UNIVERSITY COLLEGE LONDON

Faculty of Mathematics and Physical Sciences

Department of Physics & Astronomy

TIPPING SCALES IN GALAXY SURVEYS: STAR/GALAXY SEPARATION AND SCALE-DEPENDENT BIAS

Thesis submitted for the Degree of Doctor of Philosophy of the University of London

by

Maayane Soumagnac

Supervisors:
Filipe Abdalla
Ofer Lahav

Examiners:
Alan Heavens
Benjamin Joachimi

January 15, 2015

To Yael, Olivier, Tal and Sivane Soumagnac, my North stars,

and to the memory of Henia Ziv Suchar and Irmgard Klar.

I, Maayane Tamar Soumagnac, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

In the first part of this thesis, we address the problem of separating stars from galaxies in future large photometric surveys. We derive the science requirements on star/galaxy separation, for measurement of the cosmological parameters with the Gravitational Weak Lensing and Large Scale Structure probes, in chapter 2. We formulate the requirements in terms of the *completeness* and *purity* provided by a given star/galaxy classifier. In order to achieve these requirements, we propose a new method for star/galaxy separation in chapter 3, combining Principal Component Analysis with an Artificial Neural Network. When tested on simulations of the Dark Energy Survey (DES), this multi-parameter approach improves upon purely morphometric classifiers (such as the classifier implemented in SExtractor), especially at faint magnitudes. Chapter 4 is dedicated to the testing of this tool on real data, namely the recent internal release of DES Science Verification data.

In the second part and last chapter of this thesis, chapter 5, we develop a method to detect the modulation by Baryonic Acoustic Oscillations of the density ratio of baryon to dark matter across large regions of the Universe. Such a detection would provide a direct measurement of a difference in the large-scale clustering of mass and light and a confirmation of the standard cosmological paradigm from a different angle than any other measurement. We measure the number density correlation function and the luminosity weighted correlation function of the DR10 releases of the Baryon Oscillation Spectroscopic Survey (BOSS), and fit a model of scale dependent bias to our measurement. Although our measurement is compatible with previous theoretical predictions, more accurate data is needed to prove or disprove this effect.

Acknowledgements

Firstly, I would like to thank my two supervisors, Filipe Abdalla and Ofer Lahav, thanks to whom I discovered a whole new world during the last three years. Thank you Filipe for giving me the confidence that pragmatic solutions can often be found, no matter what obstacles stand in the way. Thank you Ofer for sharing your wisdom and vision, and for all the precious teachings that I will never find in any book. Thanks to the two of you for allowing me to pursue a research project in Israel during my PhD. I feel very lucky to have been your student and I look forward to many more years of collaboration and learning from you.

Thank you to Hiranya Peiris, for wise advice throughout this PHD and to all the Group A cosmologists at UCL, in particular to a group of inspiring Post Doctoral Fellows who helped me with precious explanations and advice: Manda Banerji, Donnacha Kirk, Stephanie Jouvel, Sreekumar Balan, Barnaby Rowe and Cristiano Sabiu.

I would like to thank the DES collaboration, the most exciting and inspiring place that I have discovered during the last three years. Thanks to Will Percival and Gary Bernstein for taking the time to give me precious advice and corrections. Thanks to the team of the Cerro Tololo observatory and to Jim Annis, for sharing the secrets of the Blanco telescope and conveying the excitement of observational cosmology. Thank you to my collaborators on my first DES project: Ignacio Sevilla, Emmanuel Bertin, Luis Da Costa, Enrique Gaztanaga, Mike Jarvis, Brian Yanny, Basilio Santiago and Risa Wechsler. Thanks to a group of inspiring DES and BOSS early career scientists: Joe Zuntz, Molly Swanson, Michael Busha, Ashley Ross, Sarah Hansen, Nic Ross and Carlos Cunha, for sharing your insights and advice on this new world.

Thank you to Demitri Muna and the whole SciCoder team at New York University,

who introduced me to Pyhton and gave me precious tools for research.

Thank you Avi Loeb and Rennan Barkanna, for allowing me to work on a great research project in Tel Aviv University during my last year, and to the whole Tel Aviv University Astrophysics department for hosting me. Thank you to Amir Harel, Gili Brown, Uzi and Lili Shemesh, the Kainan, Koren, Harel and Sucar families for your support during my time in Israel and your unwavering belief that I will one day be a doctor.

I was very lucky to learn from great teachers on the journey who led me to this PhD, and would like to remember especially Mr Philippe Fatus for high level astronomy lessons in primary school, Monique Guichard for late night patient explanations, Mr David, Mme Stoliaroff, Mme Cretual-Meyer, Mr Label, Mr Barbet Massin, Mr Lesgourgues, Mr Taillet, Mr Salati, Mr Zuber, Uri Elhanany and the clarinetist Basil Bratos, for believing and encouraging me during decisive years.

Thank you to the people who introduced me to research during my master degree, especially to Sebastien Chevrot from The Observatoire Midi Pyrennee, Bruce Lipschultz and Matthew Reinke from the Plasma Science and Fusion Center at MIT, Julien Devriendt, Yohan Dubois and Renee Hlozek from the Beecroft Institute of Particle Astrophysics and Cosmology in Oxford University.

Thank you to my London flatmate Joe, and to my Cambridge friends and adoptive families, who made England a happy place for me. Thank you Dr Ohad Kammar for sharing your flat and PhD tips. Thank you Reuven and Rochel Leigh, Elisheva and Yisrael Malkiel, for opening your warm houses to late night challenging discussions. Thank you Arathi for the morning runs and for helping me with the English in this thesis. Thank you Sarah, for special library sessions.

Thank you Oz for the laughs, the help and the support during the last sprint, for opening up new books and horizons, and for believing.

Thanks to my inspiring and beloved brother Tal. Thanks to my sister, Sivane, for your loving support and example, since we were little and during the last three years. I am looking forward to many other happy moments with both of you. Thanks to Yonathan and Noamotek for adding their star to our family.

And to my parents, Papa and Ima, the biggest “thank you” of them all, for giving me the opportunity to study and teaching me the value of it. Thank you for the love, for the home, for such a happy childhood and for giving me the strength to go discover the world.

אוֹדֶה לָאֵל לֵבָב חוֹקֵר בְּרָן-יַחַד כּוֹכְבֵי בֹּקֶר

I will thank G.d, who knows my heart, joining the singing of the morning stars.

R. Shmaya Kosson, 16^{th} *century Jewish poem.*

Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilization, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar," every "supreme leader," every saint and sinner in the history of our species lived there-on a mote of dust suspended in a sunbeam.

It has been said that astronomy is a humbling and character-building experience. There is perhaps no better demonstration of the folly of human conceits than this distant image of our tiny world. To me, it underscores our responsibility to deal more kindly with one another, and to preserve and cherish the pale blue dot, the only home we've ever known.

Carl Sagan, *Pale Blue Dot.*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"There will come a time when our descendants will be amazed that we did not know things that are so plain to them [...] Many discoveries are reserved for ages still to come, when memory of us will have been effaced. Our universe is a sorry little affair unless it has in it something for every age to investigate [...] Nature does not reveal her mysteries once and for all."

Seneca, Natural Questions Book 7, c., 1^{st} century

"And although people call this " the Big Bang", it was not only "big", it was "everywhere", and it was not really an explosion, it was space stretching out. It's really quite unfortunate that "the Everywhere Stretches" is not nearly as catchy as "the Big Bang"."

Henry Reich, MinutePhysics, 2013

Cosmology refers to the study of the Universe as a whole. It asks some of the most audacious questions human beings have ever been able to formulate: what is our world? what is our Universe made of? How did it start and where is it headed? Is there one or many Universes? Such audacity has led many, throughout history, to echo Sophocles thought: "Astronomy? Impossible to understand and madness to investigate."[1], or to accuse astronomers of pure heresy.

[1] Sophocles, c. 420 BCE

Cosmology has a long history in the pages of both philosophy and religion. If not the most ancient science of them all, cosmology is the most ancient motivation for the study of astronomy, which is practiced in every ancient civilization and is the only practical branch of natural philosophy to be included in the quadrivium of the medieval universities, together with arithmetic, geometry and music.

For Plato and, under his influence, medieval philosophers and theologians, the desire to understand the physical world is an instrumental step that should lead to spiritual and metaphysical progression of the individual, "for every one, as I think, must see that astronomy compels the soul to look upwards and leads us from this world to another."[2] Ptolemy, one of the most influential Greek astronomers of antiquity, perhaps best expresses the mingling of astronomy and mysticism in the ancient world: "Mortal as I am, I know that I am born for a day, but when I follow the serried multitude of the stars in their circular course, my feet no longer touch the earth. I ascend to Zeus himself to feast me on ambrosia, the food of the gods."[3] This close link between theology and astronomy is still very present in medieval philosophy, notably in the work of Maimonides, the preeminent Spanish Jewish Thorah scholar, philosopher, astronomer, and physician. Maimonides considered the study of science a necessary condition to metaphysical elevation. He saw in it a way to approach and embrace the Devine, to be "near the truth [...] in the palace in which the king lives" [4] and at the same time, a humbling way to appreciate "how he [the human being] is a tiny, lowly, and dark creature"[5].

Nowadays philosophy and astronomy occupy distinct and separate areas of research, but cosmology, through the very nature of the questions it asks, remains one of their meeting points.

In this introductory chapter, I provide some of the theoretical frameworks necessary to understand the work presented in this thesis.

The first part (section 1.1) presents the observations of the expanding Universe, which led to our current model: the ΛCDM paradigm. This "concordance cosmology" successfully explains all our current observations, while ironically pointing at the fact that we ignore what most of the Universe is made of. As explained in section 1.1.7, this model can

[2]Plato's republic, Book VII, CUP, 529a.

[3]Ptolemy, Anthol. Palat., ix, 577.

[4]Maimonides, the Guide for Perplexed, Part III, chapter LI, translated into English by Michael Friedlnder.

[5]Maimonide, Mishne Torah, Sefer Ha-Madda, Hilkhot Yesodey Ha-Torah, chapter II, translated into English by Eliyahu Touger.

be decomposed into two different problems: the background problem (section 1.2) and the inhomogeneity problem (section 1.3).

The 20^{th} century has been extremely fruitful for cosmology. It started with very little understanding of the expansion of the Universe, before the works by Einstein and Hubble, and ended with a mind blowing discovery: that not only is the Universe expanding, but it is doing so in an accelerated way. Explaining this acceleration - either with a new "dark energy" which would drive it, or by re-thinking our entire model - is one of the biggest challenges faced by cosmology in the next decades. The last section of this introduction (section 1.4) is dedicated to the different approaches adopted to probe the accelerated expansion and uncover the nature of dark energy. The detection of Baryon Acoustic Oscillations and Weak Gravitational lensing, in the first decade of our century, are only two examples of how exciting our time is for cosmology.

1.1 It's an expanding world

1.1.1 The cosmological principle

Our description of the Universe is based on the assumption that the Universe is homogeneous and isotropic when smoothed on sufficiently large scales. This central assumption is called the *cosmological principle* and has a deep philosophical meaning. At the very source of it lies a statement about the way we perceive ourselves in the Universe, which has changed throughout cultures and history. Rudnicki (1995) gives examples of different cosmological principles which were stated throughout Human history:

- **The ancient Indian cosmological principle:** *The Universe is infinite in space and time and is infinitely heterogeneous.*
- **The ancient Greek cosmological principle:** *Our Earth is the natural center of the Universe.*
- **The Copernican cosmological principle:** *The Universe as observed from any planet looks much the same.*
- **The perfect cosmological principle:** *The Universe is (roughly) homogeneous and isotropic in both time and space*

- **The (generalised) cosmological principle:** *The Universe is (roughly) homogeneous in space and time, and is isotropic in space only, i.e. anisotropic in time.*
- **The anthropic principle:** *A human being, as he/she is, can exist only in the Universe as it is.*

Until the middle ages, the Universe was thought to be organized around human kind, whereas the modern view suggests that there should be nothing special about the area of the galaxy where we live. This statement can be generalised to any region in space and formulated as an assumption about the whole Universe. Edward Arthur Milne formulated it in the following way: the Universe as a whole should be homogeneous, with no privileged point playing a particular role. The perfect cosmological model, which is more symmetric than the generalised one since it is isotropic in both time and space, led to the steady state model, which was rejected on observational grounds in favor of the Big Bang theory (section 1.1.4).

Like any other model about the physical world, this one cannot be proved, but only falsified. The methodology adopted by modern cosmology is to assume the cosmological principle, and to test for consistency with the observations. If the data proves the homogeneity hypothesis wrong, then there are other possible inhomogeneous models, which in turn need to be tested with the observations.

In all this thesis, the cosmological principle we will be referring to is the generalized one (see the list above).

1.1.2 Early observational cosmology

Two complementary discoveries were made possible at the beginning of the 20^{th} century, by the progress of observational astronomy: the discovery of the galactic structure - i.e. the fact that the Universe is populated with galaxies separated by large voids - and the discovery of the expansion of the Universe. They where both the results of a long process which started as soon as the 18^{th} century and where both made possible by the observation of the first spectral lines, at the beginning of the 19^{th} century, and the discovery of the Doppler Effect in 1842.

Spectral lines

Elements and their atoms have characteristic energy levels governed by quantum theory. An emission or absorption of a discrete and specific packet of light energy leads to a transition between these energy levels. This definite energy has a fixed corresponding wavelength known for each element and their energy levels in the rest frame. However, for a moving object, such as a galaxy containing the element, this wavelength or *spectral line* will be Doppler shifted, towards one end of the spectrum.

Doppler Effect

In 1842, Johann Christian Doppler argued that if an observer receives a wave emitted by a body in motion, the wavelength measured will be shifted proportionally to the relative speed of the emitting body (projected along the line of sight).

$$\Delta\lambda/\lambda = \vec{v}.\vec{n}/c \ , \tag{1.1}$$

where c is the celerity of the wave. Doppler suggested that this effect could be observable for sound waves, and maybe also for light. The later assumption was checked experimentally in 1868 by Sir William Huggins, who found that the spectral lines of some neighboring stars were slightly shifted toward the red or blue ends of the spectrum. So, it was possible to know the projection along the line of sight of star velocities, v_r, using the redshift z defined as

$$z \equiv \Delta\lambda/\lambda = v_r/c \ , \tag{1.2}$$

where c is the speed of light.

Early intuitions in favor of expansion

At the beginning of the 20^{th} century, progress in instrumentation allowed to go from measurements of the redshift of stars, to measurements of the redshift of fainter and diffuse objects, listed under the name of *nebulae*, which where to become what we now call galaxies. The first measurements, performed on the brightest nebulae, indicated some arbitrary distribution of red and blue shifts, just like for stars. But with more observations, it appeared that the statistics were biased in favor of red shifts, suggesting that a majority

of nebulae, unlike stars, were going away from us.

Observations by Leavitt, Shappley and Hubble brought the first confirmation of the galactic structure of the Universe: the nebulae were likely to be some distant replicas of the Milky Way, and these galaxies were separated by large voids. In this new picture of the Universe, the "local" scales, populated by stars where to be distinguished from the much larger galactic scales.

This new picture of the galactic structure, together with the fact that most distant galaxies where redshifted, led to the intuition that on the largest observable scales, the Universe was expanding. At first, this idea was not widely accepted because it seemed to conflict with the modern version of the cosmological principle. Indeed, in the most general case, a given dynamic of expansion takes place around a center: observing the Universe expanding around us seemed to be an evidence for the existence of a center, very close to our own galaxy, so how could the observed expansion of the Universe be compatible with the absence of center? The answer to this apparent paradox was to be given by the Hubble's law.

1.1.3 The Hubble's law

In fact, the expansion of the Universe does not imply the existence of a center, and is not necessarily contradictory with the cosmological principle. The homogeneity of the Universe is compatible either with a *static* distribution of galaxies, or with a very specific velocity field, obeying a linear distribution: $\vec{v} = H\vec{r}$. Along the line of sight, this translates into what was later called the Hubble's law:

$$v = H_0 d \, . \tag{1.3}$$

Vesto Melvin Slipher and then Edwin Hubble - who gave his name to the proportionality constant, H_0 - tried to check this idea. In order to measure distances, they used a particular type of variable stars, cepheids, whose intrinsic luminosity follows a well-known trend which allows to infer their distance from the observer. Such objects are called "standard candles". In 1929, Hubble published a work based on eighteen galaxies. Although this work later appeared to be quite imprecise and biased (Hubble did not use regular cepheids), this experiment was reconfirmed by many others (see figure 1.1), and is

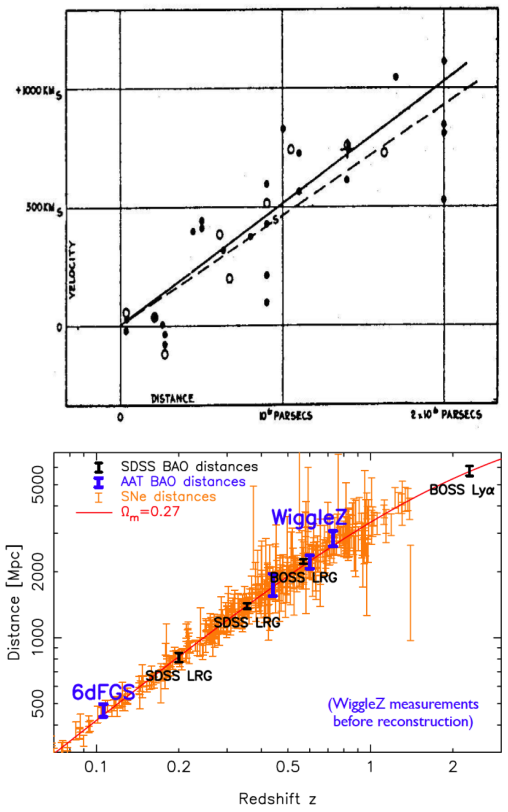

Figure 1.1: Top: The first Hubble diagram, showing the proportionality between distance and redshift. The solid line and points correspond to an analysis corrected for the suns movement. This initial diagram used 18 galaxies, and the standard-candles property of cepheids. Credit: Hubble (1929). Bottom: a recent Hubble diagram, combining a standard candles method (SNe) with other probes, such as BAOs. Credit: Blake et al. (2013)

considered as the starting point of observational cosmology.

Several comments should be made about Hubble's law:

- it only applies on large scales: nearby galaxies possess random motions known as "peculiar velocities", which do not verify Hubble's law;
- it assumes that cepheids are standard candles. The idea that supernovae at different distances, in different galaxies act in the same way is not obvious. We see in section 1.4.2 that the nuance between standard candles and standardizable objects is a complicated one.

The latest results from the Planck collaboration suggest a Hubble constant of $H_0 = 67.4 \pm 1.4 km/s/Mpc$ (Planck Collaboration XVI 2013), which is lower than previous measurements from the Hubble space telescope $H_0 = 70.1 \pm 1.3 kms^{-1}Mpc^{-1}$ (Riess

et al. 2009) where $1pc \approx 3 \cdot 10^{16}m$. Cosmologists often work in terms of h, where $H_0 = 100\ h\ kms^{-1}Mpc^{-1}$.

This relatively simple relation has far-reaching consequences. It was the first clue towards some initial conditions of the Universe, known as the "Hot Big Bang" and gave the beginning of an answer to one of the most audacious questions posed by cosmology: the age of the Universe (1.1.4). It led cosmologists to define the expansion rate, the scale factor $a(t)$, and to reformulate both distance and light emission times of the objects we observe.

1.1.4 The Hot Big Bang picture and the "age" of the Universe

If the Universe is expanding, we can in principle rewind the cosmic history and find some time at which the Universe was arbitrarily small, hot and dense. In fact, the natural conclusion of the expansion is that all the observable Universe we can see was at one point in time, more or less at one point in space. Lemaitre called this initial condition the "primeval atom". The commonly used term of "Big Bang" was introduced in 1949 by Fred Hoyle - who ironically remained a defender of the steady state theory - in a popular radio broadcast, to describe the origin of the Universe from some hot dense initial condition.

The term "Big Bang" is misleading for several reasons. It seems to imply that the Universe *started* at an initial *singularity* in space. In fact, our current Friedmann description (see section 1.2) of the early evolution of the Universe is not supposed to hold until $a(t) = 0$. It starts at the Planck time $t_p = \sqrt{G\bar{h}/c^5} = 5.39 \cdot 10^{-44}s$, which defines the scale at which current physical theories fail. On this scale, the density reaches a critical value called the Planck density, the entire geometry of space-time as predicted by general relativity (see section 1.2.1) breaks down, and as yet an undiscovered theory that combines general relativity and quantum mechanics - where the classical notions of time and space disappear - is needed to describe the laws of physics. Candidates for such theories exist, mainly in the framework of "string theories".

Another common misconception about the Big Bang is the fact it implies that the whole Universe was confined into a singular point. It is true that the *observable* Universe, (the portion of the whole Universe which we can see from Earth, i.e. which is accessible to us via causality) was shrunk into a very small portion of space with a characteristic size equal to the planck length ($l_p = \sqrt{G\bar{h}/c^3} = 1.62 \cdot 10^{-35}m$), which was not a single point, nor was the rest of the Universe confined in the same piece of space. In fact, it is possible

that the Universe was infinite in space, while the observable Universe was shrunk to the size of the Planck length.

This being said, Hubbles discovery was the first clue towards some initial conditions of the Universe we know: the one accurately described by our theory of gravity. These initial conditions, called "Hot Big Bang", are a point in time where the whole Universe was much more dense and hotter than today, and where the observable Universe was shrunk to a small (yet non-singular) location in space. The age of the Universe is defined as the time elapsed since the Big Bang i.e. since these initial conditions. Current measurements put the time to these initial conditions at 13.798 ± 0.037 billion years (Planck Collaboration I 2013).

1.1.5 The scale factor

To measure the expansion rate of the Universe, cosmologists use the scale factor $a(t)$, which links any "real", "physical" distance between two objects, x, with *comoving* coordinates, χ, which are carried along with the expansion:

$$x = a(t)\chi \ . \tag{1.4}$$

The Hubble's law (equation 1.3) can then be reformulated as

$$H(t) = \frac{\dot{a}(t)}{a(t)} \ , \tag{1.5}$$

the Hubble constant H_0 being the value of the Hubble's parameter $H(t)$ at present time, t_0.

1.1.6 Distance measures in an expanding Universe

Within the Friedmann-Lemtaitre-Robertson-Walker (FLRW) metric (the simplest solution for a homogeneous, isotropic, non-stationary Universe, which we present in section 1.2.2), infinitesimal physical distances dl are given by the scale factor $a(t)$ time the comoving line element,

$$dl_{phys}(t) = a(t)dl \ . \tag{1.6}$$

Three-dimensional spaces with constant non-zero curvature fall in two categories: 3-spheres and 3-hyperboloids. A convenient choice of polar coordinate leads to the following expression for the line elements in such spaces

$$dl^2 = \left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\Phi)\right] , \tag{1.7}$$

where k is a constant number, related to the spatial curvature, as will be detailed in section 1.2.2. Distances on macroscopic scales are obtained by integrating the line element of equation 1.6.

Let us assume that we are sitting in the origin of a spherical coordinate system, at time t_0, and that we observe a comoving object, i.e. an object whose movement is only determined by the expansion of space. What is our distance to this object? This question is ambiguous for two reasons. First, do we mean, in units of today (i.e. the distance between our position *today* and the position of the object *today*) or in units of the time when the object emitted? Secondly, not only has the object moved away from us since it emitted the light we receive, but what then was one unit of distance has now expanded! To solve the second question, we distinguish the *comoving* distance from the *physical* distance.

Comoving distance and the problem of distance measurement in an expanding Universe

The comoving coordinates are carried away with the expansion of the Universe, whereas the physical coordinates are not. By construction, the comoving distance between two comoving objects does not depend on time, unlike the physical distance between them.

A comoving object emitting light at $(t_e, r_e, \theta_e, \Phi_e)$, is now located at $(t_0, r_e, \theta_e, \Phi_e)$. If by *distance* we mean the distance computed on the constant-time hyper-surface with $t = t_0$, i.e. the distance as defined today, then it is given by

$$d = \int_0^{r_e} dl = a(t_0) \int_0^{r_e} \frac{dr}{\sqrt{1-kr^2}} . \tag{1.8}$$

The scale factor $a(t_0) \equiv a_0$ is usually defined to be 1, and the distance above coincides with the comoving distance

$$\chi(r_e) \equiv \int_0^{r_e} \frac{dr}{\sqrt{1-kr^2}} = \int_{t_0}^{t_e} \frac{cdt}{a(t)} , \tag{1.9}$$

which can be integrated to

$$\chi(r) = \begin{cases} \sin^{-1}(r) & \text{if } k = +1 \\ r & \text{if } k = 0 \\ \sinh^{-1}(r) & \text{if } k = -1 \end{cases} \tag{1.10}$$

Hence, it is useful to define the function f_k

$$f_k(x) = \begin{cases} \sin(x) & \text{if } k + 1 \\ x & \text{if } k = 0 \\ \sinh(x) & \text{if } k - 1 \end{cases} \tag{1.11}$$

which defines the relation between the comoving distance χ and the comoving coordinate r, $r = f_k(\chi)$. We will see an important use of this function in section 1.2.2.

The comoving distance χ to a comoving object emitting light at $(t_e, r_e, \theta_e, \Phi_e)$, when the scale factor was $a(t_e) \equiv a_e$, can therefore be expressed as a function of time, redshift, and scale factor as follows:

$$\chi = \int_{t_e}^{t_0} \frac{cdt}{a(t)} = \int_{a_e}^{a_0} \frac{cda}{a^2 H(a)} = \int_0^{z_e} \frac{cdz}{a_0 H(z)} \, . \tag{1.12}$$

The relation between r_e and t_e depends on the function $a(t)$ and the curvature k: the knowledge of k and $a(t)$ would allow to explicitly integrate equation 1.12. The comoving distance is a well defined notion, up to a choice of marginalization of $a(t)$. But in practice distances are not directly measurable quantities in cosmology. Astronomers have developed indirect experimental ways to probe them, each of them leading to a particular definition of distance.

Distances are usually measured in two ways:

1. from the angular diameter of standard rulers;
2. from the luminosity of standard candles.

Redshift

Because of the expansion of the Universe, light emitted by some source is seen by an observer to be shifted to a longer, or redder wavelength. This redshift can be related to

the expansion history, i.e. the scale factors at emission and observation:

$$1 + z = \frac{\lambda_{obs}}{\lambda_{emitted}} = \frac{a(t_0)}{a(t)} \tag{1.13}$$

An object "at redshift z" emitted the light we see when the scale factor was equal to the fraction $\frac{1}{1+z}$ of its value today, $a(t_0)$. If one has previous knowledge of the expansion history (i.e. the function a(t)), then the redshift gives us an information about the time t_e it emitted the light we observe, and subsequently about its comoving distance, by integrating $cdt/a(t)$ between t_0 and t_e.

This is the method used to infer the spatial distribution if galaxies from galaxy redshift surveys. It assumes an *apriori* knowledge of the function $a(t)$ which is often precisely what one wants to measure.

The relation between the redshift and the time of emission depends on our knowledge of the expansion history, $a(t)$.

Angular diameter distance

Surprisingly, there exist objects in the Universe of which we know the size in advance, due to some of their physical properties. Such objects are called *standard rulers*. This is the case of the Baryon Acoustic oscillation scale which we will introduce in section 1.4.4. In an Euclidian space, the distance d to an object on a sphere can be inferred from its angular diameter $d\theta$ and its physical diameter dl: $d \cdot d\theta = dl$. Even if the geometry is not Euclidian, we can adopt this as one possible definition of the distance: the "angular diameter distance" d_A is defined as

$$d_A = \frac{dl}{d\theta} \, . \tag{1.14}$$

From equation 1.6, we see that the size of an object perpendicular to the line of sight is related to its angular size $d\theta$ via

$$dl_{phys}(t_e) = a(t_e) r_e d\theta \, , \tag{1.15}$$

where t_e is the time when the object emitted the light which we observe, and r_e is the comoving coordinate of the object along the line of sight. This allows us to infer the

relation between the angular diameter distance d_A and the redshift:

$$d_A = \frac{dl}{d\theta} = a(t_e)r_e = \frac{a(t_0)}{1+z}r_e = \frac{a(t_0)}{1+z}f_k(\chi) \ , \tag{1.16}$$

which in turn can be expressed as a function of a, z or t with equation 1.12. The redhsift-angular diameter distance relation therefore depends on both the expansion history $a(t)$ and the curvature k.

If we know the physical size of an object (a standard ruler), we can infer information on the geometry of the Universe by measuring its redshift and its angular diameter.

Luminosity distance

Standard candles are objects of which the absolute luminosity can be physically estimated, independently to their distance and apparent luminosity (flux density, energy per unit of time and surface). For an isotropically emitting source, the energy is spread evenly over the surface of a sphere and so the observed flux f (energy per unit of time) is related to the distance of the source d, through:

$$f = \frac{L}{4\pi d^2} \ . \tag{1.17}$$

This only applies in euclidian space, but one can adopt this relation as one possible definition of distance, called the *luminosity distance*:

$$d_L = \sqrt{\frac{L}{4\pi f}} \ . \tag{1.18}$$

In euclidian space, in the absence of expansion and curvature, d_L would simply correspond to the distance to the source. In the FLRW space-time, things are more complex, and the relation between the absolute flux and the apparent luminosity is rather:

$$l = \frac{L}{4\pi a^2(t_0)r_e^2(1+z)^2} \ , \tag{1.19}$$

leading to

$$d_L = a(t_0)r_e(1+z) = a(t_0)f_k(\chi)(1+z) \ . \tag{1.20}$$

Let us explain this relation. What is always true is that the apparent luminosity (the

energy per unit of time and surface) is given by the absolute luminosity (the energy per unit of time) divided by the surface: $l = L/\mathcal{S}$. Both the numerator and the denominator take particular forms in the case of a curved expanding space. The surface over which the photons emitted at r_e are distributed today, is obtained by integrating over the infinitesimal surface element $dS^2 = a^2(t_0)r_e^2 \sin\theta d\theta d\Phi$, which gives $\mathcal{S} = 4\pi a^2(t_0)r_e^2$. The energy carried by each photons is also affected by the expansion, twice. Indeed, the energy carried by a photon is inversely proportional to its physical wavelength λ, and therefore to $a(t_e)$: the energy of each photon has therefore been divided by $(1+z)$ since it was emitted (and multiplied by $a(t_0)$). Moreover, the fact that λ scales like $(1+z)$ not only implies that the energy scales like $(1+z)^{-1}$, but also that the frequency at which the observer receives each photon (i.e. the photons per unit of time) scales like $(1+z)^{-1}$, which explains the second factor $(1+z)^{-1}$ in equation 1.19.

The knowledge of $a(t)$, the curvature and other cosmological parameters allows to calculate explicitly $r_e(t_e(z))$, and therefore the redshift-luminosity distance relation. In section 1.4, we show how this relation allows to constrain dark energy.

Limit of small redshifts

$$d_L = a(t_0)r_e(1+z) = a(t_e)(1+z)^2 r_e = d_A(1+z)^2 \ . \tag{1.21}$$

In the limit of small redshift, the three definitions of distance $\chi = a(t_0)\chi$, d_A and d_L are equivalent and reduce to the definition of distance in Euclidian space, related to the redshift through $d = z/H_o$.

1.1.7 Perturbed expansion: the background problem and the inhomogeneities problem

In order to describe the observed Universe, cosmologists had to build a model. The two ingredients of the standard cosmological model are:

- the cosmological principle (which we presented in section 1.1.1 of this introduction);
- a fully viable and working theory describing gravity, General Relativity (GR) (see section 1.2.1).

Most calculations and predictions in cosmology are obtained taking a perturbation theory approach, i.e. assuming that the exact description of the Universe can be decomposed

into two distinct problems:

- **the background problem**, independent and self-consistent. The inhomogeneities such as stars and galaxies are *smoothed*, i.e. averaged over a scale bigger than the largest distance between them. The Universe is assumed to be a smooth distribution of matter: an idealized "cosmological fluid". The goal of the background problem is to compute the evolution of the cosmological fluid (see section 1.2).

- **the inhomogeneity problem**. It consists in writing first-order (i.e. linear) perturbations within the given smooth background and solve for their evolution. This approach can be pushed to second order (quadratic) perturbations, but then the equations become very hard to solve. The goal is to understand e.g. the growth of Large Scale Structures (LSS) or the Cosmic Microwave Background (CMB) anisotropies.

How good is this description? Figure 1.2 shows the spatial distribution of galaxy positions to a comoving distance of 600Mpc. The galaxies are far from being randomly distributed: they are arranged into large voids and dense filaments, forming the delicate *cosmic web*. The homogeneous isotropic approach does not describe these large structures, neither does it work for the description of small scale structures or fully non-linear problems (such as the merging of two galaxies, which is better described by a newtonian non-linear approach).

In fact, the perturbed expansion provides an excellent description of the Universe at present time on large scales (i.e. when we smooth the picture on scales of $\approx 100 Mpc$) and on all scales at early times: we know from the Cosmic Microwave Background (CMB) that the Universe was smooth to around 1 part to 10^5 at the time of recombination.

1.2 The homogeneous, isotropic Universe and the ΛCDM paradigm

1.2.1 General Relativity

General Relativity provides a fully viable and working theory describing gravity, the most pervasive and important force in our observable Universe. The main paradigm shift introduced by GR is that gravitation is not considered a force, or a field, but is rather formulated as the curvature of space-time, sourced by matter.

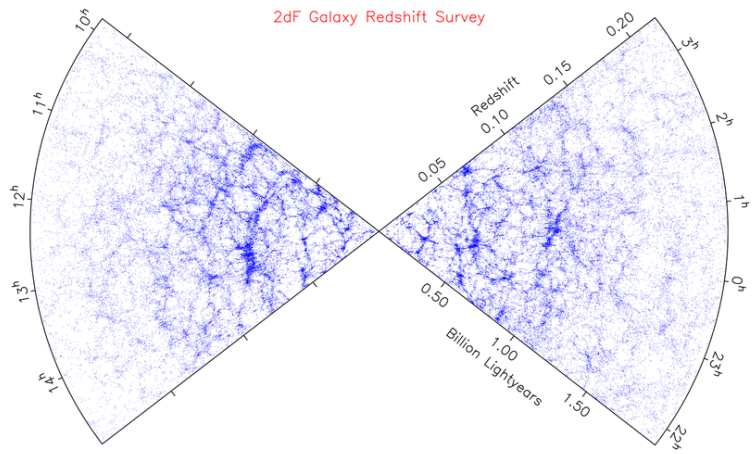

Figure 1.2: Slices through the 3D map of galaxy positions from the 2dF galaxy redshift survey. Credit: 2dF, (Colless et al. 2001).

General Relativity is partially motivated by the *equivalence principle*, which was already suggested by Galileo when he asserted that the inertial and gravitational masses of an object are identical: the effect of a gravitational field is locally equivalent to the effect of acceleration in the absence of a gravitational field. Or in other words: within a gravitational field, in each point of space, one can always find a frame within which the laws of physics are locally the same as in the absence of a gravitational field.

When Einstein tried to build a theory of gravity compatible with the invariance of the speed of light, the equivalence principle and the Newtonian limit, he found that the minimum price to pay was to abandon the principle of a gravitational potential related to the distribution of matter (and of which the gradient gives the gravitational field in any point) and to introduce the following key ideas.

1. **Our 4 dimensional space-time is curved.** In mathematical terms, the GR space-time is a pseudo-Riemanian variety: the line element ds^2 is related to any set of coordinated x^i via the *metric* $g_{\mu\nu}(x)$:

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \ . \tag{1.22}$$

2. **Free-falling objects describe geodesics in this space-time, which are bent by the curvature.** In mathematical terms, the trajectory of free-falling particles

obeys the geodesic equation:

$$\frac{d^2x^\mu}{d\tau^2} = \Gamma^\mu_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau} \ . \tag{1.23}$$

3. **The** *curvature* **of space-time in any point (as opposed to the** *gravitational potential***) is linked to the properties of matter at this point.** The relation between them is given by Einstein's equation:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \ , \tag{1.24}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the scalar curvature and $T_{\mu\nu}$ is the energy momentum tensor of the matter field.

In the GR framework, the cosmological principle has two extremely powerful consequences: one purely kinematic, on the metric of the space-time; and one on the shape of the energy momentum tensor of the matter field.

1.2.2 Kinematic consequences of the cosmological principle

Consequence on the metric : the FLRW metric

The simplest solution for a homogeneous, isotropic, non-stationary Universe is the Friedmann-Lemtaitre-Robertson-Walker (FLRW) metric, whose line element is

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + f_k^2(r)(d\theta^2 + \sin^2\theta d\Phi^2)] \ , \tag{1.25}$$

where $f_k(r)$ is the function we previously defined in section 1.11 and its possible forms are

$$f_k(r) = \sin(r), r, \sinh(r) \ , \tag{1.26}$$

(depending on the value of k:$\{+1, 0, -1\}$ respectively), r is the comoving coordinates and t is cosmic time. Via a simple change of coordinates, we can re-cast the FLRW metric into

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\Phi^2)\right] \ , \tag{1.27}$$

where k is related to the spatial curvature of the space and governs the geometrical properties of the 3-space corresponding to the hyper surfaces $t = constant$. In particular,

$k > 0$ corresponds to positive spatial curvature, i.e. a *closed* Universe where parallel light paths will converge; $k = 0$ corresponds to an ordinary flat Euclidian space, and $k < 0$ corresponds to a negative spatial curvature, i.e. an *open* Universe where parallel light paths will diverge.

Consequence on the shape of the energy momentum tensor of the matter field

The other implication of the homogeneous and isotropic assumption is that the energy momentum tensor of the matter field, $T_{\mu\nu}$ takes the form of an ideal fluid:

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu} = \begin{pmatrix} -\rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} , \tag{1.28}$$

where u_μ is the 4-velocity of the matter, ρ is the mean energy density and p is the mean pressure.

1.2.3 Dynamical consequences of the cosmological principle: from Einstein's equations to Friedman equations

It is worth reiterating that both the FLRW metric and the shape of $T_{\mu\nu}$ rely only on the assumptions of homogeneity and isotropy. They are mathematical paraphrase of the cosmological principle, within a purely kinematic description and do not assume a theory of gravity. in order to derive the dynamics of our Universe, i.e. the evolution of the scale factor $a(t)$, we need to apply Einsteins equations - which relate gravitating mass to local geometry - to the FLRW metric.

Applying equation 1.24 to the FLRW metric gives two independent equations, for the three independent functions $\rho(t)$, $p(t)$ and $a(t)$. Namely, the first Friedmann equation is

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} , \tag{1.29}$$

and the second Friedmann equation is

$$H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + \frac{3p}{c^2}) + \frac{\Lambda c^2}{3} . \tag{1.30}$$

Differentiating equation 1.29 with respect to t and then eliminating $\ddot{a}$ with equation 1.30, one can write another equation which can be interpreted as the first law of theremodynamics in the present context:

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p) \ . \tag{1.31}$$

Hence, the relation between $\dot{\rho}$ and the scale factor $a(t)$, i.e., the way in which the energy gets diluted during the expansion history, depends crucially on the pressure p, or, in other words, on the equation of state of the form $p = p(\rho)$, the additional independent equation required to solve for $a(t)$. In cosmology, the equation of state is usually assumed to be linear, i.e. it can be parametrized in the following simple form,

$$p = w\rho \ , \tag{1.32}$$

with w taking different values depending on the species. While the value of w may, in principle, change with redshift, it is often assumed, for simplicity, that z is independent of time. Then substituting this equation in equation 1.31 immediately gives

$$\rho \propto a^{-3(1+w)} \ . \tag{1.33}$$

Together with the first Friedmann equation, this gives the time evolution of the scale factor (assuming a linear equation of state)

$$a \propto t^{2/(3(1+w))} \ , \forall w \neq -1 \ . \tag{1.34}$$

The most important limiting cases are:

- **Non-relativistic matter (or dust):** $w = 0$**:** indeed, for objects with negligible velocities, $p = 0$ and equation 1.31 becomes $\rho \propto a^{-3}$.
- **Radiation or ultra-relativistic matter:** $w = 1/3$**:** in this case, equation 1.31 becomes $\rho \propto a^{-4}$. A ultra-relativistic fluid dilutes faster than a non-relativistic medium with the Universe expansion.
- **The cosmological constant:** $w = -1$.

The scaling of each constituent is summarised in table 1.1. If the Universe is made of

Component	w_i	$\rho(a)$	$a(t)$	type of expansion
Non-relativistic matter	0	$\propto a^{-3}$	$\propto t^{2/3}$	decelerated
radiation/relativistic matter	1/3	$\propto a^{-4}$	$\propto t^{1/2}$	decelerated
curvature	$-1/3$	$\propto a^{-2}$	$\propto t$	linear
cosmological constant	-1	$\propto a^{0}$	$\propto exp(Ht)$	exponnentially accelerated

Table 1.1: Scaling of each component in the Universe.

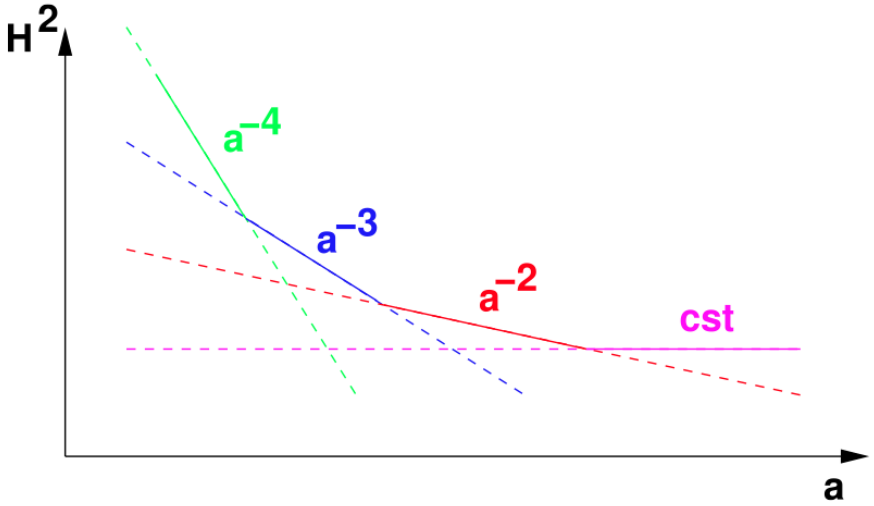

Figure 1.3: Evolution of the square of the Hubble parameter, in a scenario in which all typical contributions to the Universe expansion (radiation, matter, curvature, cosmological constant) dominate one after each other. Credit: lecture notes by Julien Lesgourgues (2009).

different fluid species, equation 1.31 still holds independently for each of the species, as long as they do not interact with each other. If the Universe consists of different fluid species with $w_i, i = 1, 2, ..., N$, and if one denotes the energy density of the i^{th} component at present time $\rho_{i,0}$, the total energy density then satisfies

$$\rho_{tot} = \sum_{i=1}^{N} \rho_{i,0} a^{-3(1+w_i)} \ , \tag{1.35}$$

where the present value of the scale factor a_0 is set to unity without loss of generality.

1.2.4 Various possible scenario for the history of the Universe

The Friedman equation 1.29 including the contribution from matter, radiation, curvature, and the cosmological constant yields:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho_R + \frac{8\pi G}{3}\rho_M - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \,, \tag{1.36}$$

where ρ_R is the radiation density and ρ_M is the matter density. Each of these contributions evolves respectively as a^{-4}, a^{-3}, a^{-2} and a^0. In the particular case where all these species are present, they will dominate the right hand side term one after the other (see figure 1.3). This does not mean that all the species have to be present, but any scenario which does not respect this order of successive dominations is excluded. During each stage, and assuming that one component strongly dominates over the other, the Universe dynamics is given by:

1. radiation domination: $a(t) \propto t^{1/2}$. Decelerated expansion.
2. matter domination: $a(t) \propto t^{2/3}$. Decelerated expansion, slower than during the radiation domination.
3. negative curvature domination: $a(t) \propto t$. Linear expansion.
4. cosmological constant domination $a(t) \propto \exp(\Lambda t/3)$. Exponentially accelerated expansion.

The Friedmann equation gives partial access to the past of the Universe. In each of these scenari, there seem to have been a time at which $a(t) \to 0$, leading to the initial conditions we described in section 1.1.4. The future of the Universe is also dictated by the Friedmann equation, and in particular, it is highly dependent on the value of the cosmological constant. If $\Lambda = 0$, the future evolution of the Universe is dictated by the curvature.

1.2.5 Cosmological parameters

It is useful to define a density parameter Ω, which depends on time and counts the energy density from all form of constituents in the Universe and is a ratio to the critical energy

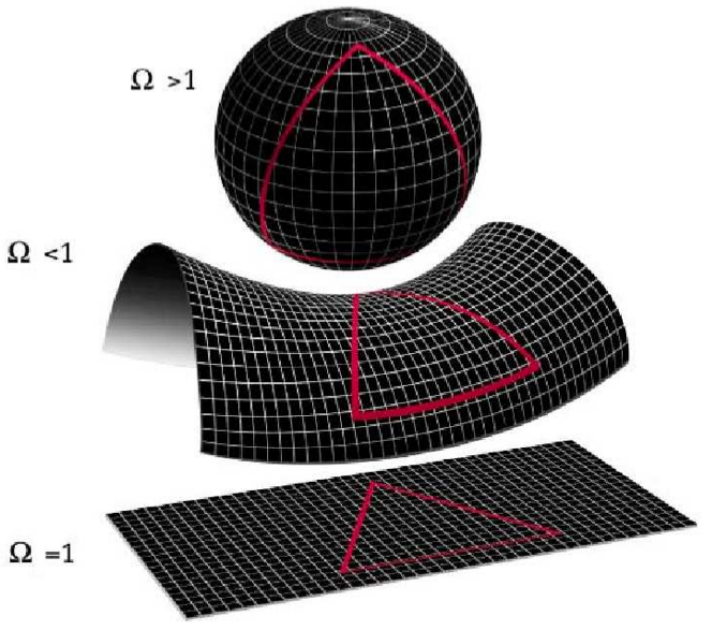

Figure 1.4: The three possible geometries for the Universe and their relation to the total density $\Omega = \Omega_M + \Omega_R + \Omega_\Lambda$. For $\Omega > 1$, $k > 0$, the Universe is closed (top). When the density is sub-critical $\Omega < 1$, $k < 1$, which corresponds to hyperbolic (open) space (middle). Finally, when $\Omega = 1$, $k = 0$, the Universe is said to be flat, which corresponds to a Euclidean, geometry (bottom). Also shown is the relation between an apparent angle and geometry. Credit: http://map.gsfc.nasa.gov/media.

density (i.e the density required for flat space geometry)

$$\Omega = \frac{8\pi G}{3H^2}\rho = \frac{\rho}{\rho_{crit}} \,, \tag{1.37}$$

where both the ρ and $\rho_{crit} = \frac{3H^2}{8\pi G}$ change with time. The Friedmann equation can then be rewritten as

$$\Omega(a) - 1 = \frac{kc^2}{H^2 a^2} \tag{1.38}$$

So the density parameter determine the curvature: $sign(\Omega - 1) = sign(k)$, and in particular:

$\rho < \rho_{crit} \Leftrightarrow \Omega < 1 \Leftrightarrow k < 0 \Leftrightarrow open$

$\rho = \rho_{crit} \Leftrightarrow \Omega = 1 \Leftrightarrow k = 0 \Leftrightarrow flat$

$\rho > \rho_{crit} \Leftrightarrow \Omega > 1 \Leftrightarrow k > 0 \Leftrightarrow close$

We recognize here one of the key ideas of GR: the curvature of the Universe $k = H_0^2(\Omega_M + \Omega_R + \Omega_\Lambda - 1)$ is entirely determined by its energy content. The density param-

eter tells us which of the three FLRW geometries describe our Universe. Often, $\Omega(a)$ is decomposed into energy density of different species - matter, radiation, and cosmological constant - $\Omega(a) = \sum_i \Omega_i(a)$, and is referred to as the *energy budget*. The contribution of the curvature can be introduced into the energy budget as a fictitious energy density, so that:

$$\Omega_i = \frac{8\pi G}{3H^2}\rho_i = \frac{\rho_i}{\rho_{crit}} \ , \tag{1.39}$$

$$\Omega_k = \frac{kc^2}{H^2a^2} = \frac{\rho_k}{\rho_{crit}} \ , \tag{1.40}$$

with $\rho_k = -\frac{3kc^2}{8\pi G a^2}$ and

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2} \ , \tag{1.41}$$

so that,

$$\sum_i \Omega_i + \Omega_k = 1 \quad \forall t \ , \tag{1.42}$$

Although this is true at any time, the cosmological parameters usually refer to the values of Ω_i at present time:

$$\Omega_i = \frac{8\pi G}{3H_0^2}\rho_{i,0} = \frac{\rho_{i,0}}{\rho_{crit}} \ , \tag{1.43}$$

$$\Omega_k = \frac{k}{H_0^2 a_0^2} \ , \tag{1.44}$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2} \ . \tag{1.45}$$

The Friedmann equation, describing the expansion history of the Universe, rewrites in terms of the cosmological parameters at present time

$$\left(\frac{H(a)}{H_0}\right)^2 = \Omega_M(a)^{-3} + \Omega_R a^{-4} + \Omega_k a^{-2} + \Omega_\Lambda \ \ \forall a > 0 \tag{1.46}$$

$$= \sum_i \Omega_{i,0} a^{-3(1+w_i)} + \Omega_{k,0} a^{-2} \ \ \forall a > 0 \ .$$

or equivalently, in terms of the redshift

$$\left(\frac{H(z)}{H_0}\right)^2 = \Omega_M(1+z)^3 + \Omega_R(1+z)^4 + \Omega_k(1+z)^2 + \Omega_\Lambda \ . \tag{1.47}$$

We will see in section 1.4 that this form of the Friedmann expansion is very useful to understand the different strategies adopted to probe dark energy.

The present time value of the energy budget Ω gives access to the entire history of cosmic expansion. This needs to be slightly nuanced, when we reach epochs where interactions allow interchanges between the densities of the different species, which is believed to have last happened at neutrino decoupling shortly before nucleosynthesis. So to probe further back into the Universes history requires assumptions about particle interactions, and perhaps about the nature of physical laws themselves. But broadly speaking, the entire cosmic history -past, present and future - can be described in terms of the value of four parameters at present time, which is extremely fortunate and is the reason why accurate measurement of the cosmological parameters is one of the holy grails of observational cosmology.

1.2.6 The content of our Universe

We have seen that the evolution of the cosmic expansion depends on the components which make up $T_{\mu\nu}$, or equivalently on the cosmological parameters Ω_i.

Recent measurements reveal an embarrassing fact: the "known" components, i.e. the components we have been able to detect so far (namely baryonic matter, radiation and neutrinos) account for less than 5% of the expanding Universe. Explaining the expansion without questioning the GR framework requires to add two "unknown" components, namely dark matter and dark energy into the energy budget of the Universe. In the current "concordance" model $\Omega_k \approx 0$, $\Omega_{b,0} \approx 0.04$, $\Omega_{cdm,0} \approx 0.23$, and $\Omega_{\Lambda,0} \approx 0.73$ for $w_\Lambda \approx -1$, the Universe is composed of about 4% baryons, 23% cold dark matter, and - within the framework of a cosmological constant model - 73% of dark energy.

Before looking at the problematic conclusions of such a budget (section 1.2.7), we detail in this section some properties of each of the components of $T_{\mu\nu}$. We start with the "dust" components, made of pressure-less non-relativistic particles (i.e. baryonic matter and cold dark matter) and continue with the ultra-relativistic contributions: radiation, and neutrinos, and will end with the dark energy.

Baryonic Matter

Baryonic matter is often referred to as "ordinary" matter: the matter that constitutes dust, stars, galaxies, ourselves, and all that surrounds us. Strictly speaking, baryons are particles made of three bounded quarks, such as neutrons and protons, but cosmologists use this term more loosely, including in it electrons and other leptons. The primordial distribution of baryons was produced during a process called *Big Bang Nucleosynthesis* and consisted of the following light elements: Hydrogen, Helium, Deuterium and Lithium.

The latest results from the Planck collaboration (Planck Collaboration XVI 2013) give $\Omega_b h^2 = 0.02207 \pm 0.00033$, with $H_0 = 67.4 \pm 1.4 km \cdot s^{-1} \cdot Mpc^{-1}$ meaning that the baryons account for about 4.9% of the total energy density.

Dark Matter

Dark matter (DM) is the name we give to non-baryonic particles interacting weakly (if at all) via electromagnetic force, which makes them invisible to our detectors. The baryonic candidate for DM, called MACHOs (for MAssive Compact Halos Objects), such as primordial black holes or brown dwarfs, are too few to account for the observed contribution of DM in the total energy density, which is the reason why DM is referred to as non-baryonic.

Historically, DM was introduced by Zwicky (1933) to explain the dynamics of galaxies. Indeed, it has been observed that a large amount of extra-matter was required to explain their hight rotation velocity. Whereas one would expect their velocity to vary like $r^{-1/2}$, where r is the radius from their galactic center, measurements showed that the velocity remains constant, which could be explained by the existence of a larger, invisible halo embedding the galaxy.

DM was introduced with a purely gravitational motivation, to explain this phenomenon which GR cannot explain. Instead of adding an unknown component, another approach consists of modifying our theory so that it does not require any unknown component. One of the most famous attempts to modify GR is the Modified Newtonian Dynamics (MOND) theory (Milgrom 1994), which proposes an acceleration scale a_0 below which Newtonian dynamics are modified in order to explain the rotation curves.

Perhaps the main "MOND killer" today is the data from the *bullet cluster* (Clowe et al. 2006), composed of two clusters who collided 150 million years ago. The collision

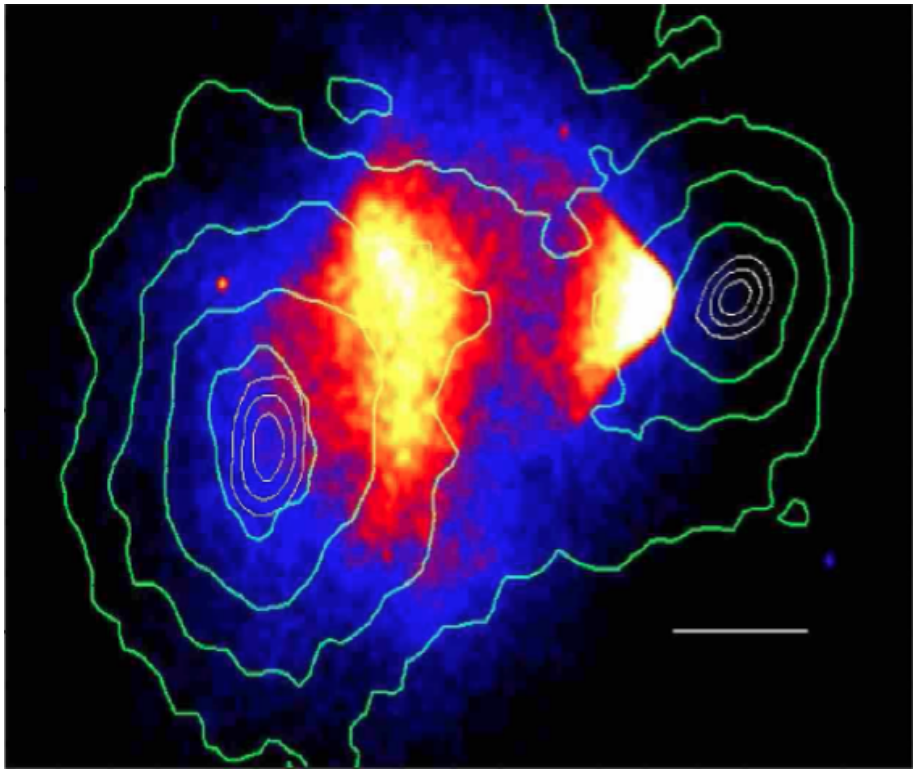

Figure 1.5: Chandra $500ks$ X-ray image of the "Bullet Cluster" $1E0657-56$. Green contours show the weak lensing κ mass reconstruction with the outer contour level at $\kappa = 0.16$ and increasing in steps of 0.07. There is definite offset between the X-ray emitting shocked ICG in the centre and the main mass concentrations which have passed through each other. Credit: Clowe et al. (2006)

induced a separation between the galaxies in the clusters and the Intra-cluster gas (ICG), whose cross-section is much higher. When mapping the concentration of mass around the galaxies (using gravitational lensing), the presence of extra-mass was measured which could not be baryonic, since the only baryonic candidate: the ICG, was left behind.

