]*, 2010.
- J. Carroll. *Spacetime and geometry: An introduction to general relativity*. Addison Wesley, 2003.
- J. Carroll, W. Press, and E. Turner. *Annu. Rev. Astron. Astrophys.*, 30:499–542, 1992.
- S. M. Carroll, D. Duvvuri, M. Trodden, and M. Turner. *Rev.Mod.Phys.*, 70:1–5, 2004.
- A. Cavaliere and R. Fusco-Femiano. *A&A*, 49:137, 1976.
- S. Chandrasekhar. *Rev. Mod. Phys.*, 15:1, 1943.
- D. Clowe, A. Gonzalez, and Maxim M. Markevitch. *Astrophys. J.*, 604:596–603, 2004.
- S. Cole, W. J. Percival, and J. A. Peacock *et al.* *MNRAS*, 362:505–534, 2005.
- P. Coles and B. Jones. *MNRAS*, 248:1, 1991.
- C. Collins, L. Guzzo, and H. Böhringer *et al.* *MNRAS*, 319:939–948, 2000.
- A. Cooray. *MNRAS*, 365:842–866, 2006.
- A. Cooray and R. Sheth. *Phys. Rep.*, 372:1–129, 2002.
- E. Copeland, M. Sami, and S. Tsujikawa. *Int.J.Mod.Phys.*, D15:1753–1936, 2006.
- M. Crocce and R. Scoccimarro. *Phys.Rev.D*, 73:635, 2006.
- M. Crocce and R. Scoccimarro. *Phys.Rev.D*, 73:23533, 2008.
- M. Crocce, P. Fosalba, J. Castander, and E. Gazta naga. *MNRAS*, 403:1353, 2010.
- H. J. Croston, G. W. Pratt, and H. Böhringer *et al.* *A&A*, 487:431–443, 2008.
- W. Cui, L. Liu, and X. Yang *et al.* *Astrophys. J.*, 687:738–744, 2008.
- P. Valageas D. Munshi and L. van Waerbeke *et al.* *Phys. Rep.*, 462:67–121, 2008.
- A. Dekel and O. Lahav. *Astrophys. J.*, 520:24, 1998.
- J. M. Dickey and F. J. Lockman. *ARA&A*, 28:215, 1993.
- S. Dodelson. *Modern Cosmology*. Academic Press, 2004.
- A. Dressler. *Astrophys. J.*, 236:351–365, 1980.
- A. Duffy, J. Schaye, and T. Kay *et al.* *MNRAS*, 390:64, 2008.
- H. Ebeling and S. Maddox. *MNRAS*, 275:1155–1159, 1995.
- H. Ebeling, A. C. Edge, and H. Böhringer *et al.* *MNRAS*, 301:881–914, 1998.
- G. Efstathiou, M. Davis, C. Frenk, and S. D. M. White *et al.* *Astrophys. J. supp*, 57:241, 1985.
- G. Efstathiou, C. Frenk, and S. D. M. White *et al.* *MNRAS*, 235:715, 1988.
- J. Einasto. *Trudy Inst. Astrofiz. Alma-Ata*, 5:87, 1965.
- J. Einasto, M. Einasto, and S. Gottlöber *et al.* *Nature*, 385:139, 1997.
- J. Einasto, M. Einasto, and E. Tago *et al.* *Astrophys. J.*, 519:441–455, 1999.

