AIX-MARSEILLE UNIVERSITÉ

ÉCOLE DOCTORALE
PHYSIQUE ET SCIENCES DE LA MATIÈRE

THÈSE

pour obtenir le titre de

Docteur en Sciences

Mention: Astrophysique et Cosmologie

Présentée et soutenue par

Johan COMPARAT

Baryonic acoustic oscillations with emission line galaxies at intermediate redshift

The large-scale structure of the universe

Thèse dirigée par Jean-Paul KNEIB et Carlo SCHIMD

préparée au Laboratoire d'Astrophysique de Marseille

soutenue le 21 juin 2013

Jury :

Rapporteurs :	L. GUZZO	-	INAF (Milano)
	H. MCCRACKEN	-	UPMC, Paris VI (IAP)
Directeurs :	J.-P. KNEIB	-	CNRS (OSU-pythéas), EPFL (LASTRO)
	C. SCHIMD	-	AMU & CNRS (OSU-pythéas)
Président :	P. AMRAM	-	AMU & CNRS (OSU-pythéas)
Examinateurs :	W. PERCIVAL	-	University of Portsmouth (ICG)
	D. SCHLEGEL	-	Lawrence Berkeley National Laboratory

Institut Pythéas
Observatoire des Sciences de l'Univers
Aix-Marseille Université

UNIVERSITÉ DE PROVENCE

ÉCOLE DOCTORALE
PHYSIQUE ET SCIENCES DE LA MATIÈRE

T H È S E

pour obtenir le titre de

Docteur en Sciences

Mention: Astrophysique et Cosmologie

Présentée et soutenue par

Johan Comparat

Observation des oscillations baryoniques primordiales des galaxies à raie d'émission à décalage vers le rouge modéré.

La structure aux grandes échelles dans l'univers

Thèse dirigée par Jean-Paul Kneib et Carlo Schimd

préparée au Laboratoire d'Astrophysique de Marseille

soutenue le 21 juin 2013

Jury :

Rapporteurs :	L. Guzzo	-	INAF (Milano)
	H. McCracken	-	UPMC, Paris VI (IAP)
Directeurs :	J.-P. Kneib	-	CNRS (OSU-pythéas), EPFL (LASTRO)
	C. Schimd	-	AMU & CNRS (OSU-pythéas)
Président :	P. Amram	-	AMU & CNRS (OSU-pythéas)
Examinateurs :	W. Percival	-	University of Portsmouth (ICG)
	D. Schlegel	-	Lawrence Berkeley National Laboratory

Institut Pythéas
Observatoire des Sciences de l'Univers
Aix Marseille Université

Abstract

The strongest force of nature on cosmological scales is gravity, thus a physical description of the structure and dynamics of the Universe mainly relies on a theory of gravity. Today the best candidate is Einstein's theory of general relativity (Einstein, 1916). Space-time is determined with respect to its matter-energy content.This yields a relationship between Einstein tensor, which accounts for the geometry of the space-time, and the energy momentum tensor, which accounts for the matter-energy fields. We skip here any consideration about the topology of space-time and work in the standard framework (Peebles, 1993).

For a homogeneous universe filled with ordinary matter or radiation, general relativity predicts that the cosmic expansion will slow down over time, in accord with Newtonian intuition. However, studies of distant supernovae found that the expansion of the universe has accelerated over the last seven billion years (Riess et al., 1998; Perlmutter et al., 1999). The understanding of this acceleration is known as the dark energy puzzle. Five main experiments are undergoing or planned to understand dark energy: more precise observations of supernovae, of the cosmic microwave background temperature and polarization anisotropies, the baryonic acoustic oscillation in the clustering of galaxies (BAO), cosmological weak lensing, and galaxy clusters counts.

After an introduction on BAO, I discuss, the properties of the galaxies that seem to be better suited to measure BAO at redshift $z > 0.6$. I test some target selection algorithms to obtain galaxies to probe BAO from the ground-based observations at $0.6 < z < 1.7$ in chapter 2, and I demonstrate the feasibility of the target selection for bright emission line galaxies. The main physical mechanisms driving the efficiency of a selection are now understood, in particular the relation to the parent photometry. A puzzling issue remains, how one can estimate quantitatively the impact of the dust on the selection efficiency. I hope to address this question with future data-sets.