The project presented in chapter 5 of this thesis could also constitute a strong argument against the MOND theory, by detecting a difference between a tracer of baryonic matter (the luminosity-weighted correlation function) and a tracer of the total matter (the number density correlation function).

Cosmologists distinguish Cold Dark Matter (CDM) - relativistic particles which become non-relativistic at early times - from Hot and Warm Dark Matter particles (HDM and WDM) which remain relativistic until late. Neutrinos are an example of HDM and are basically the only DM detected particle today. However, current observations seem to point that the major part of DM needs to be on the form of CDM.

The latest results from the Planck collaboration (Planck Collaboration XVI 2013) give

$\Omega_{cdm}h^2 = 0.1196 \pm 0.0031$, with a low $H_0 = 67.4 \pm 1.4 km \cdot s^{-1} \cdot Mpc^{-1}$, meaning that the DM accounts for about 26% of the total energy density. The total matter cosmological parameter detected by Planck is $\Omega_m h^2 = 0.1423 \pm 0.0029$, meaning that $\Omega_m = 0.314 \pm 0.020$.

Neutrinos

Neutrinos interact very weakly with electromagnetic force, making them a form of DM. In fact, they are the only type of DM which has positively been detected so far in laboratory experiments. As opposed to CDM which decouples at early times, neutrinos remain relativistic until late times, making them a form of HDM.

Neutrinos were first postulated by Pauli, in 1930, to conserve energy and momentum in β-decay, and were long assumed to be massless in the standard model of particle physics. The discovery, in the late 90's, of oscillations between neutrino flavors in the detected solar neutrinos (Fukuda et al. 1998) indicated that the neutrinos species must have finite eigenstates, and was the first evidence of physics beyond the standard model. A lot of effort in modern cosmology is put into quantifying Ω_ν, or equivalently quantifying the sum of the individual masses $\sum_i m_{\nu_i}$. Current estimates (Thomas et al. 2010; Planck Collaboration XVI 2013) give $\sum_i m_{\nu_i} < 0.28 eV$, corresponding to $\Omega_\nu < 7 \cdot 10^{-3}$.

Neutrinos have an astrophysical interest in addition to their cosmological one, as they are produced in supernovae core-collapse and hugely contribute to the energy released during such events.

Radiation

It is usually a very sensible approximation to neglect the radiation contribution to the density of energy, i.e. assume that $\Omega_R \ll 1$. Indeed, we have seen in section 1.2.3 that radiation evolves as a^{-4}. This can be qualitatively understood as the dilution with volume experienced by matter, plus an extra factor (1+z) due to the photons being redshifted with the expansion of the Universe. This strong dilution implies that, except for the very early stages of the Universe, the contribution of radiation in the overall energy density is very small.

Dark Energy

After including baryons, CDM, neutrinos, and radiation in the energy budget of equation 1.42, there is still 75% of energy which is completely unaccounted for. The most

abundant component of our Universe is this "known unknown", a type of energy which we have so little information about, that it casts a shadow on the foundations of the model which predicts it. Etienne Klein poetically noticed in an interview to the French *Revue des enseignants*, that in French, "Dark energy" ("Energy noire") is the anagram of "un-known queen" ("reine ignoree"). The next section, 1.2.7, is dedicated to the unknown Queen.

1.2.7 The "known unknown": Dark Energy and the cosmological constant

History

Einstein did not like the idea of a dynamical space-time. Originally, he introduced the cosmological constant Λ on the left hand side of the Einstein equations (i.e. as a geometric component) to produce a static solution,

$$R^{\mu\nu} - \frac{1}{2} g^{\mu}\nu R + \Lambda g^{\mu\nu} = 8\pi G T^{\mu\nu} \ . \tag{1.48}$$

After the discovery of the Hubble expansion, Einstein abandoned this concept, calling it the "biggest blunder" of his life. The community regained interest in Λ in the late nineties, with the emergence of the concept of Dark Energy, as an additional term in the right hand side of the Einstein equations (i.e. as a constituent of the energy tensor, as opposed to as geometric component).

Observational evidence

From the 1930s to the 1980s, a cosmological constant seemed unnecessary to explain cosmological observations, and the "Λ problem" was a theoretical one. In the 1980s and early 1990s, various indirect evidence started to accumulate in favor of a Ω_Λ term in the energy density budget. In particular, LSS measurements (Efstathiou et al. 1990; Maddox et al. 1990) suggested a low value for Ω_m ($\approx 0.15 - 0.4$) which gave room to a Ω_Λ term. Another striking observational evidence was that the age of the Universe, when computed only with CDM is shorter than the age of some stars observed in the Magellanic cloud (Jaffe 1996).

In this context of a rising concordance model with $\Lambda \neq 0$, two different teams, Riess et al. (1998) and Perlmutter et al. (1999), showed that the Universe is accelerating, rather than decelerating as expected in a radiation dominated or a matter dominated phase.

The questions posed by the accelerated expansion

Such an observation that $\ddot{a} > 0$ has three different possible implications, from the Friedmann equation

$$H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + \frac{3p}{c^2}) + \frac{\Lambda c^2}{3} \ ,$$

1. either the Universe is dominated by "dark energy", an unknown particle or field with negative pressure $P < -\rho/3$; or
2. there is a non-zero cosmological constant Λ; or
3. the theoretical foundations of this equation, i.e. GR, are wrong and need to be revised.

In other words, the main questions about cosmic expansion are:

1. Is the cosmic acceleration induced by a breakdown of GR on cosmological scales, i.e. a lack of the theory, or by an unknown entity, within the theory: an energy component which exerts repulsive gravity?
2. If the answer to this first question is a new energy component, what are the properties of this energy, and in particular, shall we put the cosmological constant in the r.h.s of the Einstein equation, as a new entity contributing to the stress-energy tensor or in the l.h.s, as a property of space? Is the DE energy density constant in space and time i.e. is it described by a "pure" cosmological constant in the Friedmann equation, or by a time-dependent contribution?

Dark energy as a "vacuum energy"

One option is to put the cosmological constant on the r.h.s of the Einstein's equation, i.e. to consider it as another component of the energy density of the Universe. It is then included in the stress-energy tensor, with a state equation $w = -1$, and a contribution to the energy density which does not dilute with time and is given by

$$\rho_\Lambda = \rho_{\Lambda,0} = \frac{\Lambda c^2}{8\pi G} = cst \ . \tag{1.49}$$

Pure Λ versus dynamical equation of state

An equation of state $w = -1$ corresponds to a "pure" cosmological constant model. Some models allow the equation of state to vary with redshift, and much effort is made in modern cosmology to constrain such models. Since DE only becomes important in the late Universe, the most common parametrisation of such an equation of state is a first order Taylor expansion:

$$w(a) = w_0 + (1 - a)w_a \ , \tag{1.50}$$

where a "pure Λ" model corresponds to $w_0 = -1$ and $w_a = 0$. The idea in constraining such models is that understanding the dynamics of DE would allow us to uncover its nature.

The two open questions of Dark Energy

Any theory of Dark Energy will have to explain two main problems:

1. *The fine tuning problem*: the observed value of the cosmological constant is over 100 orders of magnitude higher than the value predicted by quantum field theory for the vacuum energy ($\rho^{obs} \equiv 10^{120}\rho^{predicted}$). This is not "fine" tuning, and reveals that we basically have no good theory of dark energy.

2. *The coincidence problem*: our theory predicts that dark energy is negligible at early times and dominates in the late Universe. The ratio $\frac{\Omega_\Lambda}{\Omega_m} \propto$ scales as a^3 and there is only a very brief time when Ω_m and Ω_Λ are of comparable sizes, which seems to *coincide* exactly with our time. Figure 1.6 illustrates the coincidence problem.

Both of these problems are currently explained in the framework of anthropic arguments, which constitutes a still very controversial area of cosmology.

Alternatives to Dark Energy

There exists alternatives to a theory of DE described above, two of which are explained below:

- One alternative consists of questioning the accelerated expansion, i.e our interpretation of the observations. If we happened to live in the centre of a large cosmic void,

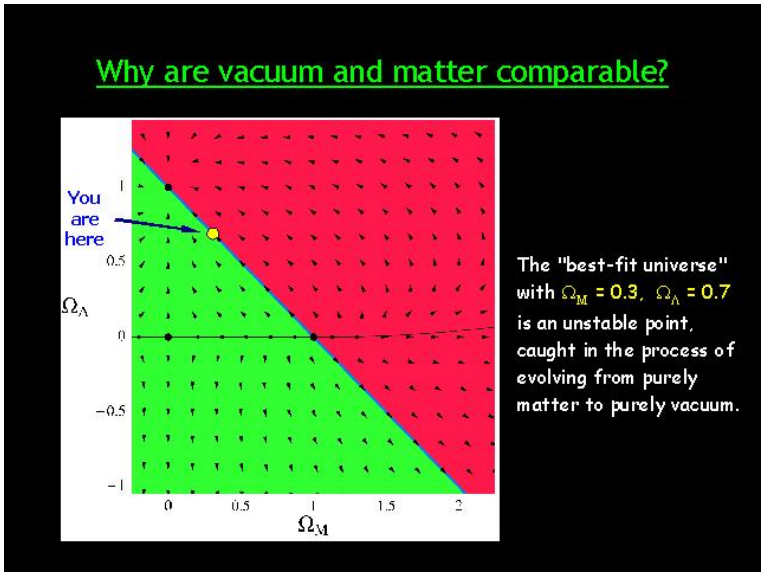

Figure 1.6: The "coincidence scandal", as seen by cosmologist Sean Carrol. Credit: Sean Carrol *: "Dark Energy and the Preposterous Universe"*

the very same observations would not imply that the Universe is expanding in an accelerated way. The price to pay for such an interpretation of the observations is to abandon the cosmological principle.

- Another alternative consists of accepting the accelerated expansion, but questioning the theory which requires DE to explain it. Modified gravity theories allow us to achieve an accelerated expansion in a Universe which contains only CDM.

1.2.8 A "simple but strange Universe"

The recent Planck's high-precision CMB map gives the most up to date values yet of the Universe's ingredients (Planck Collaboration XVI 2013). The Planck results are showing very good agreement with the ΛCDM model, and at the same time, points at a very problematic situation: we do not know what 75% of the Universe is made of. This is why figure 1.7, the cosmic recipe before and after Planck, has been humorously referred to by cosmologists as "the pizza that no one ordered". The New York Times paraphrase of Saul Perlmutter following the Planck results, qualifying the Universe of "simple but strange" seems more accurate than the "almost perfect Universe" referred to by ESA.

We are faced with two possibilities, both equally unsettling: either 75% of the Universe exists in an exotic and unknown form, or our theory to describe the Universe breaks down. Such a dilemma which is a recurrent one in the history of physics: a problem in which

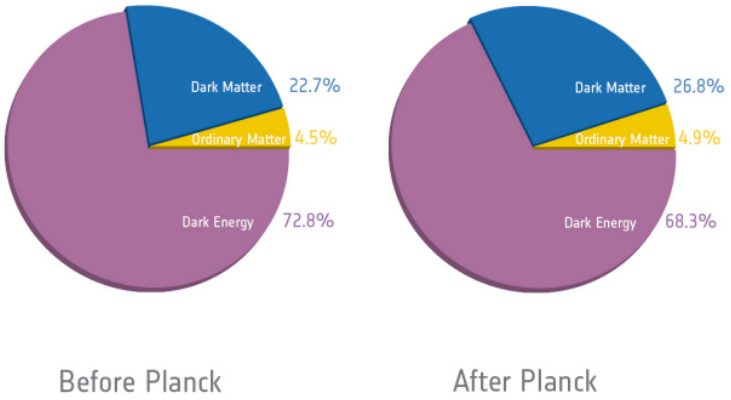

Figure 1.7: Baryonic matter occupies just 4.9. % of the Universe's mass/energy inventory. Dark matter occupies 26.8%, while dark energy accounts for 68.3%. The 'before Planck' figure is based on the WMAP 9-year data release presented by Hinshaw et al., (2012). Credit: ESA.

there is either a "legislative" solution[6] - rethinking our whole theory - or an "ontological" one - discovering a new entity that would make the theory hold. Figure 1.8 gives other examples of such situations in the history of astronomy, reminding us that there is no systematic recipe to solve such a dilemma. Facing the phenomena listed in the left column of the table, astronomers had to either come up with a new entity, or a new theory. The discovery of Neptune allowed to explain the orbit of Uranus within the framework of the Newtonian theory. In contrast, the unexpected orbit of Mercury led to a paradigm shift and the replacement of Newtonian gravity by GR. We still have to determine which of these solutions will allow to explain our current cosmological observations.

1.3 The not so smooth isotropic Universe

On scales smaller than $100Mpc$, the Universe is far from being homogeneous (see e.g. figure 1.2). Large Scale Structures (LSS) in the Universe are believed to have grown from primordial initial perturbations in a homogeneous Universe, through a process called *gravitational Instability*, to form the present structures we observe.

In the next sections, we assume the physics that generates the initial inhomogeneities and present the Newtonian approach to gravitational instability 1.3.1 without going into the details of "inflation", the mechanism which is believed to govern the growth of the primordial instabilities in the very early Universe. For reasons developed in section 1.3.3,

[6] in the terminology of Etienne Klein.

Phenomenon	New Entity	New Theory
Uranus' orbit	Neptune	-
Mercury's orbit	*(Hypothetical planet Vulcan ruled out)*	General Relativity
Beta decay	Neutrino	*(Violation of angular momentum ruled out)*
Galaxy flat rotation curves	Dark Matter?	Modified Newtonian Dynamics?
Accelerating universe (SN Ia and other data)	Dark Energy?	Modified General Relativity?

Figure 1.8: Previous versions of the cosmological puzzle. Facing the phenomena listed in the left column, astronomers had to either come up with a new entity, or a new theory. The discovery of Neptune allowed to explain the orbit of Uranus within the framework of the Newtonian theory. In contrast, the unexpected orbit of Mercury led to a paradigm shift and the replacement of Newtonian gravity by GR. We still have to determine which of these solutions will allow to explain our current cosmological observations. Credit: presentation by Ofer Lahav, 2013.

our theories only allow us to predict the statistical properties of the cosmological fields describing the inhomogeneous Universe. These properties propagate in Fourier space, e.g. into properties of matter power spectrum.

1.3.1 Dynamics of gravitational instabilities

Validity of the Newtonian approximation

General relativity is well approximated by Newtonian gravity in two cases:

1. on scales inside the Hubble radius, i.e. approximately the size of the observable Universe.

2. when describing non-relativistic matter, for which the pressure p and the energy density ρ verify $p \ll \rho$.

In particular, the newtonian approximation is valid when describing the sub-Hubble dynamics of CDM and baryons (after decoupling). It also underlies the cosmological simulations of non-linear growth of structures.

Newtonian gravitational instability equations

Within the Newtonian approximation, three equations describe the evolution of an ideal self-gravitating fluid, with density ρ, pressure $P \ll \rho$ and velocity $\vec{u}$. They are:

1. the continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla_r \cdot (\rho \vec{u}) = 0 \ , \tag{1.51}$$

2. Euler's equation:

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla_r)\vec{u} = -\frac{1}{\rho}\nabla_r p - \nabla_r \Phi \ , \tag{1.52}$$

3. the Poisson equation:

$$\nabla^2 \Phi = 4\pi\rho G \tag{1.53}$$

where $\vec{r}$ denote the Newtonian position vector, t the time, and Φ the gravitational potential, which determines the gravitational acceleration by $\vec{g} = -\nabla_r \Phi$

Recovering background cosmology

The dynamics of the homogeneous background, described by the Friedmann equations, are recovered by replacing the velocity with the Hubble's law velocity: $\vec{u} = H\vec{r}$, by considering a uniform fluid, i.e. $\partial_r \rho = 0$, and by fudging the Poisson equation including the cosmological constant: $\nabla^2 \Phi = 4\pi\rho G - \Lambda$. The Poisson equation 1.53 then satisfies

$$\Phi = \frac{1}{6}(4\pi\rho G - \Lambda) r^2 \ , \tag{1.54}$$

whereas the Euler equation 1.52 becomes

$$\frac{\partial H}{\partial t} + H^2 = \frac{1}{3}(\Lambda - 4\pi\rho G) \ , \tag{1.55}$$

which is the Newtonian limit of one of the Friedmann equations (in the relativistic limit, ρ is replaced with $\rho + 3P$). The continuity equation 1.51 rewrites

$$\partial_t \rho + 3\rho H = 0 \tag{1.56}$$

which is the usual Friedmann expression of energy conservation for $\rho \ll P$.

Perturbation equations in the comoving frame

The goal is now to derive the dynamics of any perturbation to the smooth isotropic fluid. When perturbing the density ρ and the velocity u of the fluid, as well as the gravitational potential Φ about their background value, and using comovig spatial coordinate $\vec{x} = \frac{\vec{r}}{a(t)}$,

$$\rho \to \overline{\rho}(t) + \delta\rho = \overline{\rho}(t)(1+\delta) \,, \tag{1.57}$$

$$p \to \overline{p}(t) + \delta p \,, \tag{1.58}$$

$$\vec{u} \to a(t)H(t)\vec{x} + \vec{v} \,, \tag{1.59}$$

where the peculiar velocity $\vec{v} = a\frac{d\vec{x}}{dt}$ describes changes in the comoving coordinate of a fluid element, i.e. its departure from the Hubble flow,

$$\Phi \to \overline{\Phi}(\vec{x}, t) + \phi \,. \tag{1.60}$$

The three equations describing structure formation, 1.51, 1.53 and 1.52, rewrite

$$\partial_t \delta + \frac{1}{a}\nabla \cdot [(1+\delta)\vec{v}] = 0 \,, \tag{1.61}$$

$$\partial_t \vec{v} + \frac{1}{a}(\vec{v}\cdot\nabla)\vec{v} + \frac{\partial_t a}{a}\vec{v} = -\frac{1}{\rho a}\nabla p - \frac{1}{a}\nabla\phi \,, \tag{1.62}$$

$$\nabla^2\phi = 4\pi G\overline{\rho}a^2\delta \,, \tag{1.63}$$

where ∂_t is the time derivative for a given $\vec{x}$, and ∇ is the spatial derivative with respect to the comoving position $\vec{x}$ at fixed t (which is linked to the gradient with respect to the physical position through $\nabla_r = \frac{1}{a}\nabla$, and to the time derivative through $\left(\frac{\partial}{\partial t}\right)_x = \left(\frac{\partial}{\partial t}\right)_r + H(t)\vec{x}\cdot\nabla$).

The linear approximation

A standard picture of the cosmic structure formation assumes that the initial tiny amplitude of fluctuation grow according to equations 1.61 to 1.63. Within the framework of a perturbation theory, the perturbations (and their spatial derivatives) are assumed to be small enough so that one can linearize these equations (i.e. ignore the second order part).

Moreover, the background equations 1.54 to 1.56 set the zero-th order to zero, so that the new equations describing the small perturbations in density, pressure, and velocity are

$$\partial_t \delta + \frac{1}{a} \nabla \cdot \vec{v} = 0 \ , \tag{1.64}$$

$$\partial_t \vec{v} + H\vec{v} = -\frac{1}{\overline{\rho} a} \nabla \delta P - \frac{1}{a} \nabla \phi \ , \tag{1.65}$$

$$\nabla^2 \phi = 4\pi G \overline{\rho} a^2 \delta \ . \tag{1.66}$$

Considering a barotropic fluid, i.e. with a density which is only a function of the pressure, such as $P = P(\rho)$, equation 1.65 can rewrite

$$\partial_t \vec{v} + \frac{\partial_t a}{a} \vec{v} = -\frac{c_s^2}{a} \nabla \delta - \frac{1}{a} \nabla \phi \ , \tag{1.67}$$

where $c_s^2 \equiv (\partial p / \partial \rho)$ is the sound velocity.

Taking the time derivative of the perturbed continuity equation 1.64 and combining it with the perturbed Poisson and Euler's equations, we find the equation for the growth of structures in the Newtonian theory:

$$\partial_t^2 \delta + 2H \partial_t \delta - 4\pi G \overline{\rho} \delta - \frac{1}{a^2 \overline{\rho}} \nabla^2 \delta P = 0 \ . \tag{1.68}$$

This fundamental equation shows the competition between the infall by gravitational attraction (the term $4\pi G \overline{\rho} \delta$), and the pressure support (the term $\frac{1}{a^2 \overline{\rho}} \nabla^2 \delta P$) to which an initially small fluctuation is subject. Whether such a fluctuation grows or not, depends on the balance between these two competing forces, and this balance is affected by the different phases which the Universe goes through.

This last equation also shows that the dynamics of the fluctuation depends on the cosmic expansion, through H(z), or equivalently on the present time value of the cosmological parameters.

Growth of structures in Fourier space

Within the framework of a linear perturbation theory, it is particularly interesting to derive the dynamics in the Fourier space. Thanks to the linearity, the equations that govern the evolution of each mode also govern their superposition, i.e. the overall fluctuation, and reciprocally.

We can consider the Fourier decomposition of the density fluctuations:

$$\delta_{\vec{k}}(t) = \frac{1}{(2\pi)^{3/2}} \int \delta(\vec{x}, t) d^3\vec{x} e^{-i\vec{k}\cdot\vec{x}} \ . \tag{1.69}$$

The equation for the growth of structures then translates into an equation of the evolution of the k-mode of the fluctuation,

$$\partial_t \delta + 2H\partial_t \delta + (\frac{c_s^2 k^2}{a^2} - 4\pi G\overline{\rho})\delta = 0 \ . \tag{1.70}$$

In the same way as in the real space, we can derive a criterion for a fluctuation to grow. Whether a perturbation grows or not depends on the sign of the $(\frac{c_s^2 k^2}{a^2} - 4\pi G\overline{\rho})\delta$ term. We define the Jeans wavelength $\lambda_J = c_s\sqrt{\pi/(G\overline{\rho})}$ and distinguish different cases:

- If $(c_s^2 k^2/a^2 - 4\pi G\overline{\rho}) > 0$, i.e. if the wavelength $\lambda = 2\pi a/k$ of the mode is smaller than the Jeans wavelength, $\lambda < \lambda_J$, equation 1.70 describes a damped oscillator. The pressure support counterbalance the gravitational infall, giving rise to oscillations of the fluctuation amplitude.

- If $(c_s^2 k^2/a^2 - 4\pi G\overline{\rho}) < 0$, i.e. if the wavelength of the mode is larger than the Jeans wavelength, $\lambda > \lambda_J$, gravitational collapse dominate, the perturbation is monotonically increasing and the mode is said to be unstable to gravitational accretion.

The Jeans length characterises the scale that pressure can propagate in the form of a sound wave within the infall time (i.e. the characteristic time of gravity collapse $t_{ff} \approx 1/\sqrt{(G\overline{\rho})}$). Modes larger than the Jeans mode do not have time for the pressure to resist gravitational infall, (since the time to infall is less than the time it takes to propagate a pressure disturbance across the perturbation). Below this scale, the pressure has time to counterbalance the gravitational instability, and the mode is stable and oscillates. In a Universe which is matter dominated (or radiation dominated) $w = p/\rho > -1/3$, the proper Jeans length grows faster than a, and faster than the proper wavelength of any given mode. As a result, Fourier modes that start "outside" of the Jeans length (i.e. with a higher wavelength), where they evolve by gravitational accretion, later come inside and undergo acoustic oscillations. This is related to the notion of *horizon*, which we develop in the next section.

1.3.2 Horizons

Within the framework of GR, the comoving distance travelled by a signal (sound, light, ...etc), which is not interacting or being deviated, is given by the equation of geodesics:

$$d((r_1,t_1),(r_2,t_2)) = \int_{t_1}^{t_2} \frac{v dt'}{a(t')} = \int_{r_1}^{r_2} \frac{dr}{\sqrt{1-kr^2}} \ , \tag{1.71}$$

where v is the velocity of the signal, and (r_i, t_i) are the space-time coordinates of the two points between which the signal is traveling. This distance is also the maximum possible distance between two points sharing the same signal. This allows us to define the "horizon", as the maximal distance between us (the observer) and a point sharing a given type of signal with us. The different horizons detailed below correspond to different types of signal.

As shown in section 1.3.1, equation 1.70 determines the time evolution of the modes of fluctuations. A mode, or scale, is said to "enter the horizon" when its scale length is becoming comparable to the horizon scale.

Causal horizons

Two points are said to be in causal contact if they can share the same information, or light signal. The distance between two points in causal contact is therefore given by replacing v by c, the speed of light, in equation 1.71:

$$d_H((r_1,t_1),(r_2,t_2)) = \int_{t_1}^{t_2} \frac{cdt}{a(t)}; . \tag{1.72}$$

In a flat Universe, the "causal horizon", i.e. the physical size of the space domain of all the point in causal contact, is given, at time t, by $d_H(t) = a(t)\int_0^t \frac{cdt'}{a(t')}$.

Evolution of visible scales during radiation and matter domination

During radiation and matter domination, $a(t) = t^n$ with $n < 1$ ($n = 1/2$ during radiation domination, and $n = 2/3$ during matter domination). As a result,

$$d_H = \frac{n}{1-n} R_H \ , \tag{1.73}$$

where $R_H = cH^{-1}$ is the "Hubble radius". In fact, $d_H \approx R_H$ as long as the Universe expansion is *decelerated*.

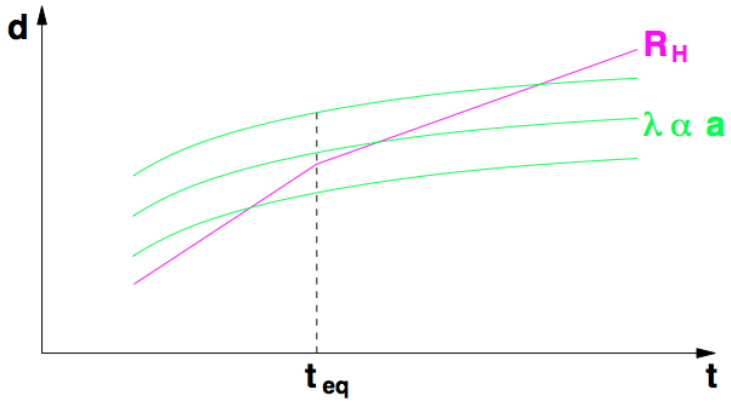

Figure 1.9: The wavelength of some observable cosmological perturbation compared with the Hubble radius, during radiation and matter domination. Since all wavelengths $\lambda(t) = a(t)2\pi/k$ grow with negative acceleration during that time, and since the Hubble radius grows linearly with time, the modes of the perturbation enter the horizon one after the other. The smaller modes enter during radiation domination ($t < t_{eq}$), whereas the larger modes enter during matter domination. Credit: lecture notes by Julien Lesgourgues (2009).

The "causality" takes a specific meaning when studying random fluctuations (as will be developed in section 1.3.3): the random properties of two points which are not in causal contact should be uncorrelated. Therefore, the correlation function $\xi(r)$ ($\langle\delta(\vec{r_1}), \delta(\vec{r_2})\rangle$, with $r = |\vec{r_1} - \vec{r_2}|$) vanishes at scales $r > d_H$, and the power spectrum should also vanish at scales k corresponding to wavelengths $\lambda(t) > d_H(t)$. Although within the inflation theory, this is not exactly the case, wavelengths smaller or higher than $d_H(t)$ correspond to two different regimes, called respectively "causal" and "acausal".

Observable Universe

The radius of the observable Universe is the size of the area which today is linked to us by causality.

$$R_{obs} = a(t_0) \int_{t_{dec}}^{t_0} \frac{cdt'}{a(t')} \,. \tag{1.74}$$

Since the time t_{dec} at which the Universe became transparent is much smaller than the age of the Universe ($t_{dec} \approx 380000 yrs$), the size of the observable Universe is, to a good approximation, the size of the causal horizon: $R_{obs}(t_0) \approx d_H(t_0)$.

Evolution of visible scales during radiation and matter domination

During radiation and matter domination, the scale factor $a(t)$ evolves as t^n with $n < 1$, so does the wavelength $\lambda(t)$ of each mode, whereas the causal horizon $d_H(t) \approx R_H(t) \approx R_{obs}$ evolves as t, as shown in figure 1.9.

As a result, the modes of the fluctuation "enter" the observable Universe one after the other.

Sound horizon

The sound horizon, i.e. the distance over which the wavefront of acoustic waves can travel between the early Universe and some later time t is obtained by replacing v by the sound velocity c_s in equation 1.71:

$$d_s^{flat} = a(t) \int_0^t \frac{c_s dt'}{a(t')} \, . \tag{1.75}$$

The sound horizon also plays a crucial role in the evolution of fluctuations.

1.3.3 A stochastic theory

The need for a stochastic theory

In order to describe the perturbed Universe, cosmologists wish to make predictions on cosmological *fields*, such as the matter over-density $\delta\rho$. For the following reasons, our theory allows us only to predict the *statistical* behavior of such fields:

- We do not have direct observational access to the primordial fluctuations, which would give us definite initial conditions for our deterministic evolution equations.
- The cosmological time scales are much longer than the time scale over which we can make observations, it is therefore very hard for us to follow the evolution of a single object.

For these reasons, we do not describe the state of the Universe by a field - in this sense our theory differs from usual fluid mechanics - but rather by the statistical properties of the fields (namely, the two points correlation function, the three point correlation function and higher moments). We consider the initial perturbations as random quantities and our theory is a theory for the evolution of these random quantities. Our Universe is considered as one realisation of a statistical ensemble of possibilities.

There are essentially two types of constraints one can draw on the statistical properties of the random fluctuations. Some come from the physics that *generates* the initial perturbations (e.g. inflation) and the others from the physics that *processes* them - linear perturbation theory - while respecting the symmetries of the background cosmology i.e. homogeneity and isotropy.

In the following, we will consider as an example field $\delta_i \equiv \delta(x_i) = \frac{\rho(x_i)-\bar{\rho}}{\bar{\rho}}$, the density contrast at the comoving position x_i.

Correlation function

If we consider a random field $\delta(\vec{x})$ (random field means that at each point $\delta(\vec{x})$ is some random number), with zero mean $< \delta(\vec{x}) >= 0$, the probability of realising some configuration of the field is a functional $Pr[\delta(\vec{x})]$.

One way to define the statistical properties of this field is to define *correlators*, i.e. expectation values of products of the the field at different spatial (or time) points. In particular, the two point correlator is given by:

$$\xi(\vec{x},\vec{y}) = \langle \delta(\vec{x})\delta(\vec{y})\rangle = \int \mathcal{D}\delta \cdot Pr[\delta] \cdot \delta(\vec{x})\delta(\vec{y}) \ . \tag{1.76}$$

Combining statistical homogeneity and isotropy ensures that $\xi(\vec{x},\vec{y})$ only depends on the absolute value $r = |\vec{x}-\vec{y}|$:

$$\xi(r) = \langle \delta(\vec{x}), \delta(\vec{x}+\vec{r})\rangle \ . \tag{1.77}$$

Power spectrum

We've seen in section 1.3.1 the advantage of decomposing $\delta(\vec{x})$ into its Fourier components $\delta(\vec{k})$,

$$\delta(\vec{k}) = \frac{1}{(2\pi)^{3/2}} \int \ d\vec{x}\delta(\vec{x}) \exp(-i\vec{k}\vec{x}) \ . \tag{1.78}$$

(Here we adopted the symmetric Fourier convention, so that $\delta(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \int \ d\vec{k}\delta(\vec{x}) \exp(+i\vec{k}\vec{x})$). We can compute the correlator in Fourier space and define the power spectrum $P(k)$ or $\mathcal{P}(k)$ so that

$$\left\langle \delta(\vec{k})\delta^*(\vec{k'})\right\rangle = P(k)\delta_D(\vec{k}-\vec{k'}) = \frac{2\pi^2}{k^3}(k)\mathcal{P}\delta_D(\vec{k}-\vec{k'}) \ . \tag{1.79}$$

P(k) is then related to the Hankel transform of the two points correlation function (c.f.)

$$\xi(r) = \frac{1}{2\pi^2}\int k^2 P(k) j_0(kr) dk = \int \frac{dk}{k}\mathcal{P}(k) j_0(kr) \ , \tag{1.80}$$

where $j_0(kr) = \sin(kr)/kr$ is a spherical Bessel function of order zero. Another way to interpret the power spectrum is by going back to the dynamics of the fluctuations (section 1.3.1). Indeed, since the k-mode of the fluctuation $\delta(\vec{x}, t)$ is a complex variable, it can be decomposed by a set of two variables, its amplitude $D_{\vec{k}}$ and its phase $\phi_{\vec{k}}$. The power spectrum is then given by

$$P(t, \vec{k}) = \left\langle D_{\vec{k}}(t)^2 \right\rangle \ . \tag{1.81}$$

Therefore $P(k)$ is a measure of the amplitude of the mode of wavelength $\vec{k}$.

Gaussian random fields

The two points c.f. and power spectrum are well defined quantities for any field. But they become extremely powerful in the particular case of a *gaussian* field. These fields are defined as such: for an arbitrary positive integer m, the m-points joint probability distribution obeys a multi-variate Gaussian:

$$P(\delta(\vec{x_1}), \delta(\vec{x_2}), \delta(\vec{x_3}), ..., \delta(\vec{x_m})) = \frac{1}{\sqrt{(2\pi)^m det(C)}} exp\left[-\sum_{i,j=1}^{m} \frac{1}{2}\delta_i C^{-1} \delta_j \right] \ , \tag{1.82}$$

where $\delta_i \equiv \delta(\vec{x_i})$, and C is the covariance matrix, defined as $C_{ij} \equiv \langle \delta_i \delta_j \rangle = \xi(|\vec{x_i} - \vec{x_j}|) = \xi_{ij}$. The m-points joint probability can be seen as the result of the discretisation of the field in m pixels: the field is then described by an m-dimensional vector $\vec{\delta} = [\delta(\vec{x_1}), \delta(\vec{x_2}), \delta(\vec{x_3}), ..., \delta(\vec{x_m})]$, and its probability distribution is a multivariate gaussian, fully specified by its correlation function. In the very particular case of a gaussian field, the covariance matrix happens to completely define the probability distribution: the gaussian field is entirely specified by the two points c.f. ξ_{ij} and its linear combinations (including its derivative and integral).

Therefore, in the very particular case of gaussian fields, the correlation function, or equivalently the power spectrum, are completely defining the field!

There are three reasons to be excited about the properties of Gaussian fields in the

context of cosmology:

1. The primordial perturbations are believed to be gaussian. This is predicted by most simple inflation models, and supported by most observations to date.

2. Any *linear* evolution theory preserves the statistical behavior of a field during time. So if the field used to be gaussian, it still is at present time and the correlation function (or the power spectrum) still entirely specifies it.

3. The properties of gaussian fields propagate into interesting properties in Fourier space. Since the Fourier transform is a linear transformation, the probability distribution for a given mode $\delta(\vec{k})$ is also a multivariate gaussian. Since the homogeneity and isotropy of the Universe impose that different modes are uncorrelated, the modes of a fluctuation form a set of *independent* multi-variate gaussians, and similarly as in Fourier space, the gaussianity is preserved through the linear evolution of the modes.

The whole evolution of the perturbations is not entirely linear throughout time and space: at late times and small scales (large k's and small r's), non-linear structure formation destroys gaussianity and leads to the filamentary structures of the cosmic web, shown e.g. in figure 1.2. The correlation function ξ takes the product of two quantities and averages them over independent realisations of the system in question. In practice we do not have access to multiple realisations of the Universe. Instead we invoke an assumption of ergodicity, that is we assume that averaging over a sufficiently large volume is equivalent to an ensemble average (Pan & Zhang 2010). In general a correlation function is often calculated as a function of separation. In this case the average is over pairs of galaxies separated by some fixed distance.

Random fields on the sphere

Astronomers often have to work with 2D maps of the sky, either for reasons inherent to the studied observable (e.g. in the case of the CMB) or because the distance of the objects is much more difficult to infer than their position in the sky. In 2D-astronomy, astronomers consider the objects as projected on a celestial sphere, a modern version of the sky dome imagined by ancient civilisations. In this context, spherical harmonics are a more natural decomposition of the cosmic random fields than the Fourier decomposition.

Spherical harmonics form a basis for square-integrable functions on the sphere:

$$f(\vec{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{lm} Y_{lm}(\vec{n}) \ , \tag{1.83}$$

where $\vec{n} = (\theta, \phi)$. The equivalent to the power spectrum in Fourier space, is the "angular power spectrum" C_l (which is related to the two point correlator, just as $P(k)$ is in Fourier space). The homogeneity and isotropy of the Universe, imply that the correlation function only depends on the separation θ on the sphere:

$$\left\langle f(\vec{n}, \vec{n'}) \right\rangle = \sum_{l} C_l \frac{2l+1}{4\pi} P_l(\vec{n} \cdot \vec{n'}) = C(\theta) \ , \tag{1.84}$$

and inversely

$$C_l = 2\pi \int_{-1}^{1} dcos(\theta) C(\theta) P_l(cos\theta) \ , \tag{1.85}$$

where P_l are Legendre polynomials.

1.4 Observational probes of cosmic acceleration

1.4.1 Strategy, observables and methods

The strategy adopted to approach the questions posed by the cosmic accelerated expansion (see section 1.2.7) is to assume the existence of a dark energy entity, and constrain its properties with increasing precision, hoping that it will exhibit a failure of GR or a time dependency of the possible cosmological constant. How do we constrain DE?

As shown in equation 1.47 , the expansion can be written as

$$H^2(z) = \left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left(\Omega_m (1+z)^3 + \Omega_r (1+z)^4 + \Omega_k (1+z)^2 + \Omega_\phi \frac{u_\phi}{u_\phi(z=0)} \right) \ , \tag{1.86}$$

where Ω_m, Ω_r and Ω_ϕ are the present day energy densities of matter, radiation, and dark energy. (Here we have changed the initial writing of equation 1.47 to account for different models of dark energy). In terms of the scale factor a, this reads

$$H^2(a) = H_0^2 [\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_k a^{-2} + \Omega_X a^{-3(1+w)}] \ . \tag{1.87}$$

Measurable	Definintion
Proper distance	$D(z) = a(t_0)r_e = a(t_0)f_k(\chi)$
Luminosity distance	$d_L(z) = a(t_0)r_e(1+z)$
Angular diameter distance	$d_A(z) = a(t_0)r_e/(1+z)$

Table 1.2: Measurable consequences of the comoving coordinate $r(z)$.

Each of the terms of the right hand side represents the time history of one kind of energy density. In particular, the term Ω_x represents the cosmological constant if $w = -1$ and dark energy if w is constant. For non constant w, the above equations generalise by replacing $a^{-3(1+w)}$ with $exp\left(3\int_a^1 \frac{da'}{a'}[1+w(a')]\right)$.

We already have very good constraints on two of the terms of this equation: $H_0^2\Omega_m$ and $H_0^2\Omega_r$ from the Cosmic Microwave background (CMB). Therefore, by measuring accurately the expansion history $H(a)$, one can determine the DE density term, i.e. the time history of its energy density (modulo uncertainties due to the curvature). Thus, the strategy adopted by cosmologists to constrain DE is to measure the expansion history $H(a)$ as accurately as possible.

When observing an astronomical source, it is straightforward to deduce the scale factor $a(t_e)$ at the time of its emission, since $a = a(t_0)/(1+z)$. But the measurement of $\dot{a}$ is much more tricky. So rather than measuring $H(a) = \dot{a}/a$ directly (as done with BAOs for example), cosmologists measure $H(a)$ indirectly, via two different observables: the distance-redshift relation and the growth-redshift relation.

First observable: the distance-redshift relation

We have shown in section 1.1.6 that all measurable distances (luminosity distance $d_L(z)$, angular diameter distance $d_A(z)$, comoving distance $D(z) = a(t_0)r_e(z)$), depend on $H(z)$ through $r_e = f_k(\chi)$ (see equation 1.12). Therefore, they depend on DE through their dependence on $H(z)$. Measuring these redshift-distance functions, summarised in table 1.2, allows an indirect measurement of $H(z)$.

Second observable: the growth of structures

A second observable consequence of DE is its effect on the growth of structures. Indeed, the competition between gravitational collapse and the accelerated expansion induced by

dark energy retards the growth of structures. The linear growth factor g is defined within a linear perturbation theory, as

$$\delta(\vec{x},t) = \frac{\rho_m(\vec{x},t) - \bar{\rho_m}(t)}{\bar{\rho_m}(t)} = \delta(\vec{x},t_i) \cdot \frac{g(t)}{g(t_i)} \ , \tag{1.88}$$

where $\delta(\vec{x},t)$ is the density fluctuation of pressureless dark matter, and t_i is an arbitrary initial time. GR provides a relation between $g(z)$ and $H(z)$:

$$\ddot{g} + 2H\dot{g} = 4\pi G\rho_m g = \frac{3\Omega_m H_0^2}{2a^3} g \ . \tag{1.89}$$

Although the exact solution of the above differential equation is not trivial (see Weinberg et al. (2013)), to a first approximation

$$\frac{g(z)}{g(z=0)} \approx exp\left[-\int_0^z \frac{dz'}{1+z'}[\Omega_m(z')]^\gamma\right] \ , \tag{1.90}$$

where γ depends only weakly on the cosmological parameters (Peebles 1980; Lightman & Schechter 1990), and where the dependency on $H(z)$ and therefore on the dark energy term of equation 1.86 is through

$$\Omega_m(z) = \frac{\rho_m(z)}{\rho_{crit}(z)} = \Omega_m(1+z)^3 \frac{H_0^2}{H^2(z)} \ . \tag{1.91}$$

Since the density fluctuations at $z = 1088$ are accurately quantified, the amplitude of matter fluctuations provides another indirect observation of DE (through $H(z)$), through the growth-redshift relation $g(z)$.

This relation has another important advantage: it provides a test of the GR theory. Indeed, GR implies a one-to-one relation between the two observables D(z) and g(z). Inconsistencies between these two observables would mean that GR is incorrect on the largest of observable scales, or that DE contributes to the growth of structures in an unexpected manner.

The two observables we just presented are illustrated in figure 1.10.

Methods

Four complementary methods are used to measure the above observables ($H(z)$, $D(z)$ and $g(z)$) with increasing precision:

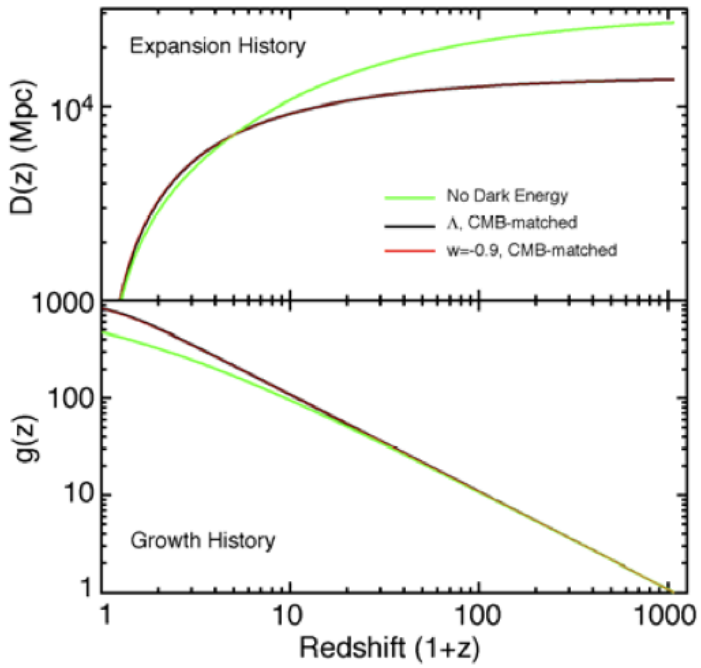

Figure 1.10: The two primary observables for DE: the distance-redshift relation $D(z)$ and the growth-redshift relation g(z) are plotted vs redshift, for three different cosmological models. The green curve is an open-Universe model with no dark energy at all. The black curve is the concordance ΛCDM model, which is flat and has a cosmological constant, i.e., $w = 1$. This model is consistent with all reliable present-day data. The red curve is a dark-energy model with $w = 0.9$, for which other parameters have been adjusted to match WMAP data. One sees that dark-energy models are easily distinguished from non-dark-energy models. Credit: DETF report (Albrecht et al. 2006).

1. Large Scale Structures (LSS) and Galaxy Cluster Counts (GC).
2. Baryon Acoustic Oscillations (BAOs).
3. Weak Gravitational Lensing (WL).
4. Type Ia supernovae (SNe).

Whereas BAOs and SNe constrain the expansion of the Universe as a whole (i.e H(a)) and are referred to as "purely geometric", WL and GCC give access to both the expansion history $a(t)$ and the growth of structures. In the following sections, we give a review of these probes.

1.4.2 Type Ia Supernovae

Supernovae (SNe) are the most straightforward way to study the accelerated expansion. In fact, they are the tool with which the accelerated expansion was originally discovered (Riess

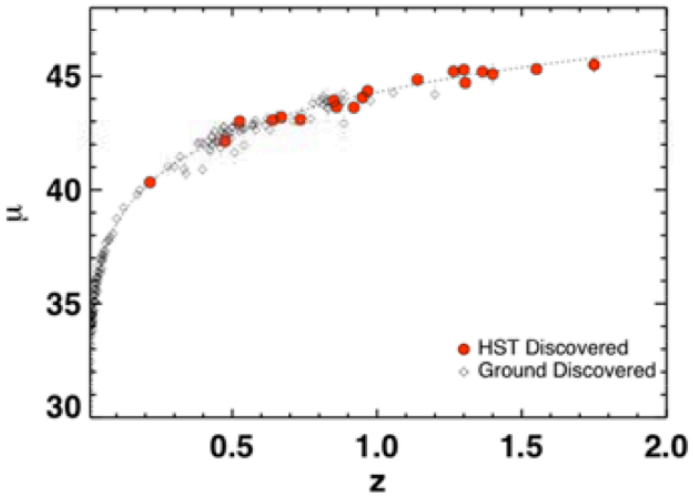

Figure 1.11: Cosmological results from the GOODS SNe. The y-axis shows the distance ($\mu = 5log10d_L + const.$) and the x-axis shows the redshift. Credit: Riess et al. (2004).

et al. 1998; Perlmutter et al. 1999). This probe uses the first observable we presented: the redshift-distance relation. Indeed, Type Ia SNe are believed to be standard candles, i.e. objects of known absolute luminosity, of which the luminosity distance d_L can be inferred via the relation $f = L/4\pi d_L^2$. Why are Type Ia SNe believed to be standard candles? Cosmologists assume that they are the explosions of objects with nearly the same mass: white dwarfs stars that accrete material until they exceed the stability limit of 1.4 solar masses derived by Chandrasekhar.

The redshift of the SNe, which is necessary to measure the redshift distance relation $d_L(z)$, is measured either on spectral lines of the SNe explosion light, or with spectral features of the host galaxy. Figure 1.11 shows the distance-redshift relation obtained with HST SNe (Riess et al. 2004).

In practice, SNe are not proper standard candles: their absolute luminosity is not completely uniform (Phillips 1993), but appears to be correlated with other distance-independent features of the event, such as the rest-frame duration of the event, or some of its spectral features. In that sense, SNe are *standardizable* objects (rather than *standard* objects), to a degree of precision which has not been entirely uncovered yet.

Other types of SNe, gravitational waves sources or gamma ray bursts, could be used in the future as standard candles, but the way they would compare to the type Ia SNe for cosmological measurements is not known yet.

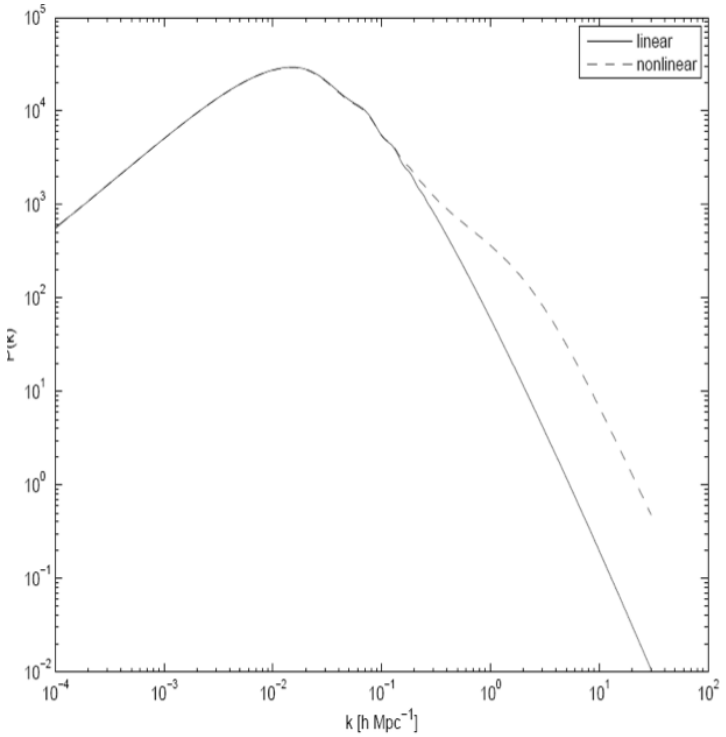

Figure 1.12: ΛCDM power spectrum, normalised to the local abundance of galaxies, for a flat Universe, $\Omega_m = 0.25$, $\Omega_b = 0.05$, $\sigma_8 = 0.8$, $h = 0.7$. Solid line shows the linear power spectrum, dotted line shows the non-linear power spectrum according to the fitting function of Smith et al. (2003). Credit: lecture notes by Julien Lesgourgues (2009).

1.4.3 Large scale structures

The growth of large scale structures (LSS) from initial fluctuations is a powerful probe of cosmology, since it highly depends on the cosmological model. LSS measurements allow us to constrain DE in various ways:

- The position of the BAOs feature provides a standard ruler to study the expansion history.
- The shape of the angular power spectrum of the galaxy density fluctuations encapsulates precious information about the clustering amplitude and the growth of structures.

It is possible to compute the theoretical matter power spectrum, for a given cosmology and to compare it with observations. In practice, the matter power spectrum is not directly measurable and cosmologists use some specific observable tracers to measure it.

qualitative explanation of the main features of the power spectrum

Let us explain qualitatively the main features of the linear power spectrum at present time (figure 1.12).

Most theories of Inflation predict that the initial (post-inflation) power spectrum is approximately scale invariant, i.e. in the form of a power law:

$$P_i(k) = k^{n_s(k)-1}, \tag{1.92}$$

where $n_s(k)$ is a constant (a non constant $n_s(k)$ means departure from the scale invariance). In a log-log space, this means that the post-inflation power spectrum is a diagonal straight line.

During radiation domination, the Jeans wavelength $\lambda_J = c_s\sqrt{\pi/(G\overline{\rho})}$ and the size of the horizon R_H scale as t, whereas each mode scales as $\lambda \propto t^{1/2}$. Since the Jeans wavelength grows faster than the modes, larger and larger modes (i.e. smaller and smaller k) "enter" the horizon. The net result is for the straight line power law to bend, at the scale corresponding to λ_J, which is the interface between the modes oscillating inside the horizon (at "small" scales and "large" k), and the growing modes, outside the horizon (at "large" scales and "small" k). With time, λ_J grows, shifting the turnover of $P(k)$ to larger and larger scales (smaller and smaller k).

After the time of matter-radiation equality, all modes grow. The bend of $P(k)$ freezes at $k = k_{eq}$, the wave number of the mode entering the horizon at the matter-radiation transition and corresponding to the size of the horizon at the matter-radiation equality (which, today, is given by $k \approx 0.01 Mpc^{-1}$).

Clearly visible in Figure 1.12 are a succession of wiggles at slightly smaller scales than the turnover. These are baryon acoustic oscillations (BAOs), a consequence of the photon-baryon fluid set up in the early Universe which we present in section 1.4.4.

At small scales, typically $k/h > 10^{-1} Mpc^{-1}$, linear theory does not apply anymore. In chapter 5 of this thesis, we present some corrective terms due to the non-linear halo collapse and the mode coupling, which both slightly affect the shape of the present days power spectrum. Such corrections have been applied e.g. in figure 1.12, where both the linear power spectrum and the non-linear power spectrum are shown.

The present shape of the power spectrum depends on a number of cosmological parameters. This makes it a powerful cosmological probe. Indeed, the overall normalisation

depends on the primordial spectrum amplitude, the age of the universe, and the cosmological constant. The slope depends on the the primordial spectrum index. The scale k_{eq} of the maximum of $P(k)$ depends on the time of the matter-radiation transition, which depends on the matter density today, i.e. on Ω_m. The shape of the spectrum at $k > k_{eq}$ depends on the spectral index, but also on the ratio Ω_b/Ω_{cdm} (a high baryon density implies a lower amplitude for $k > k_{eq}$ as well as additional oscillations).

The galaxy bias

Another obstacle to the direct measurement of the power spectrum is the galaxy bias. The matter over-density has contributions from both baryonic, luminous matter, and dark matter. We can only observe the luminous matter i.e. the baryonic contribution to the total underlying matter distribution. The main idea behind galaxy surveys is that galaxies (i.e. luminous matter) can be treated as tracers of the underlying matter distribution. By measuring the positions of galaxies (the position of a galaxy in the $3D$ space is given by its angular position on the sky and its redshift) and smoothing it on very large scales, one can construct a smooth $3D$ map of the distribution of luminous galactic matter (lgm), $\delta_{lgm}(\vec{x}, t)$ and draw from it the power spectrum of the luminous matter:

$$\left\langle |\delta_{lgmt,\vec{k})^2}| \right\rangle_{\vec{k}/k} = P_{lgm}(k) \ . \tag{1.93}$$

In order to test the prediction of the cosmological model, the function we would like to measure is not P_{lgm} but the total matter power spectrum $P_{tot}(k)$.

In practice, galaxies are *biased* tracers of matter: $\delta_{lgm}(t, k) \approx b_g \delta_{tot}(t, k)$, where b_g is called the galaxy bias. If galaxies linearly trace the total matter, i.e. if b_g is a constant, then the galaxy power spectrum $P_{lgm}(\vec{k}, z)$ relates to the (total) matter power spectrum $P_{tot}(\vec{k}, z)$ via

$$P_{lgm}(\vec{k}, z) = b_g^2 P_{tot}(\vec{k}, z) \ . \tag{1.94}$$

But in general, b_g is expected to be a function of scale and redshift, $b_g(\vec{k}, z)$. Our ignorance of the form of $b_g(\vec{k}, z)$, especially at small scales, is one of the main obstacles to the use of galaxy surveys for cosmology. There are different models for $b_g(\vec{k}, z)$, including the Halo model (Cooray et al. 2000), which assumes that the mass is concentrated in DM Haloes, with a concentration of galaxies defined by the *Halo Occupation Distribution*.