- D. Eisenstein and W. Hu. *Astrophys. J.*, 496:605, 1998.
- D. J. Eisenstein, I. Zehavi, and D. W. Hogg *et al.* *Astrophys. J.*, 633:560–574, 2005.
- V. R. Eke, S. Cole, and C. Frenk. *MNRAS*, 282:263–280, 1996.
- S. Ettori, A. Morandi, P. Tozzi, I. Balestra, S. Borgani, P. Rosati, L. Lovisari, and F. Terenziani. *A&A*, 501:61–73, 2009.
- A. E. Evrard, T. MacFarland, and H. M. P. Couchman *et al.* *Astrophys. J.*, 573:7, 2002.
- H. Feldman, N. Kaiser., and J. A. Peacock. *Astrophys. J.*, 426:23–27, 1994.
- J. E. Felten. *AJ*, 82:861–878, 1977.
- A. Finoguenov, T. J. Ponman, and J. P. F. Osmond *et al.* *MNRAS*, 374:737–760, 2007.
- K. Fisher, C. A. Scharf, and O. Lahav. *MNRAS*, 266:219, 1996.
- J. A. Frieman and E. Gaztanaga. *Astrophys. J.*, 425:392, 1994.
- M. Frigo and S. Johnson. *Proceedings of the IEEE*, 93:216–231, 2008.
- J. N. Fry. *Astrophys. J.*, 461:L65, 1996.
- J. N. Fry and E. Gaztanaga. *Astrophys. J.*, 413:447, 1993.
- P. E. Garnavich, S. Jha, and P. Challis *et al.* *Astrophys. J.*, 509:74–79, 1998.
- E. Gaztanaga, A. Cabre, and L. Hui. *preprint arXiv:0807.3551v5*, 2008.
- C. Giocoli, M. Bartelmann, and R. K. Sheth *et al.* *arXiv:1003.4740v1 [astro-ph.CO]*, 2010.
- S. Giodini, D. Pierini, and A. Finoguenov *et al.* *Astrophys. J.*, 703:982–993, 2009.
- K. Glazebrook and C. Blake. *Astrophys. J.*, 631:1–20, 2005.
- W. Greiner and B. Muller. *Gauge Theory of Weak Interactions*. Springer Verlag, 1989.
- L. Guzzo, M. Pierleoni, and B. Meneux *et al.* *Nature*, 451:541–545, 2008.
- L. Guzzo, P. Schuecker, and H. Böhringer *et al.* *AA*, 499:357–369, 2009.
- A. Hamilton. *Astrophys. J.*, 417:19–35, 1993.
- A. Hamilton. *MNRAS*, 257:257, 2000.
- A. Hamilton. *LNP*, 665:415–431, 2005.
- A. J. S. Hamilton. *Astrophys. J.*, 385:1, 1992.
- A. J. S. Hamilton. *arXiv:astro-ph/9708102v2*, page 1, 1998.
- A. J. S. Hamilton and M. Culhane. *MNRAS*, 278:73H, 1996.
- A. J. S. Hamilton and M. Tegmark. *MNRAS*, 2000.
- A. J. S. Hamilton, P. Kumar, and E. Lu *et al.* *Astrophys. J.*, 374, 1991.
- S. H. Hansen. *MNRAS*, 352:L41–L43, 2004.
- W. G. Hartley, L. Gazzola, F. R. Pierce, S. T. Kay, and P. A. Thomas. *MNRAS*, 305:2015–2021, 2008.
- E. Hawkins, S. Maddox, and S. Cole *et al.* *MNRAS*, 346:79, 2003.
- D. J. Heath. *MNRAS*, 179:351–358, 1977.
- A. Heavens. *arXiv:0906.0664v3 [astro-ph.CO]*, 2009.
- A. Heavens and J. A. Peacock. *MNRAS*, 243:133–143, 1990.
- A. Heavens and A. Taylor. *MNRAS*, 275:483–497, 1994.
- A. Heavens and A. Taylor. *MNRAS*, 290:456, 1997.
- A. Heavens, S. Matarrese, and L. Verde. *MNRAS*, 301:797–808, 1998.
- R. Hockney and J. W. Eastwood. *Computer simulation using particles*. McGraw-Hill Book Company, 1981.