In chapter 5 of this thesis, we use the luminosity of galaxies as a biased tracer of the baryonic matter, and the galaxies number density as a - differently - biased tracer of the total matter density. We then try to detect a scale-dependent difference between the two biases induced by BAOs.

1.4.4 Baryon Acoustic Oscillations

Between the end of inflation and the time of recombination, the Universe is filled with an ionized plasma, hot and dense, in which photons and baryons are coupled. Under the effect of Thompson scattering, i.e. scattering between the photons and the charged particles of the plasma, the photons are "trapped" in the plasma. During this time, the interplay between the plasma pressure and the radiation pressure results in "sound waves": spherical perturbations of the density (and pressure) propagating around each initial over-density of matter, traveling at a speed of $c_s = c/\sqrt{3}$ through the baryon-photon fluid.

As the Universe expands, the baryonic matter cools down, eventually allowing the nuclei and electrons to bind into stable, neutral atoms at the time of recombination (approximately 370000 years after the Big Bang). The mean free path of photons is then high enough for the photons to be liberated from the matter. The Universe becomes "transparent" and the baryonic shells propagating in the form of sound waves freeze, leaving an imprint in the distribution of matter.

This signature, known as Baryon Acoustic Oscillations (BAOs), is visible in the cosmic microwave background and in the large-scale distribution of galaxies. The distance traveled by the over-density shell, r_s (the sound horizon at the time of recombination), provides a feature of known physical size and as such, it is a standard ruler which makes it a precious probe of DE. The identification of the BAOs scale as a transverse angle at different redshifts determines the distance ratio $D(z)/r_s$ and allows us to infer $D_A(z)$ (and constrains DE through the observable D(z)), whereas its identification along the line of sight determines $H(z)r_s$ and provides a direct measure of $H(z)$. In chapter 5 of this thesis, we aim to detect another kind of imprint left by BAOs on the clustering of galaxies: a modulation of the density ratio of baryon to dark matter across large regions of the Universe.

Weinberg et al. (2013) give a complete overview of the history of the detection of BAOs. The first prediction of the BAOs effect, in the CMB and the late-time matter power spectrum dates back to the late 1960s (Sakharov 1966), at a time when the pure baryons cosmologies predicted a very strong effect. The BAOs effect was identified as

a standard ruler in the 1990s (Kamionkowski et al. 1994). Early results form the 2dF Galaxy Redshift Survey (2dFGRS) (Percival et al. 2001) gave the first hints for the BAOs feature in the data, followed by convincing detections in the SDSS Data Release 3 and final 2dFGRS samples (Eisenstein et al. 2005; Cole et al. 2005).

1.4.5 Weak lensing

Gravitational lensing from distant intervening mass fluctuations causes the shapes of objects to be distorted. Thus, measurement of the distortion of light allows us to get information about the distribution of mass through which the light has travelled en route to the observer. It is therefore a powerful tool both for direct mass reconstruction or as a cosmological tool. Gravitational Weak Lensing (WL) deals with very small distortions to a source image. While no single object is intrinsically round, if the intrinsic shapes of galaxies are uncorrelated with one another, one can average the apparent shapes of many thousands of such objects to extract a distortion attributed to WL. The statistical properties of this observable pattern put a constraint on the power spectrum and therefore on the cosmological model and on DE. Weinberg et al. (2013) give a complete overview of the history of the detection of WL, which was made possible by the use of large format CCDs in the beginning of our century. The first detection were made from space (Wittman et al. 2000; Bacon et al. 2000), soon followed by ground-based detections (Van Waerbeke et al. 2001). In chapter 2 of this thesis, we derive the requirement on star/galaxy separation, a problem which needs to be taken into account for use of WL as a cosmic probe. Here we present some element of the WL formalism, and some of the practicalities of cosmic shear measurement.

Weak Gravitationnal lensing

Figure 1.13 shows a typical gravitational lens system, where the distances between the observer, the lens and the source are linked through the General Relativity lens equation:

$$\alpha = \frac{D_{ds}}{D_s D_d} \frac{4\pi G M}{c^2 \theta} \,, \tag{1.95}$$

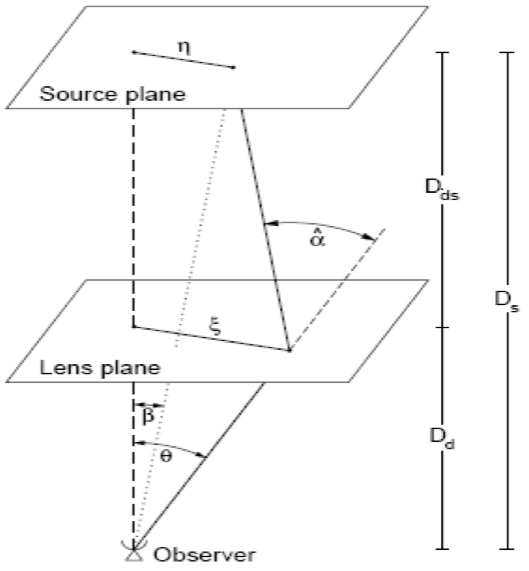

Figure 1.13: Sketch of a typical gravitational lens system showing source, lens and observer planes. Credit:Bartelmann & Schneider (2001)

where M is the mass of a point mass, D_s is the source-observer distance and D_d is the observer-lens distance.

The lensing can be described by the component of the distortion matrix:

$$\tilde{A} = \begin{bmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{bmatrix}$$

κ is the *convergence* and characterizes the *magnification* of the source. γ_1 and γ_2 are the two components of the shear field $\gamma = \gamma_1 + i\gamma_2 = \gamma e^{i2\alpha}$, which characterizes the *stretching* of the source: γ_1 describes the stretching and compression along the x-axis and γ_2 the stretching and compression along the y-axis. For non point-mass lenses with surface mass density Σ, it is useful to define the critical mass density $\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}}$, since it allows to define several regimes of gravitational lensing. The case of *strong lensing* corresponds to $\Sigma \geq \Sigma_{crit}$ and is characterized by multiple images and elongated arcs, such as those seen in the Abell 2218 cluster (Kneib et al. 1996). In the case of *weak lensing*, $\Sigma \ll \Sigma_{crit}$. In this regime, the distortion to the shape of the galaxies is very slight.

The convergence along a given line of sight, for a source with a mean redshift distri-

bution $n(z_s)$, writes:

$$\kappa = \int_0^\infty dz_s n(z_s)\kappa(z_s) = \int_0^{\chi_{max}} d\chi W(\chi)\delta(\chi) \ , \tag{1.96}$$

where $\kappa(z_s) \simeq \int_0^{\chi_s} d\chi w(\chi,\chi_s)\delta(\chi)$, is the convergence along a given line of sight up to z_s, $W(\chi) = \int_z^{z_{max}} dz_s n(z_s) w(\chi,\chi_s)$ and $w(\chi,\chi_s) = \frac{3\Omega_m H_0^2 \mathcal{D}(\chi)\mathcal{D}(\chi_s-\chi)}{2c^2\mathcal{D}(\chi_s)}(1+z_s)$ (where $\mathcal{D}(\chi)$ is the comoving angular diameter distance). W is called the lensing weight function, or lensing efficiency function. The convergence field κ is not directly observable. In practice, it can be reconstructed from measurement of the ellipticities (shear) of galaxies (see next section) from survey data. Reconstruction of the convergence field from survey shear data is still a challenging problem (see Kaiser 1998 and Hu & White 2001 for techniques in the 2D case and Taylor 2001 and Hu & Keeton 2002 for 3D mass reconstruction).

Cosmic Shear

The term "Cosmic shear" is used to describe the measurement of the very small distortions caused by WL, to study the Universe and more particularly the cosmological parameters. The two points correlation function (c.f.) and its Hankel transform, the power spectrum, were already introduced in the context of the matter distribution of galaxies. Cosmic shear offers an unbiased tracer of the underlying dark matter distribution, which makes it a powerful cosmological tool.

It is convenient to measure shear and ellipticity, not in terms of their real and imaginary parts, but as tangential and cross components:

$$\epsilon_+ \equiv -Re(\epsilon e^{-2i\phi}) \ ; \epsilon_\times \equiv -Im(\epsilon e^{-2i\phi}) \ , \tag{1.97}$$

where ϕ is the polar angle of the galaxy position relative to the lens centre, ϵ is the galaxy ellipticity and Re/Im take the real and imaginary parts of an expression respectively. $+$ denotes a component tangential to the line of sight and $\times$ denotes the component at 45 degrees to the line of sight. Similar expressions γ_+ and $\gamma_\times$ can be defined for the shear γ.

With these definitions we can define the 2 points c.f. ξ which is a measure of the cosmic shear signal,

$$\xi_\pm(\theta) \equiv \langle\gamma_+\gamma_+\rangle \pm \langle\gamma_\times\gamma_\times\rangle \tag{1.98}$$

The correlation function ξ takes the product of two quantities (in this case the shear or ellipticity of a pair of galaxies) and averages them over independent realisations of the system in question. In practice we do not have access to multiple realisations of the Universe. Instead we invoke an assumption of ergodicity, that is we assume that averaging over a sufficiently large volume is equivalent to an ensemble average (Pan & Zhang 2010). In general a correlation function is often calculated as a function of separation. In this case the average is over pairs of galaxies separated by some fixed distance.
The shear angular power spectrum writes:

$$C_l = \int \frac{d\chi}{\chi^2} W^2(\chi) P(\frac{l}{\chi}) \, , \tag{1.99}$$

where P is the matter power spectrum, and W is the lensing weight function defined previously. Note that several approximations have been made in this expression, including the Limber approximation (i.e we have assumed that the area of the sky we are interested in is small enough so that we can approximate $j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+1/2}(x)$ and $\lim_{l\to\infty} j_l(x) = \sqrt{\frac{\pi}{2l+1}} \delta_D(l + 1/2 - x))$ and $l \equiv l + 1/2$.

Cosmic shear surveys contain information about the shapes of galaxies, but also about their distance to the observer, through the redshift information. Acquiring spectroscopic redshift is a costly and time-consuming procedure, and the large number of sources used in current galaxy surveys makes it more likely for the redshift information to be from photometric sources than spectroscopic source. The redshift information is used through shear *tomography*, which consists in cutting the galaxy sample into slices in redshift and calculating the 2 points c.f. within each slice (auto-correlation) and between different slices (cross-correlation). The shear angular power spectrum between two redshift slices i and j is

$$C_{ij}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) P(\frac{l}{\chi}) \, . \tag{1.100}$$

The interest of tomography varies according to the goal of the survey. While it does not significantly improve the estimation of the matter power spectrum, it is very useful for constraining the evolution of the equation of state of dark energy, $w(a)$. An alternative way to use the redshift information is known as 3D weak lensing (Heavens 2003). This full 3D statistical analysis, using the decomposition of the power spectrum in terms of

spherical harmonics and spherical Bessel functions, allows to avoid the loss of information from binning galaxies. It is particularly promising for high-precision measurement of the dark energy equation of state parameters.

Shear Measurement

Measuring the cosmic shear signal is challenging, for two main reasons. First, extracting the shear signal from the measurement of the ellipticity of a galaxy is difficult: the induced shear on any galaxy is a small effect ($\approx 1\%$) and since each galaxy is intrinsically elliptical, it is impossible to separate the shear effect from the intrinsic ellipticity for a single object. The other challenge is observational: several observational distortions need to be corrected for, in order to convert images of small, faint galaxies into a measurement of the cosmic shear. These observational effects include CCD pixelisation and PSF, and will be summarized in more details in the next paragraph about systematics.

How is the shear extracted from galaxies images? One method is the Kaiser, Squiers & Broashurst (KSB) method (Kaiser et al. 1995). The idea is to calculate the quadrupole moments of the surface brightness distribution $I(x)$ for each source galaxy. For a galaxy with a surface brightness profile $I(\theta)$, well-defined for all angular separations θ from the centre of the image, so that:

$$\overline{\theta} \equiv \frac{\int d^2\theta w[I(\theta)]\theta}{\int d^2\theta w[I(\theta)]} \, , \tag{1.101}$$

where $w[I(\theta)]$ is a suitably chosen weight function such that the integrals converge. The tensor of second brightness moments is defined as

$$Q_{ij} = \frac{\int d^2\theta w[I(\theta)](\theta_i - \overline{\theta_i})(\theta_j - \overline{\theta_j})}{\int d^2\theta w[I(\theta)]} \, , i,j \in \{1,2\} \tag{1.102}$$

The complex observed ellipticity is linked to Q through

$$\epsilon^{obs} \equiv \epsilon_1^{obs} + i\epsilon_2^{obs} = \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}}, \tag{1.103}$$

The observed ellipticity ϵ is related to the intrinsic ellitpticity ϵ^{int} and the reduced shear $g = \gamma/(1-\kappa)$ via:

$$\epsilon^{obs} = \frac{\epsilon^{int} + g}{1 + g^*\epsilon^{int}} \, . \tag{1.104}$$

When the lensing is "weak" enough, one can do the approximation $\epsilon^{obs} = \epsilon^{int} + \gamma$. In the case where the ellipticities have been measured on a small enough patch of the sky for the shear to be constant (because light from all the sources have passed through the same mass distribution), averaging over this patch gives $\left\langle \epsilon^{obs} \right\rangle = \left\langle \epsilon^{int} \right\rangle + \gamma$. Assuming that the intrinsic ellipticities of galaxies are randomly distributed and average to zero, gives

$$\left\langle \epsilon^{obs} \right\rangle = \gamma \ . \tag{1.105}$$

Since the random intrinsic ellipticities are correlated neither with the intrinsic ellipticities nor with the shears of other galaxies,

$$\left\langle \epsilon_i^{obs} \epsilon_j^{obs} \right\rangle = \langle \gamma_i \gamma_j \rangle + \left\langle \epsilon_i^{int} \epsilon_j^{int} \right\rangle + \left\langle \gamma_i \epsilon_j^{int} \right\rangle + \left\langle \epsilon_i^{int} \gamma_j \right\rangle = \langle \gamma_i \gamma_j \rangle \ , \tag{1.106}$$

The latter assumption (of random intrinsic ellipticities) turns out to be inaccurate: *intrinsic alignment* is the main cosmic shear systematic which needs to be corrected for.

Systematics

We briefly list some of the systematics that need to be accounted for in any application of WL.

- **Intrinsic alignment.** When aiming at high-precision cosmic shear measurements, the assumption of random intrinsic ellipticities becomes inaccurate since galaxies can intrinsically align and therefore have correlated intrinsic ellipticities. Both the intrinsic ellipticities correlation and the shear-ellipticity correlation must be taken into account in equation 1.106

- **CCD effect.** Some systematics result from the properties of the CCD chips, which collect the light in galaxy surveys. These effects include nonlinear response (Van Waerbeke et al. 2006) and charge transfer inefficiency in the way electrons are read out of the CCD pixels, which can both bias the shear measurement (Massey et al. 2010).

- **Point Spread Function.** The effect of the telescope optics and the atmosphere, described by the point spread function (PSF), are particularly strong in ground-based survey, and still present in space-based mission. The distortion of the shape of the galaxies is corrected for via PSF calibration, which is usually made using star

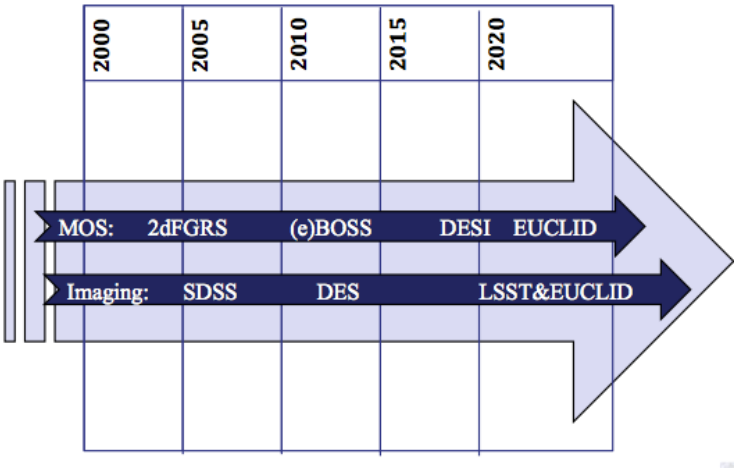

Figure 1.14: Timeline of some current observational projects in cosmology. The upper line shows the project using multi-object spectrograph, whereas the lower line shows the imagining surveys. Credit: presentation by Ofer Lahav (2013).

images, acting as point sources (Paulin-Henriksson et al. 2008). In chapter 2 of this thesis, we explore how the need for such calibration propagates into a requirement on star/galaxy separation.

1.4.6 Current landscape of galaxies surveys

The questions posed by the discovery of the cosmic acceleration inspired a number of ambitious ground-based and space-based projects, making this early 21^{st} century particularly exciting for cosmology.

The work presented in this thesis is based on the simulations and data from two of these projects: the Dark Energy Survey (DES; http://www.darkenergysurvey.org) and the The Baryonic Oscillation Spectroscopic Survey (BOSS; http://www.sdss3.org/surveys/boss.php) of SDSS-III. Together with Pan-STARRS (http://pan-starrs.ifa.hawaii.edu/public/) and HETDEX(http://hetdex.org/hetdex/), both DES and BOSS are clear examples of the current generation of stage-III dark energy experiments, as defined in the report from the Dark Energy Task Force (Albrecht et al. 2006).

The DES is notably the first of these experiments combining all the four DE probes defined by the DETF on one single facility. It will carry out large area, multi-band imaging surveys that go a factor of ten or more deeper (in flux) than the SDSS imaging survey. The BOSS survey will allow to exploit the BAOs probe, by carrying out a nearly cosmic-variance limited survey (over $10^4 deg^2$) out to $z \approx 0.7$. These experiment are only one part

of the many ambitious observational efforts which are being carried out. Figure 1.15 uses a tool called BigFoot, created by the author, Maayane Soumagnac, together with Alex Merson (UCL) during this PhD. It shows the footprints of some of the galaxy surveys overlapping with the DES footprint. Figure 1.14 shows the timeline for some of the most important observational projects of the past and coming decades.

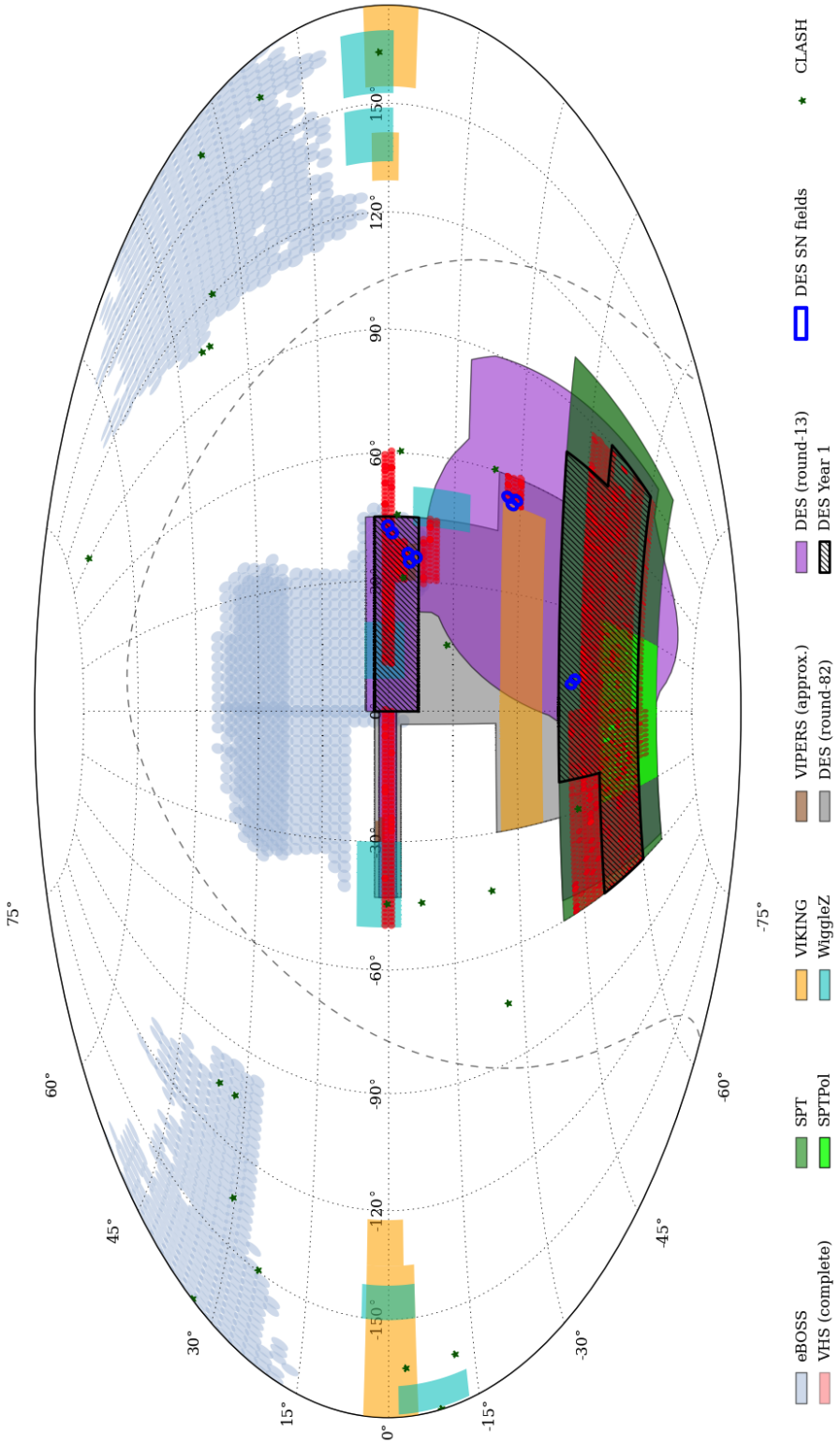

Figure 1.15: Footprints of the galaxy surveys overlapping with the footprint Dark Energy Survey. This figure has been made with the BigFoot tool, designed by the author Maayane Soumagnac and Alex Merson (UCL) during this PhD.

This page was intentionally left blank

Chapter 2

Science requirements on Star/Galaxy separation

"It seems to me that the evidence, other than the admittedly critical tests depending on the size of the galaxy, is opposed to the view that the spirals are galaxies of stars comparable with our own. In fact, there appears as yet no reason for modifying the tentative hypothesis that the spirals are not composed of typical stars at all, but are truly nebulous objects."

Harlow Shapley, *The Great Debate* (May, 1921)

"I hold, therefore, to the belief that the galaxy is probably not more than 30,000 light-years in diameter; that the spirals are not intra-galactic objects but island universes, like our own galaxy, and that the spirals, as external galaxies, indicate to us a greater universe into which we may penetrate to distances of ten million to a hundred million light-years."

Herber D. Curtis, *The Great Debate* (May, 1921)

2.1 Introduction

What makes a star look different from a galaxy in a deep image? This seemingly very simple question hides the much more complicated issue of allocating a size and a scale to objects observed in the sky, which has concerned observers and theorists throughout

the 20th century. Perhaps the most dramatic illustration of this long-standing issue is Heber Curtis's and Harlow Shapley'sx so called great debate in the 1920s, which solved the question of the size of our Galaxy in relation to cosmic scales; whereas Shapley was arguing in favor of the Milky Way embracing the entirety of the universe and spiral nebulae being part of it, Curtis saw our galaxy as one object among many other island universes.

The problem of classifying stars and galaxies in large scale surveys is a long-standing one. It has been encountered back in the early 1990's (e.g. the APM survey, Maddox et al. 1990) and poses a major challenge for all recent and large imaging cosmological surveys, including the Dark Energy Survey (DES) (http://www.darkenergysurvey.org/) and Euclid (http://sci.esa.int/euclid), which have been designed to uncover the nature of dark energy (DE). One common denominator of the wide variety of observational probes constraining DE is the necessity to select pure samples of galaxies. More specifically, all the surveys must differentiate galaxies at cosmological distances from local objects, to obtain pure, or at least well-understood, samples.

In the area of "precision cosmology", any source of systematic error is likely to play a decisive role and needs to be taken into account in order to refine the standard inflationary Big Bang picture. An example of a scientific question for which star/galaxy separation is a potentially critical systematic is the precision measurement of Primordial Non-Gaussianities (PNG). These manifest themselves by making the bias of a given type of tracers of dark matter halos strongly scale-dependent. This effect can easily be mimicked by any local systematic effect adding power at large scales and correlated with the galaxies. As the stellar distribution in the Milky Way is across large angular scales, star/galaxy separation is likely to introduce systematic errors in the measurement of PNG.

Another example is the effect of occultation of galaxies by stars of comparable magnitudes. Ross et al. (2011) showed that this effect constitutes a source of systematic error in the measurement of angular and photometric distributions of luminous red galaxies. Photometric effects associated with faint stars could therefore partially account for the excess power seen in Thomas et al. (2011) for the MegaZ-LRG survey. This work gives two other examples, in the case of Weak Lensing (WL) and Large Scale Structures (LSS) measurements, where star/galaxy separation is a key systematic, which needs to be taken into account in order to properly constrain DE.

The aim of this chapter, which covers the first part of Soumagnac et. al. 2013, is to study the impact of star/galaxy misclassification on the measurement of the cosmological

parameters in the Dark Energy Survey, in the case of the WL and LSS probes, and to show how the requirements on the statistical and systematic errors propagate into new requirements on the quality of star/galaxy separation.

The outline of this chapter is as follows. In section 2.2, we present the "DES-like" simulations which we base our analysis on, both in this chapter and in chapter 3. In section 2.3, we define the main tools used to formulate the science requirement on star/galaxy separation. In section 2.4, we derive the science requirements from the need to constrain the statistical errors, and in section 2.5, we present those derived from the constraint on the systematic errors. In section 2.6, we present additional requirement, dictated by calibration. We summarize all the derived requirements and conclude in section 2.8.

2.2 The Dark Energy Survey Simulated Catalog

As part of the process of testing and validation of the DES Data Management (DESDM) system (Mohr et al. 2012), a series of detailed simulations have been designed to serve as a test-bench for the development of the pipelines and for verifying the scientific reach of the experimental channels. Each of these iterations of the simulations are dubbed "Data Challenges" (DC). The simulation starts with the creation of galaxy catalogs stemming from an N-body simulation (Busha et al. 2013) and detailed models of the Milky Way galaxy (Rossetto et al. 2011) for the star component. These are merged and fed to an image simulator which includes atmospheric and instrumental effects. The resulting images serve as inputs for DESDM and are processed as the data will be: the code *SExtractor* (Bertin & Arnouts 1996) produces a catalogue of more than 300 parameters encapsulating information about each detected object.

The most relevant features of these simulations for our study are:

- the seeing is introduced as a function of observing time;
- the galaxy shapes have been implemented using a Sersic profile which matches the observed profile;
- the Point Spread Function (PSF) takes into consideration the seeing for that time, the optics and the distortion as a function of separation from the optical axis.

The results shown in this chapter and in chapter 3 are based on the latest release (internal to the DES collaboration) of simulated data, DC6, which covers approximately

140 square degrees to the full DES depth, corresponding to about 10 nights of observations. We select from it the objects with a model magnitude in the i band brighter than 24, as they are the ones most likely to be detected with DES.

2.3 Formalism of the science requirements on star/galaxy separation

DES will be among the first surveys to combine in a single project the observation of the four preferred dark energy probes, as identified by the Dark Energy Task Force (DETF) (Albrecht et al. 2006). SNe and Baryonic Acoustic Oscillation (BAO) constrain the expansion of the Universe as a whole and are therefore referred to as *purely geometric*. WL and GC constrain both the expansion on the Universe and the growth of Large Scale Structures (LSS) (See Weinberg et al. 2013 for a complete review).

In order to properly constrain DE, the broad variety of measures carried out within each probe must meet certain requirements defined by DES science teams. While there is no unique way to specify the constraints on dark energy experiments and probes, the Figure of Merit (FoM), defined by the DETF, provides a useful metric. If we parameterise the time evolution of DE by the equation of state $w(a) = w_o + (1-a)w_a$, where $a(t) = \frac{1}{1+z(t)}$ is the cosmic scale factor and $z(t)$ is the redshift of an object emitting at time t, the FoM is defined as the reciprocal of the area of the error ellipse enclosing 95% confidence limit in the w_o-w_a plane. Larger FoM indicates smaller errors and therefore greater accuracy on the measurement of the parameters.

In other words, reaching the FoM goals requires to minimise the error on w_o and w_a. Since the total error is the sum of the *statistical* error and the *systematic* error, we can derive two types of science requirements. More concretely, the total Mean Square Error (MSE) on a cosmological parameter p_α can be decomposed as

$$MSE[p_\alpha] = \sigma^2[p_\alpha] + \Delta^2[p_\alpha] \,, \tag{2.1}$$

where $\sigma^2[p_\alpha]$ is the statistical error variance and $\Delta[p_\alpha]$ is the parameter shift due to the systematic signals. For each probe, both of these terms needs to be controlled in order to minimise the total error.

Star/galaxy misclassification is an interesting effect because, as illustrated in figure 2.1,

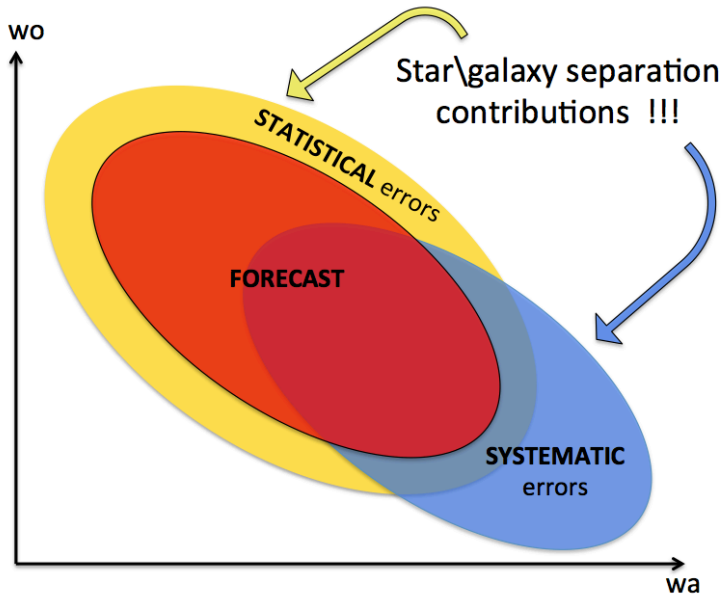

Figure 2.1: The FoM is defined as the reciprocal of the area of the error ellipse enclosing 95% confidence limit in the w_o-w_a plane, and is shown in red. The effect of statistical errors is to increase the area of the forecasted ellipse, whereas the effect of systematic errors is to shift the ellipse. Star/galaxy misclassification is an interesting effect, because it contributes to both the statistical and systematic part of the total error, for the WL and LSS probes

it contributes to both the statistical and systematic part of the total error, for the WL and LSS probes. This allows us to translate separately the requirement on the statistical term (section 2.4) and the requirements on the systematic term (section 2.5) into requirements on the quality of the star/galaxy separation. Additional requirements are specific to each probe, e.g. PSF calibration for WL (section 2.6).

We outline below a formalism to derive these requirements.

2.3.1 Completeness, contamination and purity

In table 2.1, we define the parameters used to quantify the quality of a star/galaxy classifier. For a given class of objects, X (stars or galaxies), we distinguish the surface density of well classified objects, N_X, and the density of misclassified objects, M_X.

The galaxy *completeness* c^g is defined as the ratio of the number of true galaxies

	True Galaxies	True stars
Objects classified as galaxies	N_G	M_S
Objects classified as stars	M_G	N_S

Table 2.1: Quantities used to define the purity p^X and completeness c^X of a class of objects X (stars or galaxies).

classified as galaxies to the total number of true galaxies. The stellar contamination f_s is defined as the ratio of stars classified as galaxies to the total amount of objects classified as galaxies.

$$c^g = \frac{N_G}{N_G + M_G} \,, \tag{2.2}$$

$$f_s = \frac{M_S}{N_G + M_S} \,. \tag{2.3}$$

The *purity* p^g is defined as $1 - f_s$:

$$p^g = \frac{N_G}{N_G + M_S} = 1 - f_s \,. \tag{2.4}$$

Similar parameters can be defined for a sample of stars: p^s, f_g and c^s.

We aim to formulate the requirements on the statistical and systematic errors in terms of constraints on these parameters. This will allow us to quickly compare the performance of the classifiers presented in chapter 3 and assess whether they allow us to achieve the goals of the DETF FoM.

One should note that there are some inefficiencies in the image pipeline, which are studied in DC6 and which we do not deal with in this analysis. Instead, we define the latter parameters with respect to the mock galaxy samples used to produce the image simulations. With real DES data, our results could be tested e.g. on HST data in the same fields.

2.3.2 Fisher Information Matrix

The Fisher information matrix describes how the errors on the angular power spectrum $C(l)$ (of the cosmic shear in the case of WL, and the density fluctuations of galaxies in the case of LSS) propagate into the precision on the cosmological parameters p_α . We employ this formalism (see Tegmark et al. 1997, for a review), to quantify the impact of star/galaxy misclassification on each of the terms in equation 2.1, i.e. on the statistical and systematic errors on the cosmological parameters.

The Fisher matrix can be expressed as

$$F_{\alpha\beta} = \sum_{l} \sum_{(i,j)(m,n)} \frac{\partial C_{ij}(l)}{\partial p_\alpha} Cov^{-1}[C_{ij}(l), C_{mn}(l)] \frac{\partial C_{mn}(l)}{\partial p_\beta} \ , \tag{2.5}$$

where the sum is over multipole values and redshift bins (typically five for WL). $Cov[X, Y]$ designates the covariance matrix of X and Y and is given by (Kaiser 1992; Takada & Jain 2004),

$$Cov[C_{ij}(l), C_{mn}(l)] = \frac{\{C_{im}(l)C_{jn}(l) + C_{in}(l)C_{jm}(l)\}}{f_{sky}(2l+1)\Delta l} \ , \tag{2.6}$$

where f_{sky} is the fraction of the sky covered by the survey ($f_{sky} = 0.1212$ for DES) and Δl is the width of the corresponding angular frequency bin.

2.4 Science requirements on the statistical errors

How does the need to control the statistical errors on the cosmological parameters propagate into a requirement on the quality of star/galaxy separation? In the following, we aim to answer this question in the case of the WL and LSS probes.

2.4.1 WL measurements

Gravitational lensing from distant intervening mass fluctuations causes the shapes of objects to be distorted such that they appear to be more or less elliptical. While no single object is intrinsically round, if the intrinsic shapes of galaxies are uncorrelated with one another, one can average the apparent shapes of many thousands of such objects to extract a distortion attributed to WL. The statistical properties of this observable pattern put a constraint on the power spectrum and therefore on the cosmological model and on DE.

For some concise introductions to cosmic shear, see e.g. Mellier (1999), Bartelmann & Schneider (2001) and Refregier (2003).

How do star/galaxy misclassifications affect the WL shear measurement? The predicted shear angular power spectrum $C_{ij}(l)$ depends on N_{eff}, the effective density per unit area of galaxies with reliable shape measurements,

$$C_{ij}(l) = \int_0^{r_H} dr r^2 W_i(r) W_j(r) P(l/r; r) + \delta_{ij} \frac{\sigma_e^2}{N_{eff}} \tag{2.7}$$

where $P(k = l/r)$ is the 3D matter power spectrum, $W_i(r)$ and $W_j(r)$ are the lensing efficiencies of the redshift bins (i, j), r is the comoving distance and r_H is the Universe horizon. The angular power spectrum depends on N_{eff} through the last term, i.e. the "shot noise" due to σ_e, the intrinsic ellipticity noise for the galaxy sample.

In order study the effect of N_{eff} on the statistical error $\sigma[p_\alpha]$, we compute the Fisher matrix for different values of N_{eff}. We estimate the $C_{ij}(l)$ and $\frac{\partial C^g_{mn}(l)}{\partial p_\alpha}$ terms (see Eq. 2.5) using the same code as in Laszlo et al. (2012) and Kirk et al. (2011). The setup is as follows: we use a model with eight free parameters: $\{w_o,\ w_a,\ \Omega_m,\ H,\ \sigma_8,\ \Omega_b,\ n_s\}$; we assume a Planck prior (Jochen Weller, personal communication); there are five tomographic bins of roughly equal number density between z = 0 and 3; the redshift distribution is a Smail-type distribution (e.g. equation (12) of Amara & Réfrégier 2008, with $\alpha = 2$, $\beta = 1.5$, $z_0 = \frac{0.8}{1.412}$); we compute the $C_{ij}(l)$ and $\frac{\partial C^g_{mn}(l)}{\partial p_\alpha}$ terms for $l \in [1, 1024]$, to avoid the strongly non linear regime where baryon physics will start being important and following the l-cuts performed in most recent works by the WL community (Das et al. 2012); and the photometric redshift error is $\Delta z = 5 \cdot 10^{-2}(1 + z)$.

We then compute the marginalized statistical error on the cosmological parameters by approximating them with their Cramer-Rao lower bound

$$\sigma[p_\alpha] \approx \sqrt{(F^{-1})_{\alpha\alpha}} \tag{2.8}$$

We show the results for w_o and w_a in figure 2.2 and for the other free parameters of our model in figure 2.3.

Figure 2.2 shows that larger N_{eff} translates into smaller statistical errors on w_o and w_a, i.e. larger FoM, which puts a constraint on N_{eff}: it has to be higher than a threshold value N_{thresh} which can depend on the bandpass considered,

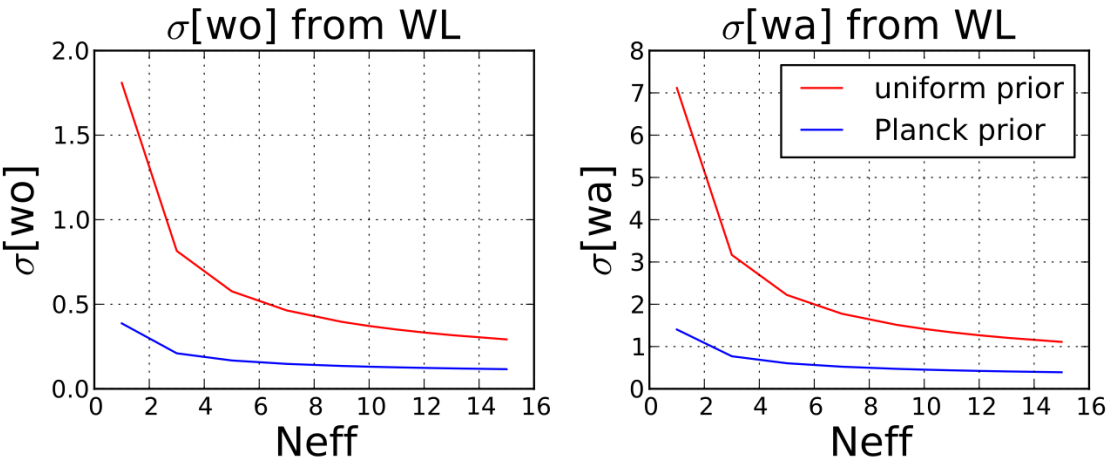

Figure 2.2: Marginalised statistical errors on the equation of state parameters w_o and w_a from the WL probe, for different values of the density of galaxies with reliable shape measurement N_{eff}. The errors are marginalised over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s\}$ and computed using the assumptions and setup described in section 2.4.1. The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

$$N_{eff} \geq N_{thresh} \ . \tag{2.9}$$

Figure 2.2 also shows asymptotes above $N_{thresh} = 10$, i.e. the effect of any variation of N_{eff} on the statistical error decreases at high N_{eff}. In practice, we require the increase of the statistical error due to star/galaxy misclassification to be smaller than 2%. If this reasonable but somewhat arbitrary goal is not achieved, it will only increase the statistical error and will not lead to a bias of the WL results. This translates into a decrease of N_{eff} smaller than 4%, i.e.

$$c^g \geq 96.0\% \tag{2.10}$$

Star-galaxy misclassification is only one among many other sources of errors leading true galaxies to be rejected from the sample of galaxies with reliable shape measurements, (e.g., shape measurement errors and photo-Z errors). To ensure that the statistical errors are controlled, this condition on c^g should be completed by constraints on the survey parameters controlling all the other sources of errors.

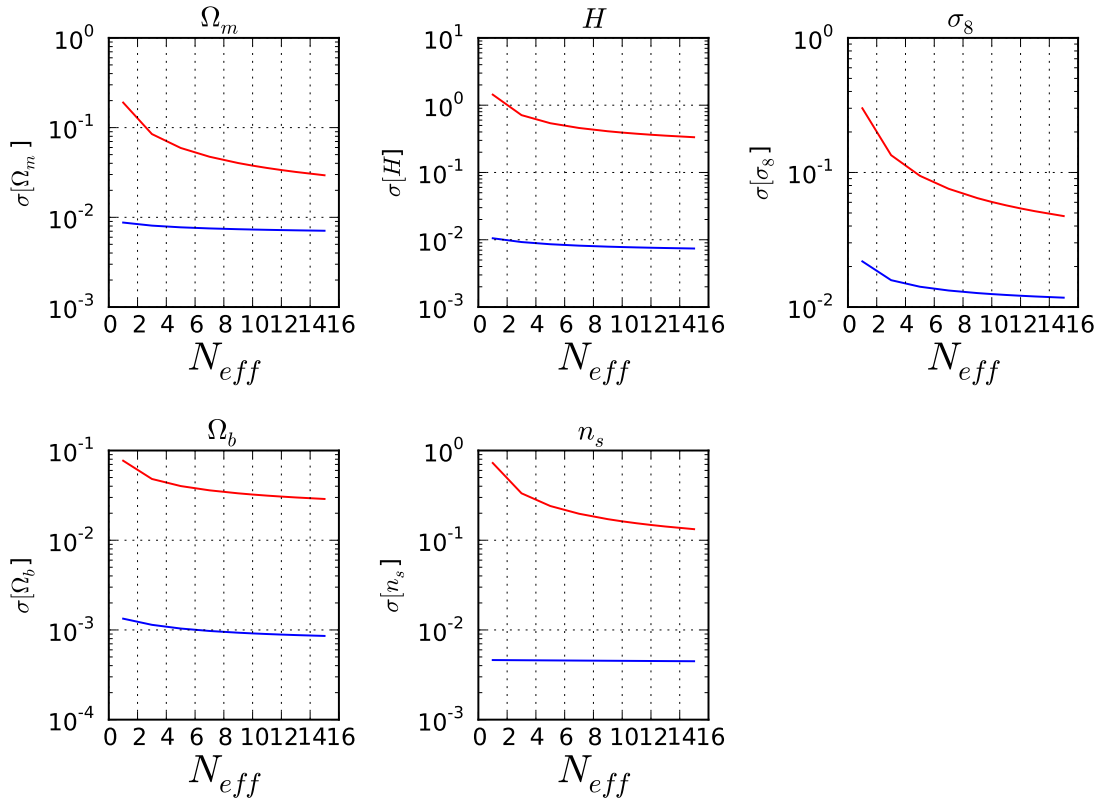

Figure 2.3: Marginalised statistical errors on $\{\Omega_m,\ H,\ \sigma_8,\ \Omega_b,\ n_s\}$ from the WL probe, computed with $l \in [1, 1024]$ in the WL case and with $l \in [10, 400]$ in the LSS case. The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

2.4.2 LSS measurements

LSS measurements allow us to constrain DE in various ways. The position of the BAO feature provides a standard ruler to study the expansion history. The shape of the angular power spectrum of the galaxy density fluctuation encapsulates precious information about the clustering amplitude and the growth of structures.

Star/galaxy misclassification affects the power spectrum measurements and the statistical error on the cosmological parameters in a similar way as in the WL case. Indeed, we can write the same equation as Eq. 2.7 for the angular power spectrum of the galaxy density fluctuations. The shot noise term is then given by $\frac{1}{N_G}$, where N_G is simply the surface density of detected galaxies. In figure 2.4, we show the evolution of the statistical errors on w_o and w_a with the density of detected galaxies, computed using the same setup as in the WL case. Figure 2.5 shows the marginalised statistical errors on the other parameters: $\{\Omega_m,\ H,\ \sigma_8,\ \Omega_b,\ n_s,\ b_g\}$ from the LSS probe.

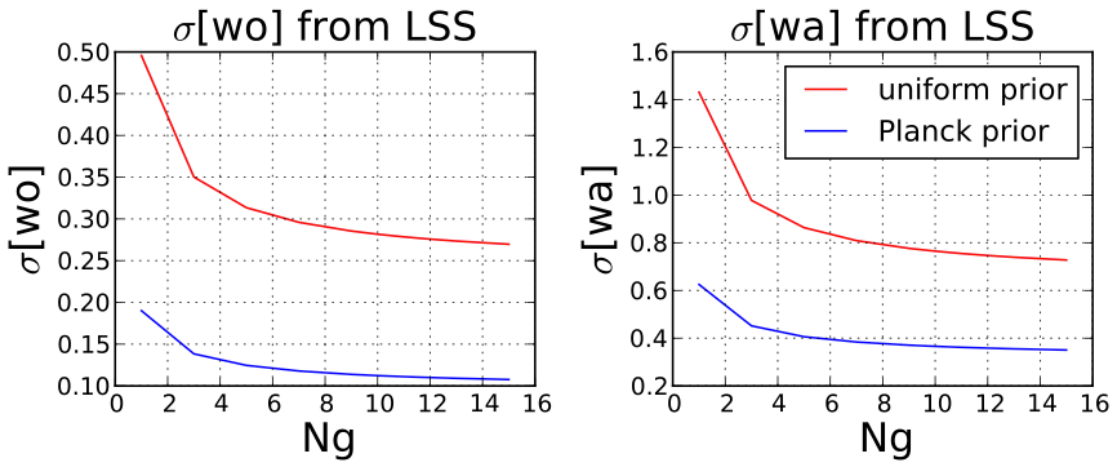

Figure 2.4: Marginalised statistical errors on the equation of state parameters w_o and w_a from the LSS probe, for different values of the density of detected galaxies N_g. The errors are marginalised over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$ and computed using the same assumptions and setup as in the WL case (see section 3.2.1), with $l \in [10, 400]$, to avoid the non linear regime and following most recent l-cuts work by the LSS community (Rassat et al. 2008). The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

In order to achieve the goals of the the LSS FoM, the 5000 sq-degrees DES survey will need to provide reliable photo-z and position measurement for about 200 millions galaxies, i.e. the number of galaxies correctly classified N_G should be higher than 11.1 per sq-arcminute (when using combined measurements from the r, i and z bandpasses). When doing the latter calculation on the truth table of DC6, for which the surface density of galaxies is $N^g_{tot} \approx 12.5$, this threshold on N_G translates into the following requirement on the galaxy completeness provided by the star/galaxy classifier: $c^g > 88.9\%$.

Note that this requirement is a necessary but not sufficient condition, as other sources of errors, apart from star/galaxy misclassification (e.g. photo-z errors), reduce the number of galaxies which can be used for LSS measurement.

2.5 Science requirements on the systematic errors

We now explore the contribution of star/galaxy misclassification as a source of *systematic* error, which need to be controlled in order for the FoM objectives to be achieved. Star/galaxy misclassifications generate a residual signal $\delta C^{sys}(l)$ in the angular power spectra (of the cosmic shear in the case of WL, and the density fluctuations of galaxies in the case of LSS), which propagates into a systematic shift $\Delta[p_\alpha]$ of the cosmological parameter p_α. We use the same formalism as in Amara & Réfrégier (2008) (see also Kirk

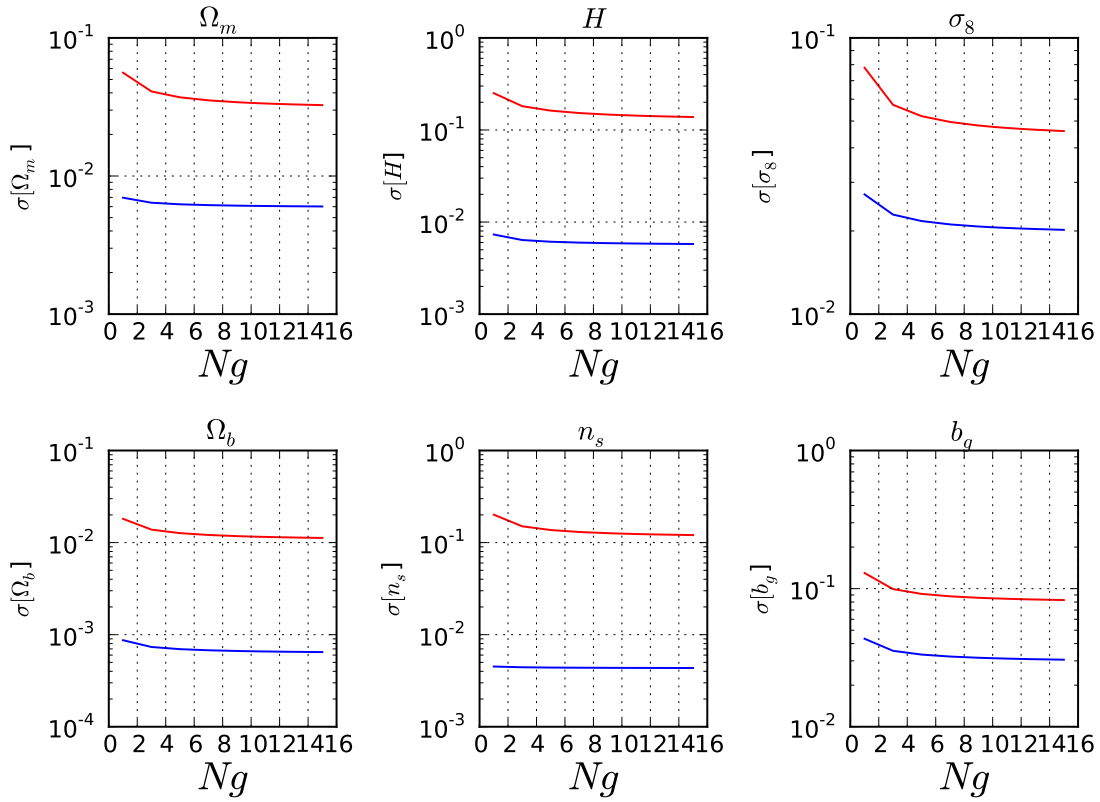

Figure 2.5: Marginalised statistical errors on $\{\Omega_m,\, H,\, \sigma_8,\, \Omega_b,\, n_s,\, b_g\}$ from the LSS probe. The errors are marginalised and computed with $l \in [10, 400]$. The red curve shows the errors computed with a non-informative prior whereas the blue curve is obtained assuming a Planck prior.

et al. 2012 and Huterer, Takada, Bernstein, & Jain 2006), to derive $\Delta[p_\alpha]$,

$$\Delta[p_\alpha] = \sum_{\beta,l,(i,j),(m,n)} (F^{-1})_{\alpha\beta} \delta C_{ij}^{sys}(l) Cov^{-1}[C_{ij}^{gal}(l), C_{mn}^{gal}(l)] \frac{\partial C_{mn}^{gal}(l)}{\partial p_\beta} \,, \quad (2.11)$$

where F^{-1} is the inverse Fisher matrix. A criterion usually used to constrain the contribution of the systematic error to the total MSE, is to define a tolerance on the systematics such that they do not dominate over statistical error. This is satisfied when

$$|\Delta[p_\alpha]| \leq \sigma[p_\alpha] \,, \quad (2.12)$$

In the following sections, we derive the systematic parameter shift for 7 cosmological

parameters $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s\}$ and the galaxy bias b_g, in the case of LSS. This allows us to translate Eq. 2.12 into requirements on the quality of the star/galaxy separation.

2.5.1 Requirement from WL measurements

In the case of WL, the systematic error $\delta C^{sys}(l)$ comes from the fact that some stars are identified as galaxies, and therefore contribute to the measured cosmic shear. We decompose the measured shear γ_m into the contribution from the true galaxies and the contamination from the misclassified stars. The galaxy shear is measured by deconvolving the observed shear and a PSF model, therefore the contamination from stars in a galaxy sample appears as a residual deconvolved shear:

$$\gamma_m = (1 - f_s)\gamma_g + f_s\gamma_{s,res} \ . \tag{2.13}$$

where $f_s = 1 - p^g$, is the stellar contamination rate (defined in Eq. 2.3) and $\gamma_{s,res}$ is the residual PSF shear, after deconvolution of the PSF model from the shape of misclassified stars. In the following analysis, we make a toy model where the residual deconvolved shears can be written as

$$\gamma_{s,res} = a\gamma_s \ , \tag{2.14}$$

where $\alpha \in [0, 1]$ and γ_s is the stellar shear. The measured two-point shear correlation function is therefore

$$< \gamma_m\gamma_m >= (1 - f_s)^2 < \gamma_g\gamma_g > +f_s^2\alpha < \gamma_s\gamma_s > \ , \tag{2.15}$$

and in terms of measured angular power spectrum, the latter equation reads

$$C^{obs}(l) = (1 - f_s)^2 C^{gal}(l) + f_s^2\alpha C^s(l) \ , \tag{2.16}$$

where $\alpha = a^2$ and where we assumed that γ_g and γ_s are uncorrelated. In practice, this is not necessarily the case. Our toy model introduces into the same term, $\alpha C_s(l)$, the auto-correlation of the residual "deconvolved star shapes" and possible cross-correlation

between them and the galaxy shear γ_g. Setting α to zero comes to neglecting both of these terms, and setting $\alpha = 1$ comes to overestimating them both. We derive the requirement on the quality of star/galaxy separation in the two limiting cases $\alpha = 1$ and $\alpha = 0$ and leave the more general case for further analysis. Equation 2.16 gives the residual systematic signal

$$\delta C^{sys}(l) = f_s^2(C^{gal}(l) + \alpha C^s(l)) - 2f_s C^{gal}(l) \ . \tag{2.17}$$

The requirement stated in Eq. 2.12 can be reformulated as a requirement on the stellar contamination rate f_s,

$$\mathcal{P}(f_s) \leq 0 \ , \tag{2.18}$$

where $\mathcal{P}$ is a second order polynomial. Indeed, when replacing $|\Delta[p_\alpha]|$ in equation 2.12 with its expression from equation 2.11 , and given the fact that f_s is positive, the constraint on f_s is: $(S)\begin{cases} -\sigma[p_\alpha] \leq T_\alpha f_s^2 - f_s \cdot 2S_\alpha \ \leq \ \sigma[p_\alpha] \\ f_s \geq 0 \end{cases}$

which comes to solve

$(S')\begin{cases} P_1(x) \ \leq \ 0 \\ P_2(x) \ \geq \ 0 \\ x \ \geq \ 0 \end{cases}$ where the two polynomial P_1 and P_2 are defined as

$$P_1(x) = T_\alpha x^2 - 2S_\alpha x - \sigma[p_\alpha] \ , \tag{2.19}$$

and

$$P_2(x) = T_\alpha x^2 - 2S_\alpha x + \sigma[p_\alpha] \ . \tag{2.20}$$

where

$$T_\alpha = \sum_\beta \sum_l \sum_{(i,j)(m,n)} (C_{ij}^g(l) + C^s(l)) Cov^{-1}[C_{ij}^g(l), C_{mn}^g(l)](F^{-1})_{\alpha\beta} \frac{\partial C_{mn}^g(l)}{\partial p_\beta} \tag{2.21}$$

and

$$S_\alpha = \sum_\beta \sum_l \sum_{(i,j)(m,n)} C_{ij}^g(l) Cov^{-1}[C_{ij}^g(l), C_{mn}^g(l)](F^{-1})_{\alpha\beta} \frac{\partial C_{mn}^g(l)}{\partial p_\beta} \tag{2.22}$$

and

$$R_\alpha = \frac{\sigma[p_\alpha]}{T_\alpha} \tag{2.23}$$

Two cases need to be distinguished:

1. If $T_\alpha > 0$, the solutions of (S') is $\begin{cases} f_s & \in & [f_{u-}, f_{u+}] \\ f_s & \geq & 0 \end{cases}$

   where

$$f_{u\pm} = \frac{S_\alpha}{T_\alpha} \pm \sqrt{\left(\frac{S_\alpha}{T_\alpha}\right)^2 + R_\alpha} \tag{2.24}$$

2. If $T_\alpha < 0$, the solutions of (S') is $\begin{cases} f_s & \in & [f_{u-}, f_{u+}] \\ f_s & \geq & 0 \end{cases}$

   where

$$f_{u\pm} = \frac{S_\alpha}{T_\alpha} \pm \sqrt{\left(\frac{S_\alpha}{T_\alpha}\right)^2 - R_\alpha} \tag{2.25}$$

As $f_s = 1 - p^g$, in terms of purity, the constraint reads (S) $\begin{cases} p^g \in [1 - f_{u+}, 1 - f_{u-}] \\ p^g \leq 1 \end{cases}$

In this analysis, we assumed that f_s is constant for all redshift tomographic bins. This assumption is violated if the redshifts are correctly measured. Indeed, since the observed stars have low redshift, the amount of true stars labelled as galaxies - and therefore f_s- should quickly drop down as the redshift increases. As a result, only part of the redshift bins taken into account in the above analysis actually contribute to he true $\Delta[p_\alpha]$. The above analysis would lead to a pessimistic constraint, since it overestimates T_α and S_α and underestimate f_{u+}, which means the lower threshold on p^g is overestimated.

In practice, photometric redshift errors lead stars to be erroneously assigned higher redshift than their true redshift, sometimes resulting in the presence of several peaks in

the redshift distribution of stars. Future work including a careful study of the evolution of f_s with redshift, should allow to refine the simplistic assumption of constant f_s for all redshift tomographic bins.

The other assumptions made to solve Eq. 2.18 are detailed below. We use the setup detailed in section 2.4.1 to compute the Fisher matrix and the marginalised statistical errors $\sigma[p_\alpha]$ on the cosmological parameters. To estimate $C^s(l)$ in Eq. 2.17, we assume it is the sum of a "shot noise" term and a term due to the correlation of stellar shapes across the field of view,

$$C^s(l) = C^s_{noise} + C^s_{tile}(l) \tag{2.26}$$

We measure $C^s_{tile}(l)$, the power spectrum of the shapes of the stars in DC6, using the same code as in Jarvis et al. (2004). The "shot noise" term is given by

$$C^s_{noise} = \frac{\sigma_s^2}{N^s_{tot}} \tag{2.27}$$

where $N^s_{tot} = N_s + M_s$ (see section 3.1.1) is the density of stars and the ellipticity of stars, σ_s, is taken as the ellipticity of the PSF. To estimate σ_s, we use the *whisker length*. Given I_{xx}, I_{yy} and I_{xy}, the second moment of the light intensity from an object in x, y coordinates, a measure of the ellipticity of the light distribution is given by $e = (I_{xx} - I_{yy})(I_{xx} + I_{yy})$. The whisker length is then defined as $w \approx \sqrt{e(I_{xx} + I_{yy})} = \sqrt{e} \cdot r_{psf}$, where r^2_{psf} is given by $(FWHM)/2.35$. $FWHM$ designates the full width at half maximum and is given by $FWHM \approx 0.94$ in DES. In addition, the hardware has been designed with a requirement on the whisker length to be lower than a threshold value of 0.2" in the r, i and z band, which we take as an estimation of $whisk$. We get $C^s \approx 1.3187 \cdot 10^{-8}$ sr. Equation 2.18 translates into a constraint on the stellar contamination $f_s < f_{s,lim}$.

Here we consider the two limiting cases $\alpha = 0$ and $\alpha = 1$ and derive the lower bounds for f_s corresponding to each of these cases, referred to as $f_{s,\alpha=0}$ and $f_{s,lim,\alpha=1}$. The true lower bound is in the interval corresponding to these limiting cases: $f_{s,lim} \in [f_{s,lim,\alpha=1}, f_{s,lim,\alpha=0}]$. In figures 2.6 and figure 2.7, we show the limiting cases $\alpha = 0$ and $\alpha = 1$ respectively. We plot the two terms of the total error $MSE[p_\alpha]$ (see equation 2.1), i.e. the systematic parameter shift $\Delta[p_\alpha]$ due to star/galaxy misclassification, and the statistical error $\sigma[p_\alpha]$, for different values of the stellar contamination f_s and for each of the cosmological parameters of our model $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s\}$. Within an experi-

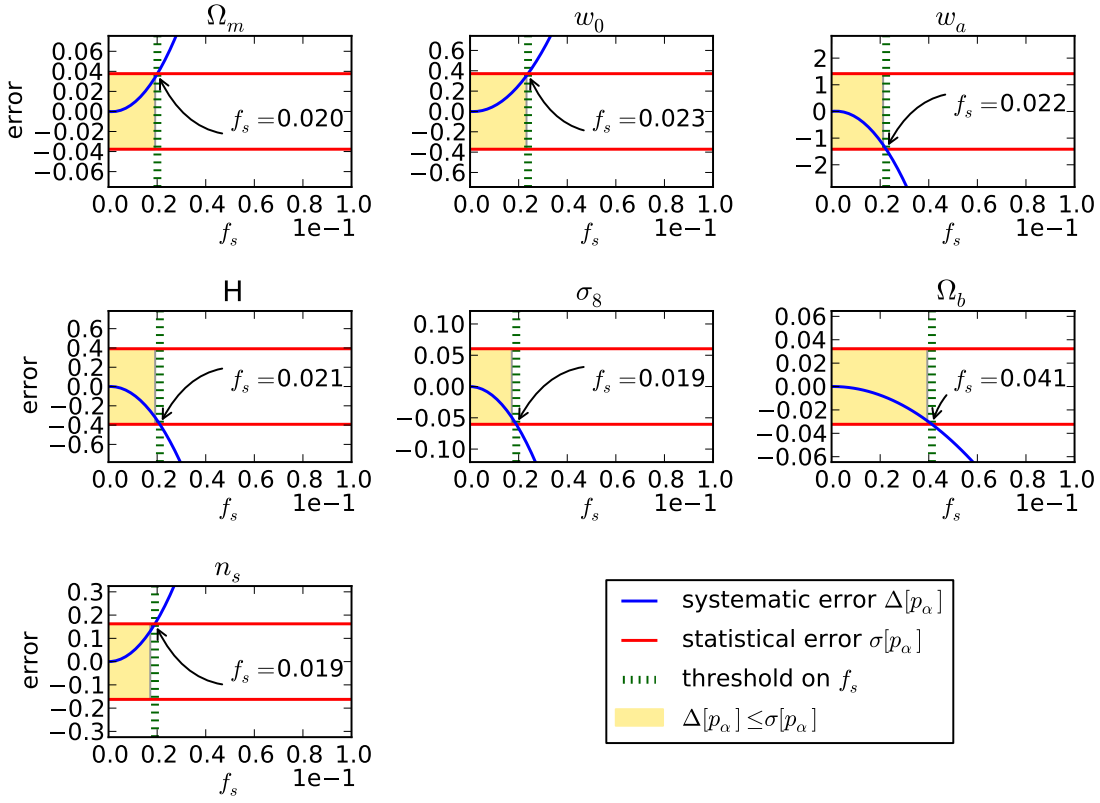

Figure 2.6: Marginalised statistical error σ (red line) and systematic parameter shift Δ (blue curve) from the WL probe, for different values of the stellar contamination f_s allowed by the star/galaxy classifier and in the limiting case $\alpha = 1$. Both σ and Δ are marginalised over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s\}$ and are computed using the setup described in section 2.4.1. The yellow area corresponds to the values of f_s for which the requirement on the systematic errors is achieved, i.e. it does not dominate over the statistical error. This requirement translates into a threshold on f_s, indicated by the green line. Unlike LSS measurements, WL measurements are not sensitive to the galaxy bias b_g, which is the reason why it does not appear above.

ment designed to constrain DE such as DES, the constraints on the quality of star/galaxy separation comes from the need to control the errors on w_o and w_a. This being said, one should keep in mind that the contamination from stars affects the precision on the measurements of other cosmological parameters, as shown in figure 2.6.

For the equation of state parameters w_o and w_a, we find that we require $f_s \leq f_{s,lim}$ with $f_{s,lim,\alpha=0} = 0.122$ and $f_{s,lim,\alpha=1} = 0.022$ (requirement driven by w_a). This translates into the following requirement on $p^g = 1 - f_s$, the purity provided by the star/galaxy classifier: $p^g \geq p^g_{lim}$ with $p^g_{lim} \in [87.7\%, 97.8\%]$. To refine this requirement, we now allow α to vary. In figure 2.8, we show the evolution of p^g_{lim} when varying α and when considering

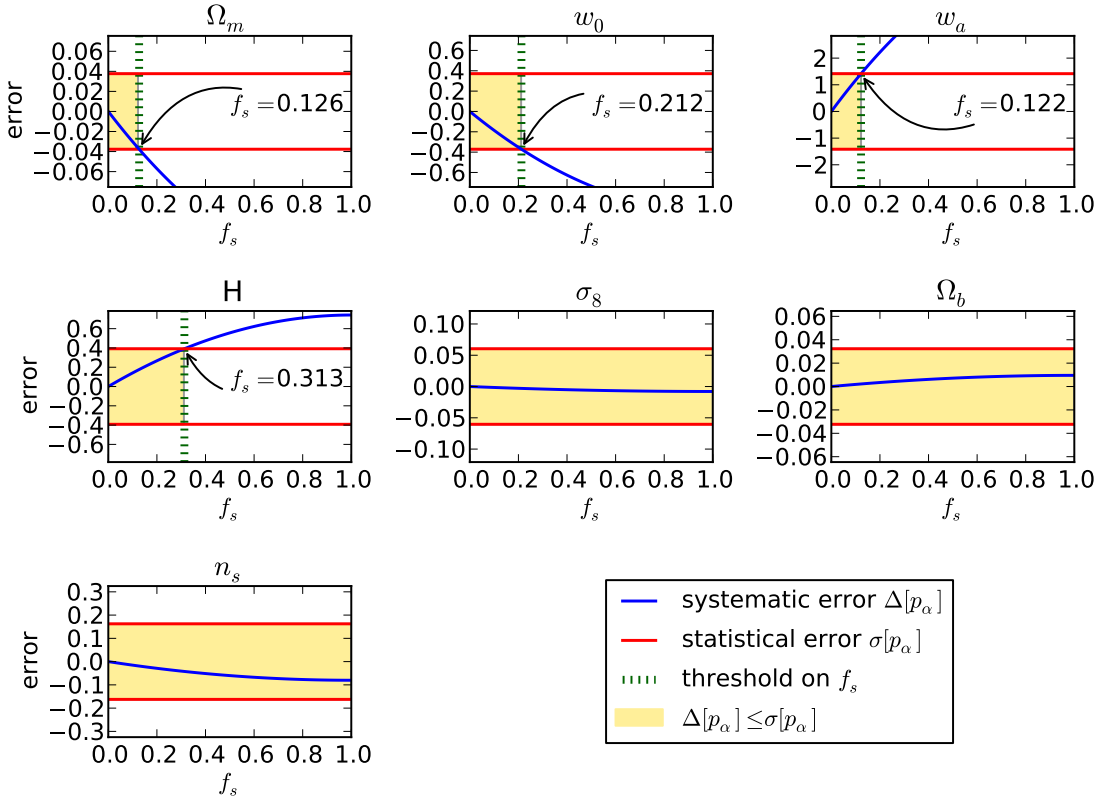

Figure 2.7: Marginalised statistical error σ (red line) and systematic parameter shift Δ (blue curve) from the WL probe, for different values of the stellar contamination f_s allowed by the star/galaxy classifier and in the limiting case $\alpha = 0$. Both σ and Δ are marginalised over $\{\Omega_m, H, \sigma_8, \Omega_b, n_s\}$ and are computed using the setup described in section 2.4.1. The yellow area corresponds to the values of f_s for which the requirement on the systematic errors is achieved, i.e. it does not dominate over the statistical error. This requirement translates into a threshold on f_s, indicated by the green line. Unlike LSS measurements, WL measurements are not sensitive to the galaxy bias b_g, which is the reason why it does not appear above.

the requirement on the parameters w_o and w_a. The threshold is driven by w_a (since the requirement to constrain the bias on w_a leads to a more stringent value of p^g_{lim}). The value of p^g_{lim} quickly grows, from the limiting case $\alpha = 1$. From $\alpha = 0.4$, p^g_{lim} grows slower and stays above 96%.

2.5.2 Requirement from LSS measurement

Like for the WL probe, achieving the objectives of the LSS FoM requires the systematic error induced by star/galaxy misclassification to be smaller than the statistical error on w_o and w_a, and we can rewrite Eq. 2.12 in the case of LSS measurements. The shape

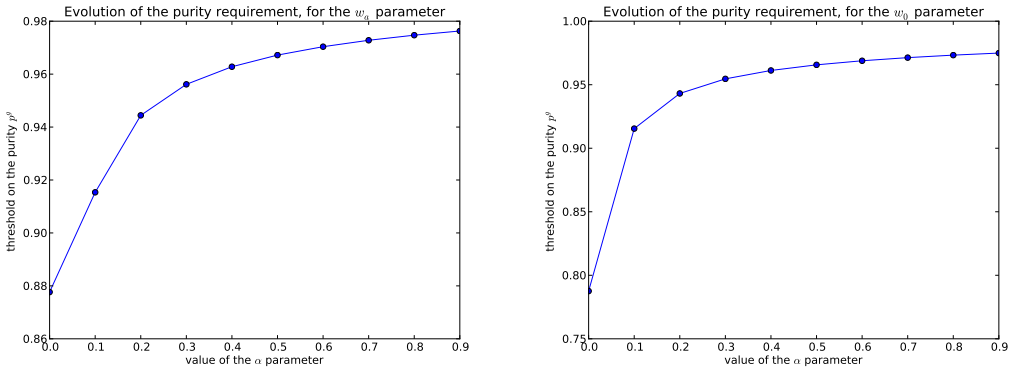

Figure 2.8: Evolution with the coefficient α of the value of p^g_{lim}, from the constraint on the bias of the equation of star parameter w_a (left) and w_o (right).

of the residual systematic signal due to star/galaxy misclassification, δC^{sys}, is obtained following the same methodology as in the WL case, by decomposing the measured density fluctuation into the contribution from the true galaxies and the contamination from the stars identified as galaxies,

$$\delta_m = (1 - f_s)\delta_g + f_s\delta_s \ . \tag{2.28}$$

Replacing the shear angular power spectrum with the density fluctuation angular power spectrum like in Eq. 2.17:

$$C^{obs}(l) = (1 - f_s)^2 C^{gal}(l) + f_s^2 C^s(l) \ , \tag{2.29}$$

we get the same requirement on the stellar contamination rate f_s as in Eq. 2.18. To estimate $C^s(l)$, we use the same stellar catalogue as used for the DES simulated sky survey produced by Busha et al. (2013). We then calculate $C^s(l)$ using the approach from Thomas et al. (2010) and an adaptation of the HEALPix code (Górski et al. 2005). We estimate the $C_{ij}(l)$ and $\frac{\partial C^g_{mn}(l)}{\partial p_\alpha}$ terms using the same code and setup as for the WL case. Figure 2.9 shows the systematic parameter shift induced by the stellar contamination, for each of the cosmological parameters of our model $p_\alpha = \{w_o, w_a, \Omega_m, H, \sigma_8, \Omega_b, n_s, b_g\}$. In particular, for the equation of state parameters w_o and w_a, we find that we require $f_s \leq 0.015$. This translates into the following requirement on $p^g = 1 - f_s$, the purity provided by the star/galaxy classifier: $p^g \geq 98.5\%$.

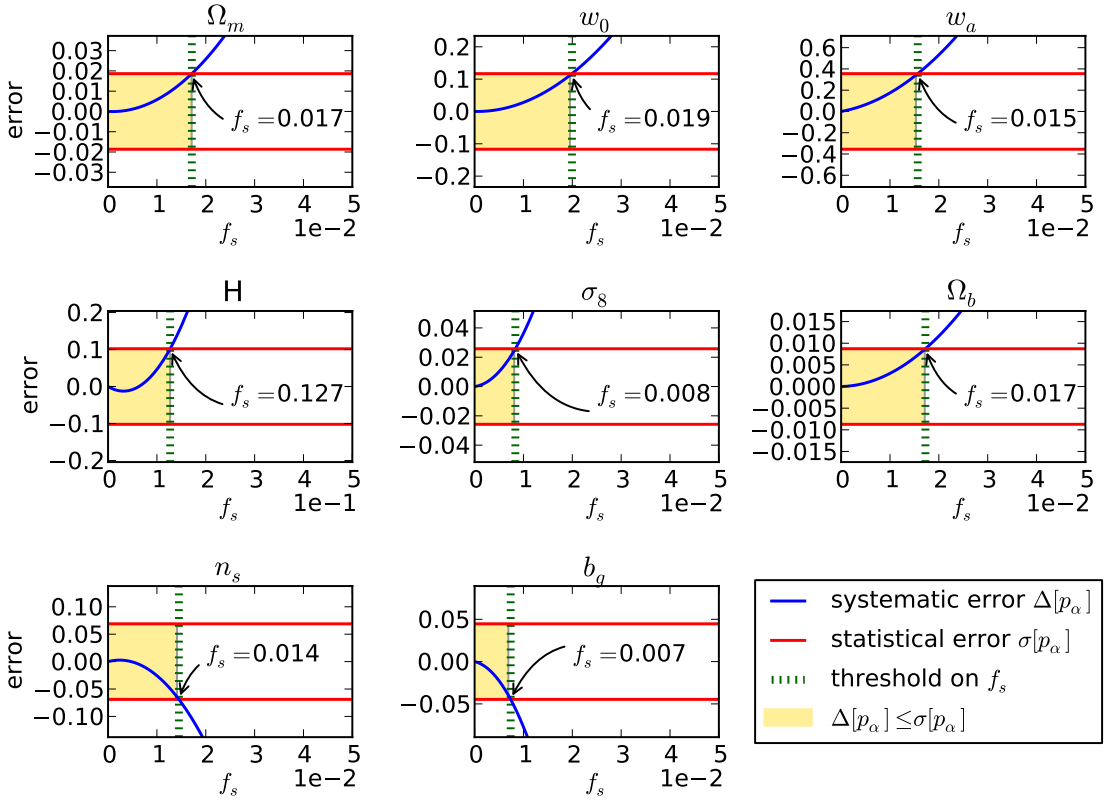

Figure 2.9: Marginalised statistical error σ (red line) and systematic parameter shift Δ (blue curve) from the LSS probe, for different values of the stellar contamination f_s allowed by the star/galaxy classifier. Both σ and Δ are marginalised over $\{\Omega_m,\ H,\ \sigma_8,\ \Omega_b,\ n_s,\ b_g\}$ and are computed using the setup described in section 2.4.1, with $l \in [10, 400]$, to avoid the non linear regime. The yellow area corresponds to the values of f_s for which the requirement on the systematic errors is achieved, i.e. it does not dominate over the statistical error. This requirement translates into a threshold on f_s, indicated by the green line. Unlike WL measurements, LSS measurements are sensitive to the galaxy bias b_g, as shown on the last panel.

The requirement on star/galaxy separation in a DE experiment is dictated by the need to accurately measure w_o and w_a. This being said, figure 2.9 demonstrates that these two parameters are not the most sensitive to the contamination by stars, which we leave for further analysis.

2.6 Stellar PSF calibration for WL

In this section, we derive two additional requirements on the quality of the star/galaxy separation, from calibration constraints specific to the WL probe. The measured shapes of galaxies include a component due to the PSF of the combined telescope, atmosphere,

and instrument which is correlated among galaxies. Removing this contribution requires careful measurement of the PSF, which is done using isolated stars. Therefore, additional requirements on star/galaxy separation come from PSF calibration for WL.

2.6.1 Requirement on c^s

In order to determine the interpolation pattern of the PSF, one needs to find enough stars to adequately cover the area of the CCD chip. Based on preliminary studies of the DES science verification data, we believe that between 100 and 200 stars per DES CCD is enough to adequately cover the area of the CCD chip and determine the interpolation pattern of the PSF. This requirement can be written as

$$N_{CCD} > \frac{200}{S_{CCD}} , \tag{2.30}$$

where S_{CCD} is the surface of a CCD. Each DECam CCD contains $2048 * 4096$ pixels, each covering 0.27arcsec. Assuming a constant surface density N_{CCD} of stars over the different CCD (i.e. $N_{CCD} = N_S$), the later equation translates into

$$c^S > \frac{200}{S_{CCD} * (N_S + M_S)} , \tag{2.31}$$

From the truth tables, we know the number of true stars $N_S + M_S$, which allows us to derive the technical constraint on the completeness of the stars samples: $c^s \geq 25\%$. In this analysis, we assumed that all stars can be used for PSF estimation. In practice, the latter lower limit on the completeness could be more stringent because of several reasons:

- Only non-saturated (i.e. I magnitude fainter than 15) stars with a signal-to-noise higher than 50 are used, in DES, for PSF calibration.
- Detector non-linearities lead to the "blooming" effect, or the "brighter-bigger" effect. Indeed the voltages induced by the photons reaching the detector, leads brighter objects to appear larger than faint objects. This effect can lead to variations of the PSF between bright and faint stars, and therefore affects the PSF calibration.

Both of these constraints reduce the number of stars available for PSF calibration. This is the reason why we made our calculation assuming the hard constraint of 200 stars per

CCD, whereas 100 stars is enough, in practice (the $c^s \geq 25\%$ requirement allows PSF calibration, even if only half of the catalog stars have flux and signal-to-noise that make them usable for PSF calibration).

2.6.2 Requirement on p^s

The upper limit on the contamination in a sample of stars comes from the fact that galaxies misclassified as stars will bias the inferred PSF, which in turn will bias the galaxy shapes. We use a toy model to estimate the bias on the shear estimate as a function of $f_g = M_g/(N_s + M_g)$, the galaxy contamination rate in the sample of stars.

We first consider the sample of objects classified as stars, used for the calibration of the PSF. Such a sample contains two types of objects: true stars and true galaxies which have been misclassified as stars. The PSF model derived from this sample can be approximated as the weighted average of both types of objects:

$$\chi^{psf}_{biased}(f_g) = f_g \chi^{mis,gal} + (1 - f_g)\chi^{psf}_{true} \,, \tag{2.32}$$

where χ is the polarisation, and is related to the observed major and minor axis a and b of the image produced by a circular source via

$$|\chi| = \frac{a^2 - b^2}{a^2 + b^2} \,, \tag{2.33}$$

and to the shear and convergence fields via

$$\chi = \frac{2\gamma(1 - \kappa)}{(1 - \kappa)^2 + |\gamma|^2} \,, \tag{2.34}$$

so that $|\chi| \approx 2|\gamma|$

We now consider a sample of galaxies of which we would like to measure the shear. The observed polarisation χ^{obs}, i.e. the polarisation after convolution with the PSF model, is linked to the true polarisation of a galaxy through the following relation (Viola et al. 2014):

$$\chi^{obs}_{gal} = \frac{\chi^{true}_{gal}}{1 + 1/R} + \frac{\chi^{psf}_{true}}{1 + R} \,. \tag{2.35}$$

The *resolution* R in the above equation is the ratio of the galaxy to PSF size. In the absence

of misclassified galaxy contaminating the sample used to measure χ_{psf} (and neglecting the other sources of errors in the PSF calibration), the measured polarisation is:

$$\chi_{gal}^{true} = (1 + 1/R) \left(\chi_{gal}^{obs} - \frac{\chi_{true}^{psf}}{1 + R} \right) . \tag{2.36}$$

However, the contamination from galaxies biases the PSF model, and the measured galaxy polarisation is rather

$$\chi_{gal}^{measured} = (1 + 1/R') \left(\chi_{gal}^{obs} - \frac{\chi_{biased}^{psf}}{1 + R'} \right) , \tag{2.37}$$

where χ_{biased}^{psf} is given by equation 2.32

As a result, the measured polarisation ca be written as

$$\chi_{gal}^{measured} = (1 + m)\chi_{gal}^{true} + c , \tag{2.38}$$

where

$$m = \frac{R/R' - 1}{R + 1} , \tag{2.39}$$

and

$$c = \left(1 + \frac{1}{R'} \right) \left(\frac{\chi_{psf}^{true}}{1 + R} - \frac{\chi_{psf}^{biased}}{1 + R'} \right) \tag{2.40}$$

The same relation can be written for the shear:

$$\gamma_{gal}^{measured} = (1 + m')\gamma_{gal}^{true} + c' , \tag{2.41}$$

where $m' = m$ and $c' = 2c$.

Here, we use the SExtractor parameters A_{image} and B_{image} to estimate the typical polarisations of the different objects in the DES Science Verification (SV) data. In particular, we compute χ_{psf}^{biased} as

$$\chi_{biased}^{psf}(f_g) = f_g \chi^{mis,gal} + (1 - f_g)\chi_{true}^{psf} , \tag{2.42}$$

where

- $\chi^{psf}_{true} \approx 2\gamma_{psf}$ and γ_{psf} is estimated as the ellipticity of the PSF,
- $\chi^{mis,gal} = \frac{A^2_{image}-B^2_{image}}{A^2_{image}+B^2_{image}}$, for the misclassified galaxies.

In figure 2.12, we show the true and biased PSF polarisation χ^{psf}_{true} and χ^{psf}_{biased}, as a function of the contamination from galaxies f_g. As the contamination from galaxies grows, the measured PSF polarisation departs from the true PSF polarisation, and approaches the polarisation $\chi^{mis,gal}$, of the misclassified galaxies which contaminate the stellar sample.

We use the SExtractor parameter $Flux\_Radius$ to compute R and R' as

$$R = \frac{\overline{Flux\_Radius^{gal.}}}{\overline{Flux\_Radius^{stars}}} \tag{2.43}$$

where $Flux\_Radius^{stars}$ is the $Flux\_Radius$ parameters for the true stars, and

$$R' = \frac{\overline{Flux\_Radius^{gal.}}}{\overline{Flux\_Radius^{stars+gal_{mis}}}} \tag{2.44}$$

where $Flux\_Radius^{stars+gal_{mis}}$ is the $Flux\_Radius$ parameter for all the objects in the sample labelled as stars (i.e. true stars and misclassified galaxies). The $Flux\_Radius$ parameters for the stars, galaxies, and all objects classified as stars (i.e. true stars and misclassified galaxies) is shown in figure 2.10. As f_g grows, more and more misclassified galaxies contribute to $\overline{Flux\_Radius^{stars+gal_{mis}}}$, the average size of all the objects classified as stars grows, which explains why $\overline{Flux\_Radius^{stars+gal_{mis}}}$ grows. The computed resolutions R and R' are shown in figure 2.11. R is larger than 1, because the average size of galaxies is larger than the average size of stars, as shown in figure 2.10. Similarly, R' is larger than 1, because the average size of galaxies is larger than the average size of the objects classified as stars (stars and misclassified galaxies), as shown in figure 2.10. As f_g grows, $\overline{Flux\_Radius^{stars+gal_{mis}}}$ grows, which explains why R' decreases when f_g increases.

Using equations 2.36 and 2.42, we can compute $\chi^{measured}_{gal}$ and χ^{true}_{gal}, shown in figure 2.13, as well as the multiplicative and additive biases, m and c, shown in figure 2.14. Previous work by the DES collaboration led to the formulation of requirements on the value of m and c, which we also show on figure 2.14. These requirements translate into requirements on the contamination from galaxies. In particular, the requirement $m < 0.004$ translates into $f_g < f_{g,lim}$ with $f_{g,lim} \in [0.06, 0.09]$ and therefore $p^s > p^s_{lim}$ with $p^s_{lim} \in [91\%, 94\%]$. The requirement on the additive bias parameter $c < 8 \cdot 10^{-4}$

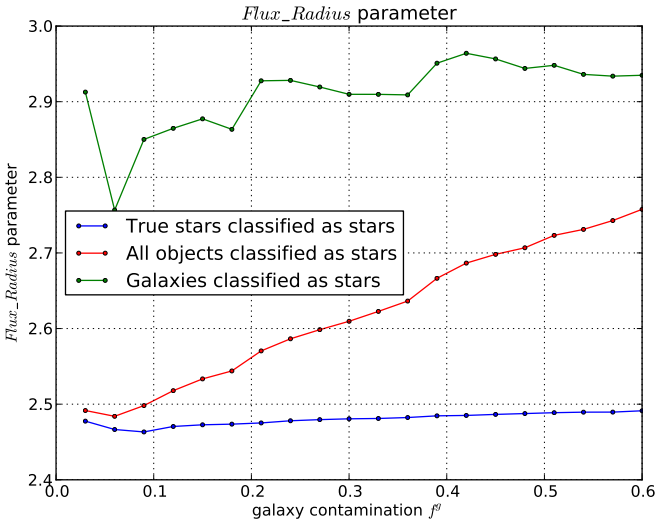

Figure 2.10: Mean *Flux_Radius* parameter for stars (blue), galaxies (green), and for all the objects classified as stars (i.e. true stars and misclassified galaxies) (red). As f_g grows, more and more misclassified galaxies contribute to $\overline{Flux\_Radius}^{stars+gal_{mis}}$, the average size of all the objects classified as stars grows.

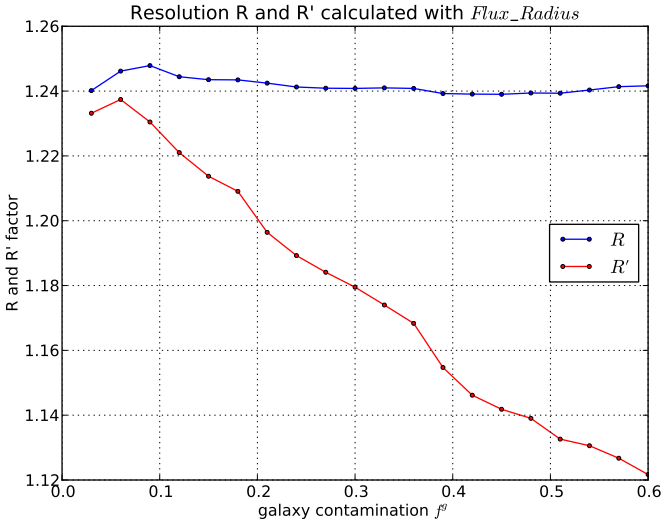

Figure 2.11: Resolution factors, R and R'. As the number of misclassified galaxies grows, the average size of all the objects classified as stars grows. This explains why R' decreases when f_g increases. Both R and R' are larger than 1, because the average size of galaxies is larger than the average size of stars and misclassified galaxies.

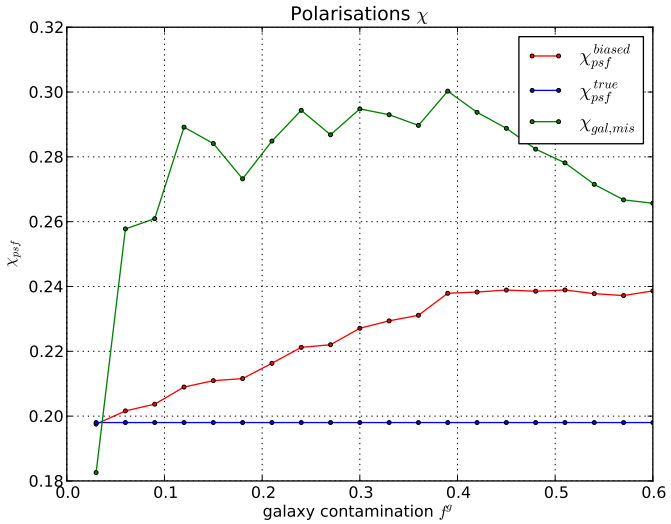

Figure 2.12: True PSF polarisation χ^{psf}_{true} (blue), polarisation of the misclassified galaxies $\chi^{mis,gal}$ (green) and biased measured PSF polarisation χ^{psf}_{biased} (red), as a function of the contamination from galaxies f_g. As the contamination from galaxies grows, the measured PSF polarisation departs from the true PSF polarisation, and approaches the polarisation of the misclassified galaxies which contaminate the stellar sample.

leads to a more stringent requirement on the contamination: $f_{g,lim} \in [0.03, 0.06]$, i.e. $p^s_{lim} \in [94\%, 97\%]$.

In practice, shear codes have the ability to sharpen the classification of stars and galaxies. Indeed, a shear measurement code convolves a model for the galaxy with the measured PSF function, and then adjusts the parameters of this model to best fit the observed data. If the best-fit values for the parameters characterising the size of the model are too small, then it is likely that the observed object is a star (or a very small galaxy). This allows to perform additional cuts of the sample of objects, using the output of the shear measurement code as an additional indication about the class of the object. For this reason, using the derived verbatim as a requirement on the star/galaxy separation is conservative.

2.7 Alternative approach: f_s and σ_s as nuisance parameters

The previous approach consisted of solving equation 2.18, to derive a constraint on the stellar contamination f_s. An alternative approach consists of considering f_s and σ_s as nuisance parameters and marginalising over them in the Fisher analysis. Indeed, f_s and σ_s

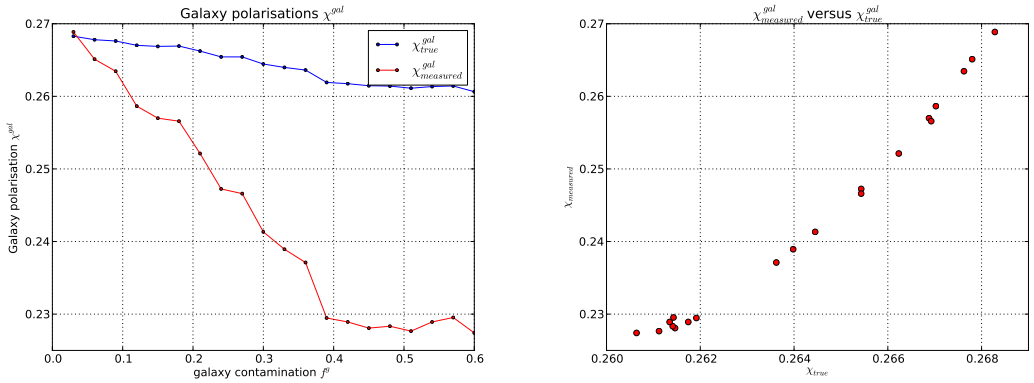

Figure 2.13: The left panel shows the true galaxy polarisation χ^{true}_{gal} (blue) and the biased measured galaxy polarisation $\chi^{measured}_{gal}$ (red), as a function of the contamination from galaxies f_g. As the contamination from galaxies grows, the measured galaxy polarisation departs from the true galaxy polarisation. The right panel shows $\chi^{measured}_{gal}$ versus χ^{true}_{gal}.

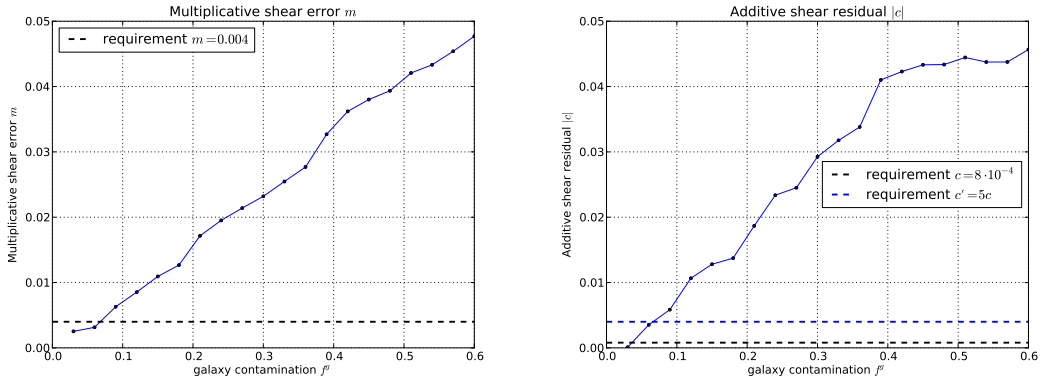

Figure 2.14: Multiplicative bias m (left) and additive biase c, m and c, shown in figure 2.14. Previous work by the DES collaboration led to the formulation of requirements on the value of m and c. These requirements translate into requirements on the contamination from galaxies. In particular, the requirement $m < 0.004$ translates into $f^g < 0.3$ and therefore $p^s > 70\%$. We show two example of requirements on c. The "conservative" requirement set by the DES collaboration leads to a stringent requirement on the contamination: $f_g \in [0.03, 0.06]$, i.e. $p^s \in [94\%, 97\%]$

enter the Fisher matrix through their contribution to the projected power spectrum $C^{ons}(l)$ defined in equations 2.17 and 2.29 for the WL probe and the LSS probe respectively. The dark energy FoM is defined as the reciprocal of the area of the error ellipse enclosing 95% confidence limit in the w_o-w_a plane. It is linked to the Fisher Matrix (FM) through the Cramer-Rao inequality (equation 2.8). The raw FM represents the ability of the probes in question to constrain cosmology and nuisance parameters. In addition we can include a prior which encodes our knowledge about those parameters before a particular experiment is undertaken. A wider prior means we are more ignorant about that parameter.

Here we calculate the degradation of a prior dark energy FoM of 0.1, for several values of the width of a gaussian prior on f_s and σ_s, for both the LSS probe and the WL probe. For simplicity, we vary the width of both the prior simultaneously, and set $\alpha = 1$ in equation 2.17. The results are shown in figure 2.15. The survey requirements set a threshold on the value of the degradation FoM/FoM_{prior}, (i.e. a horizontal line on figure 2.15), which in turn translates into a constraint on the width of the gaussian prior on f_s, i.e. a into a requirement on the quality of star/galaxy classification. For example, for the WL probe, a width of 7.5% of the gaussian prior $f_s = 10\%$ is the maximum allowed width, if we tolerate a 5% degradation of the prior FoM.

2.8 Concluding remarks

The requirements on the quality of the star/galaxy separation derived in this section are summarised in table 2.2.

A dedicated sample of stars is only needed when calibrating the PSF. Therefore, the two requirements on the samples of stars are only required for WL science. As far as samples of galaxies are concerned, LSS science requires purer samples than WL science. This is due to star contamination affecting the corresponding measured "observable" in different ways. The contribution of misclassified stars to the measured shear is dominated by the shot noise term (see Eq. 2.26), which is approximately scale independent, whereas they mimic a l-dependent density fluctuation of galaxies and therefore contribute to the LSS measurement in a more complicated way. On the other hand, WL requires a more complete samples of galaxies. This is because a "usable" object means something different for LSS and WL. In order to be usable for LSS measurement, a galaxy needs to be detected

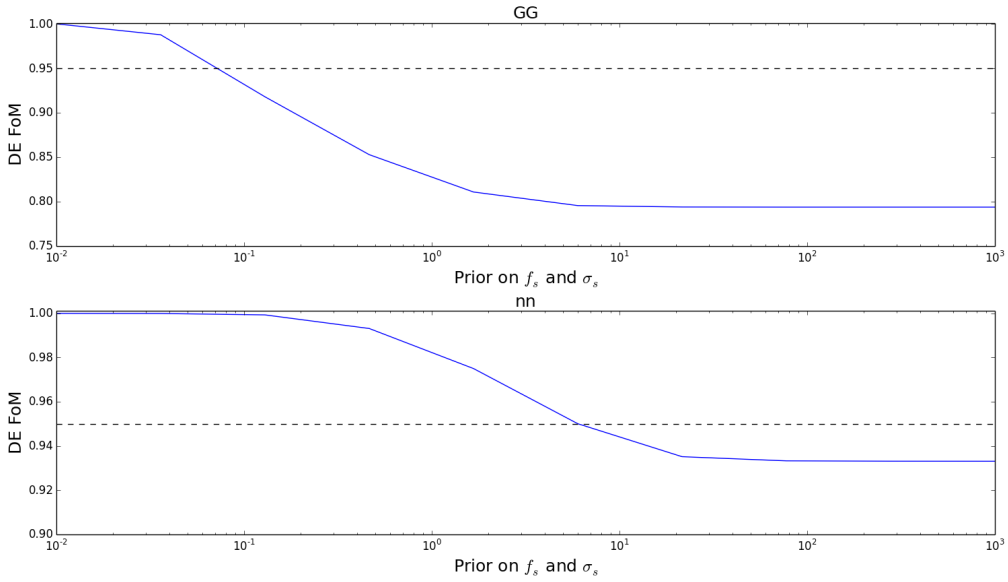

Figure 2.15: Evolution of the dark energy FoM (the reciprocal of the area of the error ellipse enclosing 95% confidence limit in the w_o-w_a plane) with the width of a gaussian prior of f_s and σ_s, for the WL probe (top panel) and the LSS probe (lower panel). The FoM is normalized to the prior FoM. A requirement on the FoM can be translated on a requirement on the width of the gaussian prior on f_s, i.e. on the quality of the star/galaxy classifier. For example, the horizontal lines correspond to a 5% degradation of the prior FoM.

Table 2.2: Summary of the science requirements on the quality of star/galaxy separation.

	requirement from LSS	**from WL**
p^g	$\geq 98.5\%$ (requirement on the systematic error)	$\geq p^g_{lim}$, with $p^g_{lim} \in [87.7\%, 97.8\%]$ (requirement on the systematic error)
p^s	-	$\geq p^s_{lim}$, with $p^s_{lim} \in [94\%, 97\%]$ (requirement on the PSF calibration)
c^g	$> 88.9\%$ (requirement on the statistical error)	$> 96.0\%$(requirement on the statistical error)
c^s	-	$\geq 25\%$ (requirement on the PSF calibration)

with a reliable photometric redshift but WL also needs the shape of the galaxy to be measurable.

In this chapter, we showed that star/galaxy misclassification contributes to both the statistical and systematic error on the measurement of cosmological parameters. In particular, it affects the measurement of the DE equation of state parameters, w_o and w_a, which future large photometric surveys such as DES aim to measure accurately. In the case of WL and LSS measurements, we translated the DETF FoM requirements on the statistical and systematic errors and the constraints from PSF calibration into the corresponding science requirements on the quality of star/galaxy separation. We formulated these requirements using two parameters: the purity and completeness of classified samples of stars and galaxy. In the next chapter, we will use these requirements to assess the performance of a new classifier, multi_class, and compare it to other classifiers currently used in galaxy surveys.

Chapter 3

Desiging a new tool for Star/Galaxy separation

"Artificial Intelligence is defined as the opposite of natural stupidity."
Woody Allen.

In this chapter, we aim to achieve the science requirements on star/galaxy separation defined in chapter 2. We present both the existing tools for star/galaxy separation, and a new method we designed and tested for the DES. In section 3.1, we summarise the current methods for star/galaxy classification and the motivations for our multi-parameter approach. The details of our method are presented in section 3.2. In section 3.3, we compare our star/galaxy classification tool to the ones provided by other methods and confront these results to the science requirements derived in chapter 2. Finally, we summarise our main conclusions in section 3.4. The results shown in this chapter are using the DC6 DES catalog, which was presented in section 2.2 of chapter 2.

3.1 Current tools for star-galaxy Separation

Different strategies have been adopted to classify stars and galaxies in large sky surveys. The morphometric approach (e.g. Kron 1980; Yee 1991; Vasconcellos et al. 2011; Sebok 1979; Valdes 1982) relies on the separation of point sources (the ones most likely to be stars) from resolved sources (presumably galaxies).

This approach is challenged at the faint magnitudes reached by the next generation of wide-field surveys, including the DES, due to the vast number of poorly resolved galaxies.

Another strategy consists of using training algorithms. Machine learning distinguishes several types of learning strategies, Artificial Neural Network (ANN) being one successfully implemented example of *supervised learning*. ANN has previously been applied to the star/galaxy separation problem (e.g. Odewahn et al. 1992, Naim 1995, Bertin & Arnouts 1996, Oyaizu et al. 2008). Indeed, star/galaxy separation shares with many other classification problems the three criteria which usually make neural computing applications particularly successful:

- The task is well-defined in that we know precisely what we want, i.e. classify objects in two distinct classes.
- There is a sufficient amount of data available to train the network to acquire a useful function based on what it should have done in these past examples.
- No simple parametrization for the output (the class of the object) as a function of the input (the parameters derived from the images) is known, and we would like to leave it to the algorythm to determine the optimal classification scheme.

Other supervised classifiers, such as Support Vectors Machine (SVM), have been more recently used for the star/galaxy separation problem, as well as unsupervised tools such as Hierarchical Bayesian techniques (e.g. Fadely et al. 2012).

Throughout this section, we will use the following notations to define:

- *classification tools* - class_star; spread_model and multi_class
- *classification output* - X_{class_star}; X_{spread_model} and X_{multi_class} .

As described below (in sections 3.1.1 and 3.1.2), class_star and spread_model are two classifiers currently implemented in *SExtractor* (Bertin & Arnouts 1996) and in the next sections we present a new method for star/galaxy separation called "multi_class", designed to achieve the science requirements derived in chapter 2 at the faint magnitudes reached by DES.

Both the morphometric and the training approaches are implemented in *SExtractor* (Bertin & Arnouts 1996), with two classifiers, class_star and spread_model.

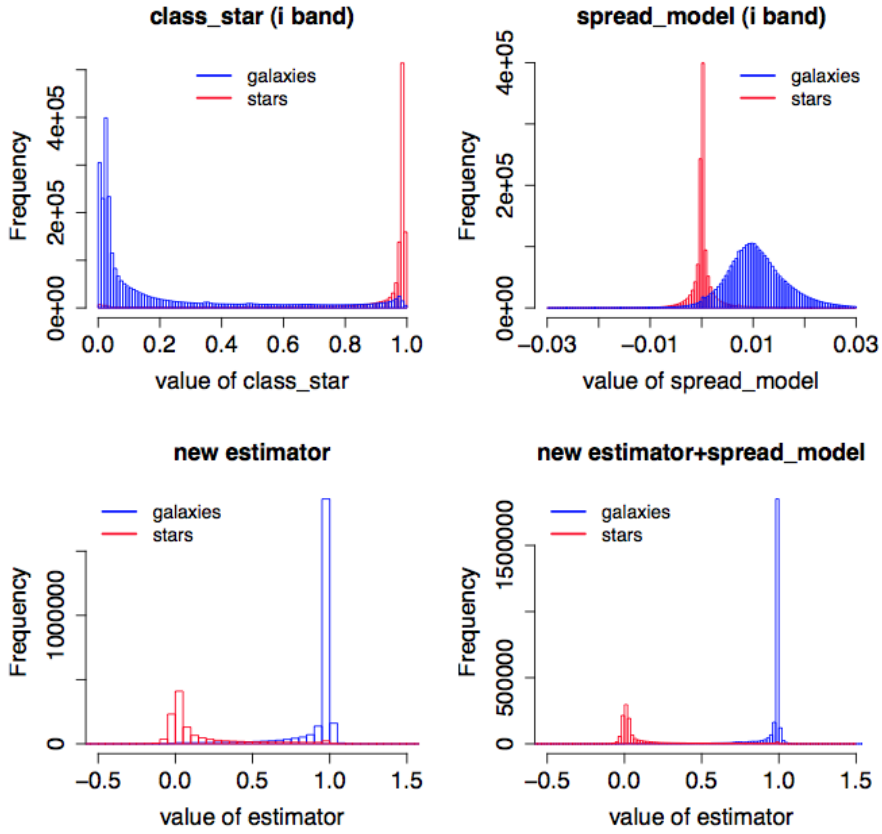

Figure 3.1: Distribution of the output of all the classifiers presented in this chapter. The two upper histograms show the classification performed by class_star and spread_model. The lower histograms show the classification performed by our new estimator, multi_class. On the right one, we incorporate X_{spread_model} in the input parameters of the ANN. The advantages of plugging X_{spread_model} into our tool are explained in section 3.2.3. This allows an increase of the purity for a given completeness, as shown in figure 3.6.

3.1.1 The training approach - class_star

The first classifier to be implemented in *SExtractor* was class_star. Its performance on our example sub-survey is shown in figure 3.1. It uses a set of features of the objects as the input space for a built-in previously trained ANN. These parameters are:

1. eight isophotal areas, at regular intervals spanning from the detection threshold to the intensity peak;
2. the intensity peak;
3. the local value of the seeing.

This specific pre-defined set of inputs, chosen mainly for historical reasons, is the main weakness of the class_star estimator. The choice of training the ANN on isophotal areas

(normalised to the local PSF footprint area) makes it sensitive to close pairs of objects (star-star, star-galaxy, galaxy-galaxy) either blended or de-blended. Since star-star pairs are common on the bright end of the source population, the classifier has a tendency to miss bright, compact galaxies.

More generally speaking, given the large amount and diversity of information encapsulated in the parameters provided by *SExtractor*, this specific choice of inputs has become hard to justify as it is using a very small part of the available information. The photometry, the shape or the size of an object should also be useful indicators of whether it is a star or a galaxy.

Class_star has the advantage of making use of several parameters and combining the information they contain. In this sense it is a "multi-parameter" estimator. However, it does not use the most relevant parameters. A more flexible and sensible choice of the inputs is likely to give much better results. This is the main motivation for the new approach tested in this chapter.

3.1.2 The morphometric approach - spread_model

The morphometric approach was used in several photometric surveys in the past. One possible implementation of this approach, adopted in the SDSS pipeline and in early versions of the DES pipeline, consists of comparing a "model magnitude", i.e. the optimal measure of the magnitude obtained by fitting a galaxy model to the object, to the "PSF magnitude" , i.e. the optimal measure of the magnitude determined by fitting a PSF model to the object. A similar strategy was adopted in the CFHTLS pipeline, where classes are assigned to objects according to their half-light radius (HLR), i.e. the circular radius which encloses half the light of an object.

The classifier implemented in recent development versions of *SExtractor*, spread_model (Desai et al. 2012; Bouy et al. 2013), carries out diverse operations directly on the image pixels with no use of the parameters generated by *SExtractor*. The newest version of spread_model is defined as

$$\mathrm{X}_{\mathrm{spread\_model}} = \frac{\phi^T W x}{\phi^T W \phi} - \frac{G^T W x}{G^T W G} \,, \tag{3.1}$$

where x is the image centered on the source, W is the inverse of the covariance matrix of the pixel noise, which is assumed to be diagonal, ϕ is the PSF and G is the circular

exponential model convolved with the PSF. It acts as a linear discriminant in the pixels space, between the best fitting local PSF model ϕ and a slightly "fuzzier" version made from the same PSF model, convolved with a circular exponential model with scale length given by $FWHM/16$ (FWHM being the Full-Width at Half-Maximum of the local PSF model). Spread_model is normalized to allow for comparison of sources with different PSFs throughout the field. By construction, spread_model is close to zero for point sources (most likely to be stars), positive for extended sources (most likely to be galaxies) and negative for detections smaller than the PSF, such as cosmic ray hits.

The performance of this late version of spread_model on our example sub-survey is shown in figure 3.1. Although this morphometric approach is quite efficient, it is not entirely satisfying as it does not make use of any of the 300 *SExtractor* parameters, which are likely to encapsulate a lot of relevant information for star/galaxy separation.

3.2 The multi_class method

3.2.1 Motivation and principle

Our goal is to combine the assets of both the morphometric approach and the training approach. We adopt the multi-parameter approach allowed by the training method and focus on making the optimal choice of input parameters. The steps of the method are as follows:

(1). Optimal choice of input parameters using a Principal Components Analysis (PCA);

(2). Training and running an ANN.

3.2.2 Step 1- optimal choice of input parameters using Principal Component Analysis

Principal Component Analysis

Principal Component Analysis (PCA) is a method which comes down to diagonalising the covariance matrix of the data. It allows to re-express the observed data (i.e. the values of a set of possibly correlated variables) in a sometimes more meaningful basis of orthogonal, i.e. linearly uncorrelated variables called *principal components*. The first principal component is chosen to have the highest possible variance and thus to account for most of the data variability. Then each succeeding principal component has the highest

possible variance under the constraint of being orthogonal - that is uncorrelated - to the preceding one. In this section, we explain its mathematical implementation.

Let N be the number of parameters in the catalogue $DC6$, and K the number of objects. This sample of N random variables can be put in a matrix form:

$$M = \begin{bmatrix} X_{1,1} & . & . & . & X_{1,N} \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ X_{K,1} & . & . & . & X_{K,N} \end{bmatrix}$$

Each parameter $X_n = (X_{1,n}, ..., X_{K,n})$ has a mean value $\bar{X_n}$ and a standard deviation $s_N = \sigma_{X_n}$. We assume the realizations have all the same probability. The vector $(\bar{X_1}, \bar{X_2}, ..., \bar{X_N})$ is called $\mathbf{g}$ and is such that $\mathbf{g} = \mathbf{M^T} \cdot \mathbf{D} \cdot \mathbf{1}$, where $D = \frac{1}{K} \cdot Id$ and $\mathbf{1}$ is the vector of $\mathbb{R}^K$ of which all components are equal to 1.

$\bar{M}$ is the *centered matrix*:

$$\bar{M} = \begin{bmatrix} X_{1,1} - \bar{X_1} & . & . & . & X_{1,N} - \bar{X_N} \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ X_{K,1} - \bar{X_1} & . & . & . & X_{K,N} - \bar{X_N} \end{bmatrix}$$

$$= M - \mathbf{1}g^T \, . \tag{3.2}$$

And $\tilde{M}$ is the *rescaled* matrix:

$$\tilde{M} = \begin{bmatrix} \frac{X_{1,1} - \bar{X_1}}{\sigma(X_1)} & . & . & . & \frac{X_{1,N} - \bar{X_N}}{\sigma(X_N)} \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ \frac{X_{K,1} - \bar{X_1}}{\sigma(X_1)} & . & . & . & \frac{X_{K,N} - \bar{X_N}}{\sigma(X_N)} \end{bmatrix}$$

$$\tilde{M} = \bar{M} \cdot D_{1/s} \ ,$$

where $D_{1/s}$ be the diagonal matrix with the inverse standard deviations

$$D_{1/s} = \begin{bmatrix} 1/s_1 & . & . & . & 0 \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ 0 & . & . & . & 1/s_N \end{bmatrix}$$

The variance-covariance matrix V is given by:

$$V = 1/K \cdot \bar{M}^T \cdot \bar{M} = M \cdot D \cdot M^T - g \cdot g^T = \bar{M} \cdot D \cdot \bar{M}^T \ ,$$

Whereas the correlation matrix is:

$$C = 1/K \cdot \tilde{M}^T \cdot \tilde{M} = \tilde{M} \cdot D \cdot \tilde{M}^T \ .$$

In this work, we chose to use the rescaled matrix $\tilde{M}$, which allows us to reduce the noise effects.

Performing a PCA aims at looking for a vector **u** such that the *projection* of the data on this vector has maximal variance. The projection of the data on **u** is:

$$\pi_u(\tilde{M}) = \tilde{M} \cdot u \ .$$

The variance of $\pi_u(\tilde{M})$ is thus:

$$\pi_u(\tilde{M})^T \cdot 1/K \cdot \pi_u(\tilde{M}) = u^T \cdot \tilde{M}^T \cdot 1/K \cdot \tilde{M} \cdot u$$

$$= u^T \cdot C \cdot u \ .$$

C is diagonalizable in an orthonormal base. We call P the change of base and Δ the

diagonal matrix associated:

$$\pi_u(\tilde{M})^T \cdot 1/K \cdot \pi_u(\tilde{M}) = u^T \cdot P^T \cdot \Delta \cdot P \cdot u$$

$$= (Pu)^T \cdot \Delta \cdot (Pu)$$

$$= v^T \cdot \Delta \cdot v \ ,$$

where $v = P \cdot u$. We look for the unitary vector v which maximizes $v^T \cdot \Delta \cdot v$, where $\Delta = Diag(\lambda_1, ..., \lambda_N)$ (the values on the diagonal are ordered in descending order). By using the Lagrange multiplier α,

$$L(u, \alpha) = u^T \cdot C \cdot u - \alpha(u^T u - 1)$$

we show that:

1. **u** is the eigenvector of C with the eigenvalue λ_1.
2. **u** has a norm equal to 1.