- D. W. Hogg. *arXiv:astro-ph/9905116v4*, 1999.
- W. Hu and N. Sugiyama. *Astrophys. J.*, 471:542–570, 1996.
- K. Huang. *Statistical Mechanics*. John Wiley & Sons, 1987.
- G. Hütsi. *A&A*, pages 1–7, 2007.
- G. Hütsi. *MNRAS*, 401:2477–2489, 2010.
- G. Hütsi and O. Lahav. *A&A*, 492:355–365, 2008.
- B. Jain and E. Bertschinger. *Astrophys. J.*, 431:495–505, 1994.
- A. Jenkins, C. Frenk, and S. D. M. White *et al.* *MNRAS*, 312:372, 2001.
- D. Jeong and E. Komatsu. *Astrophys. J.*, 691:569–595, 2009.
- Y. P. Jing and Y. Suto.
- Y. P. Jing, H. J. Mo, and G. Brner. *Astrophys. J.*, 564:1–15, 2002.
- N. Kaiser. *Astrophys. J.*, 284:L9–L12, 1984.
- N. Kaiser. *MNRAS*, 227:1–21, 1987.
- A. Y. Kamenshchik, U. Moschella, and V. Pasquier. *Phys.Lett. B*, 511:265, 2001.
- G. Kauffmann, J. M. Colberg, and A. Diaferio *et al.* *MNRAS*, 303:182, 1999.
- I. Kayo, A. Taruya, and Y. Suto. *Astrophys. J.*, 561:22–34, 2001.
- M. Kerscher, M. Mecke, and P. Schuecker *et. al.* *A&A*, 377:1, 2001.
- T. Kitching, A. F. Heavens, and A. Taylor *et al.* *MNRAS*, 376:771–778, 2008.
- A. A. Klypin and I. Kopylov. *Sov. Astron. Lett.*, 9:41–44, 1983.
- B. P. Koester, T. A. McKay, and J. Annis *et al.* *ApJ*, 660:239–255, 2007.
- E. Komatsu, J. Dunkley, and M.R. Nolta *et al.* *Astrophys. J. Supp.*, 180:330–376, 2009.
- E. Komatsu, K.M Smith, and J. Dunkley. *arXiv:1001.4538v2 [astro-ph.CO]*, 180, 2010.
- M. Kowalski, D. Rubin, and G. Aldering *et al.* *ApJ*, 686:749–778, 2008.
- C. Lacey and S. Cole. *MNRAS*, 262:627–649, 1993.
- O. Lahav. *MNRAS*, 333:961, 2002.
- O. Lahav and Y. Suto. *Living Reviws in relativity*, 7:8, 2004.
- S. Landy. *Astrophys. J.*, 567:L1–L4, 2002.
- S. D. Landy and A. S. Szalay. *Astrophys. J.*, 412:64, 1993.
- G. Lemson and G. Kauffmann. *MNRAS*, 302:111, 1999.
- A. Lewis and S. Bridle. *Phys. Rev. D.*, 103511, 2002.
- Antony Lewis, Anthony Challinor, and Anthony Lasenby. *Astrophys. J.*, 538:473–476, 2000.
- A. R. Liddle and D. H. Lyth. *Cosmological inflation and Large Scale Structure*. Cambridge University Press, 2000.
- H. Lin, R. P. Kirshner, and S. A. Schectman *et al.* *Astrophys. J.*, 471:617–633, 1996.
- A. Linder. *Astropart.Phys.*, 29:336–339, 2008.
- M. Longair. *Formation of galaxies*. Springer, 2005.
- Z. Lukic, D. Reed, S. Habib, and K. Heitman. *Astrophys. J.*, 692:217, 2010.
- A. Maccio, A. Duttin, and *et al.* *arXiv:0805.1926v2 [astro-ph]*, 2008.
- N. Makino, S. Sasaki, and Y. Suto. *Astrophys. J.*, 497:555–558, 1998a.
- N. Makino, S. Sasaki, and Y. Suto. *Astrophys. J.*, 497:555–558, 1998b.