The eigenvalue λ_1 is the variance on the first PCA axis. The next axis **v** is found in the same way, under the constraint of being orthogonal to **u**.

Application to DC6

We make a broad pre-selection of all the parameters likely to be relevant for star/galaxy classification. These parameters are listed in table 3.1. They include:

1. **photometry** in 5 bands (g,r,i,z and y);
2. the **size** of objects;
3. the **shape** of objects;
4. the **surface brightness** of objects;
5. qualifiers of the **fitting** procedure;
6. the **output of the class_star classifier**, X_{class_star};
7. additional **analysis-dependent information**.

Table 3.1: DC6 pre-selected parameters, grouped as defined in section 3.2.2, by type of information they provide: (i): photometry; (ii) size; (iii): shape; (iv): surface brightness; (v): qualifiers of the fitting procedure; (vi): output of the class_star classifier; (vii): additional analysis-dependent information. It should be noted that all of these parameters are distance-dependent. The need for K-correction to the magnitudes is therefore dealt with by including the photometric redshift in this pre-selected parameters space.

	Parameters	**Description**
(i)	mag_aper_ in 5 bands	Fixed aperture magni- -tude with 6 different apertures
	mag_auto in 5 bands	Kron-like elliptical aperture magnitude
	mag_iso in 5 bands	Isophotal magnitude
	mag_model in 5 bands	Magnitude from model- fitting
	mag_petro in 5 bands	Petrosian-like elliptical aperture magnitude
	mag_psf in 5 bands	Magnitude from PSF- fitting
	mag_spheroid in 5 bands	Spheroid total magn- -itude from fitting
(ii)	kron_radius (from the de- tection image)	Kron apertures
(iii)	ellipticity (from the de- tection image)	$1 - B_{image}/A_{image}$
(iv)	isoarea_world in 5 bands	Isophotal area above analysis threshold
	FWHM_world in 5 bands	FWHM assuming a gaussian core
(v)	chi2_model in 5 bands	Reduced chi-square of the fit
	chi2_psf in 5 bands	Reduced chi-square from PSF-fitting
	niter_model in 5 bands	Number of iterations for model-fitting
(vi)	X_{class_star} in 5 bands	Output from class_star
(vii)	nlowdweight_iso	Number of pixels with low detection weight over the isophotal profile
	photoZ	photometric redshift

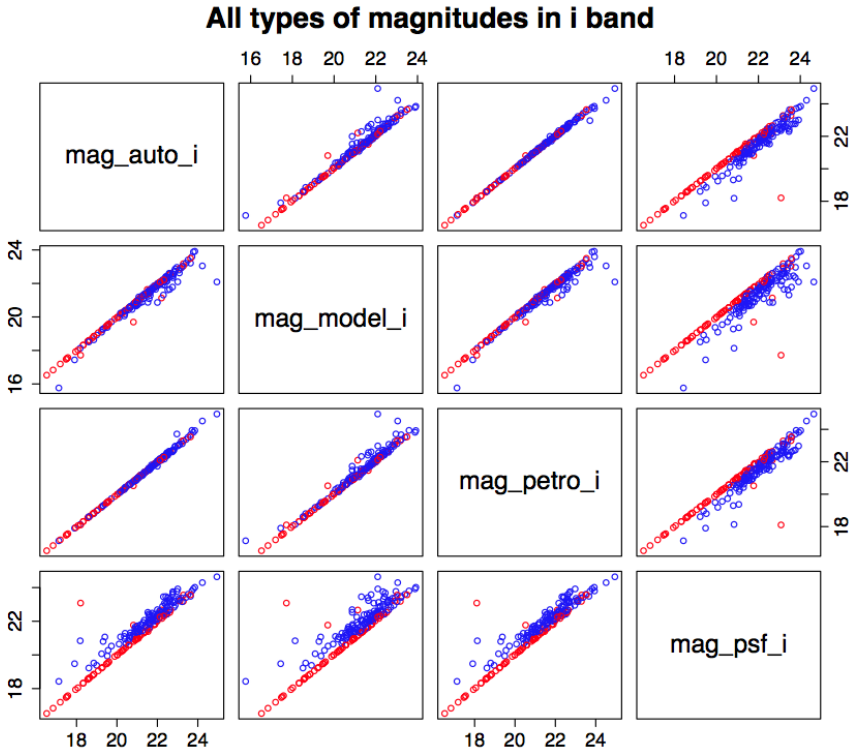

Figure 3.2: Scatter plots for stars (red markers) and galaxies (blue markers), for four different types of magnitudes in the i band. The magnitudes are strongly correlated and PCA is therefore well adapted to re-express them in a new basis of independent variables.

Ideally, we could run an ANN with this full set of relevant inputs. In practice, training the ANN is a non-linear iterative process, which becomes more time consuming and less robust as the number of input parameters increases. In fact, defining an optimal set of input parameters consists of minimising its size while maximising the amount of relevant information it contains.

Our initial set of parameter is redundant, as many of the parameters within each sub-group are dependent variables. For example, we show in figure 3.2 the dependencies between four types of magnitudes parameters measured in a given band. In order to reveal the redundancies within the data and compress it, we use a PCA. This statistical method, which, as explained in section 3.2.2 comes down to diagonalising the covariance matrix of the data, allows us to re-express the pre-selected parameters detailed above in a basis of *principal components*.

We run several "well-informed" PCAs on sub-ensembles of parameters, rather than a "blind" PCA on the full set of initial parameters. We choose to group in these sub-ensembles parameters which have the same units (or measure) and which are linearly

dependent on each other (such as the magnitudes in a given band, as shown in figure 3.2). Indeed, when the parameters are linearly dependent, PCA is successful at finding a new basis of meaningful independent variables.

Our new set of parameters includes uni-band parameters from the initial set (such as the photometric redshift or the ellipticity), as well as the principal components from the PCAs listed below:

- PCA on the five bands of each multi-band parameter;
- PCA on the six fixed-aperture magnitudes in each band;
- PCA on the six other types of magnitudes in each band (i.e. *mag_auto*, *mag_iso*, *mag_model*, *mag_petro*, *mag_spheroid* and *mag_psf*).

Figure 3.3 shows the variances of the principal components of these six types of magnitudes in each band as a function of their index. Each of these PCAs shows that most of the variance of the data is encapsulated in a reduced number of principal components.

In many cases, using PCA for data reduction consists of selecting only the principal components with the highest variance and approximating the data by its projection on this smaller set of variables. This encompasses the assumption that the important information is represented by the components with the highest variances. In the case of star/galaxy separation, this assumption is too simplistic. Indeed, the class of an object is only one possible source of variance and high variance could also be due to differences between objects in a given class. Therefore, when looking for the most relevant components for star/galaxy separation, we need another criterion to quantify their aptitude to separate between the classes. We calculate the *Fisher discriminant* (Fisher 1936) for each of the new parameters, defined as the inter-class variance over the intra-class variance,

$$\mathcal{F}_i = \frac{(\overline{X_{G,i}} - \overline{X_{S,i}})^2}{\sigma_{G,i}^2 + \sigma_{S,i}^2} \,, \tag{3.3}$$

where $\overline{X_{A,i}}$ is the empirical mean value of the i^{th} parameter for class A and $\sigma_{A,i}^2$ is its empirical variance. Figure 3.4 shows the classification performed by the three parameters with the highest Fisher discriminant. The fifteen parameters with the highest Fisher discriminant form our final set of input parameters for the ANN (more than twenty input parameters make the ANN less robust, so we limit the basic set to fifteen parameters, in anticipation of the other five that will be added in section 3.2.3).

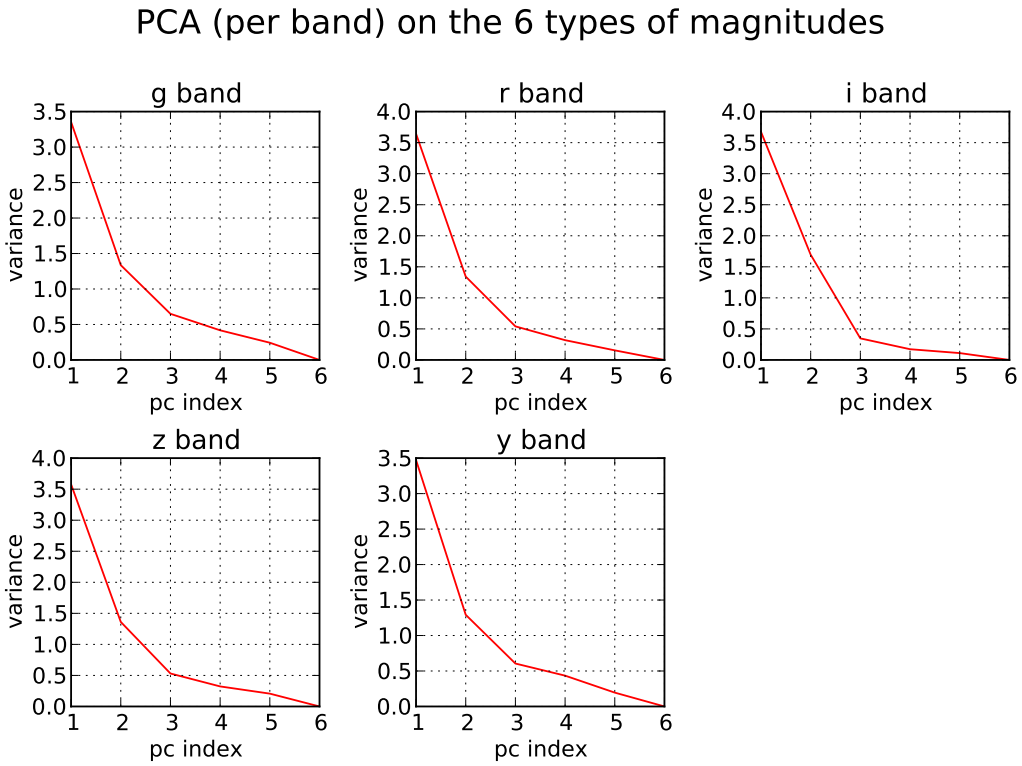

Figure 3.3: Value of the variance of the principal components as a function of their index for the fives (per-band) PCAs performed on the six types of magnitudes: *mag_auto*, *mag_iso*, *mag_model*, *mag_petro*, *mag_spheroid* and *mag_psf*.

3.2.3 Step 2 - running an Artificial Neural Network on the optimal inputs space

Once a set of optimal parameters is defined, the next step consists of mapping these parameters to the class of the objects. This mapping is performed by training an ANN.

ANN: principle and advantages

In essence, an ANN is a highly-flexible, fully non-linear fitting algorithm. During the training phase, it receives a set of input patterns and a given property (in our case the class of the object), which needs to be fitted to them. The training consists of several iterations during which a number of free parameters known as *weights* are adjusted so as to minimise the difference between the outputs of the neural network for each pattern and the desired property. The algorithm then learns how to link the inputs to the desired property. After the training phase, the ANN can be used to infer this property from a set of input objects for which it is unknown. For our analysis, we train an ANN to map the

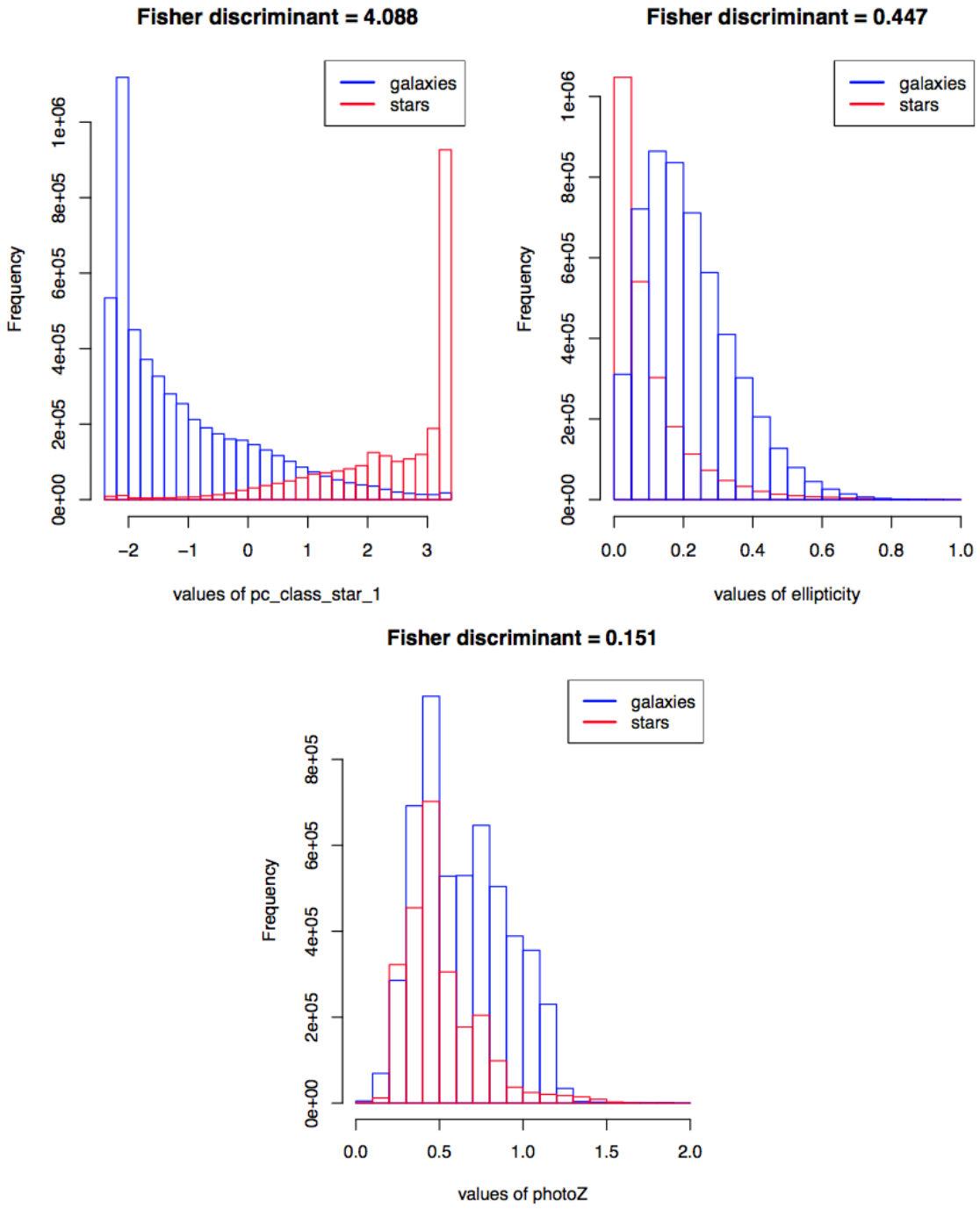

Figure 3.4: Distribution of the three parameters with the highest Fisher discriminant, for stars and galaxies as indicated in the figure. *pc_class_star_1* (top left) is the first principal component from a PCA performed on the five bands of X_{class_star}. The two other parameters shown, *ellipticity* (top right) and *photoZ* (bottom) have not gone through any PCA.

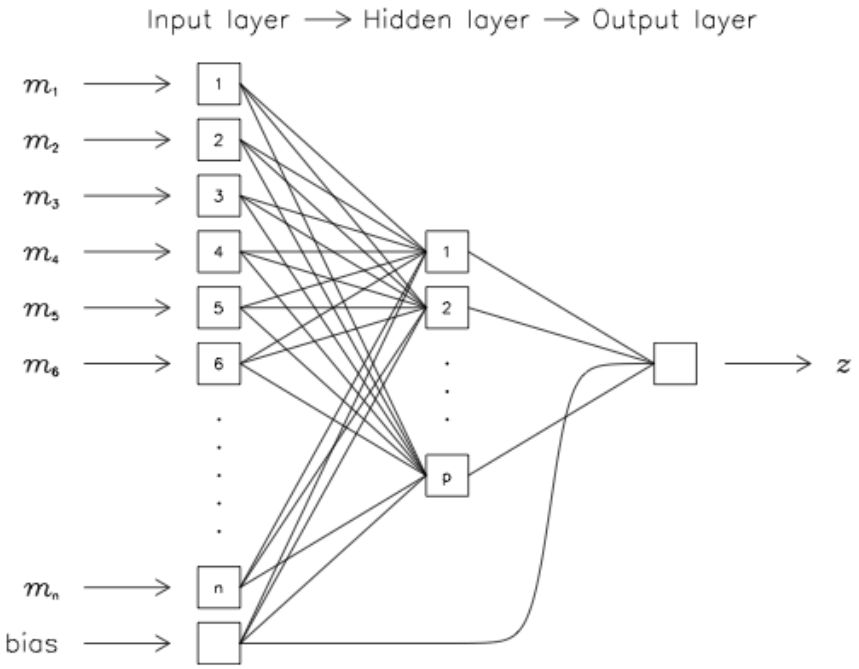

Figure 3.5: Schematic diagram of neural network as implemented by ANNz from Collister & Lahav (2004). When used for photometric redshift measurement, the input layer of ANNz consists of nodes that take magnitudes in the different filters used for photometry, but we use a different set of input parameters, carefully defined and selected according to the procedure of section 3.2.2. A single hidden layer consisting of p nodes is shown here although more hidden layers could be used. The output layer has a single node that gives e.g. the photometric redshift. In our case the output property is not the redshift, but the class of an object (i.e. the object being a star or a galaxy). Each connecting line between nodes carries a weight, w_{ij} . The bias node allows for an additive constant when optimising weights.

set of optimal input parameters selected in section 3.2.2 to the class of the object (star or galaxy) on a sample of objects for which the answer is known (the training is made on the DC6 simulations for which we know the true class of each object). The ANN is then used to deduce the class of a distinct set of objects.

An ANN is made of computing units called *neurons*, arranged in several layers and connected by synapses in which the information flows in a single direction. The complexity of the network depends on the number of layers and neurons in each layer. We chose to use the ANNz photometric redshift code (Collister & Lahav 2004) , which was originally designed for photometric redshift measurements, but can be effectively and straightforwardly applied to our classification problem.

Figure 3.5 shows the ANNz setup, when used for photometric redshift measurements.

The first layer receives the parameters selected in section 3.2.2 as inputs and the last layer outputs a number between 0 and 1, which represents the probability of the object to be a star or a galaxy. All the nodes in the hidden layers in between the input layer and the output are interconnected and connections between nodes i and j have an associated weight, w_{ij} . Each node i is assigned a value u_i and an activation function $g_i(u_i)$

$$g_i(u_i) = \frac{1}{1 + exp(-u_i)} \tag{3.4}$$

The value of a subsequent node j is then calculated as the summation of the weighted values of the activation functions of all nodes i pointing to it:

$$u_j = \sum_i w_{ij} g_i(u_i) \ . \tag{3.5}$$

The ANN requires a training set that is used to minimize the cost function, E, with respect to the free parameters w_{ij},

$$E = \sum_k (\mathcal{C}(w_{i,j}, m_k) - \mathcal{C}_{train,k})^2 \ , \tag{3.6}$$

where m_k is the k^{th} input vector of the training set, $\mathcal{C}(w_{i,j}, m_k)$ is the class computed by the network for this input vector, and $\mathcal{C}_{train,k}$ is the known class. To avoid an over-fitting, every network is tested on a validation set of galaxies, for which the output $\mathcal{C}$ are also known. The network with the lowest value of E as calculated on the validation set is selected and the sample of unknown classes is run through it for class estimation. This algorithm is fully implemented within ANNz (Collister & Lahav 2004).

The trade-off between the complexity of the architecture (the number of hidden layers and nodes in each hidden layer) of the network and its performance has been investigated by Firth et al. (2003). For the same number of parameters, adding extra hidden layers is found to give greater gains than widening existing layers. As the network complexity is increased, the accuracy eventually converges so that no further improvement is gained by adding additional nodes. We chose a network architecture with an input layer of fifteen parameters (or twenty, as explained in the next section) and two hidden layers of twenty (twenty-five, respectively) nodes, which turns out to be sufficiently complex for such convergence to be achieved.

Training on real data, as opposed to simulations, is preferable, yet more challenging.

One option would be to use data from space-based surveys, as in space the PSF is not affected by the seeing. Data from the Hubble Space Telescope could be used to train our tool for the real DES survey data. This is the aim of chapter 4.

Plugging other classifiers in the method

Using an ANN brings flexibility to the training approach. It allows us not only to choose which inputs to use, but also in what number. In particular, we can take the output of other classifiers as inputs to our method.

We run a PCA on the five X_{class_stars} (in the five bands). Not surprisingly, the first principal component has a high Fisher discriminant (as shown in figure 3.4) and is therefore included in the 15 input parameters selected in section 3.2.2. As the the five bands of X_{spread_model} are less clearly linearly dependent, we choose not to run a PCA on them and add the five X_{spread_model} to the set of fifteen input parameters, which amounts to twenty input parameters.

Figure 3.6 presents the purity level at a given completeness for these two different configurations of our method. The performance of our method with fifteen input parameters (orange curve) can be compared to the performance when plugging in X_{spread_model} (pink curve). Including X_{spread_model} in the inputs allows an increase in the level of the purity by 2% at faint magnitudes. Running the ANN on the fifteen preselected parameters (orange curve) already gives better results than spread_model (blue curve) for most of the magnitude range (except for the very faint magnitudes, in the galaxies case). However, the best results are obtained by combining the two, i.e by running the ANN on a hybrid input space combining the 15 selected parameters and X_{spread_model}.

3.3 Classification results

We showed that we can optimise our classifier performance by using a "well-informed" PCA strategy (section 3.2.2), and by incorporating X_{spread_model} into the method (see above). We now compare our classifier performance to the one of the other classifiers. We will focus on comparing multi_class to spread_model, as the performance of class_star is widely surpassed by both spread_model and multi_class for most of the magnitude range (as shown in Figure 3.6).

For LSS, our new classifier allows us to achieve requirements which cannot be fulfilled

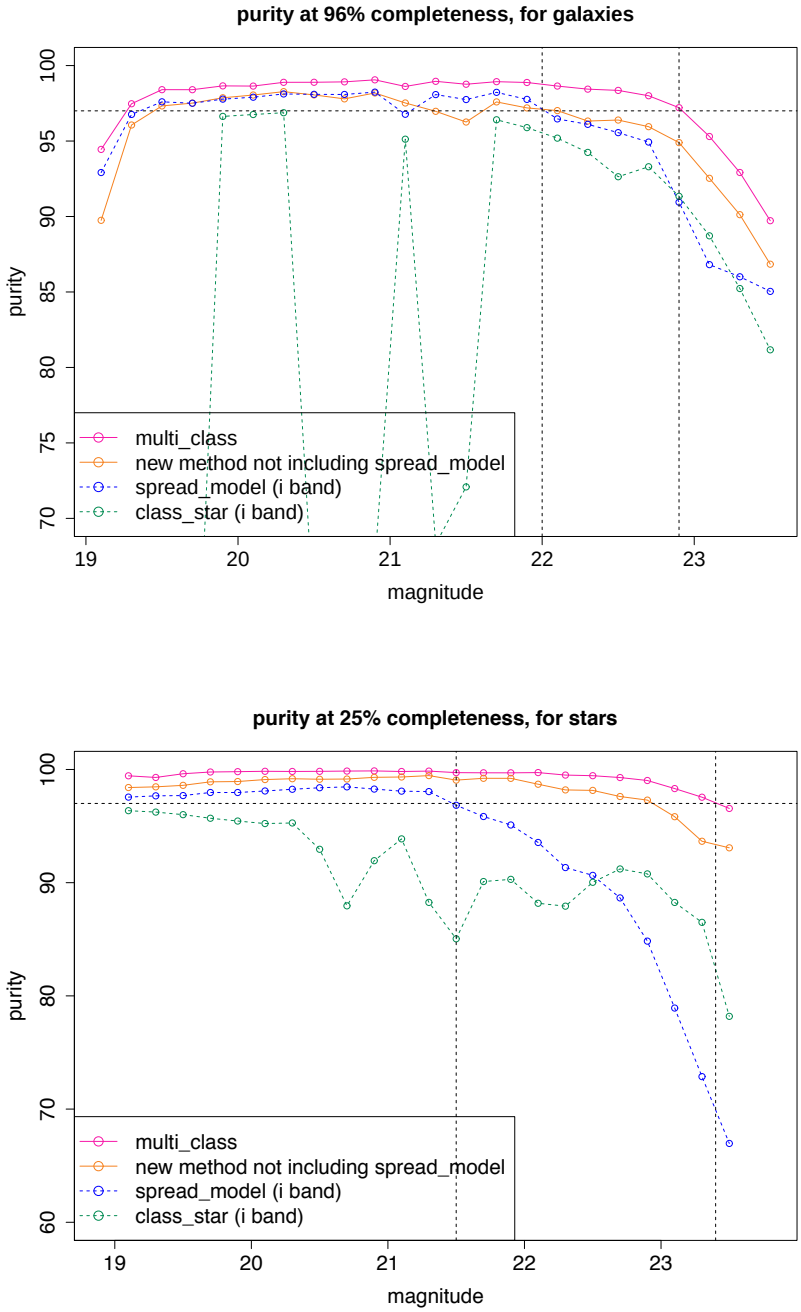

Figure 3.6: Purity level at the required completeness, for the WL probe, as a function of magnitude in the i band. The orange and pink curves correspond to different versions of our method: one case (orange), we ran ANNZ on the set of parameters selected following the PCA; in the other case (pink), we added spread_model (in 5 bands) to this set of parameters, which increase the level of purity. It appears that "plugging" spread_model to the ANN inputs increases the purity level.The blue and green ones show the performance of the classifiers class_star and spread_model. The orange curve is obtained when running the ANN on the 15 parameters selected in section 3.2.2 and the pink curve, the final version of multi_class, is obtained when adding spread_model in five bands to this set of inputs. The dashed horizontal line shows the science requirement from WL science on p^g (97.8%, section 2.5.1 of chapter 2) and p^s (97.0%, section 2.6 of chapter 2). The requirement on p^g is achieved by multi_class up to magnitudes of 22.9, whereas spread_model only allows us to reach 22.0. The requirement on p^s is achieved up to magnitudes of 23.4 with multi_class, versus 21.5 with spread_model.

by spread_model. Figure 3.7 shows that the 98.5% limit on p^g (derived in section 2.5.2 of chapter 2 and shown in purple on the figure) cannot be reached by spread_model, whereas multi_class allows us to reach it up to magnitudes of 22.9 (at the required 88.9% completeness level, derived in section 2.4.2 of chapter 2).

For WL, multi_class allows us to increase the magnitude limit below which the science requirements are achieved. Figure 3.6 shows that this magnitude limit increases from 21.5 to 23.4 for the requirement on the stars purity p^s, and from 22.0 to 22.9 for the requirement on the galaxy purity p^g. Figure 3.7 and figure 3.8 generalise this to a broad range of completenesses. In figure 3.9, we consider the improvement in the purity of a sample of stars and a sample of galaxies, as a function of magnitude, for a large range of completenesses. At faint magnitudes - typically fainter than 23 - multi_class improves the purity achieved by spread_model by up to 12% for galaxies and by up to 20% for stars.

3.4 Conclusions

In order to meet the requirements defined in chapter 2, we built an efficient method for star/galaxy classification, called multi_class, which combines a PCA with a learning algorithm. Our multi-parameter approach allows us to make use of the huge amount of information provided by *SExtractor*. In particular, the use of PCA allows us to better understand the correlations in the data, and to implement this physical knowledge in the classifier.

In ground-based surveys such as DES, the image quality is not constant with position and therefore any purely morphometric method gives limited performance, especially at faint magnitudes. The flexibility of using an ANN allows us to consider the morphometry as one input parameters among many others and to integrate the performance of other classifiers to our new tool. Our new classifier, multi_class, significantly improves the performance of the morphometric classifier implemented in *SExtractor* (spread_model), which cannot achieve the LSS science requirements on star/galaxy separation. For both the LSS and WL probes, it allows us to widen the range of both magnitude and completeness where the derived science requirements are achieved. For magnitudes fainter than 23, multi_class improves the purity achieved by spread_model by up to 12% for galaxies and by up to 20% for stars.

The faint magnitudes reached by this new classifier constitute an important asset,

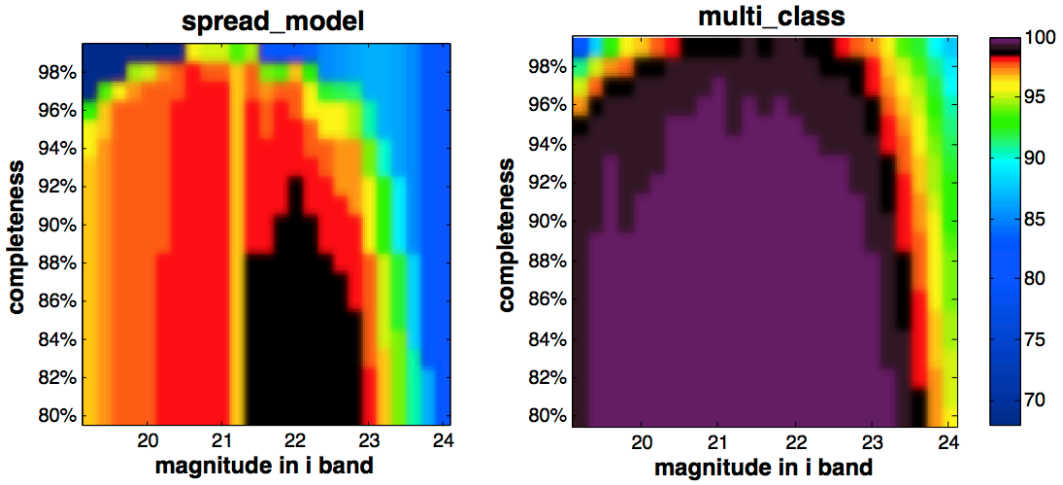

Figure 3.7: Level of purity for a sample of galaxies p^g, for different magnitudes and values of the completeness. The 98.5% level requirement from LSS (section 2.5.2 of chapter 2) is shown in purple, and the 97.8% limit required for WL (section 2.5.1 of chapter 2) is shown in black. Spread_model does not allow to achieve the LSS requirement, which multi_class can reach. Multi_class also allows us to achieve the requirement from WL at fainter magnitudes than spread_model.

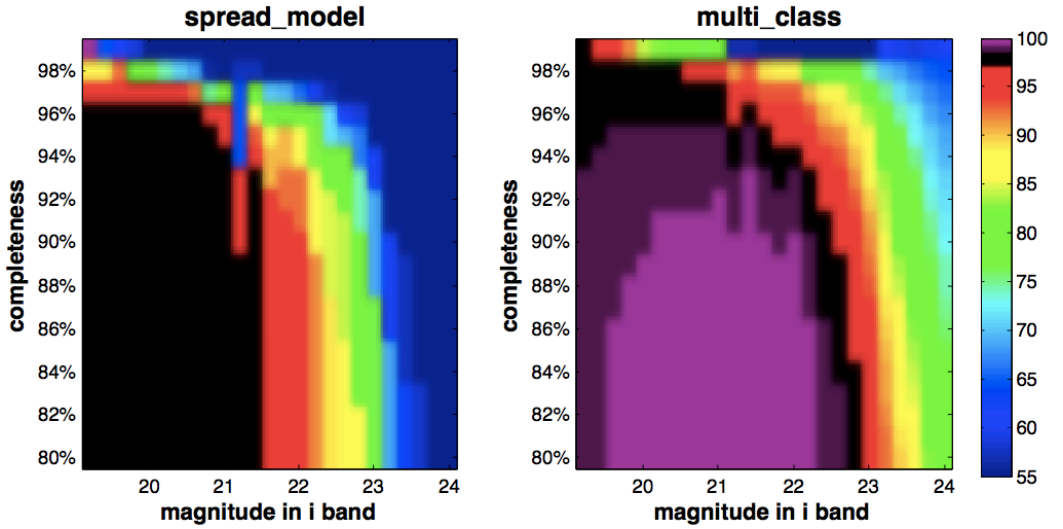

Figure 3.8: Level of purity for a sample of stars p^s, for different magnitudes and values of the completeness. The 97% science requirement (from WL, derived in section 2.6 of chapter 2) is shown in black. Higher purity levels are shown in purple and light purple. Our new estimator, multi_class, allows us to widen the range of both magnitude and completeness where this requirement is achieved.

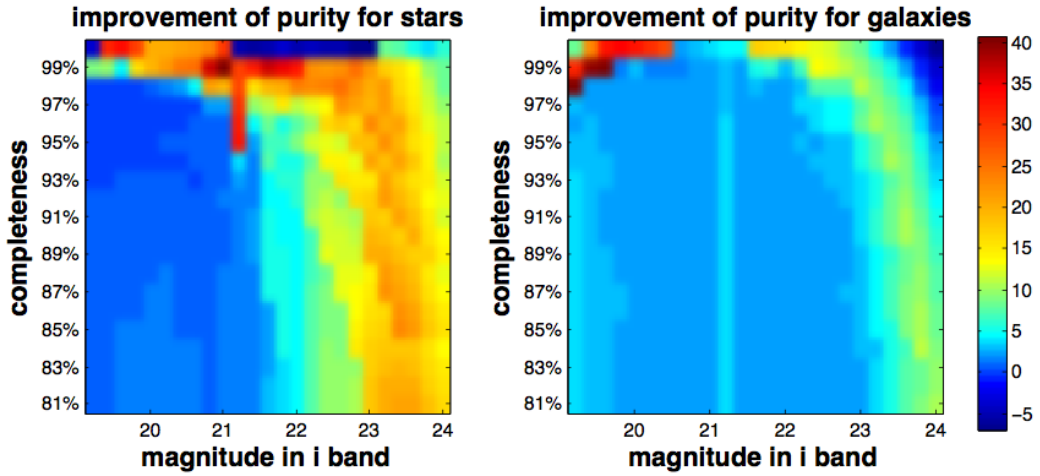

Figure 3.9: Difference of the purity level achieved by multi_class and spread_model, $p_{multi\_class} - p_{spread\_model}$ for stars (left) and galaxies. At faint magnitudes (ranging from 23 to 24), multi_class allows us to increase the level of p^s achieved by spread_model by up to 20%, and p^g by up to 12%.

which should allow to achieve the science requirements on star/galaxy separation in the next generation of wide-field photometric surveys. The recent beginning of the DES survey operations in September, 2013, will allow us to continue the testing and optimisation of multi_class on real data. A first attempt is presented in chapter 4 of this thesis.

Chapter 4

Star/Galaxy separation in the DES Science Verification data

"To write it, it took three months; to conceive it three minutes; to collect the data in it all my life."

F. Scott Fitzgerald

"Errors using inadequate data are much less than those using no data at all."

Charles Babbage

The classifier we built in chapter 3, multi_class, uses a neural network, and therefore requires to be trained on a set of objects for which the class - i.e. the object being a star or a galaxy - is known. When using simulations as in chapter 3, the class of the object is provided by the truth tables used to generate the simulated data. However, when using real data, obtaining a reliable training set is a more challenging problem. The equivalent of the truth table can be obtained by cross-matching the available catalog with two types of data:

1. Data from *space-based* facilities. In this case, the seeing is good enough so that the class of the object is not ambiguous.

2. *Spectroscopic* data. In this case, the spectra of the objects, which are very different for stars and for galaxies, provide a straightforward answer to the classification problem.

The recent internal release of the first DES science verification data (section 4.1) and its cross-matching with data from the Hubble Space Telescope (HST) and the VIMOS VLT Deep Survey (VVDS), gave us an opportunity to play on both these aspects, and to perform a first testing of multi_class on real data.

4.1 The year one science verification data

The DES first year of operation has been dedicated to "Science verification" (SV), i.e. an "end-to-end test of all the systems needed to actually conduct the survey, both on the mountain and elsewhere"[1]. Successfully passing the formal acceptance requirements has been a necessary condition for DES to begin taking survey data, in september 2013. In this chapter, we use the latest release of the DES-SV data: the SVA1 Gold Catalog v1.0 (January 2014), to start testing and optimising our classifier.

4.1.1 The science verification catalog

The idea in the creation of this catalog was to be "as generous as possible" with the cuts consistent with a "science-ready" galaxy catalog. The footprint selection was made for coverage with at least one CCD depth in all each of the bands g, r, i and z. A cut was made with $Dec > -61^{\circ}$ to remove the Large Magellanic Cloud (LMC). Although this does remove approximately $40 deg^2$ from the SPTE region, tests performed by the DES Data Management team showed that the LMC cannot be accurately calibrated to the same scale as the rest of the survey and deblending is hard, which is the reason it has been removed. However, we are planning to use the LMC to calibrate our classifier in future work. This cut also has the advantage of removing $5 deg^2$ contaminated by stray light from *R Doradus* , the second brightest star in the infrared sky. Additional footprint cuts were made to remove regions with:

1. very high density of "crazy color" objects (due to satellites, airplanes, and stray light)

[1] DES Data Management website definition.

Catalog	Field	Area	Type	Nb. of good quality stars	Magnitude range	Associated paper	Comments
ACS COSMOS	COSMOS	~1 deg^2	Space imaging	~15000	mag_auto_i<25	Leauthaud07	
ACS GOODS-S	SN-C	~160 arcmin^2	Space imaging	~300	mag_auto_z < 27	Giavalisco04	Conservative cut in mag_auto<24
VVDS-DEEP-02	SN-X	0.6 deg^2	Spectroscopy	~600	17.5<mag_auto_i<24.75	LeFèvre13	Includes ultra-deep
VVDS-CDFS	SN-C		Spectroscopy	~100	17.5<mag_auto_i<24		
ACES	SN-C		Spectroscopy	~300	18<mag_auto_i<23	Cooper12	
VIPERS-W1	SN-X		Spectroscopy	~900	mag_auto_i<22.5	Guzzo13	

Figure 4.1: Detail of the fields matched to the SVA1-Gold catalog and included in the catalog we use to test multi_class. Credit: Nacho Sevilla.

2. very low density of galaxies (primarily at the edges and near other masked regions).

In total, the cuts removed 1.5% of the remaining area. Overall, the SVA1 Gold Catalog v1.0 covers $254.4 deg^2$ with g, r, i, z coverage north of $Dec > -61$. A total of $\approx 223.6 deg^2$ of area also has Y band. Figure 4.2, created with the BigFoot tool - designed by the author and Alex Merson (UCL) - shows the footprint of the SVA1 Gold Catalog v1.0, together with relevant footprints.

4.1.2 The cross-matched catalog

In this work, we do not use all the SVA1 Gold Catalog v1.0, but only the objects of this catalog which have been cross-matched to objects from spectroscopic data sets and data taken from space. Figure 4.1 shows the list of fields which have been cross-matched with the SVA1 Gold Catalog v1.0, to produce the final catalog we used.

Figure 4.2 shows the footprint of these fields, together with DES year-one footprint and final footprint.

4.1.3 Size of the training set

In the DC6 simulation used in chapter 3, the stars-to-galaxies ratio in the training set was approximately 46%, with 2405280 stars and 5227909 galaxies. The cross-matching with spectroscopic data leads to a catalog which is biased in favor of galaxies. Indeed, the stars-to-galaxies ratio in the cross-matched catalog is approximately 10%. Figure 4.3 shows the number of stars and galaxies per magnitude bin, for both DC6 and the training part of the cross-matched catalog.

The artificial network implicitly takes the distribution of objects as a prior, therefore if we believe that the stars-to-galaxies ratio is higher in the survey than in our available

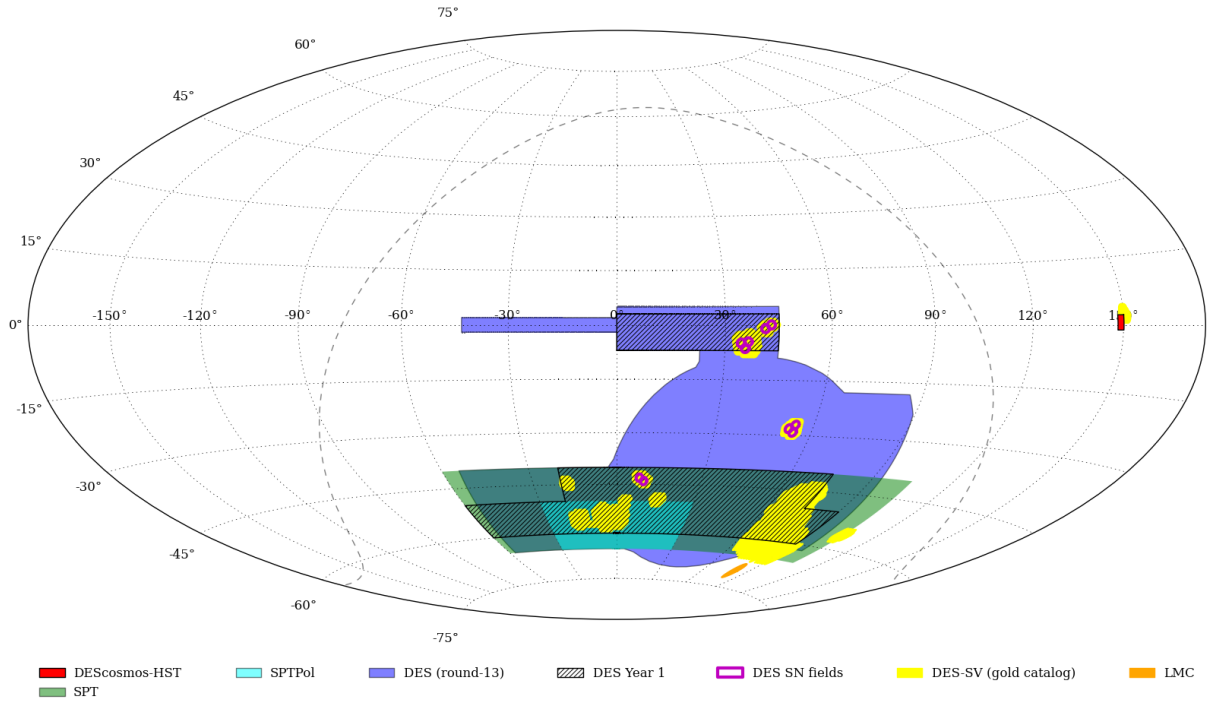

Figure 4.2: Footprint of DES (blue) the SVA1-Gold catalog (yellow), and the cross-matched catalog (red). We also show the SPT area, overlapping with the SV catalog, and the LMC, which has been willingly removed from it.These footprint have been created with the BigFoot tool, developed by the author together with Alex Merson (UCL).

catalog, something should be done to increase the neural network prior for the density of stars. We solve this by "replicating" the stars in the training set, in a way we explain below. In addition, such replication makes the training more robust by increasing the size of the training set.

In order to generate random stars in the training set, we use the error bars on the parameters of each objects and assuming they are gaussian. A widely used method for drawing a random vector $\vec{x}$ from the N-dimensional multivariate normal distribution with known mean vector μ and covariance matrix is using the Cholesky decomposition of the given covariance matrix. Given the Cholesky decomposition $AA^T = C$ of the given covariance matrix C, and $\vec{z} = (z_1, z_2, ..., z_N)^T$ a vector which components are N independent standard normal variates, then the vector $\vec{x} = \vec{\mu} + A\vec{z}$ has the desired distribution.

In all the following, we use a catalog with a number of replications of the stars which allows us to reach a ratio of stars-to-galaxies of about 80%. Figure 4.3 shows the distribution of stars and galaxies in the training set before and after the replication.

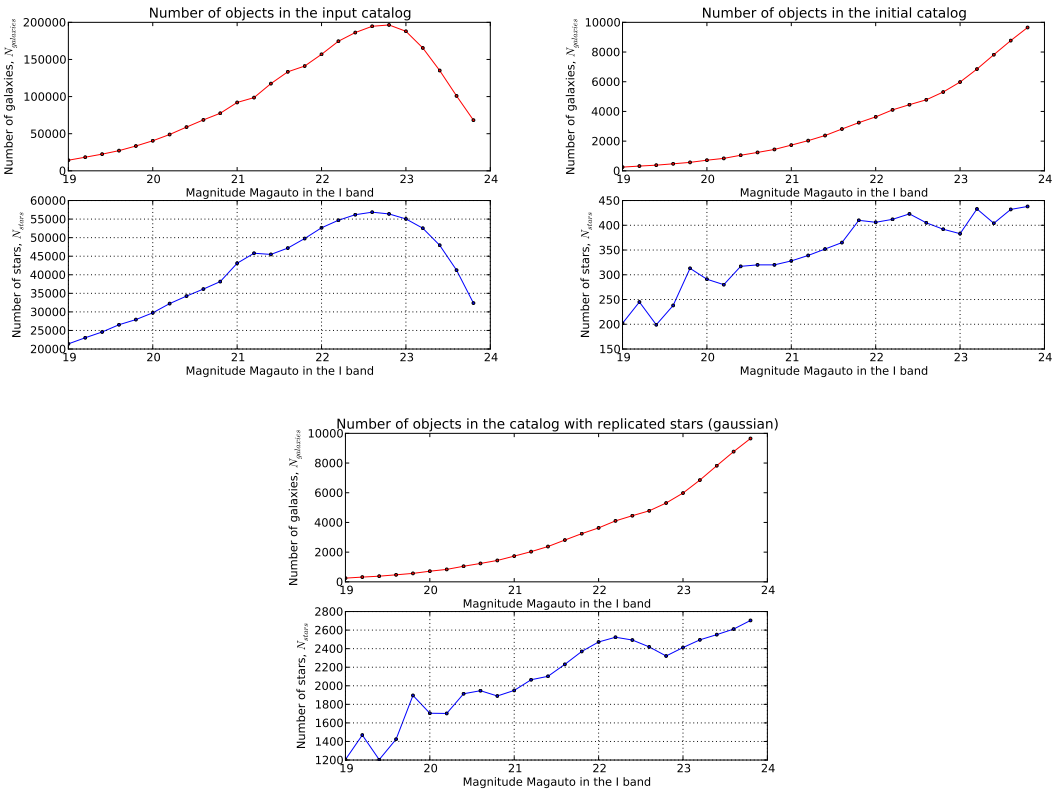

Figure 4.3: Number of objects in DC6 (top left), in the initial SV catalog (top right), and in the SV catalogue after replication of the stars (lower panel).

4.2 The method

4.2.1 Summary of the method terminology

Below is a summary of the main steps of the method presented in chapter 3.

- We pre-select a broad set of parameters which are likely to be relevant for star/galaxy separation;
- We group these parameters into *baskets*;
- The multi_class algorithm performs the following steps
  1. PCAs within each basket;
  2. selection of a number w of "winners", i.e. the w principal components with the highest Fisher discriminant as defined in equation 3.3 of chapter 3;
  3. training of a neural network taking as inputs the w "winners" as well as the five bands of the classifier spread_model presented in chapter 3;

4. projection of a testing set on the w "best" principal components space;

5. testing of the neural network on the projected testing set.

4.2.2 The different configurations of the method

We test different setups of multi_class. In particular, we explore the influence of two aspects of the method on the performance of multi_class:

1. the way the parameters are grouped in each "pca basket" (section 4.3.1);

2. the number of "winners" (section 4.3.2).

Other characteristics of the method that could be optimised include the network architecture, or the optimal combination of spread_model bands used as input to the neural network together with the w "winners".

We adopt the following notation for a given setting: $\{group_i; nw\}$, where $group_i$ is the name of the grouping strategy ($group_1$ or $group_2$) and n is the number of "winners".

4.3 Results

Like in chapter 3, the performance of the different configurations of multi_class is quantified by measuring the purity level as a function of magnitude, at fixed completeness (figures **??**, 4.6 and 4.7), as well as for a large range of completnesses (figures 4.10 and 4.11). We add to this a measurement of the purity-completness domains (figures 4.4, **??**, 4.5 and **??**).

4.3.1 Effect of varying the grouping strategy

Performing sub-PCAs on groups of parameters, grouped into "baskets", as opposed to performing one PCA on the entire set of parameters, gave better results on the DC6 simulations. Here, we try two different ways of grouping the parameters, referred to as $group_1$ and $group_2$. The first method, $group_1$, consists of making one basket per multiple-bands parameter, each basket containing the five bands of a given parameter (e.g. one basket contains a specific magnitude parameter in the five bands). The second method, $group_2$, is a grouping "per band", i.e. one basket consists of different types of magnitudes in the I band.

In figure 4.4 we show the purity-completeness domain $\{p^g, c^g\}$, obtained with both the $\{group_1; 15w\}$ and $\{group_2; 15w\}$ configurations of multi_class, and we compare them to the purity-completeness domains obtained with spread_model, and class_star. The purity-completeness domain $\{p^g, c^g\}$ are shown for different magnitude ranges: $[18.0, 21.0]$, $[21.0, 23.0]$, $[23.0, 24.0]$ and the full range $[18.0, 24.5]$, and each point of a given curve corresponds to a value of the cut on the corresponding classifier. For bright objects with magnitudes in the range $[18.0, 21.0]$, the $\{group_1; 15w\}$ and $\{group_2; 15w\}$ configurations show very similar performances. However, for faint objects in the range of magnitudes $[23.0, 24.0]$, the $\{group_1; 15w\}$ configuration allows to out-perform spread_model (i.e. to achieve points with both higher p^g and higher c^g, which on the figure means that the multi_class curve is above the spread_model one). Similarly, figure 4.5 shows the purity-completeness domain $\{p^s, c^s\}$. The difference between the performance of the $\{group_1; 15w\}$ and $\{group_2; 15w\}$ configurations appears to be more important in this case. For objects in the $[21.0, 23.0]$ magnitude range, the purity-completeness domain does not overlap the WL science requirement in the case of the $\{group_2; 15w\}$ configuration, whereas it does in the case of the $\{group_2; 15w\}$ configuration. For faint objects, in the magnitude range $[23.0, 24.0]$, the curve corresponding to the $\{group_1; 15w\}$ configuration is above the curve corresponding to spread_model, which means it allow to reach higher purity and completeness than spread_model. However, spread_model out-performs the $\{group_2; 15w\}$ configuration. Overall, the $\{group_1; 15w\}$ configuration seems to lead to better performances than the $\{group_2; 15w\}$ configuration. However, in the future, a deeper study of the limiting Fisher discriminant in each case will be necessary to completely understand and take into account this difference (see section 4.4).

4.3.2 Effect of varying the number of "winners"

We try configurations with different numbers of "winners". Since $group_1$ is found to give better results on the purity-completeness domain (see section 4.3.1 above), we focus on comparing $\{group_1; 5w\}$, for which the number of "winners" is $w = 5$ and $\{group_1; 15w\}$, for which the number of "winners" is $w = 15$. In Figures 4.6, we show the purity level p^g at a given completeness, as a function of the magnitude in the i band. The $group_1$ configuration, with $w = 5$ and $w = 15$ are both compared to the performance of the classifiers class_star and spread_model. The figure shows a very slight improvement when $w = 15$. In particular, the $\{group_1; 15w\}$ configuration of multi_class allows to reach

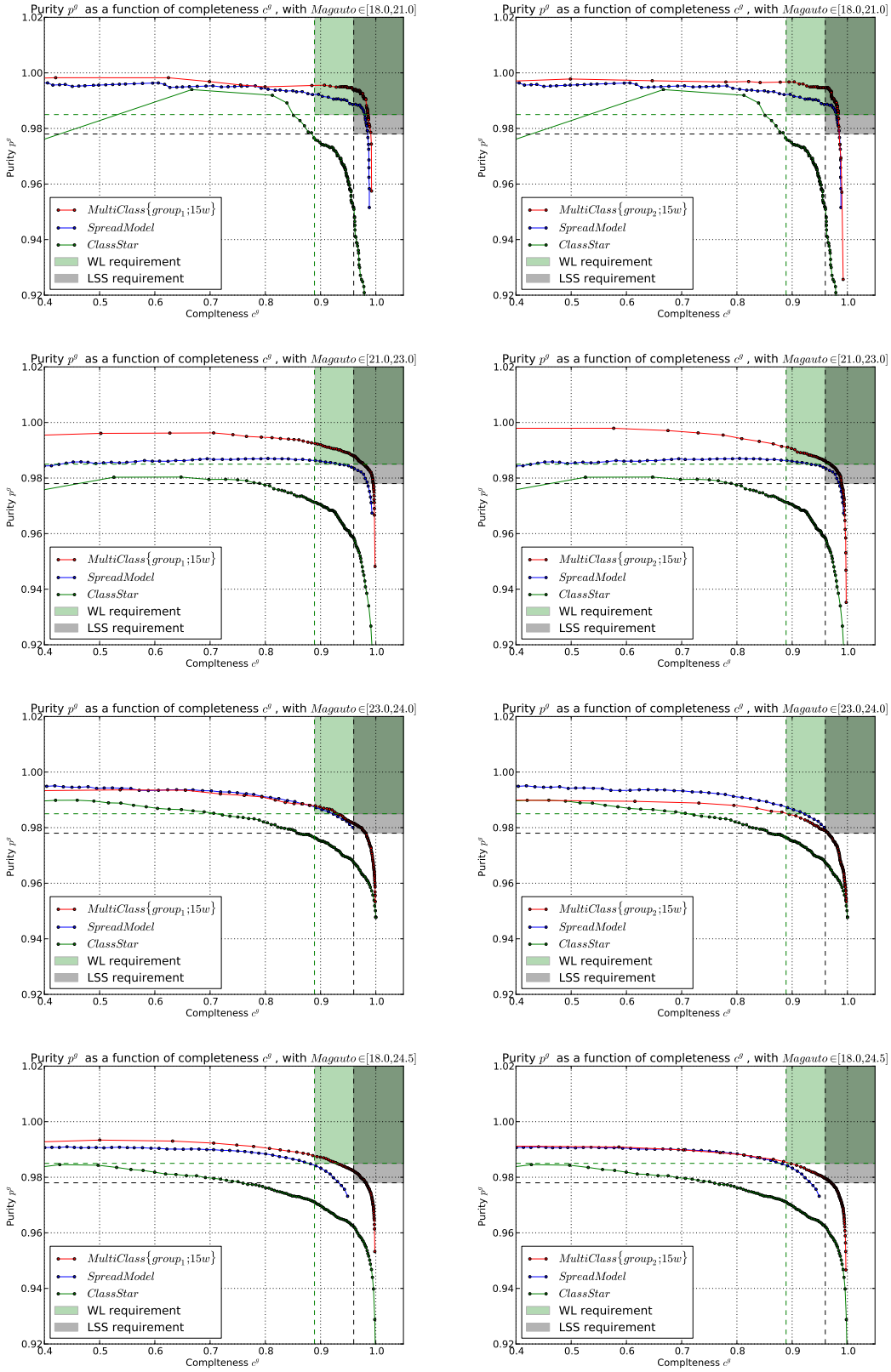

Figure 4.4: Purity-completeness domain for galaxies, with spread_model, class_star, and the $\{group_1; 15w\}$ configuration (left panels) and the $\{group_2; 15w\}$ configuration (right panels) of multi_class, for different magnitude ranges: $[18.0, 21.0]$, $[21.0, 23.0]$, $[23.0, 24.0]$ and the full range $[18.0, 24.5]$ (from top to bottom). Each point of a given curve corresponds to a value of the threshold on the corresponding classifier. The threshold on multi_class is in $[0, 1]$, with bins of 0.01. The threshold on spread_model is in $[0, 0.02]$, with bins of 0.0002. The threshold on class_star is in [0,1], with bins of [0.01].