- R. Mandelbaum, U. Seljak, and T. Boltauf et al. *MNRAS*, 405:2078–2010, 2010.
- M. Manera and E. Gaztanaga. *arXiv:0912.0446v2 [astro-ph.CO]*, 2009.
- M. Manera and D. F. Mota. *MNRAS*, 371:1373–1380, 2006.
- A. Mantz, S. Allen, H. Ebeling, D. Rapetti, and A. Drlica-Wagner. *MNRAS*, 406:1773–1795, 2010.
- S. Matarrese, L. Verde, and A. Heavens. *MNRAS*, 290:651–662, 1997.
- B. J. Maughan. *Astrophys. J.*, 668:772, 2007.
- A. Meiksin and M. White. *MNRAS*, 308:1179, 1999.
- A. Meiksin, M. White, and J. A. Peacock. *MNRAS*, 304:851, 1999.
- R. Mewe, E.H.B.M. Gronenschild, and G.H.J. van den Oord. *A&AS*, 62:197, 1985.
- C. Miller and D. Batuski. *ApJ*, 551:625–642, 2001.
- C. Miller, R. Nichol, and D. Batuski. *ApJ*, 555:68–73, 2001.
- H. J. Mo and S. D. M. White. *MNRAS*, 282:347–361, 1996.
- F. Montesano, A. G. Sánchez, and S. Phleps. *MNRAS*, 408:2397, 2010.
- B. Moore, F. Governato, and .Quinn et al. T. *Astrophys. J.*, 499:L5–L8, 1998.
- L. Moscardini, S. Materrese, and F. Lucchini et al. *MNRAS*, 316:283–298, 2000.
- J. S. Mulchaey. *Annu. Rev. Astron. Astrophys.*, 38:289–335, 2000.
- J. S. Navarro, C.S Frenk, and S. D. M. White. *Astrophys. J.*, 462:563, 1996.
- J. S. Navarro, A. Ludlow, and V. Springel et al. *arXiv:0810.1522v2 [astro-ph]*, 2008.
- M. C. Neyrink, A. J. S. Hamilton, and N. Y. Gnedin. *MNRAS*, 362:337–348, 2005.
- H. Nishioka and K. Yamamoto. *asto-ph/9902218*, 1999.
- K. Nock, W. J. Percival, and A. Ross. *MNRAS*, 2010.
- P. Norberg, C. Baugh, and E. Hawkins et al. *MNRAS*, 462:563, 2001.
- A. Ortiz-Gil, L. Guzzo, and P. Schuecker et al. *MNRAS*, 348:325, 2004.
- P.J. Outram, T. Shanks, and B. J. Boyle et al. *MNRAS*, 348:745–752, 2004.
- T. D. Kitching P. G. Castro, A. F. Heavens. *Phys.Rev.D*, 72:023516, 2005.
- T. Padmanabhan. *High energy astrophysics*. Cambridge University Press, 1993a.
- T. Padmanabhan. *Structure Formation in the Universe*. Cambridge University Press, 1993b.
- J. A. Peacock. *Cosmological Physics*. Cambridge University Press, 1999.
- J. A. Peacock and S. J. Dodds. *MNRAS*, 267:1020, 1994.
- J. A. Peacock and M. J. West. *MNRAS*, 259:494–504, 1992.
- P. J. E. Peebles. *Astrophys. J.*, 185:413–440, 1973.
- P. J. E. Peebles. *The Large-Scale structure of the Universe*. Princeton University Press, 1980.
- P. J. E. Peebles and B. Ratra. *Rev.Mod.Phys.*, 75:559–606, 2003.
- W. J. Percival. *AA*, 3637, 2005.
- W. J. Percival, W. Southerland, and J. A. Peacock et al. *MNRAS*, 337:1068, 2002.
- W. J. Percival, D. Buerkey, and A. Heavens et al. *MNRAS*, 353:1201–1220, 2004a.
- W. J. Percival, L. Verde, and J. A. Peacock. *MNRAS*, 347:645–653, 2004b.
- W. J. Percival, R. C. Nichol, and D. J. Eisenstein et al. *Astrophys. J.*, 657:645–663, 2007a.
- W. J. Percival, S.Cole, and D. J. Eisenstein. *MNRAS*, 381:1053–1066, 2007b.