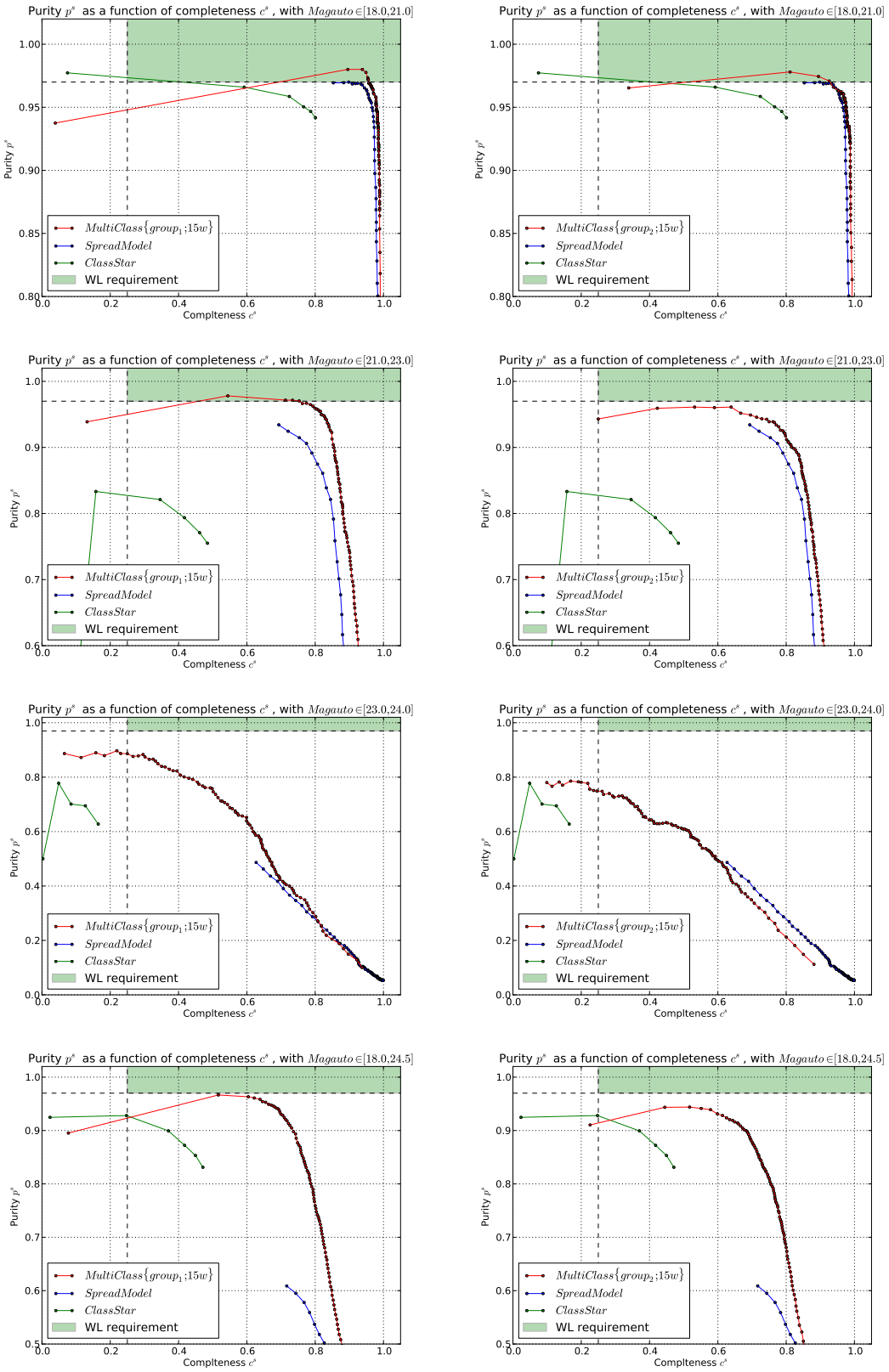

Figure 4.5: Purity-completeness domain for stars, with spread_model, class_star and the $\{group_1; 15w\}$ configuration (left panels) and the $\{group_2; 15w\}$ configuration (right panels) of multi_class, for different magnitude ranges: $[18.0, 21.0]$, $[21.0, 23.0]$, $[23.0, 24.0]$ and the full range $[18.0, 24.5]$ (from top to bottom). Each point corresponds to a different threshold on the classifier. The threshold on multi_class is in $[0, 1]$, with bins of 0.01. The threshold on spread_model is in $[0, 0.02]$, with bins of 0.0002. The threshold on class_star is in [0,1], with bins of [0.01].

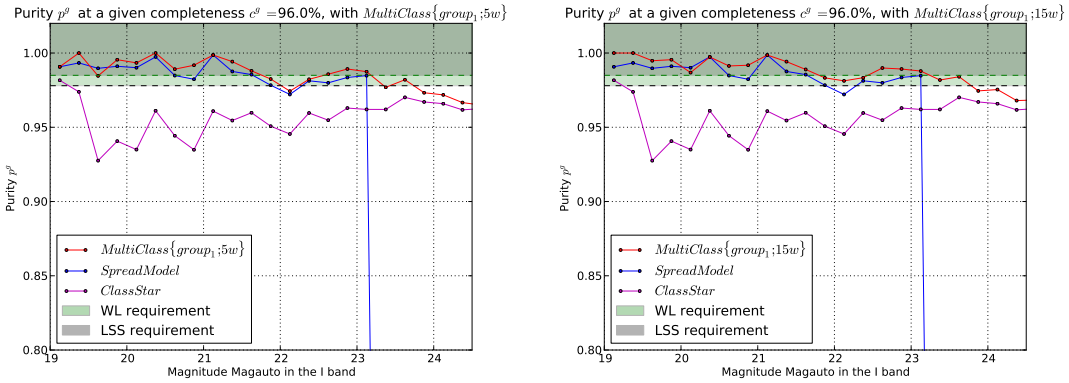

Figure 4.6: Purity level p^g at the required completeness, for the WL and LSS probe, as a function of magnitude in the i band. The left panel corresponds to the $group_1$ configuration, with $w = 5$. The right panel shows the same configuration $group_1$ with $w = 15$. The red curve corresponds to multi_class, whereas the blue and green ones show the performance of the classifiers class_star and spread_model. The magnitude bin size is 0.25. The green domain corresponds to the stringent limit of the science requirement from WL, on p^g (97.8%, section 2.5.1 of chapter 2), whereas the grey domain corresponds to the science requirement from LSS science on p^g (98.5%, section 2.5.2 of chapter 2). The purity is set to zero if a classifier does not allow to reach the required completeness $c^g = 96.0\%$: here, spread_model does not allow to reach the $c^g = 96.0$ beyond a magnitude of 24. The completeness reached by spread_model on $[24, 24.25]$, $[24.25, 24.5]$, $[24.5, 24.75]$, $[24.75, 25.0]$ are 84.4%,77, 0%, 65.6% and 57.3% respectively. The $\{group_1; 15w\}$ configuration of multi_class allows to reach the $p^g \geq 98.5$ requirement from LSS up to magnitudes similar to spread_model (up to $[23.0, 23.25]$). But it allows to increase the magnitude limit below which the most stringent WL requirement is achieved ($p^g \geq 97.8$) at least two magnitude bins (0.5).

the requirement from LSS, $p^g \geq 98.5$ up to magnitudes similar to spread_model (up to $[23.0, 23.25]$), but it allows to increase the magnitude limit below which multi_class stays consecutively in the domain where the WL requirement (its higher limit $p^g = 97.8\%$, section 2.5.1 of chapter 2) is achieved, by at least two magnitude bins (0.5). Given that this requirement is a higher limit, this improvement is not very significant. Similarly, figure 4.7 shows the purity level p^s at a given completeness, as a function of magnitude in the i band. For bright objects, the case $w = 5$ and $w = 15$ show very similar results. However, for objects fainter than magnitude 22, the case $w = 15$ seems to reach slightly higher purity levels than $w = 5$. Here again, a deeper study of the limiting Fisher discriminant in each case will be necessary to completely understand and take into account this difference (see section 4.4).

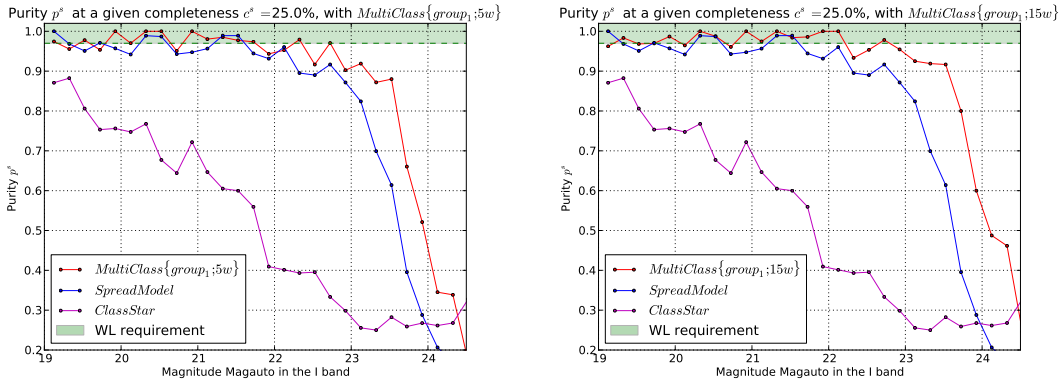

Figure 4.7: Purity level p^s at the required completeness, for the WL and LSS probe, as a function of magnitude in the i band. The left panel corresponds to the $group_1$ configuration, with $w = 5$. The right panel shows the same configuration $group_1$ with $w = 15$. The magnitude bin size is 0.2. The red curve corresponds to the given configuration of multi_class, whereas the blue and green ones show the performance of the classifiers class_star and spread_model. The green domain corresponds to the stringent limit of the science requirement from WL science on p^s (97%, section 2.5.1 of chapter 2). The $\{group_1; 15w\}$ configuration of multi_class allows to reach the WL requirement up to $[22.6, 22.8]$, versus $[21.4, 21.6]$, versus $[21.4, 21.6]$ reached by spread_model. (One should note that the star purity p^s is affected by the high galaxies-to-stars high ratio of our sample).

4.3.3 Comparision of multi_class with other classifiers

Figure 4.8 and 4.9 show the histograms, in linear and logarithmic scale respectively, for two configurations of multi_class, ($\{group_1; 15w\}$ and $\{group_2; 15w\}$), and the two other classifiers presented in chapter 3, spread_model and class_star.

Since all our results show that class_star is outperformed by both spread_model and multi_class on most of the magnitude range considered, so we focus on comparing multi_class and spread_model.

As far as galaxies are concerned, figures **??** and 4.6 show that multi_class reaches higher purity than spread_model, on the entire magnitude range. On both these figures, the purity provided by each classifier is set to zero if the classifier does not allow to reach the fixed completeness (namely $c^g = 88.9$ in figure **??** and $c^g = 96\%$ in figure 4.6). Spread_model does not allow to reach the requirement on the completeness up to a magnitude of 24.5, whereas multi_class does. In particular, spread_model does not reach the $c^g = 88.9\%$ requirement from LSS (derived in chapter 3) in the $[23.75, 24.0]$ bin (see figure **??**), and the $c^g = 96\%$ requirement from WL beyond the $[23, 23.25]$ bin (see figure 4.6).

At both these levels of completeness, the $\{group_1; 15w\}$ configuration of multi_class al-

lows to reach the $p^g \geq 98.5$ requirement from LSS up to magnitudes similar to spread_model (it is reached up to $[23.5, 23.75]$ for $c^g = 88.9\%$, and up to $[23.0, 23.25]$ for $c^g = 96\%$). But it allows to increase the magnitude limit below which the WL requirement is achieved ($p^g \geq 97.8$) by one magnitude bin (0.25 magnitude) in the $c^g = 88.9\%$ case, and by two (0.5 magnitude) in the $c^g = 96\%$ case.

This improvement is more visible in the case of stars: the $\{group_1; 15w\}$ of multi_class allows to reach the WL requirement $p^s \geq 97.0$) up to $[22.6, 22.8]$, versus $[21.4, 21.6]$ reached by spread_model. (One should note that the star purity p^s is affected by the high galaxies-to-stars ratio of our test sample).

Figures 4.10 and 4.11 are the generalisation of this, to a broad range of completenesses. The required threshold on p^g and p^s are shown as the darkest color of the color maps, showing that multi_class allows to widen the range of both completeness and magnitude at which the requirements are achieved. For galaxies (figure 4.10), the main asset of multi_class is that it allows us to reach completenesses which are inaccessible to spread-model. In the domain of completenesses which is accessible to both classifiers, multi_class only allows us to increase the purity reached by spread_model by up to 2%. For stars, multi_class allows us to increase the purity achieved by spread_model at faint magnitudes (typically higher than 23) by up to 46.2%.

The purity-completeness domains showed in figures 4.4 and 4.5, show that the $\{group_1; 15w\}$ configuration of multi_class outperforms spread_model in each of the magnitude ranges $[18.0, 21.0]$, $[21.0, 23.0]$, $[23.0, 24.0]$, as well as on the the full range of the catalog, $[18.0, 24.5]$.

4.4 Conclusion

The beginning of the testing of multi_class on the DES SV data allowed us to find one configuration of our method which outperforms spread_model and class_star on the full range of magnitudes of the catalog. DES began survey operations in September, 2013, and will be running for five years. Therefore, we should be able to continue the testing and optimisation of multi_class. More cross-matching, and especially the planned addition of stars from the LMC in the cross-matched catalog, should allow a better training of our classifier.

Future works should include an optimization of the grouping strategy, through a careful

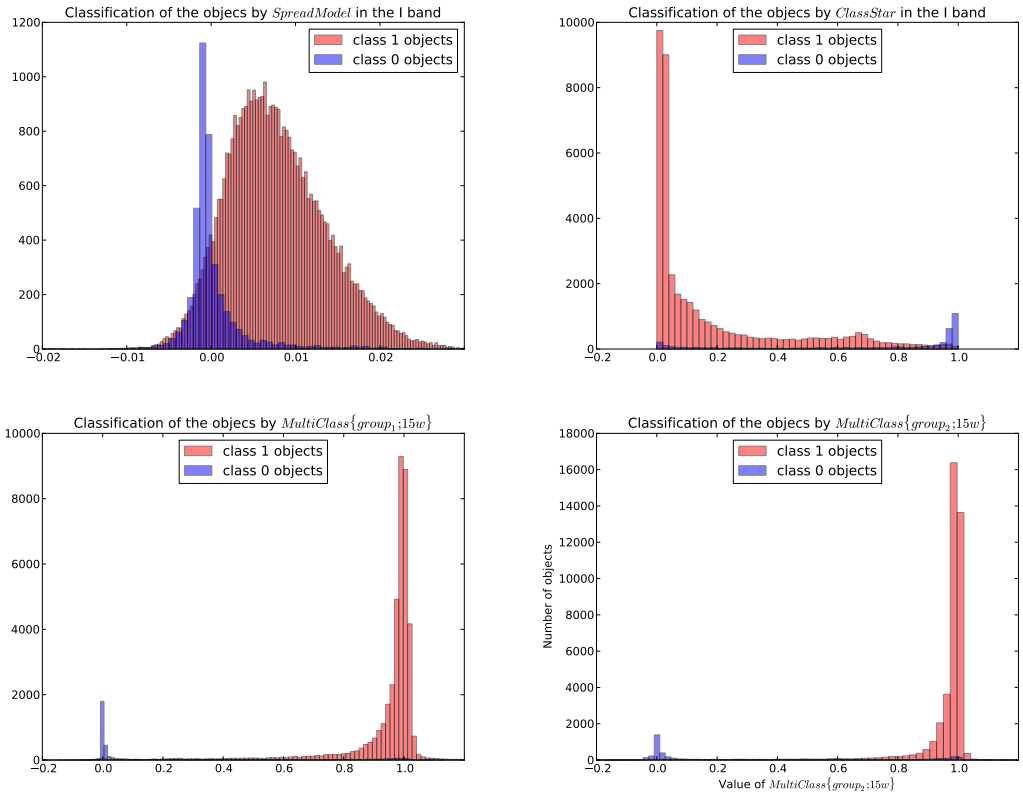

Figure 4.8: Histogram of some of the classifiers compared in this chapter. The upper panels show spread_model (left) and class_star (right), whereas the two lower panels show two different configurations of multi_class: $\{group_1; 15w\}$ (left) and $\{group_2; 15w\}$ (right).

study of the threshold value of the Fisher discriminant of the principal components which optimises the performance of multi_class. In other words, we should study the way in which the optimal number of winners $w(f)$ depends on the Fisher discriminant of the selected principal component, which would allow to make the method even more flexible, by setting the threshold on the Fisher discriminate, rather than the number of "winners".

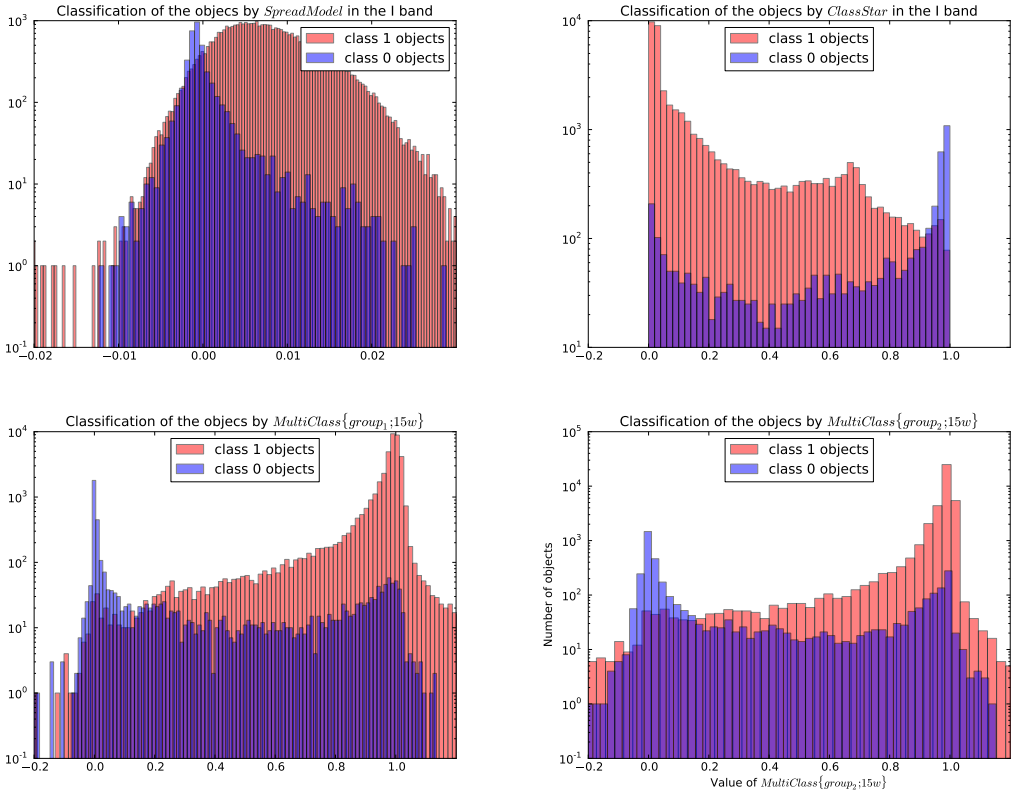

Figure 4.9: Histogram of some of the classifiers compared in this chapter, in logarithmic scale. The upper panels show spread_model (left) and class_star (right), whereas the two lower panels show two different configurations of multi_class: $\{group_1; 15w\}$ (left) and $\{group_2; 15w\}$ (right).

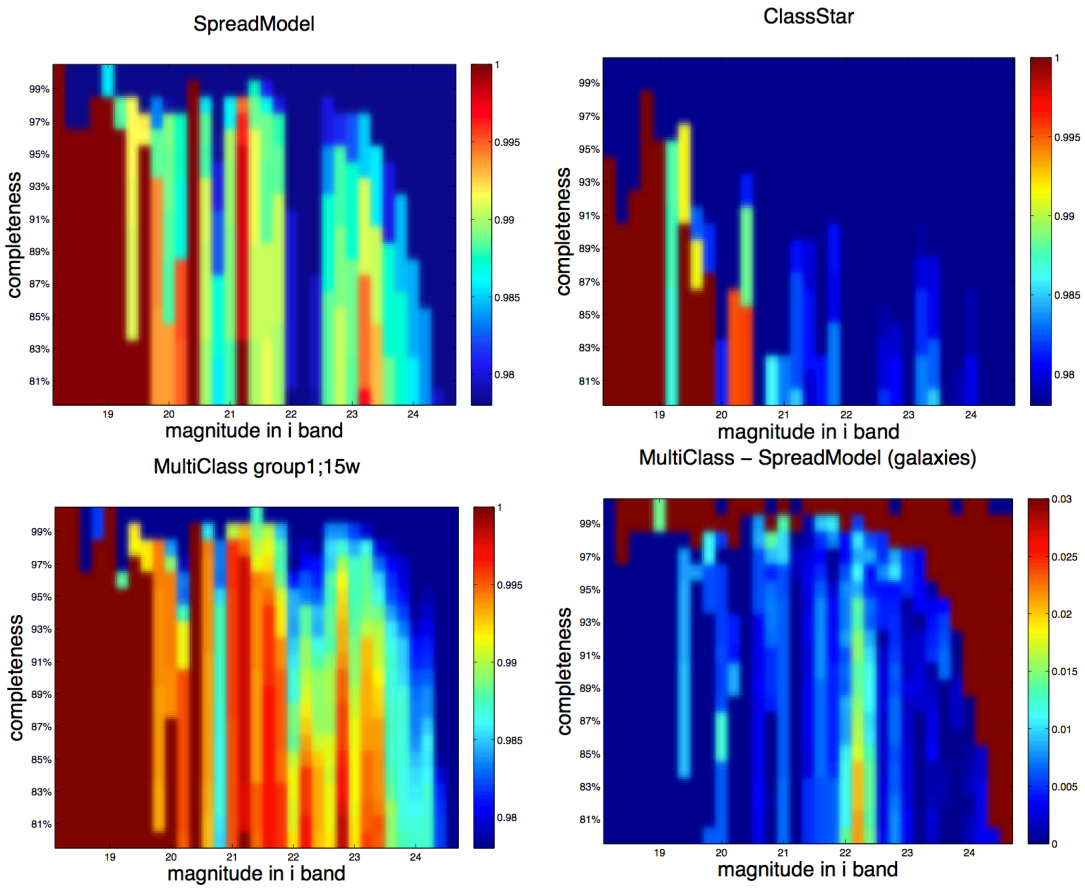

Figure 4.10: Level of purity for a sample of galaxies p^g, for different magnitudes and values of the completeness $c^g \in [80\%, 100\%]$, with spread_model (top left), class_star (top right) and the $\{group_1; 15w\}$ configuration of multi_class (lower left). Any purity below the 97.8% level requirement from WL (section 2.5.1 of chapter 2) is shown in dark blue. At the 88.9% completeness level, class_star does allows us to achieve the required on the purity above 21.9, although not consecutively. Spread_model allows to reach 24, and Multi_class allows us to achieve the requirement up 24.2. The lower right panel shows the improvement by multi_class with respect to spread_model. In the red area, the completeness is achieved by multi_class and not by spread_model. This constitutes the main asset of multi_class. In the domain of completenesses which is accessible to both classifiers, multi_class only allows us to increase the purity reached by spread_model only by up to 2%.

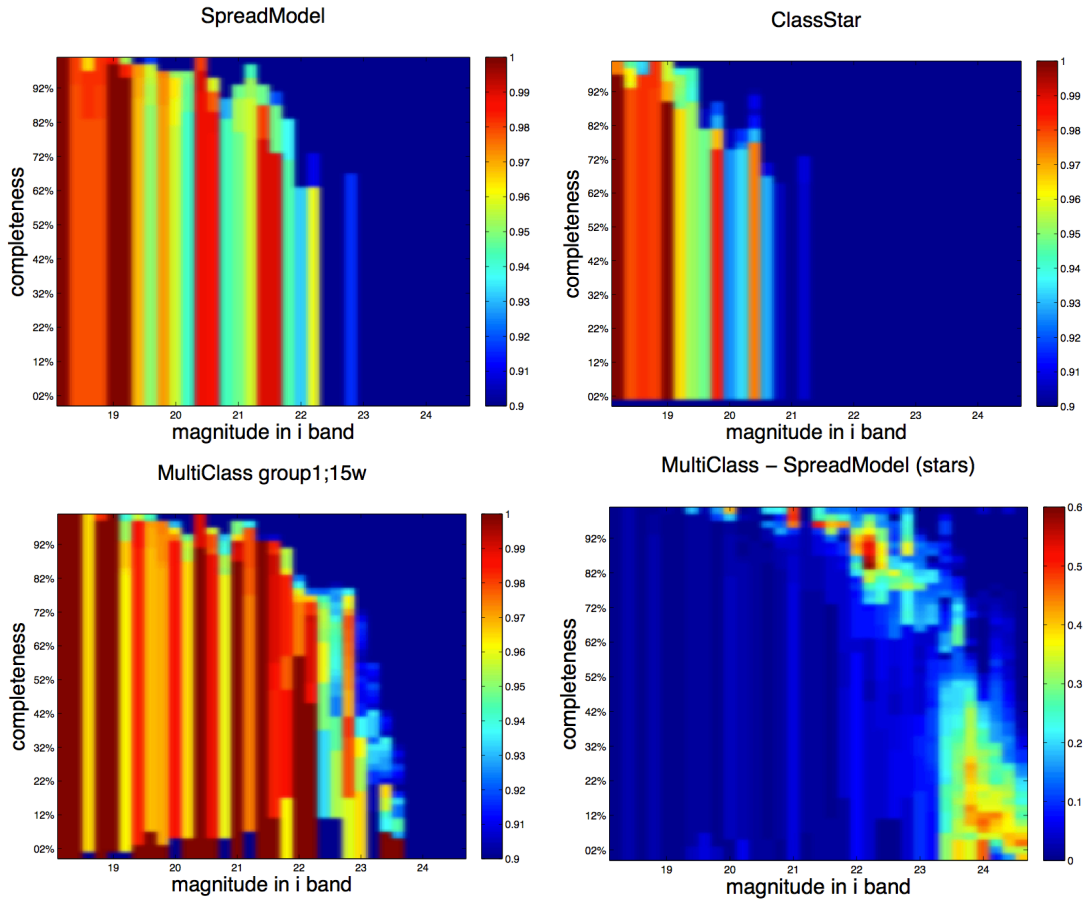

Figure 4.11: Level of purity for a sample of stars p^s, for different magnitudes and values of the completeness $c^s \in [0\%, 100\%]$, with spread_model (top left), class_star (top right) and the $\{group_1; 15w\}$ configuration of multi_class (lower panel). Any purity below the 90% level is shown in dark blue, and the 97% requirement from WL (section 2.5.1 of chapter 2) is shown in orange. class_star does not allow to achieve the WL requirement above a magnitude of 20.4, versus 21.6 for spread_model. Multi_class allows us to reach 22.8, at the $c^s \geq 25\%$ level required. Multi_class widen the range of both completeness and magnitude at which high purity levels ($\geq 90\%$) are achieved. The lower right panel shows the improvement by multi_class with respect to spread_model. At faint magnitude, typically higher than 23, the improvement by multi_class reaches 46.2%.

Chapter 5

Scale Dependent Bias from Baryon Acoustic oscillation: detection in the BOSS DR9 and DR10 data releases

"It remains a miracle (to us) that the optimization of the chi-square objective [...] has a linear solution. One can attribute this to the magical properties of the Gaussian distribution, but the Gaussian distribution is also the maximum-entropy distribution [...] constrained to have zero mean and a known variance; it is the limiting distribution of the central limit theorem. That this leads to a convex, linear algebra solution is something for which we all ought to give thanks."

David Hogg, Jo Bovy, Dustin lang

"The chance which now seems lost may present itself at the last moment."

Jules Verne, *Around the World in Eighty Days*

In this chapter we conduct the first observational search for an effect partially modeled in Barkana & Loeb (2011): the modulation, from Baryon Acoustic Oscillations (BAOs) of the baryon-to-matter ratio of density fluctuations across large regions of the Universe.

In section 5.1, we present this effect. The specific way in which we choose to formulate the model for this effect is explained in section 5.2. We then explain our measurement in section 5.3. In section 5.4, we present our model-fitting strategy. We show the results of simulations we conducted to test the feasibility of our search in section 5.5. The result of the model fitting on real data are shown in 5.6. We present our model selection strategy and its main results in section 5.7. We give concluding remarks in section 5.8.

5.1 Introduction: a model for a scale dependent bias from BAOs

5.1.1 Predictions from Barkana & Loeb (2011)

The acoustic waves which propagated in the baryon-radiation fluid before the time of recombination, known as Baryon Acoustic Oscillations (BAOs), where only felt by the baryonic part of the total matter and not followed by Dark Matter (DM). We cannot measure directly δ_b and δ_{tot}, but rather observable *tracers* of these quantities. Comparing any observable tracing the *total* matter density fluctuation δ_{tot}, with an observable tracing *only* the baryons δ_b should allow us to measure this effect.

The idea of the work conducted here is to use:

- the number density δ_n of galaxies as a tracer of the total matter density fluctuation δ_{tot},
- the absolute luminosity density of galaxies as a tracer of the baryonic density fluctuation δ_b.

We now detail these two assumptions.

Number density δ_n as a tracer of the matter density fluctuation δ_{tot}

Galaxies sample the high peaks of the total matter density. The number density fluctuations δ_n are driven by the underlying total matter density fluctuation δ_{tot}, with a bias b_n, which should be approximately constant on large scales. In the case of flux-limited surveys, a scale-dependent bias is added, as will be explained later in this section, but for now we treat the more general case:

$$\delta_n = b_n \cdot \delta_{tot} \tag{5.1}$$

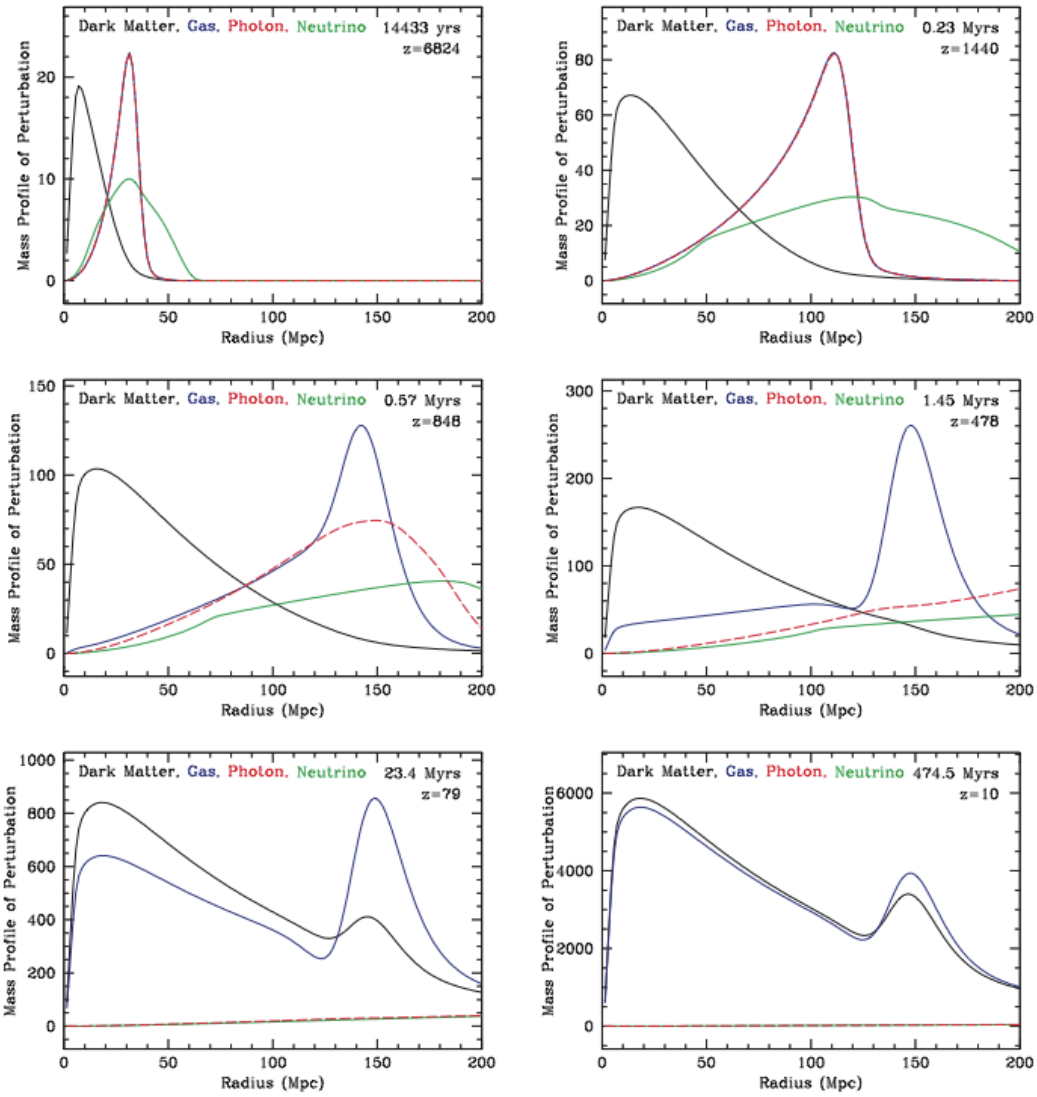

Figure 5.1: Snapshots of evolution of the radial mass profile vs. comoving radius of an initially point-like over-density located at the origin. All perturbations are fractional for that species; moreover, the relativistic species have had their energy density perturbation divided by 4/3 to put them on the same scale. The black, blue, red, and green lines are the dark matter, baryons, photons, and neutrinos, respectively. The redshift and time after the big bang are given in each panel. The units of the mass profile are arbitrary but are correctly scaled between the panels for the synchronous gauge. Top left: Near the initial time, the photons and baryons travel outward as a pulse. Top right: Approaching recombination, one can see the wake in the cold dark matter raised by the outward-going pulse of baryons and relativistic species. Middle left: At recombination, the photons leak away from the baryonic perturbation. Middle right:With recombination complete, we are left with a CDM perturbation toward the center and a baryonic perturbation in a shell. Bottom left: Gravitational instability now takes over, and new baryons and dark matter are attracted to the over-densities. Bottom right: At late times, the baryonic fraction of the perturbation is near the cosmic value, because all of the new material was at the cosmic mean. These figures were made by suitable transforms of the transfer functions created by CMBFAST (Seljak & Zaldarriaga 1996; Zaldarriaga & Seljak 2000). Credit: Eisenstein et al. (2007b).

Absolute luminosity fluctuation δ_L as a tracer of the baryon density fluctuation δ_b

An area with a higher baryonic mass fraction δ_b/δ_{tot} than average is expected to produce more stars per unit total mass, hence more luminous matter, and to result in galaxies with lower mass-to-light ratio. The scale-dependency of δ_b/δ_{tot} induced by BAOs, should translate into a scale dependency of δ_L/δ_{tot}, where δ_L is the absolute luminosity weighted density fluctuation.

This being said, the mean luminosity of a given galaxy population relates to the baryonic content of the surrounding in a non-trivial way. The link between them is a combination of

1. the way in which the luminosity of a galaxy depends on the baryon fraction of the host halo,

2. the way in which the baryonic content of the host halo reflects the underlying baryonic contribution to the total matter density.

The luminosity density ρ_L, for a given population of galaxies, is given by

$$\rho_L = n_{gal} \langle L \rangle \ , \tag{5.2}$$

where $\langle L \rangle$ is the mean absolute luminosity of the galaxies.

In terms of fluctuations $\delta_L = (\rho_L - \overline{\rho_L})/\overline{\rho_L}$, this translates into

$$\frac{d\rho_L}{\rho_L} = \frac{dn_{gal}}{n_{gal}} + \frac{d\langle L \rangle}{\langle L \rangle} \tag{5.3}$$

We have seen that $\delta_n = \frac{dn_{gal}}{n_{gal}} = b_n \delta_{tot}$ (when not considering a flux-limited survey). In fact, the mean luminosity $\langle L \rangle$ of a population of galaxies may also depend on δ_{tot}, through its merger history. We model this dependency with a different bias $b_n + b_{L;t}$:

$$\delta_L = (b_n + b_{L;t}) \cdot \delta_{tot} \tag{5.4}$$

This would be right if $\langle L \rangle$ only depended on the large scale matter density. However, $\langle L \rangle$ also depends on the baryon fraction in the host halo, f_b. Following Barkana & Loeb (2011), we assume that $\langle L \rangle \propto (f_b)^{b_{L;f}}$, where $b_{L;f} \approx 1.4$ is the bias factor of the luminosity

density with respect to the halo baryon fraction. Hence equation 5.4 becomes

$$\delta_L = (b_n + b_{L;t}) \cdot \delta_{tot} + b_{L;f}\delta_f \tag{5.5}$$

The link between δ_f - the baryonic content of the halo - and δ_b - the baryonic content of the surrounding - is complex because of the non-linearity of halo collapse. It is derived in Barkana & Loeb (2011) as,

$$\delta_f = \frac{A_r}{\delta_c}[r(k) - r_{lss}]\delta_{tot} \ , \tag{5.6}$$

We do not re-derive this equation, which was taken as a presupposition in our work, but explain each term:

- $r(k)$ is the fractional baryon deviation $r(k) = \delta_b/\delta_{tot} - 1$, shown in figure 5.2 as a function of the scale k, and at various redshifts. $r(k)$ approaches a constant r_{lss} which depends on the redshift, on scales below the BAOs.
- δ_c is the critical total matter density δ_{tot} of the halo at which the critical density of collapse is independent of mass and is equal to $1,69$ in the Eistein-De Siter limit, valid over a wide range of redshifts, (Naoz & Barkana 2007).
- A_r is a corrective amplification factor coming from the use of the linear $r(k)$ in the non-linear halo collapse problem, and is expected to be $A_r \approx 3$, from simulations computed in Barkana & Loeb (2011).

Hence, the final equation for δ_L is

$$\delta_L = (b_n + b_{L;t}) \cdot \delta_{tot} + b_{L,\Delta}(r(k) - r_{lss}) \cdot \delta_{tot} \ , \tag{5.7}$$

where $b_{L,\Delta}$ is a bias factor measuring the overall dependence of galaxy luminosity and the underlying difference between the baryon and total density fluctuations and is predicted in Barkana & Loeb (2011) to be around $b_{L,\Delta} \approx 2.5$.

Flux-limited surveys

In Barkana & Loeb (2011), the authors show that, in the case of a flux limited survey, both equations 5.1 and 5.7 must be slightly rethought. In such surveys, observed samples

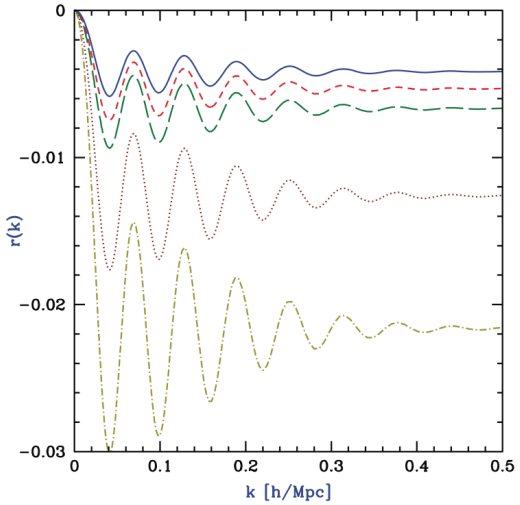

Figure 5.2: The fractional baryon derivation $r(k) = (\delta_b/\delta_{tot}) - 1$, as a function of the scale k, at various redshifts (z = 0,0.5,1,3 and 6 from top to bottom). Credit: Barkana & Loeb (2011).

are limited by flux, or equivalently by luminosity if, for simplicity, we consider samples at a single redshifts. The number of observable galaxies per unit of volume is given by

$$F(L) = \int_{L'=L}^{\infty} \phi(L')dL' \ , \tag{5.8}$$

where ϕ is the luminosity function. The observed luminosity density of these galaxies becomes

$$\rho_{obs} = \langle L \rangle \, F(L), \tag{5.9}$$

where the mean luminosity of the sample $\langle L \rangle$ is now defined as

$$\langle L \rangle = \frac{1}{F(L)} \int_{L'=L}^{\infty} L'\phi(L')dL' \ . \tag{5.10}$$

One can then show that equations 5.1 and 5.7 can be rewritten as

$$\delta_n = (b_n + C_{min} b_{L,t})\delta_{tot} + C_{min} b_{L,\Delta}[r(k) - r_{lss}]\delta_{tot} \ , \tag{5.11}$$

$$\delta_L = (b_n + (1 + D_{min})b_{L,t})\delta_{tot} + (1 + D_{min})b_{L,\Delta}[r(k) - r_{lss}]\delta_{tot} \ , \tag{5.12}$$

where $C_{min} = \frac{L_{min}\phi(L_{min})}{F(L_{min})}$ and $D_{min} = \frac{L_{min}}{\langle L \rangle} C_{min}$ with $\langle L \rangle$ evaluated for $L = L_{min}$.

5.1.2 Compensated Isocurvature Perturbations

The measures of the relation between dark matter and baryons, is related to the search for Compensated Isocurvature Perturbations (CIPs) (see Grin et al. (2014)).

Measurements of primordial density perturbations are consistent with adiabatic initial conditions, for which the ratios of neutrino, photon, baryon and DM number densities are initially spatially constant. On the one hand, the simplest inflationary models predict adiabatic fluctuations (Guth & Pi 1982; Linde 1982). On the other, hand, more complex inflationary scenari (Brandenberger 1994; Linde 1984; Axenides et al. 1983) predict fluctuations on the relative number densities of different species, known as Isocurvature Perturbations (IPs). CMB temperature anisotropies limit the contribution of both baryons and DM to the total IPs amplitude. "Compensated Isocurvature Perturbations" (CIPs) are perturbations in the baryons density δ_b which are compensated for by corresponding fluctuations in the DM δ_{DM}, so that the total density is unchanged. Such fluctuations are very hard to detect, since gravity and all its effects measurable by galaxy surveys, including galaxy numbers, only depend on the total density. Galaxy clusters gas fractions observations (Holder et al. 2010) have led to a weak constraint of the CIPs. 21cm absorption observations would hardly allow a better constraint (Gordon & Pritchard 2009) of such perturbations.

Under the the standard assumption of a scale-invariant power spectrum for this field, equations 5.12 and 5.11 are modified to

$$\delta_n = (b_n + C_{min} b_{L,t})\delta_{tot} + C_{min} b_{L,\Delta}[(r(k) - r_{lss})\delta_{tot} + \delta_{CIP}] \,, \tag{5.13}$$

$$\delta_L = (b_n + (1 + D_{min}) b_{L,t})\delta_{tot} + (1 + D_{min}) b_{L,\Delta}[(r(k) - r_{lss})\delta_{tot} + \delta_{CIP}] \,, \tag{5.14}$$

where δ_{CIP} is a separate field that is uncorrelated with δ_{tot}. We hope to improve the 10^{-2} current constraint on the amplitude of a scale invariant CIPs power spectrum (Grin et al. 2014).

Assumptions and limitations of the Barkana & Loeb (2011) model

The model presented in Barkana & Loeb (2011) is based on

1. the standard theoretical understanding of galaxy formation;

2. a standard observationally-based result for star formation (linking between the luminosity and the gas fraction in haloes $\langle L \rangle \propto (f_b)^{b_{L;f}}$);

3. a result from numerical simulations (linking between the baryon fraction in the linearly extrapolated halo perturbation and the baryon fraction in the actual virialised halo).

The limitations of this model is that the above ingredients of the model are fairly complex, and the specific effect predicted is subtle and has not been directly tested in either simulations or observations.

5.1.3 Three reasons to be excited

Detecting the imprint left by BAOs in the δ_b/δ_{tot} would be important for three reasons:

1. The *detection* of the effect would provide a direct measurement of a difference in the large-scale clustering of mass and light and a confirmation of the standard cosmological paradigm from a different angle than any other measurement. In this sense, the detection of this effect will help rule out alternative theories of gravity such as non-DM models such as MOND (Milgrom 1994). This would provide evidence as significant as the bullet cluster evidence, with the additional advantage that this effect happens on linear scales.

2. The *amplitude* of the effect would allow to calibrate the dependence of the characteristic mass-to-light ratio of galaxies on the baryon mass fraction of their large scale environment.

3. The resulting measurement of A_{CIP} could provide a much better constant on CIPs than the current limits, (10^{-2} for the amplitude of a scale-invariant CIPs power spectrum, Grin et al. (2014)).

5.1.4 Outline of the project

To summarise, we aim to measure the δ_b/δ_{tot} ratio, through a measurement of δ_n and δ_L, the link between them being described by equations 5.11 and 5.12. In order to measure $b_{L,\Delta}$, we proceed in several steps:

- In section 5.2, we reformulate the whole model in terms of relevant observables: the two points correlation functions ξ_L and ξ_n of the tracers density fluctuations δ_n and δ_L.
- In section 5.3, we measure ξ_L and ξ_n
- In section 5.5, we create some simulations of ξ_L and ξ_n, to check the feasibility of such a detection and whether our model fitting procedure allows to recover a fiducial model.
- We then conclude on the significance of our detection with a model selection calculation, in section 5.7.

5.2 Model in terms of correlation function

In the previous section we presented our model for the tracers of δ_b and δ_{tot}: the stochastic fluctuations δ_n and δ_L. However, the observable quantities in galaxy surveys are the two point statistics of such fluctuations, namely the power spectrum or the two-point correlation functions (2PCF). We take a different approach than in Barkana & Loeb (2011), where the observational proposal to verify equations 5.11 and 5.12 is formulated in terms of power spectra, and reformulate them in term of the 2PCF:

The 2PCF is defined as

$$\xi(\mathbf{x},\mathbf{y}) \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{d^3\mathbf{k}'}{(2\pi)^{3/2}} \left\langle \delta(\mathbf{k})\delta(\mathbf{k}') \right\rangle e^{i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}\cdot\mathbf{y}} = \frac{1}{2\pi^2} \int k^2 P(k) j_0(kr) dk \ , \quad (5.15)$$

where $P(k)$ is the power spectrum defined by $\langle \delta(k)\delta(k') \rangle \equiv P(k)\delta^D(k-k')$. In real space, and assuming $|r(k) - r_{lss}| << 1$, equation 5.11 and equation 5.12 translate into the following,

$$\xi_n = b_1^2 \cdot \xi_{tot} + 2b_1 b_2 \cdot \xi_{add} + b_2^2 \cdot \xi_{CIP} \quad (5.16)$$

$$\xi_L = b_3^2 \cdot \xi_{tot} + 2b_3 b_4 \cdot \xi_{add} + b_4^2 \cdot \xi_{CIP} \ , \quad (5.17)$$

$$= \xi_n + (b_3^2 - b_1^2) \cdot \xi_{tot} + 2(b_3 b_4 - b_1 b_2) \cdot \xi_{add} + (b_4^2 - b_2^2) \cdot \xi_{CIP} \ ,$$

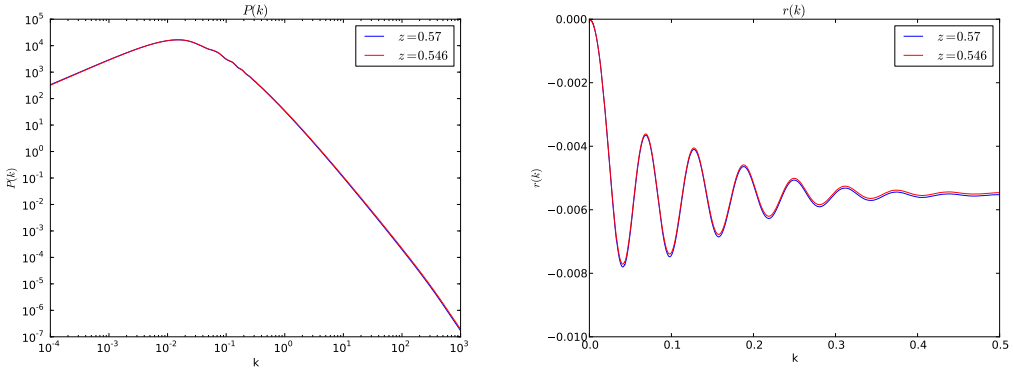

Figure 5.3: $P(k)$ and $r(k)$ computed with CAMB (Lewis et al. 2000), shown at the DR9 median redshift ($z = 0.546$) and the DR10 median redshift ($z = 0.570$).

with

$$b_1 = b_n + C_{min} b_{L,t} \tag{5.18}$$

$$b_2 = C_{min} b_{L,\Delta} \tag{5.19}$$

$$b_3 = b_n + (1 + D_{min}) b_{L,t} \tag{5.20}$$

$$b_4 = (1 + D_{min}) b_{L,\Delta} \tag{5.21}$$

The correlation functions, and the acoustic signatures imprinted in it evolve with time, and a key issue in order to model them is to understand how. We first adopt a simplistic approach and model ξ_{tot} and ξ_{add} with a linear perturbation theory, in section 5.2.1. We then correct for the non-linearity of the clustering of galaxies in section 5.2.2.

5.2.1 Linear-regime matter correlation

Power spectrum of δ_{tot}

To model ξ_{tot} and ξ_{add} in equations 5.16 and 5.17, we first compute a linear power spectrum $P(k)$ and a linear fractional baryon deviation $r(k)$ using CAMB (Lewis et al. 2000). We assume the same fiducial ΛCDM+GR, flat cosmological model with $\Omega_m = 0.274$, $h = 0.7$, $\Omega_b h^2 = 0.0224$, $n_s = 0.95$ and $\sigma_8 = 0.8$, matching that used by the BOSS collaboration in Anderson et al. (2013). P(k) and r(k) are computed for the median redshifts of the two samples we use, namely BOSS DR9 and DR10 releases (see section 5.3).The computed $P(k)$ and $r(k)$ are shown in figure 5.3

Power spectrum of δ_{CIP}

To model ξ_{CIP}, we make the standard assumption that the power spectrum of the CIP field is of the form $P_{CIP}(k) = A_{CIP}k^{-3}$ (Grin et al. 2014). Since the corresponding correlation function

$$\xi_{CIP}(r) = \frac{1}{2\pi^2}\int k^2 P_{CIP}(k)\cdot j_0(kr)dk = \frac{A_{CIP}}{2\pi^2}\int \frac{j_0(kr)}{k}dk$$

diverges, we start the integral at $k = 10^{-4}h/Mpc$, which is also the minimum value of the CAMB linear power spectrum we use to model ξ_{tot}.

5.2.2 Corrections to the linear correlation function

The matter correlation ξ_{tot} predicted by linear perturbation theory does not exactly describe the clustering of galaxies. Nonlinear gravitational collapse and redshift distortions modify galaxy clustering relative to that of the linear-regime matter correlations, changing the shape of the correlation function. In particular, according to linear perturbation theory, the acoustic signature increases in amplitude but its spatial pattern remains static, i.e. the characteristic scale imprinted in the early universe remains unaltered, whereas non linear growth of structure leads to a shift of the acoustic peak.

These corrections are particularly important when using the BAO peak as a standard ruler and as a probe of the universe expansion history: a shift in the acoustic scale of one percent generates systematics in the deduced dark energy equation of state parameter w of about five percent (Eisenstein et al. 2005), which is not negligible compared to the expected statistical errors in the next generation of galaxy surveys.

For our work , which consists of looking for a particular signature of the BAOs, these corrections are likely to be important too. We account for two systematic effects due to nonlinear clustering: damping of the BAO peak and mode coupling.

Damping

Simulations have shown that nonlinear structure formation and, to a lesser extent, redshift distortions erase the higher harmonics of the acoustic oscillations. This degrades the measurement of the acoustic scale. This effect is accounted for by "damping" the linear theoretical BAO on small scales. The damping term is often approximated by a Gaussian smoothing (Percival et al. 2010). The corrected correlation function convoluted with the

damping term, which, in the Fourier space, comes to multiplying the power spectrum by the damping term transform,

$$\xi_{tot}(r) = \xi(r) \otimes e^{-(k_* \cdot s)^2} = \frac{1}{2\pi^2} \int k^2 P(k) e^{-(k_* \cdot k)^2} j_0(kr) dk \ , \tag{5.22}$$

The damping is also applied to ξ_{add} and ξ_{CIP}. The effect of the damping on $P(k)$, ξ_{tot} and ξ_{add} is shown in figure 5.4, its effect on our model for ξ_n and ξ_L is shown in figure 5.15.

Mode coupling

Mode coupling generates additional oscillations that are out of phase with those in the linear spectrum, leading to shifts in the scales of oscillation nodes defined with respect to a smooth spectrum. When Fourier transformed, these out-of-phase oscillations induce percent-level shifts in the acoustic peak of the two-point correlation function. The corresponding correction to the damped linear correlation function is given in Crocce & Scoccimarro (2008), as:

$$\xi_{tot}(r) = \xi(r) \otimes e^{-(k^* \cdot s)^2} + A_{MC} \xi'(r) \xi^{(1)}(r) \ , \tag{5.23}$$

where $\xi(r)$ denotes the linear correlation function of equation 5.15 and

$$\xi'(r)\xi^{(1)}(r) = \int \frac{d^3k}{k} P(k) j_1(kr) \ , \tag{5.24}$$

where j_1 is the first order Bessel function. The effect of the mode coupling term on our model is shown in figure 5.16.

5.2.3 Systematics

The systematic effects which have been found in the BOSS data cause changes that are quite close to constant shifts and there are theoretical systematics (e.g., the integral constraint) that cause close to constant shifts. A simple way to account for systematics that would affect ξ_L and ξ_n differently is to add a constant to the model in equation 5.17. Thus, equation 5.17 becomes

$$\xi_L = b_3^2 \cdot \xi_{tot} + 2 b_3 b_4 \cdot \xi_{add} + b_4^2 \cdot \xi_{cip} + b_{sys} \ , \tag{5.25}$$

New model equations

Our final model equations are:

$$\xi_n = b_1^2 \cdot \xi_{tot} + 2b_1 b_2 \cdot \xi_{add} + b_2^2 \cdot \xi_{CIP} \tag{5.26}$$

$$\xi_L = b_3^2 \cdot \xi_{tot} + 2b_3 b_4 \cdot \xi_{add} + b_{sys} + b_{CIP} \cdot \xi_{cip} \ , \tag{5.27}$$

where

$$\xi_{tot}(r) = \xi(r) \otimes e^{-(k^* \cdot s)^2} + A_{MC}\xi'(r)\xi^{(1)}(r) \ . \tag{5.28}$$

$$\xi(r) = \frac{1}{2\pi^2} \int k^2 P(k) j_0(kr) dk \tag{5.29}$$

$$\xi_{add}(r) = \left(\int k^2 (r(k) - r_{lss}) P(k) \cdot j_0(kr) dk \right) \otimes e^{-(k^* \cdot s)^2}$$

$$= \frac{1}{2\pi^2} \int k^2 (r(k) - r_{lss}) P(k) e^{-(k_* \cdot k)^2} j_0(kr) dk \ , \tag{5.30}$$

$$\xi_{CIP}(r) = \left(\frac{1}{2\pi^2 A_{CIP}} \int k^2 P_{CIP}(k) \cdot j_0(kr) dk \right) \otimes e^{-(k^* \cdot s)^2}$$

$$= \frac{1}{2\pi^2} \int \frac{j_0(kr)}{k} e^{-(k_* \cdot k)^2} dk \ . \tag{5.31}$$

where $b_{CIP} = (b_4^2 A_{CIP})$ and $P(k)$ is the linear matter power spectrum.

In order to compute the oscillatory integral ξ_{tot}, ξ_{add} and ξ_{CIP}, we wrote a Python wrapper for the fftlog code from Hamilton (2000).

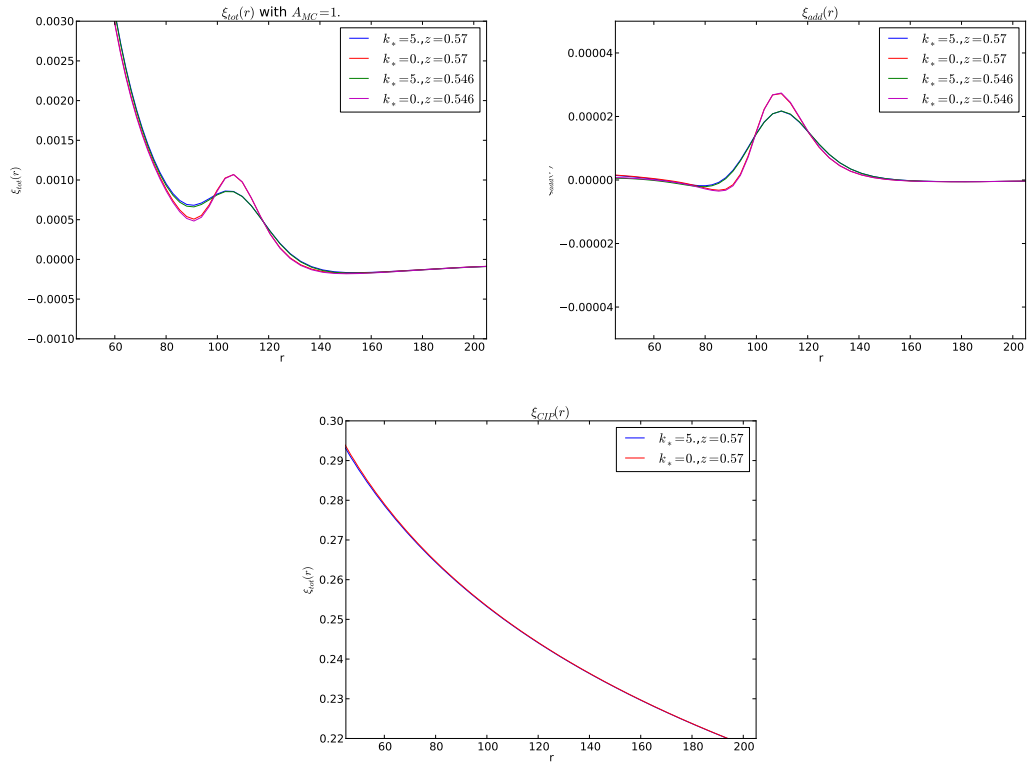

Figure 5.4: The three ingredients of the model, ξ_{tot}, ξ_{add} and ξ_{cip}, shown at the DR9 median redshift ($z = 0.546$) and the DR10 median redshift ($z = 0.57$). ξ_{tot} and ξ_{add} are shown for two different values of the damping parameter k_*.

5.3 Measurement

5.3.1 The BOSS DR9 and DR10 samples

In all this analysis, we use the latest data releases from the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS), DR9 and DR10 (Ahn et al. 2014). The BOSS collaboration latest release, DR11, is to be publicly released with the final BOSS data set, but the covariance matrix and correlation function are already public, and we use them to to check the consistency of some of our measurements with the BOSS ones.

The Sloan Digital Sky Survey (York et al. 2000), divided into SDSS I, II (Abazajian et al. 2009), and III (Eisenstein et al. 2011), used a drift-scanning mosaic CCD camera (Gunn et al. 1998) to image over one third of the sky (14 555 square degrees) in five photometric bandpasses (Fukugita et al. 1996; Doi et al. 2010) to a limiting magnitude of $r \approx 22.5$ using the dedicated 2.5-m Sloan Telescope located at Apache Point Observatory in New Mexico.

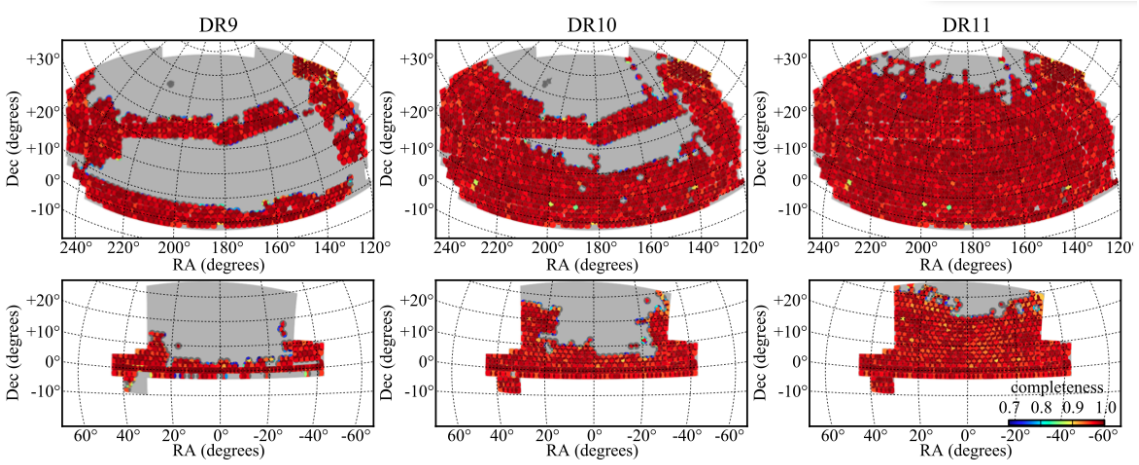

Figure 5.5: Evolution of the BOSS sky coverage from DR9 to DR11. Top panels show the observations in the North Galactic Cap (NGC) while lower panels show observations in the South Galactic Cap (SGC). Colors indicate the spectroscopic completeness within each sector as indicated in the key in the lower right panel. Gray areas indicate the BOSS expected footprint upon completion of the survey. The total sky coverage in DR9, DR10, and DR11 is $3275 deg^2$, $6161 deg^2$, and $8377 deg^2$, respectively. Credit: Anderson et al. (2013).

BOSS is primarily a spectroscopic survey, which is designed to obtain spectra and redshifts for 1.35 million galaxies over an extragalactic footprint covering 10 000 square degrees. These galaxies are selected from the SDSS DR8 imaging. Together with these galaxies, 160 000 quasars and approximately 100 000 ancillary targets are being observed. The method for obtaining the spectra (Smee et al. 2013) ensures a homogeneous data set with a high redshift completeness of more than 97 per cent over the full survey footprint. Redshifts are extracted from the spectra using the methods described in Bolton et al. (2012). A summary of the survey design appears in Eisenstein et al. (2011), and a full description is provided in Dawson et al. (2013). Figure 5.5 shows the evolution of the sky coverage between the data releases DR9 and DR11.

Two classes of galaxies were selected by BOSS to be targeted for spectroscopy using SDSS DR8 imaging. The "LOWZ" algorithm is designed to select red galaxies at $z < 0.45$ from the SDSS DR8 imaging data. The CMASS sample, which is the one used in this analysis, is designed to be approximately stellar-mass-limited above $z = 0.45$.

The CMASS Data Release 9 (DR9) contains 264283 massive galaxies covering 3275 square degrees with a median redshift $z = 0.546$, whereas DR10 has 501844 galaxies covering 6161 square degrees, almost three times the coverage of DR9, and a median redshift z = 0.57.

A full description of DR9 and DR10 can be found in Anderson et al. (2012) and Anderson et al. (2013) respectively.

5.3.2 Measurement of the correlation functions

Practical issues

There are many practical problems which inhibit our ability to accurately measure the statistics of the galaxy distribution, as defined in equation 5.15. In theory we usually consider smooth density fields, $\delta(x)$, however in reality the underlying density field is discretely sampled by individual galaxies. This leads to shot noise problems on small scales. Also, the observed galaxies are not contained within a regular shaped contiguous region of space. It is typical for galaxy redshift surveys to be irregularly shaped in angular sky coverage due to dust extinction, bright stars, tracking of the telescope, etc. In taking these issues into account, we must use statistical estimators which can deal with such problems and are optimised for our purposes (see Percival (2007), for a review of correlation function practicalities).

Estimator

The correlation functions are based on our ability to compare the distribution of data (galaxies) to a random sampling. The 2PCF compares the number of pairs of galaxies (DD) with pairs of random points (RR) at some fixed separation, r, (see Kerscher et al. 2000, for a good review of correlation estimators). In our analysis, the two-point correlation functions ξ_n and ξ_L, are computed using the optimal Landy-Salay estimator (Landy & Szalay 1993) which requires the creation of a catalog of random positions. The computation then uses pair counts between the galaxy-galaxy samples (DD), the random-random (RR) and the cross counts between galaxy-random points (DR).

$$\xi = \frac{DD - 2DR + RR}{RR}, \tag{5.32}$$

In the above expression we have adopted the following definitions:

$$DD = \frac{1}{N_D(N_D-1)} \sum_{i}^{N_D} \sum_{j}^{N_D} \sqcap(|\vec{x}_i^{data} - \vec{x}_j^{data}|) \ , \tag{5.33}$$

$$DR = \frac{1}{N_D(N_R)} \sum_{i}^{N_D} \sum_{j}^{N_D} \sqcap(|\vec{x}_i^{rand} - \vec{x}_j^{data}|) \ , \tag{5.34}$$

$$RR = \frac{1}{N_R(N_R-1)} \sum_{i}^{N_D} \sum_{j}^{N_D} \sqcap(|\vec{x}_i^{rand} - \vec{x}_j^{rand}|) \ , \tag{5.35}$$

where $\vec{x}^{rand}$ and $\vec{x}^{data}$ are the position vectors of the randoms and data respectively and $\sqcap$ is the rectangular step function defined as,

$$\sqcap(t) = \begin{cases} 0 & \text{if } |t-r| > dr/2 \\ 1 & \text{if } |t-r| < dr/2 \end{cases} \tag{5.36}$$

This function selects only pairs of points separated by a distance r in a bin of width dr. Then we can see that in equation 5.32, DD, DR and RR represent the number of pairs separated by r and normalised by the total number of possible pairs.

Computation

In our analysis, the counts DD, DR and RR are computed using an efficient tree-based, parallel, search algorithm. The code is based upon the structure known as "kd-trees" which is a way of organizing a set of data in k-dimensional space in such a way that once built, any query requesting a list of points in a neighborhood can be answered quickly without going through every single point.

Measurement of $\xi_n(r)$

Our measurement of the number density of galaxies correlation function ξ_n is compared to the measurement performed by the BOSS collaboration (pre-reconstruction) in figure 5.6, using DR9 and DR10 positions, and the published DR11 correlation function.The computation of the luminosity weighted correlation function ξ_L required more steps which are detailed in the next section.

5.3.3 Measurement of $\xi_L(r)$

Absolute magnitude and absolute luminosity

We calculate the two-point correlation function of the absolute luminosity density fluctuations, ξ_L, using the same estimator and algorithms for ξ_n, and weighting each object with

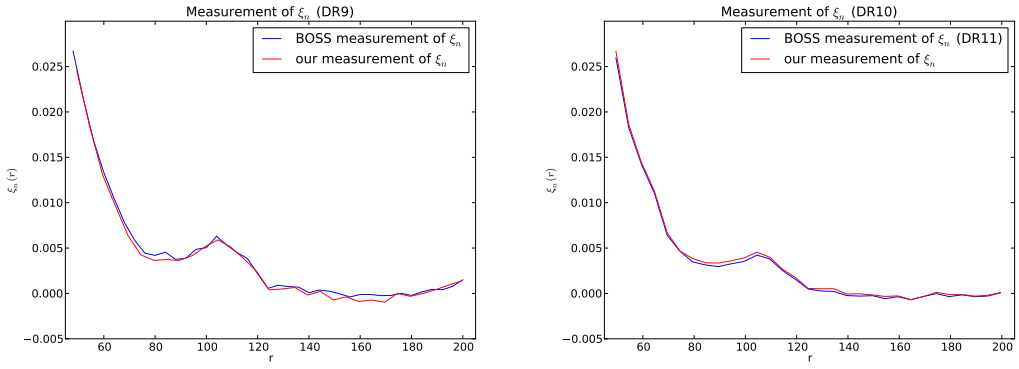

Figure 5.6: Our measurement of ξ_n compared to the BOSS collaboration measurement, when using the CMASS DR9 data release (left) and when using the latest release (right; our measurement, in red, was performed with the published DR10 positions, and the BOSS measurement, in blue, corresponds to the published DR11 correlation function, for which the positions are not yet public).

its absolute luminosity. The absolute luminosity is calculated using the i and g bands photometric data, from the CMASS DR10 catalogs. We first compute the absolute magnitudes, using a combination of the "cmodel" magnitude parameter, referred to as m_{cm}, and the extinction parameter, e:

$$M_{abs} = m_{cm} - e - (5log_{10}(D_L) + 25) \ , \tag{5.37}$$

where the luminosity distance D_L (in Mpc) is linked to the comoving distance D_M via $D_L = (1+z) \cdot D_M$. The m_{cm} magnitude is a parameter in the DR10 catalogs derived from the composite flux $F_{composite} = f \cdot F_{dev} + (1-f) \cdot F_{exp}$ which is the best fitting linear combination of the exponential fit and the de Vaucouleurs fit in each band.

The parameter e encapsulates the extinction correction, i.e. the account for the absorption and scattering of electromagnetic radiation by dust and gas between the observed galaxies and us. It has been computed following Schlegel et al. (1998). The distribution of the i and g absolute magnitudes is shown in figure 5.7.

The absolute luminosities are then computed using

$$L_{abs}/L_{sun} = 10^{-(M_{abs}-M_{sun})/2.5} \ , \tag{5.38}$$

where $M_{sun} = 4.83$ is the absolute magnitude of the sun. The distribution for the absolute luminosity in the i and g bands is shown in figure 5.7

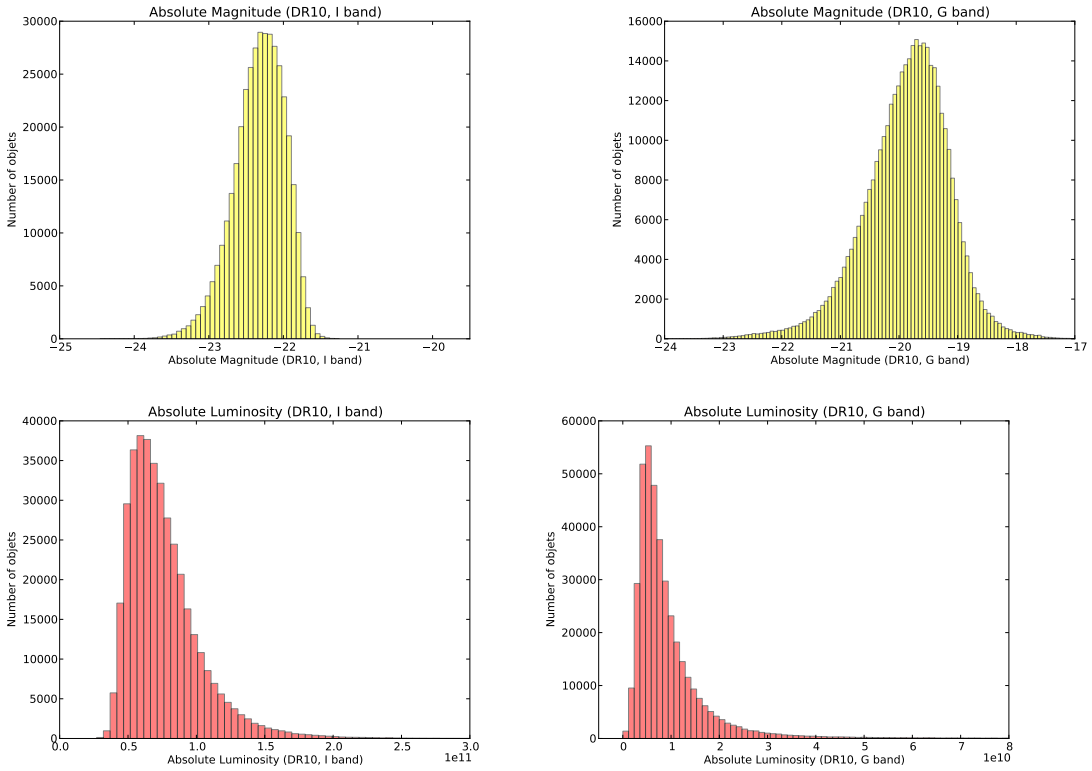

Figure 5.7: Distribution of the absolute magnitudes (top panels) and absolute luminosity (lower panels) in the North part of CMASS-DR10, using the i band (left panels) and the g band (right panels).

Measurement of C_{min} and D_{min} in the i band

Our model equations are simplified when we can do the approximation $C_{min} = D_{min} = 0$, where $C_{min} = \frac{L_{min}\phi(L_{min})}{F(L_{min})}$ and $D_{min} = \frac{L_{min}}{\langle L \rangle} C_{min}$ with $\langle L \rangle = \frac{1}{F(L)} \int_{L'=L}^{\infty} L'\phi(L')dL$ evaluated for $L = L_{min}$.

The calculated values for C_{min} and D_{min} (four the i band) are summarised in the table below:

	$\langle L \rangle$	C_{min}	D_{min}
DR9	$7.680 \cdot 10^{10}$	$8.689 \cdot 10^{-4}$	$3.072 \cdot 10^{-4}$
DR10	$7.660 \cdot 10^{10}$	$1.694 \cdot 10^{-3}$	$5.826 \cdot 10^{-4}$

In all the following analysis, we allow b_2 to vary, but the fact that C_{min} and D_{min} are close to zero orient our choice of priors.

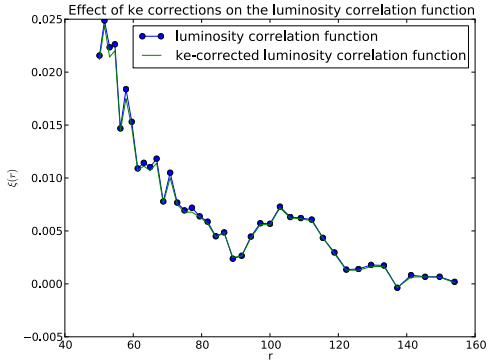

Figure 5.8: ξ_L, measured on the North part of the CMASS DR9 catalog, with and without ke-corrections.

k-e corrections

In addition to the galactic extinction, we explore the impact of two types of corrections due to the fact we observe samples at various redshifts:

- K-corrections, due to redshift effects: Light emitted between λ_e and $\Delta\lambda_e$ becomes light observed between $\lambda_e(1+z)$ and $\Delta\lambda_e(1+z)$, both the wavelength and the bandpass change.
- Evolutionary correction (e): changes in the galaxys luminosity and color between the time the light was emitted and today

The galactic k-corrections and e-corrections in magnitudes at the position of each object using are taken from Tojeiro et al. (2012). As shown in figure 5.8, the ke-corrections do not affect the measurement of the correlation function in a noticeable way, and for simplicity we ignore them in the rest of this analysis.

Covariance matrix for ξ_n and ξ_L

We compute the covariance matrix for ξ_n and for ξ_L, using 100 Jackknife samples.Our covariance and inverse covariance matrix are shown in figure 5.9, and compared to the covariance and inverse covariance published by the BOSS collaboration (in the case of ξ_n). From the comparison with the BOSS-DR10 covariance, it seems like we are slightly over-estimating our covariance, which is not surprising given that the JK is a rather crude method considering the accuracy of the measurement. We show, in section 5.6.1, that this

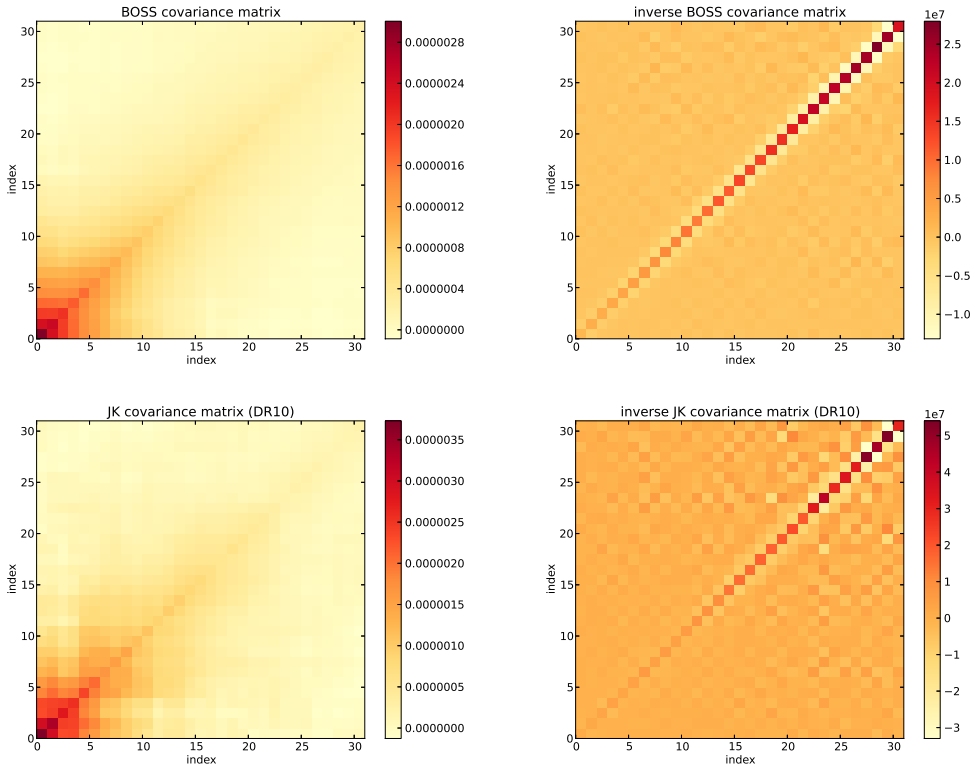

Figure 5.9: Covariance matrix (left) and inverse covariance matrix (right) for the ξ_n measurement, measured by the BOSS collaboration (upper panel; published DR11 covariance matrix), and by us (lower pannel)

our fit for ξ_n is in good agreement with a fit using the BOSS ξ_n and covariance matrix, which indicates that this overestimation does not affect the fit.

Joint covariance matrix for ξ_n and ξ_L

Since the uncertainties of the measurements of $\xi_n(r)$ and $\xi_L(r)$ at a given point are expected to be correlated, we need to compute the full covariance matrix for the joint measurement of $\xi_n(r)$ and $\xi_L(r)$. The full covariance matrix is shown in figure 5.10. It is far from being diagonal, or even block-diagonal, which shows the importance of fitting ξ_n and ξ_L jointly.

5.4 Model Fitting

Before we fit the model presented in section 5.1 to our measurement, we show the effect of each parameter of our model equations, in figures 5.11, 5.12, 5.13, 5.14, 5.15 and 5.16.

We present the model fitting basic formalism in section 5.4.1, and the algorithm we used to perform Monte Carlo Markov Chains (MCMC) in section 5.4.2. We then show

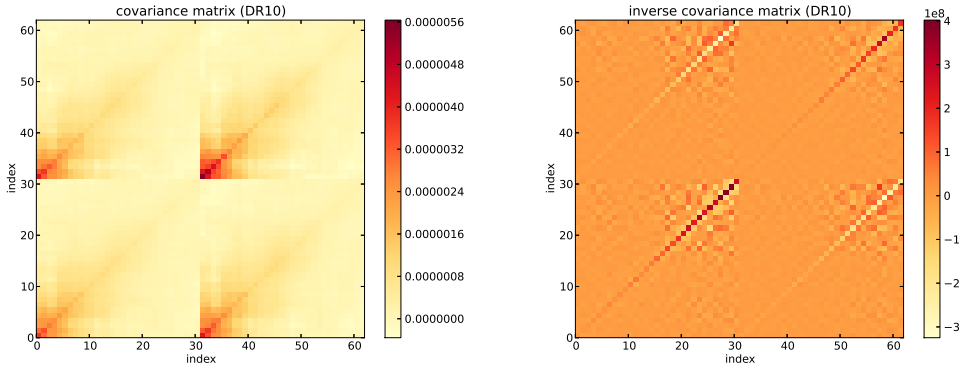

Figure 5.10: Our measurement of the joint covariance matrix (top panel) and inverse joint covariance matrix (lower panel), using the CMASS-DR10 sample. It is not diagonal: the uncertainties on ξ_n and ξ_L are correlated, which underlines the importance of performing a joint fit of ξ_n and ξ_l.

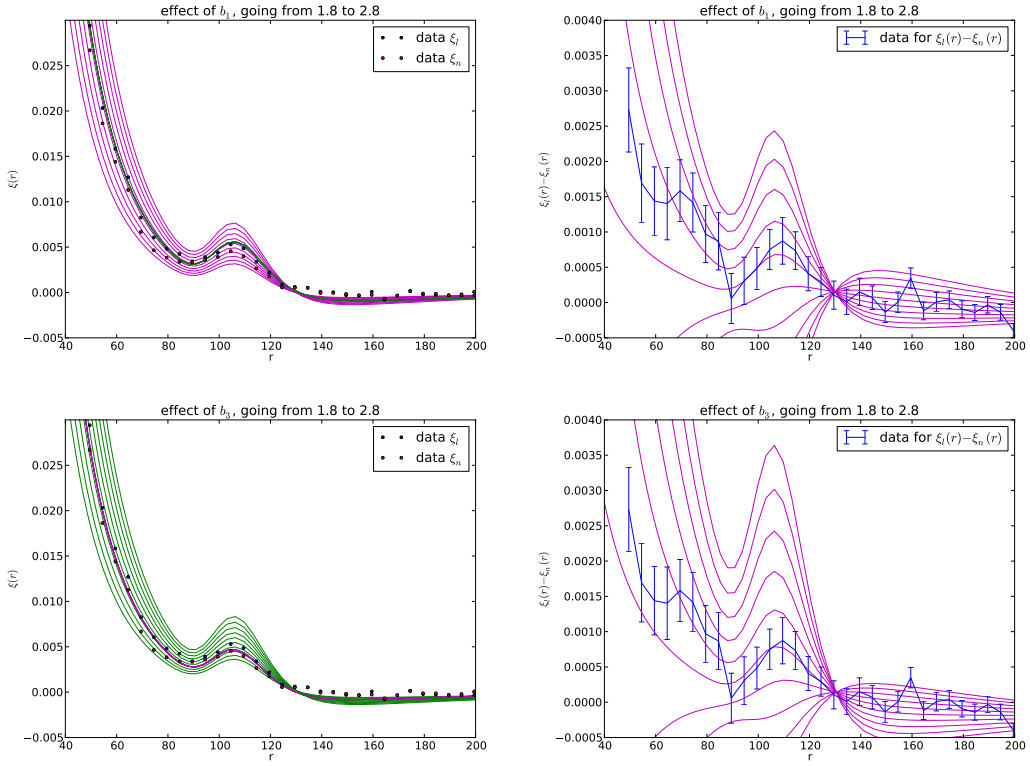

Figure 5.11: Effect of the b_1 (top panels) and b_3 (low panels) parameters on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). When not varying, the parameters are fixed at $b_1 = 2.193$, $b_2 = 0$, $b_3 = 2.269$, $b_4 = 4.300$, $b_{sys} = -0.002$, $k_* = 1.638$, $A_{MC} = 2.826$, $b_{cip} = 0.006$.When varying, b_1 (b_3 respectively) takes ten linearly spaced values from 1.8 to 2.8, the $\xi_n(r)$ functions ($\xi_l(r)$ respectively) with lower values (around $r = 100Mpc/h$) corresponding to the lower values of b_1 (b_3 respectively). As expected from equations 5.26 and 5.27, b_1 only affects the shape of ξ_n and b_3 the shape of ξ_l.

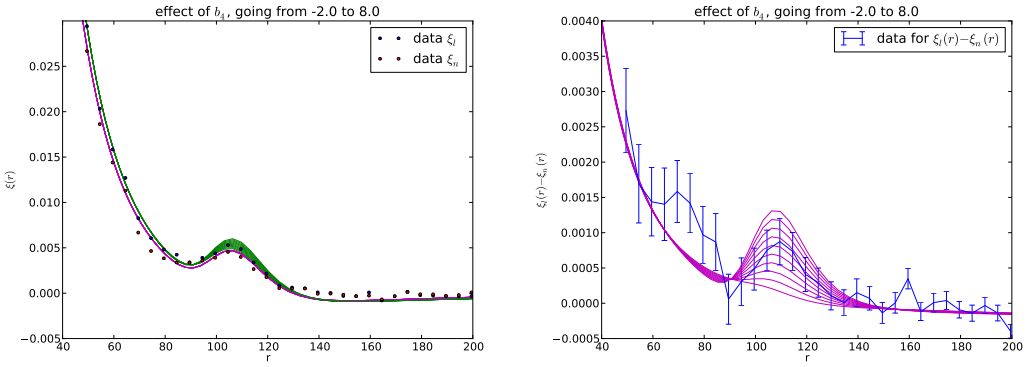

Figure 5.12: Effect of the b_4 parameter on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). When not varying, the parameters are fixed at $b_1 = 2.193$, $b_2 = 0.000$, $b_3 = 2.269$, $b_{sys} = -0.002$, $k_* = 1.638$, $A_{MC} = 2.826$, $b_{cip} = 0.006$. b_4 takes ten linearly spaced values from -2 to 8, the flatter $\xi_l(r)$ functions (i.e. with lower values around $r = 100Mpc/h$) corresponding to the lower values of b_4. As expected from equation 5.27, b_4 describes a discrepancy between ξ_n and ξ_l which is scale dependent, and appears at the BAO's scales.

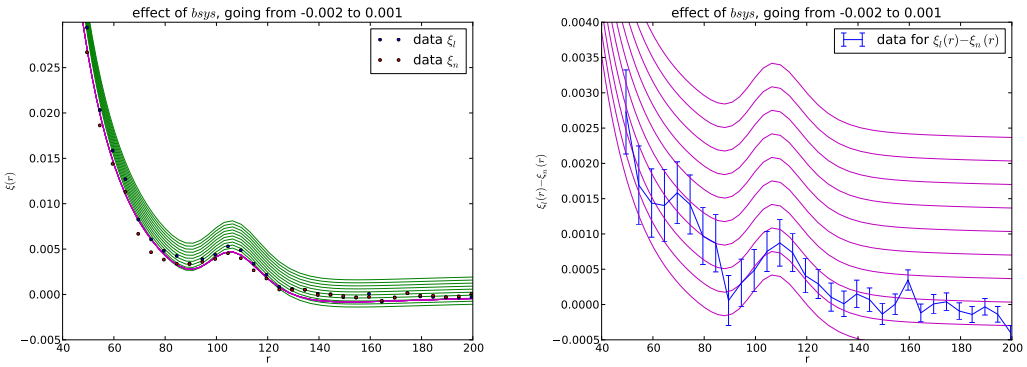

Figure 5.13: Effect of the b_{sys} parameter on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). The other parameters are fixed at $b_1 = 2.193$, $b_2 = 0.000$, $b_3 = 2.269$, $b_4 = 4.300$, $k_* = 1.638$, $A_{MC} = 2.826$, $b_{cip} = 0.006$. The parameter b_{sys} takes ten linearly spaced values from -0.002 to 0.001, the $\xi_l(r)$ functions with lower values (around $r = 100Mpc/H$) corresponding to the lower values of b_{bsys}. As expected from equation 5.27, b_{sys} only affects the shape the shape of ξ_l.

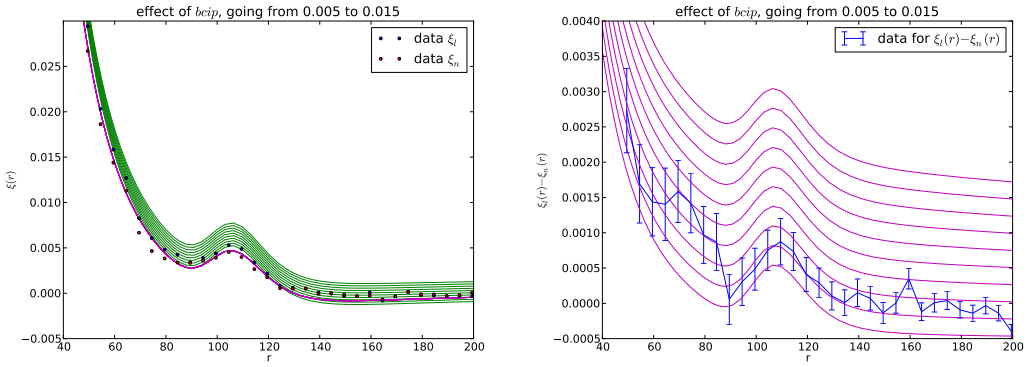

Figure 5.14: Effect of the b_{cip} parameter on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). The other parameters are fixed at $b_1 = 2.193$, $b_2 = 0.000$, $b_3 = 2.269$, $b_4 = 4.300$, $k_* = 1.638$, $A_{MC}, = 2.826$, $b_{sys} = -0.002$. The parameter b_{cip} takes ten linearly spaced values from 0.005 to 0.015, the $\xi_l(r)$ functions with lower values corresponding to the lower values of b_{cip}. As expected from equation 5.27, b_{cip} only affects the shape of ξ_l.

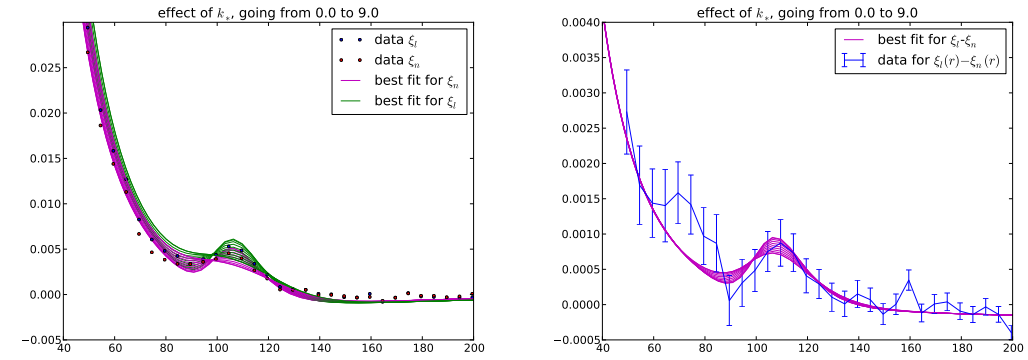

Figure 5.15: Effect of the k_* parameter on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). The other parameters are fixed at $b_1 = 2.193$, $b_2 = 0.000$, $b_3 = 2.269$, $b_4 = 4.300$, $A_{MC} = 2.826$, $b_{sys} = -0.002$, $b_{cip} = 0.006$. The parameter k_* takes ten linearly spaced values from 0 to 9. Lower values of k_* corresponds to more peaked functions $\xi_n(r)$ and $\xi_l(r)$, whereas higher values of k_* damp the peak between $r \approx 70$ and $r \approx 120$.

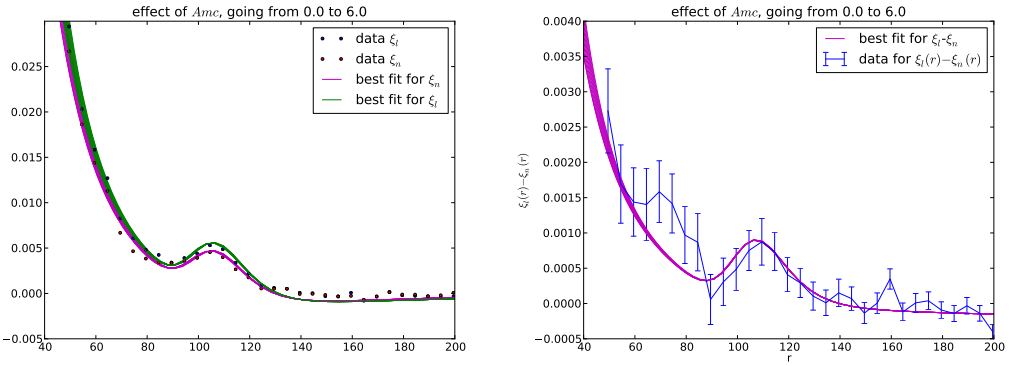

Figure 5.16: Effect of the A_{MC} parameter on the model. We show the model for ξ_n, ξ_l (left) and $\xi_l - \xi_n$ (right). The other parameters are fixed at $b_1 = 2.193$, $b_2 = 0.000$, $b_3 = 2.269$, $b_4 = 4.300$, $b_{sys} = -0.002$, $k_* = 1.638$, $b_{cip} = 0.006$. The parameter A_{MC} takes ten linearly spaced values from 0 to 6. The way A_{MC} affects the shape of $\xi_n(r)$ and $\xi_l(r)$ is complex. Here, higher values of A_{MC} correspond to the functions $\xi_l(r)$ (and $\xi_l(r) - \xi_n(r)$) with lower values at small r's.

the results of simulations testing the feasibility of our detection in section 5.5. We show our results, first when fitting only the DR10 data for ξ_n, and then when fitting jointly the DR10 data for ξ_n and ξ_L, in sections 5.6.1 and 5.6.2 respectively.

5.4.1 Formalism

We adopt the terminology of Hogg et al. (2010) for our likelihood functions calculations. In particular, we try to define our "generative model" and an "objective scalar" as carefully as possible. The steps of our method are as follows:

1. We define the data vectors and model vectors.
2. We define the "generative model" for our data.
3. We compute the "objective scalar" i.e. the likelihood function.

Vectors definition

Equations **??** and **??**, which we rewrite below,

$$\xi_n = b_1^2 \cdot \xi_{tot} \ ,$$

$$\xi_L = b_3^2 \cdot \xi_{tot} + 2b_3 \cdot b_4 \cdot \xi_{add} + b_{sys} + b_{CIP} \cdot \xi_{cip} \ ,$$

can be rewriten in a matrix form as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \tag{5.39}$$

where Y is a $2N \times 1$ vector (N=31 is the number of bins):

$$\mathbf{Y} = \begin{bmatrix} \xi_n(r_1) \\ . \\ . \\ \xi_n(r_N) \\ \xi_l(r_1) \\ . \\ . \\ \xi_l(r_N) \end{bmatrix} \tag{5.40}$$

A is a $2N \times 5$ matrix:

$$\mathbf{A} = \begin{bmatrix} \xi_{tot}(r_1) & 0 & 0 & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ \xi_{tot}(r_N) & 0 & 0 & 0 & 0 \\ 0 & \xi_{tot}(r_1) & \xi_{add}(r_1) & 1 & \xi_{cip}(r_1) \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & \xi_{tot}(r_N) & \xi_{add}(r_N) & 1 & \xi_{cip}(r_N) \end{bmatrix} \tag{5.41}$$

And X is the 5×1 vector we need to determine:

$$\mathbf{X} = \begin{bmatrix} b_1^2 \\ b_3^2 \\ 2 \cdot b_3 \cdot b_4 \\ b_{sys} \\ b_{CIP} \end{bmatrix} \tag{5.42}$$

Generative model

In the terminology of Hogg et al. (2010), a generative model is a parametrized quantitative description of a statistical procedure that could reasonably have generated the data. In our case, we assume the data really do come from equations **??** and **??** and that the only reason that our data point deviate from our model is an offset in the ξ direction, drawn from a gaussian distribution of zero mean and known variances σ_ξ.

Objective scalar

In a model fitting problem, the objective scalar we wish to optimize is the posterior probability for our model: we wish to get the optimal set of parameters $\theta = \{b_1, b_3, b_4, b_{sys}, b_{CIP}, A_{MC}, k_*\}$ which maximize the probability of our model $\mathcal{M}$ given the data $\mathcal{D}$, i.e. the posterior probability $Pr(\theta|\{\mathcal{D}, \mathcal{M}\})$. Bayes' theorem relates the posterior probability distribution to the likelihood $\mathcal{L} \equiv Pr(\mathcal{D}|\theta, \mathcal{M})$, via the prior $\pi \equiv Pr(\mathcal{D}|\mathcal{M}, \theta)$:

$$Pr(\theta|\{\mathcal{D}, \mathcal{M}\}) = \frac{Pr(\mathcal{D}|\{\theta, \mathcal{M}\}) \cdot Pr(\theta|\mathcal{M})}{Pr(\mathcal{D}|\mathcal{M})} = \frac{\mathcal{L} \cdot \pi}{E} \ , \tag{5.43}$$

where the evidence $E = Pr(\mathcal{D}|\mathcal{M})$ is the probability of getting the data $\mathcal{D}$, given the model $\mathcal{M}$. It can be seen as the likelihood averaged over all the possible parameters within a model. In the model fitting problems, one tries to get the optimal set of parameters θ *within the framework of one specific model*. In this case, E is a constant and is ignored, since it does not change the conclusion of the optimization. We will see in section 5.7 that this is no longer true when adopting a model selection approach to our problem, but for now, we aim at optimizing the objective scalar $Pr(\theta|\{\mathcal{D}, \mathcal{M}\}) \propto Pr(\mathcal{D}|\{\theta, \mathcal{M}\}) \cdot Pr(\theta|\mathcal{M})$.

The likelihood of our generative model is :

$$\mathcal{L} \propto exp\left[-\frac{1}{2} R^T \cdot C^{-1} \cdot R\right] \tag{5.44}$$

where $\mathbf{R} = \mathbf{Y} - \mathbf{AX}$, and C^{-1} is the inverse covariance matrix of the data $\mathbf{Y}$, shown in figure 5.10. We apply the following uniform (not "informative"), priors for the seven parameters of our model:

- $b_1 \in [0, 5]$
- $b_2 \in$ [0,10]

- $b_3 \in [0, 5]$
- $b_4 \in [-10, 10]$
- $b_{sys} \in [-0.01, 0.01]$
- $k_* \in [0, 10]$
- $A_{MC} \in [0, 6]$
- $b_{CIP} \in [-1, 1]$

We believe this is a conservative choice of priors. The prior on b_4 is willingly taken to be broad, although Barkana & Loeb (2011) forecasted it to be around $b_4 \approx 2.6$ (in a case where $C_{min} = D_{min} = 0$, $b_n = 2$ and $b_{L;t} = 1$). The prior on b_{CIP} is taken to be broader than the upper limit of 10^{-2} set in Grin et al. (2014) for $A_{CIP} = b_{CIP}/b_4^2$. The priors on b_1, k_* and A_{MC} are taken to be consistent with previous works on the BOSS data (Crocce & Scoccimarro 2008; Anderson et al. 2012, 2013). The prior on b_2 is also chosen to be broad, although preliminary estimations of this parameter that we made, indicate it should be close to zero in the case of the CMASS DR10 sample.

5.4.2 Computation

We first numerically optimize the likelihood function, which corresponds to the maximum a posteriori value (m.a.p.) in the case of an uninformative prior. The problem then becomes to estimate the uncertainties on the m.a.p. values of each parameter. In fact, rather than the m.a.p. values, we want an estimate of the posterior probability function, i.e. the distribution of parameters that is consistent with our data, and to be able to marginalise over it to get the distribution of each parameter. This is made possible by Monte Carlo Markov Chain (MCMC) sampling.

We used the MCMC emcee algorithm (Foreman-Mackey et al. 2013) to sample from the posterior probability distribution, and quote the uncertainties based on the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, corresponding to $1 - \sigma$ in the case of a gaussian.

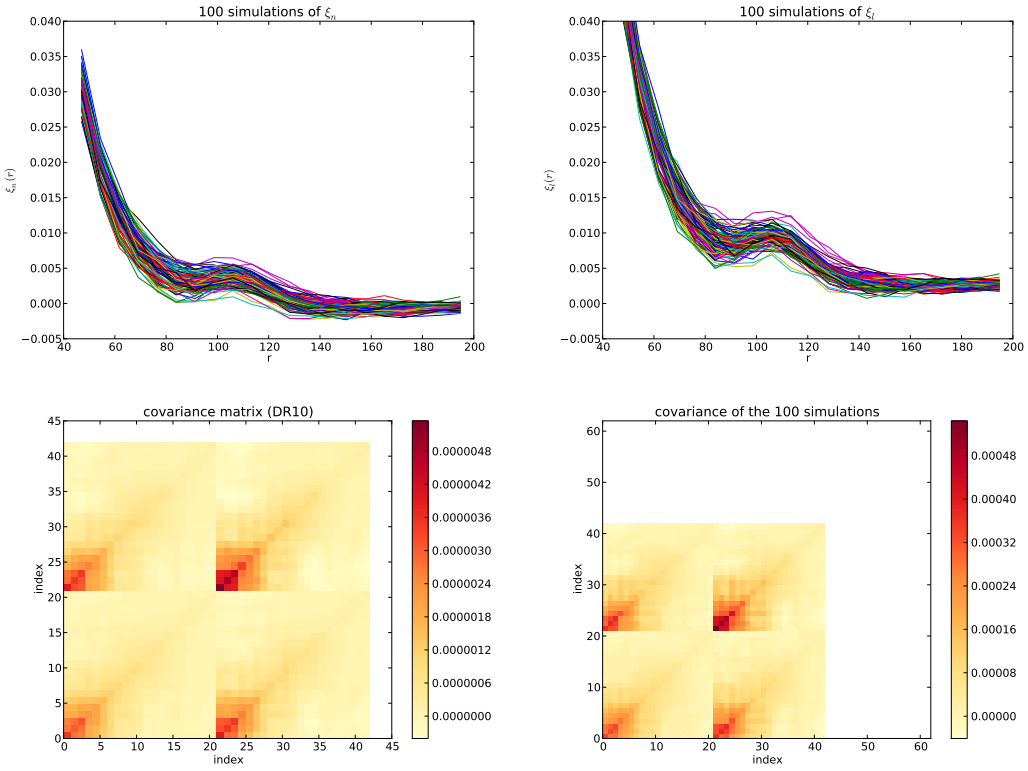

Figure 5.17: The 100 simulated signals $\xi_n(r)$ (top left) and $\xi_l(r)$ (top right) are drawn from the data covariance matrix.

5.5 Simulations

In order to control the feasibility of our detection, we generate a signal ξ_{fid} with the following fiducial values of the parameters: $b_1 = 2.1$, $b_3 = 2.5$, $b_4 = 5$, $b_{sys} = 10^{-4}$, $A_{MC} = 1$ and $k_* = 5$. We then draw noise from our computed covariance matrix (shown in figure 5.10) and add it to the fiducial signal, in order to obtain a noisy signal. A widely used method for drawing a random vector $\vec{x}$ from the N-dimensional multivariate normal distribution with known mean vector μ and covariance matrix is using the Cholesky decomposition of the given covariance matrix. Given the Cholesky decomposition $AA^T = C$ of the given covariance matrix C, and $\vec{z} = (z_1, z_2, ..., z_N)^T$ a vector which components are N independent standard normal variates, then the vector $\vec{x} = \vec{\mu} + A\vec{z}$ has the desired distribution. In figure 5.17, we show 100 simulated signals generated in this way. We also control that the covariance matrix from which we draw the signal, and the covariance matrix of these 100 signals are consistent with each other (see figure 5.17).

The maximum of the likelihood function corresponds to a recovered signal shown in

	Fiducial parameters	Maximum likelihood	Median of the marginalised distribution	1σ Confidence limits
b_1	2.1	2.111	2.143	$^{+0.064}_{-0.057}$
b_3	2.5	2.488	2.525	$^{+0.073}_{-0.067}$
b_4	5	6.874	6.164	$^{+2.236}_{-2.573}$
$b_{sys} \cdot 10^3$	0.1	-0.317	-0.362	$^{+2.045}_{-2.024}$
k_*	5.0	4.608	4.673	$^{+0.776}_{-0.827}$
A_{MC}	1.0	0.911	2.111	$^{+2.045}_{-1.443}$
$b_{CIP} \cdot 10^2$	1.5	1.689	1.713	$^{+0.876}_{-0.883}$
χ^2/dof	-	1.455	-	-

Table 5.1: **Simulations:** Fiducial values of each parameter (left), and median value of the marginalised posterior probability distributions computed with the MCMC. The error bars correspond to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the median value and the 1σ values (in the case of a gaussian).

figure 5.18, with the corresponding $\chi^2(r)$ and residuals $|model - data|(r)$. the value corresponding to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the median value and the 1σ values (in the case of a gaussian) are confronted to maximum likelihood values and the fiducial values in the table 5.1. In figure 5.19, we show the marginalised posterior probability distribution for each parameter, confronted to the fiducial value. In figure 5.20 we add the two dimensional projections of the posterior probability distributions of our parameters, to show the correlations between the parameters.

The fiducial values are recovered within 2σ. In the work presented here, the simulations are mostly designed to assess that our model fitting procedure is working, but in further analysis we aim to use simulations in order to check the feasibility of our detection. In particular, it could be used to evaluate:

1. for a given amount of noise, the minimum value of b_4 that would allow a 5σ detection of $b_4 \neq 0$.

2. for a given value of b_4, the maximum amount of noise that would allow a 5σ detection of $b_4 \neq 0$.

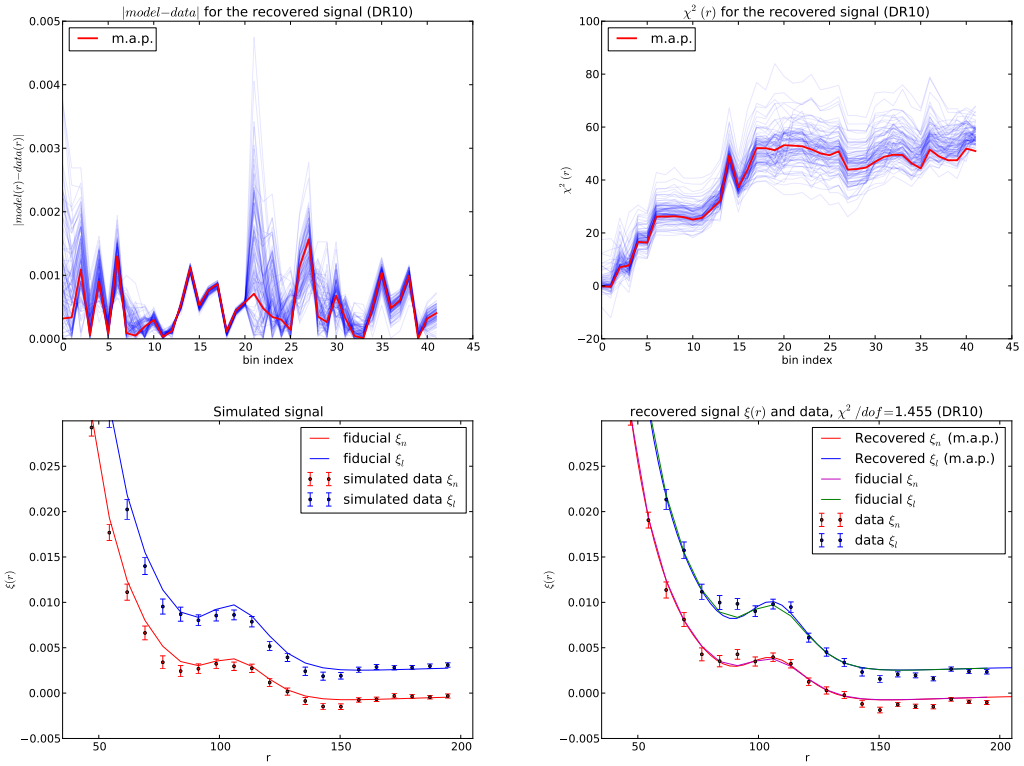

Figure 5.18: **Simulation:** The lower panels show the fiducial signal (right) and the recovered signal (left), corresponding to the maximum likelihood values, i.e. the maximum a posteriori (m.a.p) values (since the priors are uniform). The upper panels show the residuals $|model - data|(r)$ (left), and the value of $\chi^2(r)$ (right), for the maximum likelihood (in red) and for 100 samples of the MCMC chain (in blue). Note that the maximum likelihood corresponds to the minimum$\chi^2(r)$ value, as expected.

5.6 Fits on real data

5.6.1 Fitting $\xi_n(r)$ only

In order to assess the quality of our model and check the consistency of our model-fitting procedure, we fit our model for $\xi_n(r)$ (equation 5.26) to both our measured ξ_n and the ξ_n measured and published by the BOSS collaboration.

The maximum-likelihood fits are shown in figure 5.21. In figure 5.22, we show the marginalised posterior probability distribution for each parameter, and the two dimensional projections of the posterior probability distributions. The value corresponding to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the median value and the 1σ values (in the case of a gaussian) values are shown in table 5.2.

When fitting our model to our measurement of ξ_n and to the BOSS measurement,

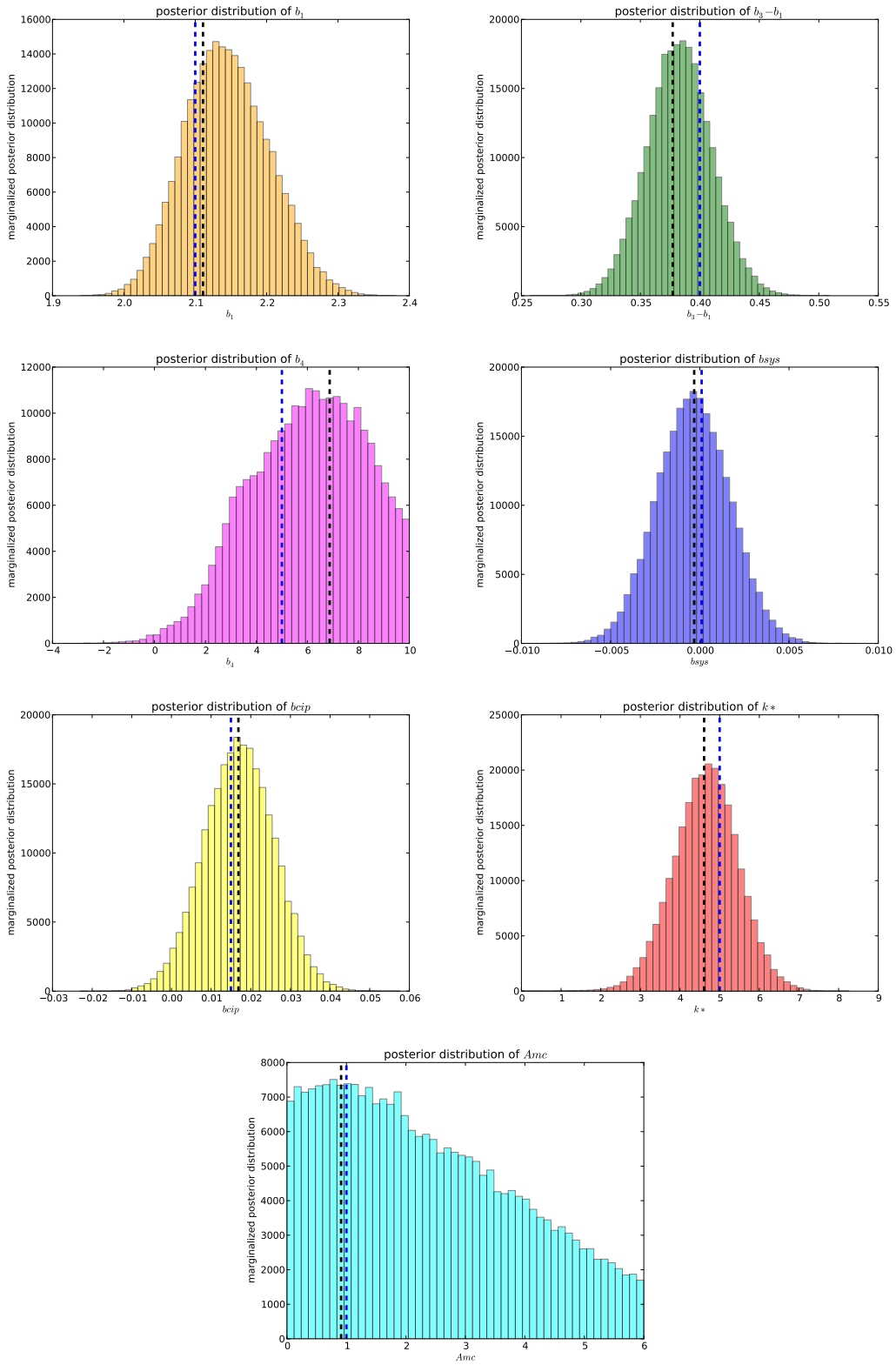

Figure 5.19: **Simulation:** Marginalised distribution for each parameter. The black dashed line shows the m.a.p, value of the distribution, whereas the blue one shows the fiducial value.