- W. J. Percival, B. A. Reid, and D. J. Eisenstein *et al.* *MNRAS*, 401:2148–2168, 2010.
- S. Perlmutter, G. Aldering, and M. della Valle *et al.* *Nature*, 391:51, 1998.
- S. Perlmutter, G. Aldering, and G. Goldhaber *et al.* *Astrophys. J.*, 517:565, 1999.
- S. Phleps, C. Wolf, and J. A. Peacock *et al.* *A&A*, 426:23–27, 2006.
- A. Pillepich, C. Porciani, and O. Hahn. *MNRAS*, 402:191–206, 2010.
- T. Plagge, B. A. Benson, and P. A. R. Ade *et al.* *Astrophys. J.*, 716:1118–1135, 2010.
- C. Porciani, S. Matarrese, and F. Luccini. *MNRAS*, 298:1097, 1998.
- G. Pratt, H. J. Croston, and M. Arnaud *et al.* *arXiv:0809.3784v2*, 2009.
- H. Press, S. Teukolski, W. Vetterling, and B. Flannery. *Numerical Recipes in C*. Cambridge University Press, 2002.
- W. H. Press and P. Schechter. *Astrophys. J.*, 187:425, 1974.
- E. Puchwein, D. Sijacki, and V. Springel. *Astrophys. J.*, 687:L53, 2008.
- B. A. Reid, W. Percival, and D. J. Eisenstein *et al.* *MNRAS*, 404:60, 2009.
- F. Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill Book Company, 1965.
- T. Reiprich and H. Böhringer. *Astrophys. J.*, 567:716–740, 2002a.
- T. Reiprich and H. Böhringer. *Astrophys. J.*, 567:716–740, 2002b.
- J. Retzlaff, S. Borgani, and S. Gottlöber *et al.* *NewA*, 3:631–646, 1998.
- A. G. Riess, L. G. Strolger, and J. Tonly *et al.* *Astrophys. J.*, 607:665, 2004.
- C. Rimes and A. J. S. Hamilton. *MNRAS*, 371:1205–1215, 2006.
- A. Ross, W. J. Percival, and R. J. Brunner. *MNRAS*, 2010.
- E. Rozo, S. Dodelson, and J. Friemann. *Phys. Rev D*, 70:1–16, 2004.
- A. G. Sánchez, P. Schuecker, and H. Böhringer. *MNRAS*, 362:1225–1232, 2005.
- A. G. Sánchez, C. Baugh, and *et al.* *MNRAS*, 366:189, 2006.
- A. G. Sánchez, C. M. Baugh, and R. Angulo. *MNRAS*, 4:1470–1490, 2008.
- A. G. Sánchez, C. M. Baugh, and E. Gaztanaga *et al.* *MNRAS*, 400:1643–1664, 2009.
- M. A. Sánchez-Conde, J. Betancort-Rijo, and F. Prada. *MNRAS*, 378:339–352, 2007.
- C. L. Sarazin. *X-ray Emission from Clusters of Galaxies*. Cambridge University Press, Cambridge, 1988.
- P. Schuecker. *astro-ph/052234v1*, 2005.
- P. Schuecker, H. Böhringer, and L. Guzzo *et al.* *AA*, 425:367–383, 2001.
- P. Schuecker, H. Böhringer, and C. Collins *et al.* *AA*, 398:867–878, 2003a.
- P. Schuecker, H. Böhringer, and C. Collins *et al.* *A&A*, 398:867–878, 2003b.
- R. Scoccimarro. *Phys. Rev. D*, 70:1–19, 2004.
- R. Scoccimarro and J. Frieman. *Astrophys. J.*, 105:37–73, 1996.
- U. Seljak and M. Zaldarriaga. *Astrophys. J.*, 496:437–444, 1996.
- H. J. Seo, E. R. Siegel, and D. J. Eisenstein *et al.* *Astrophys. J.*, 686:13–24, 2008.
- R. K. Sheth. *MNRAS*, 364:796, 2005.
- R. K. Sheth and B. Jain. *MNRAS*, 345:529–538, 2003.
- R. K. Sheth and G. Tormen. *MNRAS*, 308:119, 1999.
- R. K. Sheth, H. J. Mo, and G. Tormen. *MNRAS*, 323:1–12, 2001.
- R. K. Sheth, A. J. Connolly, and R. Skibba. *arXiv:astro-ph/0511773v1*, 2005.

- J. Silk. *Astrophys. J. Supp*, 151:459–471, 1968.
- R. Skibba and R. Sheth. *MNRAS*, 392:1080–1091, 2009.
- R. Skibba, R. K. Sheth, and A. Conlolly et al. *MNRAS*, 369:69–76, 2006.
- R. Smith. *arXiv:0810.1960v2*, 2008.
- R. Smith, J. A. Peacock, and A. Jenkins et al. *MNRAS*, 341:1311–1332, 2003.
- R. Smith, R. Scoccimarro, and R. K. Sheth. *Phys Rev D*, 2007.
- R. Smith, R. Scoccimarro, and R. K. Sheth. *Phys Rev D*, 77:043525, 2008.
- R. E. Smith, P. I. R. Watts, and R. K. Sheth. *MNRAS*, 365:214–230, 2006.
- D. N. Spergel, R. Bean, and O. Dorea et al. *Astrophys. J.*, 170:377–408, 2007.
- R. Stanek, A. E. Evrard, and H. Böhringer et al. *Astrophys. J.*, 648:956–968, 2006.
- R. Stanek, E. Rasia, A. E. Evrard, F. Pierce, and L. Gazzola. *ApJ*, 715:1508, 2010.
- R. Sunyaev and Y. A. Zeldovich. *Astrophysics and Space Sciences*, 7:3–19, 1980.
- Y. Suto, H. Magira, and K. Yamamoto. *PASJ*, 52:249–257, 2000.
- A. S. Szalay, T. Matsubara, and S. Landy. *Astrophys. J.*, 498:L1–L14, 1998.
- H. Tadros and G. Efstathiou. *MNRAS*, 282:1381–1396, 1996.
- H. Tadros, G. Efstathiou, and G. Dalton. *MNRAS*, 296:995–1003, 1998.
- H. Tadros, W. E. Ballinger, and A. N. Taylor. *MNRAS*, 305:527–546, 1999.
- R. Takayashi, N. Yoshida, and M. Takada. *arXiv:0902.0371v1 [astro-ph.CO]*, 2008.
- A. Taruya and Y. Suto. *Astrophys. J.*, 542:559–577, 2000.
- M. Tegmark. *Phys. Rev. Lett*, 79:3806, 1997.
- M. Tegmark and P. J. E. Peebles. *Astrophys. J.*, 500:79, 1998.
- M. Tegmark, M. R. Blanton, and M. A. Strauss et al. *Astrophys. J.*, 606:702, 2004.
- J. Tinker. *MNRAS*, 374:477–492, 2007.
- J. Tinker and A. Wetzel. *arXiv:0909.1325v1 [astro-ph.CO]*, 2009.
- J. Tinker, D. Weinberg, and Z. Zheng et al. *ApJ*, 631:41–58, 2005.
- J. Tinker, A. Kravtsov, and A. Klypin A. et al. *arXiv:0803.2706v1 [astro-ph]*, 2008.
- J. Tinker, B. Robertson, and A. Kravtsov et al. *arXiv:1001.3162v2 [astro-ph.CO]*, 2010.
- J. Truemper. *Science*, 260:1769–1771, 1993.
- C. Tsallis. *Introduction to non extensive statistical mechanics*. Springer Verlag, New York, 2009.
- F. C. van den Bosch, X. H. Yang, and H. J. Mo. *MNRAS*, 352:1302, 2003.
- F. C. van den Bosch, X. H. Yang, and H. J. Mo et al. *MNRAS*, 376:841–860, 2007.
- L. Verde. *arXiv:0911.3105v1 [astro-ph.CO]*, 2009.
- L. Verde and A. Heavens. *Astrophys. J.*, 553:14–24, 2001.
- A. Vikhlinin, R. A. Burenin, and H. Ebeling et al. *Astrophys. J.*, 692:1033–1059, 2009a.
- A. Vikhlinin, A. V. Kravtsov, and R. A. Burenin et al. *Astrophys. J.*, 692:1060–1074, 2009b.
- M. Vogelsberger and S. D. M. White. *arXiv:1002.3162v1 [astro-ph.CO]*, pages 1–17, 2009.
- W. Voges, B. Aschenbach, and Th. Boller et al. *A&A*, 349:389–405, 1999.
- C. Wagner, V. Mueller, and M. Steinmetz. *A&A*, 487:63–74, 2008.
- D. Wake, R. K. Sheth, and R. C. Nichol et al. *MNRAS*, 387:1045–1062, 2008.