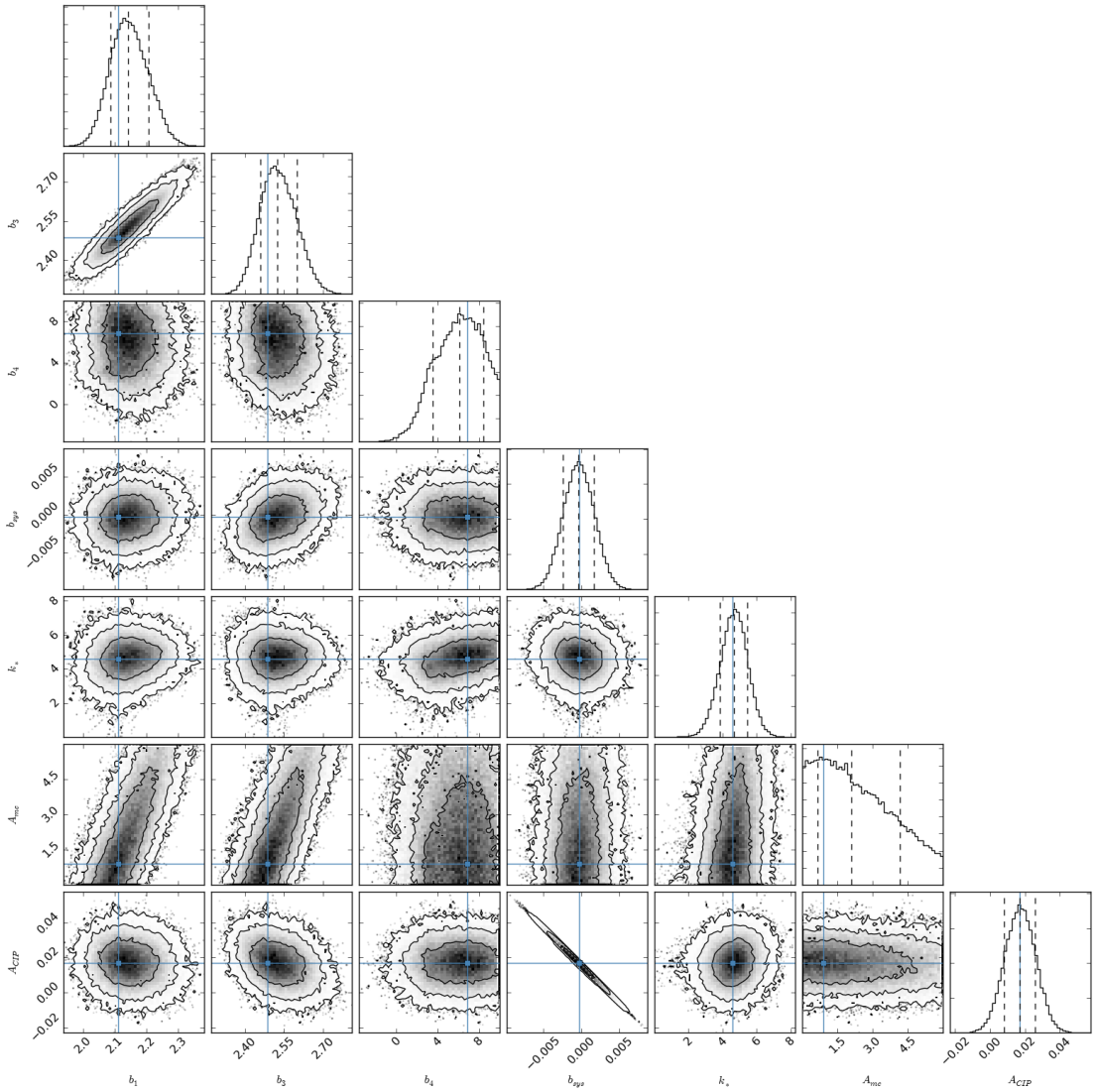

Figure 5.20: **Simulation:** All the one and two dimensional projections of the posterior probability distributions of our parameters, $\{b_1, b_3, b_4, b_{sys}, k_*, A_{MC}, b_{CIP}\}$. This quickly demonstrates all of the covariances between parameters. The fiducial values of the parameter are shown in figure 5.19.

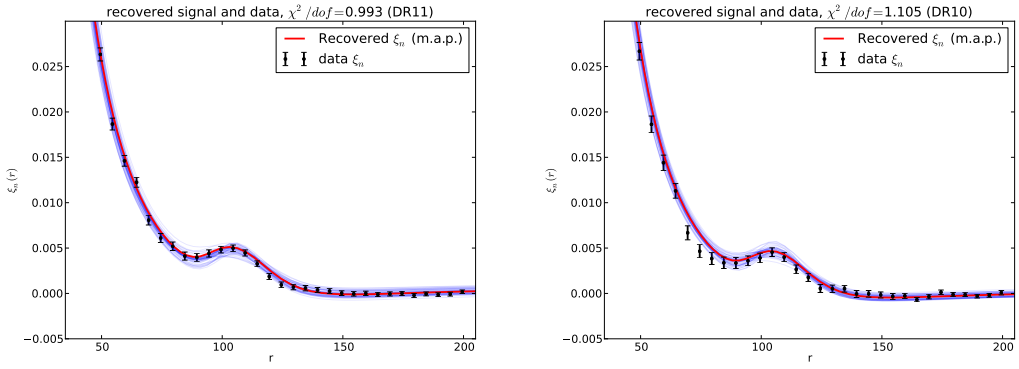

Figure 5.21: **:Fitting ξ_n.** The left panels have been obtained with the published correlation function and covariance matrix measured by the BOSS collaboration (using the un-published DR11 data), whereas the right panels have been obtained with our measurement of ξ_n using the CMASS-DR10 sample. The figures show, in red, the maximum likelihood fit, i.e. the maximum a posteriori (m.a.p) values (since the priors are uniform) and 100 samples of the MCMC chain, in blue.

the values of the parameters are generally consistent, apart from A_{MC} and b_{CIP}, which are prior-driven (see figure 5.22). Apart for these two prior-driven parameters, the value of the parameters are also consistent, within 1σ, between DR10 and DR11. A deeper understanding of the difference between them will be possible once the DR11 data is publicly released. Moreover, the values of χ^2/dof (dof is for "degrees of freedom", and is obtained by subtracting the number of parameters to the total amount of data points, i.e. the size of the joint covariance matrix) that we obtain when fitting our model to the BOSS measurement converge with the range of values published by the BOSS collaboration (Anderson et al. 2013), where the same data and a different model were used for ξ_n. This shows that our model reasonably describes the correlation function ξ_n measured by BOSS. The two conclusions of fitting only ξ_n, are that our measurement is consistent with the BOSS measurement and that our model for ξ_n provides a good description of both these measurements.

5.6.2 Joint fit of $\xi_n(r)$ and $\xi_l(r)$

We now present the results for the joint fit of ξ_n and ξ_l, using the CMASS-DR10 sample. The maximum likelihood fit is shown in figure 5.23.

In figures 5.24, we show the marginalised posterior probability distribution for each parameter and in figure 5.26 the two dimensional projections of the posterior probability distributions. Both show results of our analysis in the i band, whereas the same figures

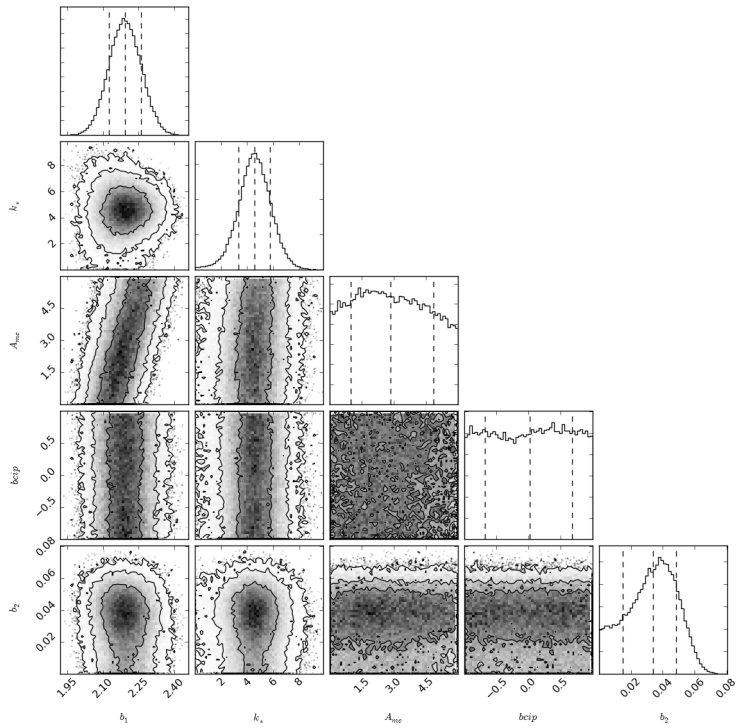

Figure 5.22: **Fitting ξ_n.** All the one and two dimensional projections of the posterior probability distributions of our parameters, $\{b_1, b_3, b_4, b_{sys}, k_*, A_{MC}, b_{CIP}, b_2\}$. This quickly demonstrates all of the covariances between parameters. The contours correspond to the 1σ, 2σ and 3σ percentiles. The dashed lines show the 1σ percentile of the marginalized distributions.

for the g band analysis are shown in figure 5.25 and 5.27. The value corresponding to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the median value and the 1σ values (in the case of a gaussian) values are shown in table 5.3 for both the i and g band.

The median and m.a.p values of the marginalised distribution of b_4 are consistent with Barkana & Loeb (2011), in which the authors forecasted $b_4 \approx 2.5$ with similar b_1, and $b_2 \approx 0$. This being said, a zero value for b_4, which would mean a non-detection of the effect we are searching for, is within the 2σ limit (for both the i and g band).

In the case of the i band, our model-fitting procedure seems to fail to fit the shape of our data for $\xi_l - \xi_n$ (see figure 5.23) around the BAO peak, i.e. between $r = 90Mpc$ and $r = 140Mpc$, which is the area described by b_4 (see figure 5.12). One source of a systematic difference between the model and data in this area is the presence of non-linear effects

	DR10	BOSS (DR10)	BOSS (DR11)
b_1	$2.202^{+0.069}_{-0.067}$	$2.082^{+0.076}_{-0.074}$	$2.212^{+0.068}_{-0.069}$
k_*	$4.608^{+1.211}_{-1.224}$	$4.633^{+1.331}_{-1.311}$	$4.587^{+1.011}_{-1.008}$
A_{MC}	$3.052^{+2.001}_{-1.851}$	$0.000^{+2.094}_{-1.661}$	$6.199^{+1.630}_{-2.130}$
$b_{cip} \cdot 10^{-2}$	$1.850^{+66.26}_{-70.05}$	$2.980^{+66.19}_{-69.52}$	$-0.317^{+67.10}_{-66.87}$
$b_2 \cdot 10^{-2}$	$3.918^{+1.441}_{-1.905}$	$2.891^{+1.845}_{-1.767}$	$5.540^{+1.173}_{-1.607}$
χ^2/dof	1.105	1.119	0.993

Table 5.2: **Fitting ξ_n:** Maximum likelihood values, when fitting our model for ξ_n (equation 5.26) to (from left to right): our measurement of ξ_n with the CMASS-DR10 sample; the published BOSS collaboration measurement of ξ_n with the CMASS-DR10 sample; the published BOSS collaboration measurement of ξ_n with the CMASS-DR11 sample. The error bars correspond to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the 1σ values (in the case of a gaussian). The values of the parameters obtained with our measurements are generally consistent with the values obtained by fitting our model to the BOSS measurements, apart from A_{MC} and b_{cip}, which is prior driven (see figure 5.22), and for which we use the same prior as in Sánchez et al. (2012). The values of χ^2/dof are consistent.

	i band			g band		
	median	max. lik.	conf.lim.	median	max. lik.	conf.lim.
b_1	2.203	2.274	$^{+0.070}_{-0.075}$	2.128	2.140	$^{+0.067}_{-0.064}$
$b_2 \cdot 10^2$	3.408	3.956	$^{+1.290}_{-1.813}$	4.392	4.762	$^{+1.110}_{-1.635}$
b_3	2.241	2.303	$^{+0.081}_{-0.083}$	2.377	2.388	$^{+0.103}_{-0.102}$
b_4	2.575	2.426	$^{+2.545}_{-2.100}$	3.100	4.688	$^{+4.103}_{-4.484}$
$b_{sys} \cdot 10^3$	-8.156	-9.021	$^{+1.839}_{-1.248}$	-3.897	-3.823	$^{+4.252}_{-3.649}$
k_*	4.553	5.019	$^{+0.990}_{-0.970}$	2.499	2.552	$^{+1.260}_{-1.453}$
A_{MC}	3.893	6.000	$^{+1.548}_{-2.553}$	2.737	3.168	$^{+2.115}_{-1.820}$
$b_{CIP} \cdot 10^2$	3.732	4.147	$^{+0.538}_{-0.795}$	2.894	2.869	$^{+1.632}_{-1.893}$
χ^2/dof	1.443			0.990		

Table 5.3: **Fitting ξ_n and ξ_l:** Median value of the marginalised posterior probability distributions computed with the MCMC (left column), maximum likelihood value (middle column), and confidence limit (right column) for each parameter. The error bars correspond to the 16^{th}, 50^{th}, and 84^{th} percentiles of the samples in the marginalised distributions, i.e. the median value and the 1σ values (in the case of a gaussian). The parameters are obtained using our measurement on the CMASS-DR10 sample.

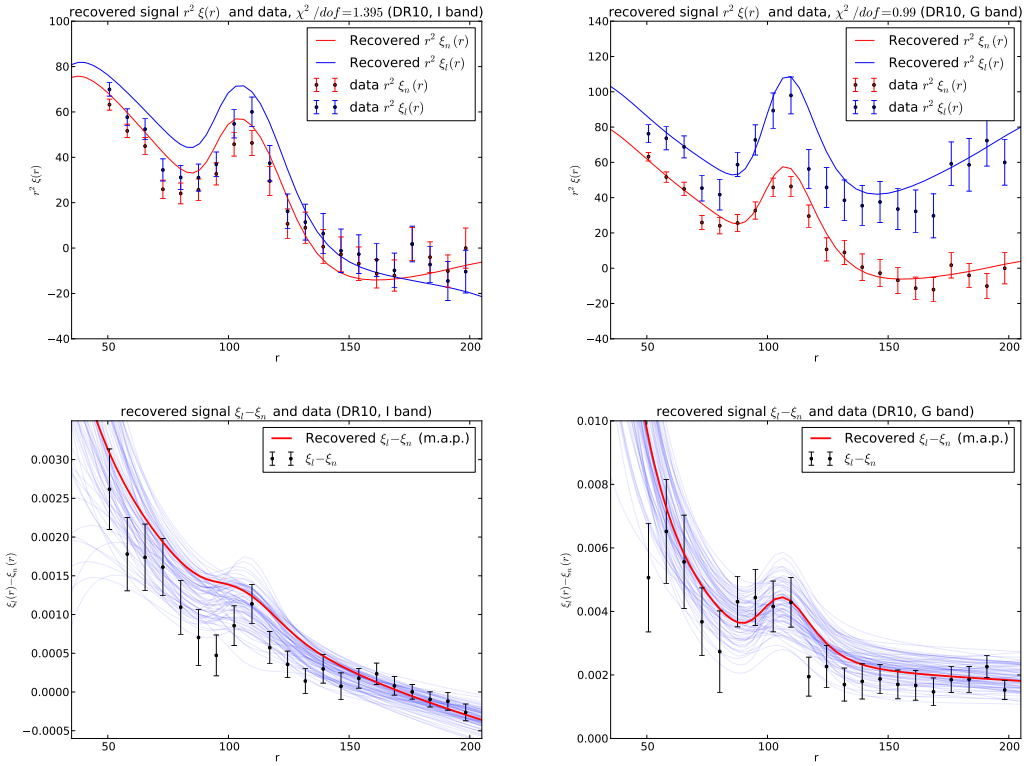

Figure 5.23: **Joint fit of ξ_n and ξ_l.** The left panels are for the i band and the right ones are for the g band. The top panels show the data and the maximum likelihood fit for $r^2 \cdot \xi_n$ and $r^2 \cdot \xi_l$, also corresponding to the maximum a posteriori (m.a.p) fit (since the priors are uniform). The lower panels show the data and the best fit for $\xi_l - \xi_n$, together with 100 samples of the MCMC chain (in blue).

which, in previous works on BOSS data, are mostly corrected by a procedure known as *reconstruction*. Presented in Eisenstein et al. (2007a), reconstruction has been successfully applied to the BOSS data in Anderson et al. (2013) and Tojeiro et al. (2014) to deal with the effect of non-linearities on the BAO peak. The idea is to partially reverse the effects of non-linear growth of structure and large-scale peculiar velocities from the data. Rather than modifying the model to account for the non-linear effects, reconstruction acts on the data itself. It reduces the anisotropy in the clustering, reverses the smoothing of the BAO feature due to second-order effects, and significantly reduces the expected bias in the BAO distance scale that arises from these same second-order effects. In future work, reconstruction may allow a better account for the non-linear effects and improve the goodness of the fit and increase the evidence for a non-zero b_4.

Since $A_{CIP} = b_{CIP}/b_4^2$, the case of $b_4 = 0$ leads to an unconstrained A_{CIP}, and needs to be treated separately. We leave this for further analysis.

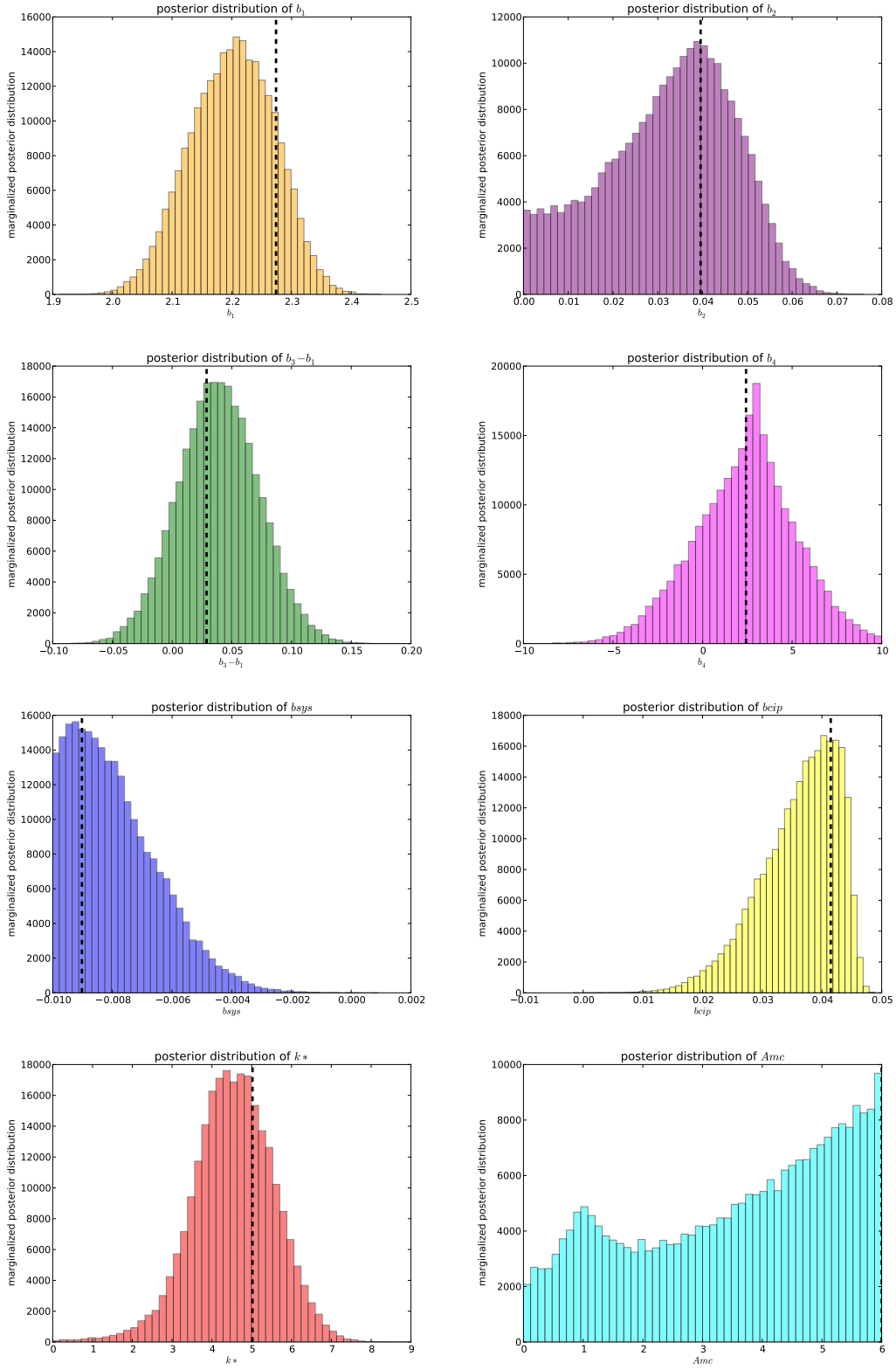

Figure 5.24: **Marginalised distribution for each parameter, i band.** The black dashed line shows the maximum likelihood value of each parameter.

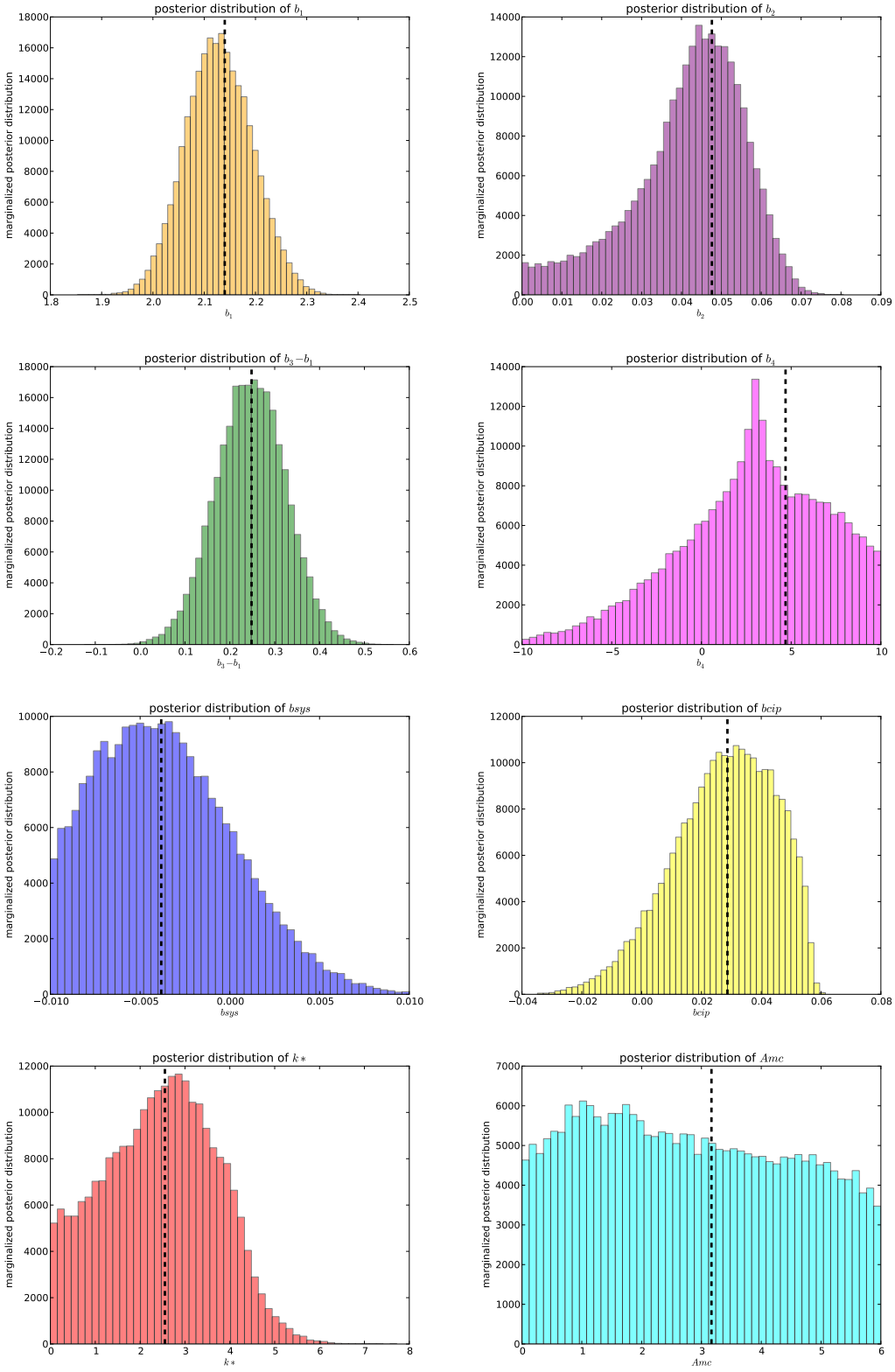

Figure 5.25: **Marginalised distribution for each parameter, g band.** The black dashed line shows the maximum likelihood value of each parameter.

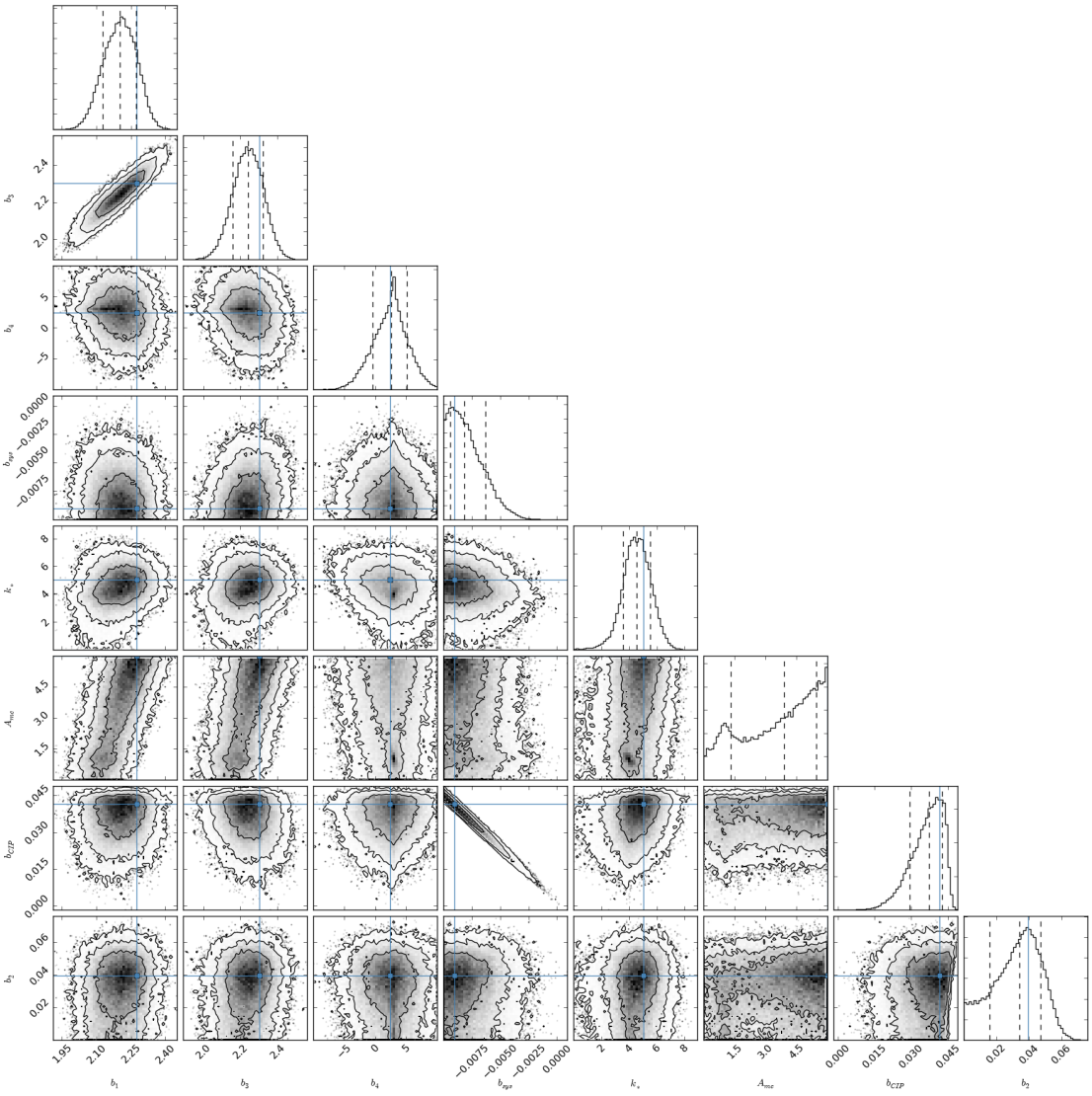

Figure 5.26: **Joint fit of ξ_n and ξ_l, i band:** all the one and two dimensional projections of the posterior probability distributions of our parameters, $\{b_1, b_3, b_4, b_{sys}, k_*, A_{MC}, b_{CIP}, b_2\}$. This quickly demonstrates all of the covariances between parameters. The contours correspond to the 1σ, 2σ and 3σ percentiles. The blue line corresponds to the maximum likelihood value of each parameter, which is also the maximum a posteriori value (m.a.p.). The dashed lines show the 1σ percentile of the marginalized distributions.

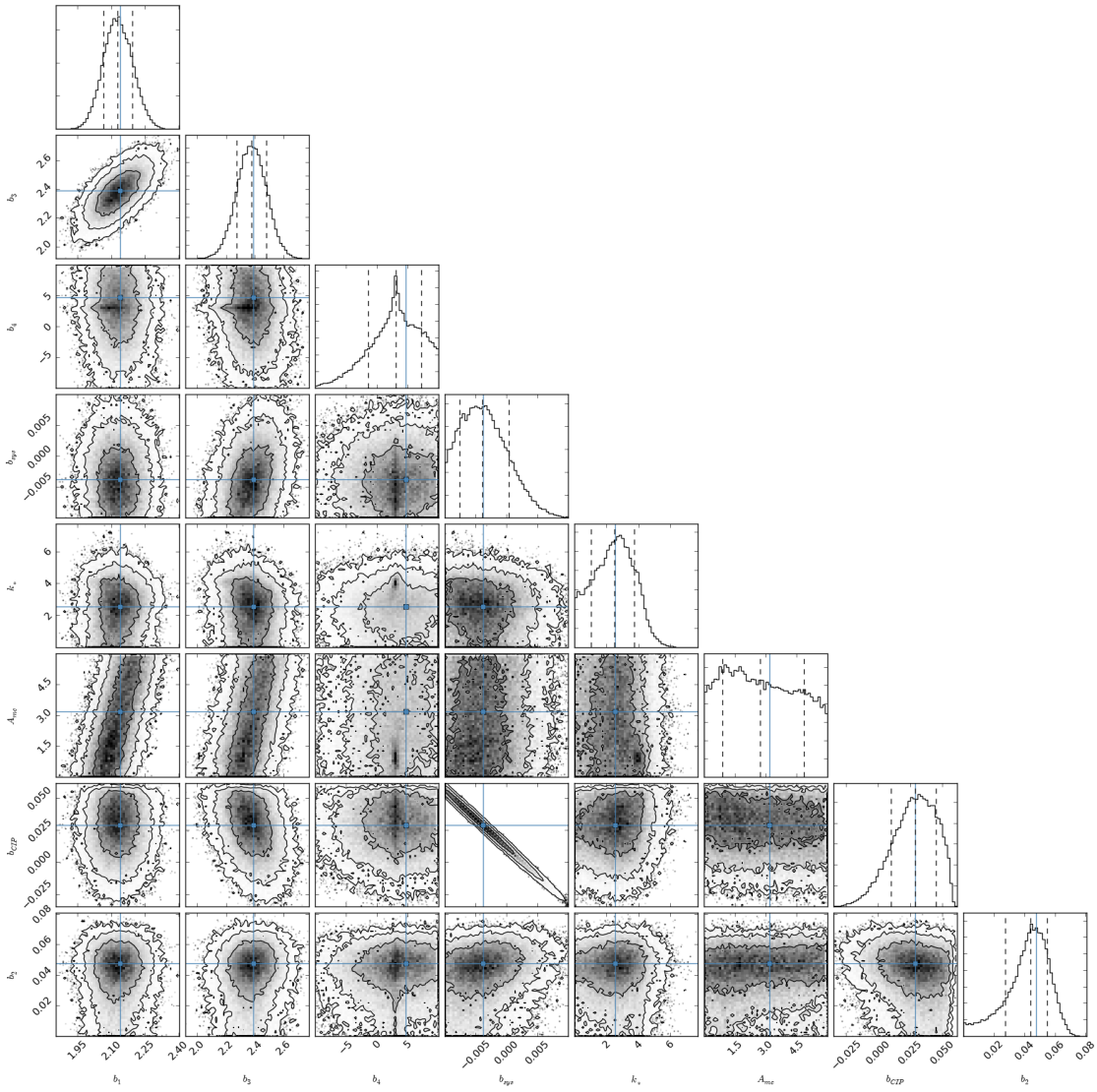

Figure 5.27: **Joint fit of ξ_n and ξ_l, g band:** all the one and two dimensional projections of the posterior probability distributions of our parameters, $\{b_1, b_3, b_4, b_{sys}, k_*, A_{MC}, b_{CIP}, b_2\}$. This quickly demonstrates all of the covariances between parameters. The contours correspond to the 1σ, 2σ and 3σ percentiles. The blue line corresponds to the maximum likelihood value of each parameter, which is also the maximum a posteriori value (m.a.p.). The dashed lines show the 1σ percentile of the marginalized distributions.

5.6.3 Effect of the errors on the photometry

In order to test the effect of the photometric errors on the final result, we repeat the above analysis on a catalog with synthetic magnitudes. The synthetic magnitudes are generated using the error bars on the m_{cm} parameter of each object and assuming it is gaussian. A widely used method for drawing a random vector $\vec{x}$ from the N-dimensional multivariate normal distribution with known mean vector μ and covariance matrix is using the Cholesky decomposition of the given covariance matrix. Given the Cholesky decomposition $AA^T = C$ of the given covariance matrix C, and $\vec{z} = (z_1, z_2, ..., z_N)^T$ a vector which components are N independent standard normal variates, then the vector $\vec{x} = \vec{\mu} + A\vec{z}$ has the desired distribution.

We also explore the impact of k-corrections and evolutionary corrections (e-corrections), both due to the fact that we observe samples at various redshifts. K-corrections account for the change in the wavelength and bandpass: light emitted between λ_e and $\Delta\lambda_e$ becomes light observed between $\lambda_e(1+z)$ and $\Delta\lambda_e(1+z)$. Evolutionary corrections account for the change in the luminosity and color of each galaxy between the time at which the light was emitted, and the present time. We use k-corrections and e-corrections in magnitudes at the position of each object from Tojeiro et al. (2012).

In both cases, we find that the best fit value of the parameter of interest, b_4, is shifted by less than 1σ. The overall result is unchanged, i.e. a zero value for b_4, which would mean a non-detection of the effect we are searching for, is still within the 2σ limit. If future data sets allow to reduce the error bars on b_4, care should be taken to account for the shift of b_4 caused by both the error on the photometry and k-e corrections.

5.7 Model selection

5.7.1 Formalism

Nested Models

Answering whether we detect a scale dependent bias of the luminosity correlation function comes to answering the following question: does the data support the inclusion of a non-zero extra parameter b_4? Rather than a question of parameter estimation (i.e. the determination of the most probable values for the extra parameters within the context of a single model) this is a question of model comparison. The models we wish to confront

are as follows:

- In the first model, $\mathcal{M}_1$, there is no scale dependent bias, and therefore $b_4 = 0$.
- In the second model, $\mathcal{M}_2$, there is a scale dependent bias, therefore $b_4 \neq 0$.

The parameter space $\Theta = \{b_1, b_3, b_4, b_{sys}, A_{MC}, k^*, b_{cip}\} = \{\phi, \psi\}$ is partitioned into the common parameters, $\phi = \{b_1, b_3, b_{sys}, A_{MC}, k^*, b_{CIP}\}$, and the extra parameters, $\psi = \{b_4\}$, describing the scale dependent bias. The two models are nested, as defined in Verde et al. (2013).

Bayes Factor

Within a Bayesian framework, the key model comparison quantity is the Bayes factor, which is the ratio of the Evidence values for two different models and is defined below.

The probability distribution (Pr) for a set of parameters θ, given a model $\mathcal{M}$, and data $\mathcal{D}$, is the posterior, $P = Pr(\theta|(\mathcal{D}, \mathcal{M}))$. Bayes' theorem relates the posterior to the likelihood $\mathcal{L} \equiv Pr(\mathcal{D}|\theta, \mathcal{M})$, via the prior $\pi \equiv Pr(\mathcal{D}|\mathcal{M}, \theta)$:

$$Pr(\theta|\{\mathcal{D}, \mathcal{M}\}) = \frac{Pr(\mathcal{D}|\{\theta, \mathcal{M}\}) \cdot Pr(\theta|\mathcal{M})}{Pr(\mathcal{D}|\mathcal{M})} = \frac{\mathcal{L} \cdot \pi}{E} \ , \tag{5.45}$$

where the evidence $E = Pr(\mathcal{D}|\mathcal{M})$ is the probability of getting the data $\mathcal{D}$, given the model $\mathcal{M}$ and can be seen as the likelihood averaged over all the possible parameters within a model. In the model fitting problems, one tries to get the optimal set of parameters θ *within* the framework of one specific model. In this case, E is a constant and is ignored, since it does not change the conclusion of the optimization. As soon as one leaves the framework of one specific model for doing model selection, E becomes the value of interest. In that case, the aim is to confront our degree of belief in two different models in the light of the data, i.e. to compare $Pr(\mathcal{M}_1|\mathcal{D})$ and $Pr(\mathcal{M}_2|\mathcal{D})$. Developing each term with Bayes' theorem $Pr(\mathcal{M}|\mathcal{D}) = Pr(\mathcal{D}|\mathcal{M}) \cdot Pr(\mathcal{M})/Pr(\mathcal{D})$, we can write

$$\frac{Pr(\mathcal{M}_1|\mathcal{D})}{Pr(\mathcal{M}_2|\mathcal{D})} = \frac{Pr(\mathcal{D}|\mathcal{M}_1) \cdot Pr(\mathcal{M}_1)}{Pr(\mathcal{D})} \cdot \frac{Pr(\mathcal{D})}{Pr(\mathcal{D}|\mathcal{M}_2) \cdot Pr(\mathcal{M}_2)} \ . \tag{5.46}$$

Since we do not have any prior preference toward one of the models, $\frac{Pr(\mathcal{M}_1)}{Pr(\mathcal{M}_2)}$ is typically set to 1 and the above ratio simplifies to

$$\frac{Pr(\mathcal{M}_1|\mathcal{D})}{Pr(\mathcal{M}_2|\mathcal{D})} = \frac{E_1}{E_2} \ . \tag{5.47}$$

$ln(E_{b_4\neq0}/E_{b_4=0})$	interpretation	betting odds
< 1	not worth a bare mention	$< 3:1$
$1 - 2.5$	substancial	$3:1$
$2.5 - 5$	strong	$> 12:1$
> 5	highly significant	$> 150:1$

Table 5.4: The slightly modified Jeffrey's scale we use to interpret the Evidence ratio $ln(E_{b_4\neq0}/E_{b_4=0})$.

5.7.2 Computation and results

In order to calculate the evidence for our nested models, we use a multimodal nested sampling algorithm, *MultiNest* (Feroz & Hobson 2008). *MultiNest* provides Evidence for each model, and the evidence ratio is $ln(E_{b_4\neq0}/E_{b_4=0}) = -0.162$ for the i band analysis, and $ln(E_{b_4\neq0}/E_{b_4=0}) = +0.165$ for the g band analysis.

We use the slightly modified Jeffreys' scale, shown in table 5.4 (**??**Verde et al. 2013), to interpret these value. This scale classifies Evidence ratios from not worth a bare mention to highly significant. Both values calculated for the Evidence ratios correspond to non-significant ratio, suggesting that neither the i band data nor g band data privilege one model over the other. Therefore, the current data do not allow us to conclude on the detection or non-detection of the scale-dependent bias we search for.

5.8 Conclusion

We have developed a method for the detection of the modulation, from BAOs, of the large scales ratio of baryonic matter to total matter, with a data set containing 3-D positions and photometry. We have investigated the sensitivity of the BOSS CMASS DR10 data to this effect, via the parameter b_4, and obtained a null detection consistent with both $b_4 = 0$ and the theoretical b_4 predicted by Barkana & Loeb (2011). This lack of evidence is reflected by the evidence ratio we measure in section 5.7. We expect more accurate data to prove or disprove the prediction from Barkana & Loeb (2011), if our error bars on b_4 decrease by a factor of 5. A better account for the non-linear effects, e.g. with reconstruction, could improve the goodness of the fit and increase the evidence for a non-zero b_4. Future developments of this work should also include a deeper study of the feasibility of the detection, with current and future data sets, e.g. with simulations.

An important "sub-product" of the detection of a non zero b_4 would be the constraint on b_{CIP}. However, since $A_{CIP} = b_{CIP}/b_4^2$, the case of $b_4 = 0$ leads to an unconstrained A_{CIP}, and needs to be treated separately in order to constrain A_{CIP}. Therefore another immediate extension to this work would be to compute the distribution of A_{CIP} assuming $|b_4| > 0.01$.

The method we have explored is for the detection of the scale dependent bias with a data set containing 3-D positions and photometry. Our method will be ready to use when the DR11 release will be made public, and we are also planning to adapt it to data set containing 2-D positions and photometric redshift such a the DES.

This page was intentionally left blank

Chapter 6

Concluding remarks and Future work

One of the most important things I have learnt during this PhD has been to accept and appreciate that all research is, to some extent, open-ended. In this last section of this thesis, I summarize the main concluding remarks of each chapter and highlight some of their planned extensions.

Chapter 2, chapter 3 and chapter 4 addressed the problem of separating stars from galaxies in future large photometric surveys, and in particular in the Dark Energy Survey. Chapter 2 established the main requirements on the quality of the star/galaxy separation, for measurement of the cosmological parameters with the WL and LSS probes. We translated the DETF FoM requirements on the statistical and systematic errors and the constraints from PSF calibration into the corresponding science requirements on the quality of star/galaxy separation. We formulated these requirements using two parameters: the purity and completeness of classified samples of stars and galaxies.

This formulation served as a basis to assess the performance of multi_class, the classifier we designed in chapter 3. The aim of multi_class is to use as much as possible of the precious information encapsulated in the numerous parameters of the DES catalog. In this sense, it takes a multi-parameter approach to star/galaxy classification. We first use PCA to outline the correlations between the objects parameters and extract from it the most relevant information. We then use the reduced set of parameters which we call

the "winners", as input to an Artificial Neural Network. When tested on simulations of the Dark Energy Survey (DES), this multi-parameter approach improves upon purely morphometric classifiers such as the classifier implemented in SExtractor, spread_model and class_star. This is a valuable asset, especially at the faint magnitudes reached by the DES.

The improvement which multi_class allows on simulations, was confirmed on real data in chapter 4, where we used the recent release of DES Science Verification data to test the performance of our classifier. Testing of multi_class on the DES SV data allowed us to find one configuration of the method which outperforms spread_model and class_star on the full range of magnitudes of the catalog. The recent beginning of the DES survey operations in September, 2013 will allow us to continue the testing and optimisation of multi_class. More cross-matching, and especially the planned addition of stars from the LMC in the cross-matched catalog, should allow a better training of our classifier. Future works on multi_class should include a careful study of the threshold value of the Fisher discriminant of the principal components which optimise the performance of multi_class. In other words, we should study the way in which the optimal number of winners $w(f)$ depends on the Fisher discriminant of the selected principal component, which would allow to make the method even more flexible, by setting the threshold on the Fisher discriminant, rather than the number of "winners". Other aspects of the method, such as the network architecture, could also be optimised in the future.

In chapter 5, we investigated a completely different topic: the sensitivity of the two latest releases of the Baryon Oscillation Spectroscopic Survey (BOSS) data to a scale dependent bias predicted by Barkana & Loeb (2011): the modulation by Baryonic Acoustic Oscillations of the density ratio of baryon to dark matter across large regions of the Universe. The detection of this effect would provide a direct measurement of a difference in the large-scale clustering of mass and light and a confirmation of the standard cosmological paradigm from a different angle than any other measurement. We measured the number density correlation function and the luminosity weighted correlation function of the BOSS DR10 CMASS sample, and fit a model of scale dependent bias to our measurement.

Our current measurement does not allow us to conclude on a detection, in spite of the fact that the maximum likelihood value we measure for b_4 - the parameter characterising the effect - is consistent with predictions by Barkana & Loeb (2011). We expect more accurate data to prove or disprove the prediction from Barkana & Loeb (2011), if our

error bars on b_4 decrease by a factor of 5. A better account for the non-linear effects, e.g. with reconstruction, could increase the evidence for a non-zero b_4. Future developments of this work should also include a deeper study of the feasibility of the detection, with current and future data sets, e.g. with simulations. Moreover, an immediate extension to this work would be to add the LOWZ samples to this analysis.

An important "sub-product" of the detection of a non zero b_4 would be the constraint on the Compensated Isocurvature Perturbation amplitude A_{CIP}, characterized in our model my the b_{CIP} parameter. However, since $A_{CIP} = b_{CIP}/b_4^2$, the case of $b_4 = 0$ leads to an unconstrained A_{CIP}, and needs to be treated separately in order to constrain A_{CIP}. Therefore another immediate extension to this work would be to compute the distribution of A_{CIP} assuming $|b_4| > 0.01$.

We have developed a method for the detection of the scale dependent bias with a data set containing 3-D positions and photometry. Our method will be ready to use when the DR11 release will be made public, and we are also planning to adapt it to data set containing 2-D positions and photometric redshift such a the DES.

This page was intentionally left blank

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJSS, 182, 543

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2014, ApJSS, 211, 17

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, ArXiv Astrophysics e-prints

Amara, A. & Réfrégier, A. 2008, MNRAS, 391, 228

Anderson, L., Aubourg, E., Bailey, S., et al. 2013, ArXiv e-prints

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435

Axenides, M., Brandenberger, R., & Turner, M. 1983, Physics Letters B, 126, 178

Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625

Barkana, R. & Loeb, A. 2011, MNRAS, 415, 3113

Bartelmann, M. & Schneider, P. 2001, , 340, 291

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Blake, C., Baldry, I. K., Bland-Hawthorn, J., et al. 2013, MNRAS, 436, 3089

Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, AJ, 144, 144

Bouy, H., Bertin, E., Moraux, E., et al. 2013, A&A, 554, A101

Brandenberger, R. H. 1994, International Journal of Modern Physics A, 9, 2117

Bridle, S., Shawe-Taylor, J., Amara, A., et al. 2009, in Annals of Applied Statistics, American Astronomical Society Meeting Abstracts

Busha, M. T., Wechsler, R. H., Becker, M. R., Erickson, B., & Evrard, A. E. 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts, 341.07

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJL, 648, L109

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Collister, A. A. & Lahav, O. 2004, PASP, 116, 345

Cooray, A., Hu, W., & Miralda-Escudé, J. 2000, ApJL, 535, L9

Crocce, M. & Scoccimarro, R. 2008, , 77, 023533

Das, S., de Putter, R., Linder, E. V., & Nakajima, R. 2012, , 11, 11

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, ApJ, 757, 83

Doi, M., Tanaka, M., Fukugita, M., et al. 2010, AJ, 139, 1628

Efstathiou, G., Sutherland, W. J., & Maddox, S. J. 1990, Nature, 348, 705

Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007a, ApJ, 664, 675

Eisenstein, D. J., Seo, H.-J., & White, M. 2007b, ApJ, 664, 660

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Fadely, R., Hogg, D. W., & Willman, B. 2012, ApJ, 760, 15

Feroz, F. & Hobson, M. P. 2008, MNRAS, 384, 449

Firth, A. E., Lahav, O., & Somerville, R. S. 2003, MNRAS, 339, 1195

Fisher, R. A. 1936, 179

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fukuda, Y., Hayakawa, T., Ichihara, E., et al. 1998, Physical Review Letters, 81, 1562

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gordon, C. & Pritchard, J. R. 2009, , 80, 063535

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Grin, D., Hanson, D., Holder, G. P., Doré, O., & Kamionkowski, M. 2014, , 89, 023006

Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, AJ, 116, 3040

Guth, A. H. & Pi, S.-Y. 1982, Physical Review Letters, 49, 1110

Hamilton, A. J. S. 2000, MNRAS, 312, 257

Heavens, A. 2003, MNRAS, 343, 1327

Hogg, D. W., Bovy, J., & Lang, D. 2010, ArXiv e-prints

Holder, G. P., Nollett, K. M., & van Engelen, A. 2010, ApJ, 716, 907

Hu, W. & Keeton, C. R. 2002, , 66, 063506

Hu, W. & White, M. 2001, ApJ, 554, 67

Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101

Jaffe, A. 1996, ApJ, 471, 24

Jarvis, M., Bernstein, G., & Jain, B. 2004, MNRAS, 352, 338

Kaiser, N. 1992, ApJ, 388, 272

Kaiser, N. 1998, ApJ, 498, 26

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Kamionkowski, M., Spergel, D. N., & Sugiyama, N. 1994, ApJL, 426, L57

Kerscher, M., Szapudi, I., & Szalay, A. S. 2000, ApJL, 535, L13

Kirk, D., Laszlo, I., Bridle, S., & Bean, R. 2011, ArXiv e-prints

Kirk, D., Rassat, A., Host, O., & Bridle, S. 2012, MNRAS, 424, 1647

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643

Kron, R. G. 1980, ApJSS, 43, 305

Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64

Laszlo, I., Bean, R., Kirk, D., & Bridle, S. 2012, MNRAS, 423, 1750

Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473

Lightman, A. P. & Schechter, P. L. 1990, ApJSS, 74, 831

Linde, A. D. 1982, Physics Letters B, 116, 335

Linde, A. D. 1984, Soviet Journal of Experimental and Theoretical Physics Letters, 40, 1333

Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990, MNRAS, 243, 692

Massey, R., Stoughton, C., Leauthaud, A., et al. 2010, MNRAS, 401, 371

Mellier, Y. 1999, Ann. Rev. Astr. Astrophys., 37, 127

Milgrom, M. 1994, Annals of Physics, 229, 384

Mohr, J. J., Armstrong, R., Bertin, E., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Naim, A. 1995, in Bulletin of the American Astronomical Society, Vol. 27, American Astronomical Society Meeting Abstracts, 1416

Naoz, S. & Barkana, R. 2007, MNRAS, 377, 667

Odewahn, S. C., Stockwell, E. B., Pennington, R. L., Humphreys, R. M., & Zumach, W. A. 1992, AJ, 103, 318

Oyaizu, H., Lima, M., Cunha, C. E., et al. 2008, ApJ, 674, 768

Pan, J. & Zhang, P. 2010, , 8, 19

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., & Bridle, S. L. 2008, A&A, 484, 67

Peebles, P. J. E. 1980, The large-scale structure of the universe

Percival, W. 2007, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 720, The Invisible Universe: Dark Matter and Dark Energy, ed. L. Papantonopoulos, 157

Percival, W. J., Baugh, C. M., Bland-Hawthorn, J., et al. 2001, MNRAS, 327, 1297

Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148

Perlmutter, S., Turner, M. S., & White, M. 1999, Physical Review Letters, 83, 670

Phillips, M. M. 1993, ApJL, 413, L105

Planck Collaboration I. 2013, ArXiv e-prints

Planck Collaboration XVI. 2013, ArXiv e-prints

Rassat, A., Amara, A., Amendola, L., et al. 2008, ArXiv e-prints

Refregier, A. 2003, Ann. Rev. Astr. Astrophys., 41, 645

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Macri, L., Casertano, S., et al. 2009, ApJ, 699, 539

Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, ApJ, 607, 665

Ross, A. J., Ho, S., Cuesta, A. J., et al. 2011, MNRAS, 417, 1350

Rossetto, B. M., Santiago, B. X., Girardi, L., et al. 2011, AJ, 141, 185

Rudnicki, K. 1995, The cosmological principles.

Sakharov, A. D. 1966, Soviet Physics JETP, 22, 241

Sánchez, A. G., Scóccola, C. G., Ross, A. J., et al. 2012, MNRAS, 425, 415

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Sebok, W. L. 1979, AJ, 84, 1526

Seljak, U. & Zaldarriaga, M. 1996, ApJ, 469, 437

Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Takada, M. & Jain, B. 2004, MNRAS, 348, 897

Taylor, A. N. 2001, ArXiv Astrophysics e-prints

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22

Thomas, S. A., Abdalla, F. B., & Lahav, O. 2010, Physical Review Letters, 105, 031301

Thomas, S. A., Abdalla, F. B., & Lahav, O. 2011, Physical Review Letters, 106, 241301

Tojeiro, R., Percival, W. J., Wake, D. A., et al. 2012, MNRAS, 424, 136

Tojeiro, R., Ross, A. J., Burden, A., et al. 2014, MNRAS, 440, 2222

Valdes, F. 1982, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 331, Instrumentation in Astronomy IV, 465–472

Van Waerbeke, L., Mellier, Y., Radovich, M., et al. 2001, A&A, 374, 757

Van Waerbeke, L., White, M., Hoekstra, H., & Heymans, C. 2006, Astroparticle Physics, 26, 91

Vasconcellos, E. C., de Carvalho, R. R., Gal, R. R., et al. 2011, AJ, 141, 189

Verde, L., Feeney, S. M., Mortlock, D. J., & Peiris, H. V. 2013, , 9, 13

Viola, M., Kitching, T. D., & Joachimi, B. 2014, MNRAS, 439, 1909

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, , 530, 87

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143

Yee, H. K. C. 1991, PASP, 103, 396

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zaldarriaga, M. & Seljak, U. 2000, ApJSS, 129, 431

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

ד. כִּי אֶרְאֶה שָׁמֶיךָ מַעֲשֵׂה אֶצְבְּעֹתֶיךָ יָרֵחַ וְכוֹכָבִים
אֲשֶׁר כּוֹנָנְתָּה:

ה. מָה אֱנוֹשׁ כִּי תִזְכְּרֶנּוּ וּבֶן אָדָם כִּי תִפְקְדֶנּוּ:

"When I see Your heavens, the work of Your fingers, the moon and stars that You have established, what is man that You should remember him, and the son of man that You should be mindful of him?"
Psalms, Chapter 8