- Y. Wang. *Astrophys. J.*, 536:531, 2000.
- M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro. *Astrophys. J.*, 646:881, 2006.
- S. Weinberg. *Rev. Mod. Phys.*, 61, 1989.
- C. Wetterich. *Physics Lett B.*, 594:17, 2001.
- M. White. *AA*, pages 37–32, 2001.
- M. White. *Astrophys. J.*, 143:241–255, 2002.
- M. White and N. Padmanabhan. *MNRAS*, 395:2381–2384, 2009.
- M. White, Z. Zheng Z., and M. Brown *et al.* *Astrophys. J.*, 655:L69–L72, 2007a.
- M. White, Z. Zheng, and M. J. I. Brown *et al.* *Astrophys. J.*, 655:L69–L72, 2007b.
- S. D. M. White and C. Frenk. *Astrophys. J.*, 391:52, 1991.
- S. D. M. White and M. Vogelsberger. *MNRAS*, 392:281–286, 2009.
- S. D. M. White, M. Davies, and C. Frenk. *MNRAS*, 209:27P–31P, 1984.
- H. Y. Wu, R. Andrew, and R. H. Wechsler. *Astrophys. J.*, 713:856, 2010.
- H. W. Wyld. *Annals of Physics*, 14:143–165, 1964.
- X. Xu, M. White, and N. Padmanaban *et al.* *arXiv:1001.2324v3 [astro-ph.CO]*, 2010.
- K. Yamamoto and H. Nishioka. *asto-ph/9908006*, 1999.
- X. H. Yang, H. J. Mo, and F. C. van den Bosch. *MNRAS*, 339:1057, 2003.
- A. Zandivares, A. G. Abadi, and D. Lambas. *MNRAS*, 326:147–154, 2001.
- H. Zhan, L. Knox, and J. A. Tyson *et al.* *Astrophys. J.*, 640:8–17, 2006.
- Z. Zheng, A. Berlind, and D. Weinberg *et al.* *Astrophys. J.*, 633:791–809, 2005.
- F. Zwicky. *Helv. Phys. Acta* 6:110, 1933.

Acknowledgments

Contrary to what is written in the cover, this is not the work of one man, for undoubtedly many people were involved helping at their best for me to finish this thesis. I apologize beforehand if I look now ungrateful with those who took part on it and whose names will not appear in these ending credits.

I am deeply grateful with my wife Pilar and my daughter Celeste. Not only they embodied a reason to carry on and finish my Ph.D, but also a reason to be tied to this planet. Their support was vital when the strength was low and the future looked uncertain (more than usual), and dedicate this thesis to them is just a tiny approach to thank them fairly.

How can I thank Peter Schücker for having given me the chance, whatever the reason he had in mind was, to start my Ph.D at MPE?. I wish I could have met him much better, instead of still being wondering “what if?”. His words of wisdom, applicable both for life and astronomy, guided me to be in the position of being typing the last pages of this thesis. Many things in this thesis were possible after having read his papers, and I still need to read them again and again.

I am indebted with Prof. Ralph Bender, who gave me the chance to finish my Ph.D at MPE. I acknowledge the Ph.D fellowship of the International Max Planck Research School. and the economical support provided by the DFG cluster of excellence “Origin and Structure of the Universe”. I thank Raúl Angulo and Carlton Baugh for providing us with the L-BASICC II simulations. This work contains observational data obtained at the ESO La Silla observatory. We wish to thank the ESO Team for their support. I thank Francesco Montesano and Holger Schlagenhauer for providing me with measurements of matter power spectrum and correlation function.

Hans Böhringer opened the gates to get out from the (interesting though) maze I started my Ph.D with, by letting me be involved in the clustering analysis of the REFLEX II catalogue. He allowed me to work, interact and learn from great scientist and human beings like Luigi Guzzo, Chris Collins and of course, Hans himself. They supported my work with words of encouragement and optimism, even when the work was pretty raw. I have found in Hans a good example of what an astronomer is, and I thank him for having introduced me to the exciting subject of galaxy clusters.

Such escape from that maze wouldn't be possible without the advent of Ariel Sánchez to MPE. I deeply thank Ariel for his help, advice and patience with this cosmology apprentice that I am, and for having shared his experience in the large-scale structure business with me.

I thank Stephanie Phleps for her invaluable support during these four years, not only in academic side of this period. Many things were possible due to her support, constant

encouragement and advice.

Special thanks go to Payel Das and Hourii Ziaepour for having spent time reading a couple of chapters of the manuscript. I thank again Payel for having helped me to improve my English, my German, and my communication skills.

I thank Boris and Tanya for the friendly and lovely kaotish atmosphere they provided, amazingly overcoming the barriers of language and making me feel by their side as if I were at home. I thank Ximena Mazzalay for providing me with some cigarrets and sending me all the offers she got to fly with the A380. Thanks to Eva and Inma for ... being what they are. I thank Michael and Yaneth for having hosted me during the final months of this work. Life is easier when one has friends like you all. To you all my gratitude.

I thank our parents, relatives and friends who supported us with their thoughts, prayers, tithes and other expressions of solidarity and love that traveled thousand miles to fill us with the necessary strength to stand a foreign way of life. I feel kind-of blessed by the providence for having placed me in such a fabulous institute as MPE. It is hard still for me consider myself an astronomer (mainly when no title I have obtained has that word written in it), but the fact of having been closer to astronomical data and have tried to do science with it, places me closer to that childhood's dream. To all those who made it possible, vielen Dank.

Andrés Balaguera Antolínez
Garching
September 29th 2010

En memoria de mi padre Freddy Enrique Balaguera Barrera (12.02.1951-20.10.2010)

Curriculum Vitae

Andrés Balaguera Antolínez

Birth date October 2nd 1977
Place of birth Bogotá D.C., Colombia
Nationality Colombian

Education

1983-1988 Concentración Distrital el Jazmin (Primary school)
1989-1995 Instituto Técnico Salesiano de Cundinamarca (High School)
1996-2002 Studies of physics at Universidad Nacional de Colombia.
2003-2006 Master in physics at Universidad de los Andes
2006-2010 Ph.D student at Max Planck Institute for extraterrestrial physics
2010- Post-doctoral researcher at Argelander Institute for Astronomy, Bonn

Publications

- A. Balaguera-Antolínez, A. Sánchez, H. Böhringer et al., *The REFLEX II Galaxy Cluster Survey: power spectrum analysis*, submitted to MNRAS.
- A. Balaguera-Antolínez, A. Sánchez, H. Böhringer et al., *Cosmological constraints from the REFLEXII power spectrum*, in preparation.
- A. Sánchez, A. Balaguera-Antolínez, H. Böhringer et al., *The REFLEX II Galaxy Cluster Survey: mock catalogs*, in preparation.
- A. Sánchez, A. Balaguera-Antolínez, H. Böhringer et al., *The REFLEX II Galaxy Cluster Survey: correlation function*, in preparation.
- A. Balaguera-Antolínez, Mota, D., F., Nowakowski, M., 2007, MNRAS 382, 621-640
- A. Balaguera-Antolínez, M. Nowakowski, 2007, Class. Quant. Grav.24, 2677-2688
- A. Balaguera-Antolínez, Mota, D., F., Nowakowski, M., 2006, Class. Quant.Grav. 23, 4497-4510
- A. Balaguera-Antolínez, M. Nowakowski, Astron. & Astrophys. 2005, 441, 23-29

- A. Balaguera-Antolínez, M. Nowakowski, C. Böhmer, 2006, *Class. Quant. Grav.*23, 485-496
- A. Balaguera-Antolínez, M. Nowakowski, C. Böhmer, 2005, *Int. J. Mod. Phys. D*14 (2005)

Attended meetings and schools

- 2010: Congreso Colombiano de Astronomía. IV Escuela Colombiana de Astronomía, Bogotá, Colombia.
- 2010: Third Science Week of the Excellence Cluster, October 2009, Garching, Germany.
- 2009: Annual Fall Meeting and 82nd General Assembly of the AG (German astronomical society), Potsdam, Germany.
- 2008: SIGRAV School in cosmology, GGI, Firenze, Italy.
- 2008: Leopoldina Conference of Dark Energy, Munich, Germany.
- 2007: TransRegio Winter School on Astrophysics, Passo del Tonaly, Italy.
- 2006: International Max Planck School in Astrophysics, Garching, Germany.
- 2005: 11th Latin American Regional IAU Meeting, Pucón, Chile.
- 2005: Latin American Astronomy Summer School. Santiago de Chile, Chile.
- 2005: Congreso Nacional de Física. Barranquilla, Colombia.
- 2005: III Escuela Colombiana de Astronomía. Bogotá, Colombia.
- 2005: Cosmology Meeting. Granada, Spain.
- 1999: Congreso Nacional de Física. Bogotá, Colombia.