Then I investigate the two main systematic errors on the determination of the BAO scale we expect due to using emission line galaxies as tracers of the matter. In the third chapter, I show that the incompleteness in the redshift distribution, due to the measurement of the redshift with [OII] , is related to the instrumental resolution. I find there are two interesting regimes. For an observation of the brightest [OII] emitters, a moderate resolution is sufficient, whereas for a fainter survey, the highest the resolution the best. In chapter 4, I measure the clustering of these tracers on small and intermediate scales to determine their galaxy bias and their cross-correlation coefficient *i.e.* the relation between the galaxies acting as BAO tracers and the dark matter halos they inhabit. The linear galaxy bias of the selections discussed in the second chapter and I find they are highly biased ($b \sim 1.8$ at $z = 0.8$). As the time required to observed at a given signal to noise the galaxy power spectrum decreases with the square of the bias, using such tracers it requires less observing time to observe the BAO. Although, it constitutes a new challenge for reconstruction algorithms aiming at reducing the impact of non-linear effects of the growth of structures on the BAO scale and for the elaboration of mock catalogs.

Note de l'auteur

Les peintures presentés au début de chaque chapitre représentent l'état d'esprit général à adopter pour aborder chaque chapitre. J'ai eu la chance de pouvoir admirer les originaux lors de visite de musées non loin des lieux des conférences auxquelles j ai été convié. Ce sont les Figures 1, 2.1, 4.1, 5.1.
The paintings presented at the beginning of each chapter represent the general state of mind to adopt to read each chapter. I had the chance to admire the original copy when visiting museum not far from conferences I was part of during my PhD. See Figures 1, 2.1, 4.1, 5.1.

Contents

Figure 1: Joan Miró, la naissance du monde, 1925.

Chapter 1

Baryonic acoustic oscillation: an introduction

The distribution of the galaxies in the Universe is a direct probe of the structure formation history in the Universe. The imprint in the late-time clustering of matter from the baryon acoustic oscillations in the early Universe, and observed in the cosmic microwave background (Peebles and Yu, 1970; Sunyaev and Zeldovich, 1970), has emerged as an enticing way to measure the distance scale and expansion history of the Universe (Eisenstein and Hu, 1998; Seo and Eisenstein, 2003; Matsubara, 2004; Blake and Bridle, 2005; Dolney et al., 2006). In particular, the acoustic peak in the three dimensional correlation function constitutes a robust standard ruler to probe the geometry of the Universe at different redshifts.
The acoustic length scale r_s is defined as the comoving distance that the sound waves can travel from the Big Bang until recombination at redshift $z \sim 1100$, it is given by equation (1.1); see Holtzman (1989); Hu and Sugiyama (1996); Eisenstein and Hu (1998)

$$r_s = \int_{z_{\rm rec}}^{\infty} dz \frac{c_s(z)}{H(z)} \tag{1.1}$$

Using current cosmic microwave background Planck data, the measurement of the acoustic scale is $r_s = 144.71 \pm 0.60$ Mpc with $H_0 = 67.3$ km s^{-1} Mpc^{-1}. It corresponds to an angle $\theta_s = 0.596724° \pm 0.00038°$ on the microwave background sky (Planck Collaboration et al., 2013). Perturbation theory and numerical simulations argue that the accuracy of the acoustic scale throughout cosmic time is below 1% , making it an excellent standard ruler (Eisenstein et al., 2007b). The measurement of the baryonic acoustic oscillation feature (BAO) is of prime interest for cosmology as it provides a nearly systematic-free geometrical information that is complementary to other cosmological probes such as the cosmic microwave background and weak lensing. This standard ruler offers a robust route to the study of dark energy. Though to interpret the length of this ruler, a theory concerning the formation of structure is needed.

1.1 Gravitational clustering of matter in the linear regime: setup of BAO

The distance r_s that acoustic waves can propagate in the first million years of the Universe is a characteristic scale, detectable not only in the CMB anisotropies (Bennett et al., 2012; Hinshaw et al., 2012; Planck Collaboration et al., 2013), but also in the late-time clustering of galaxies (Beutler et al., 2011; Blake et al., 2011c; Padmanabhan et al., 2012; Anderson et al., 2012).

1.1.1 Matter power spectrum

Under the assumption of a Universe filled with baryons, dark matter, dark energy, photons, and massless neutrinos; and if gravitation has no preferred scale, the initial power spectrum of matter fluctuations, denoted $P_i(k)$, is proportional to k^{n_s}, where n_s is the scalar spectrum power-law index at $k_0 = 0.05$ Mpc^{-1}. Harrison and Zeldovich showed that an equipartition of the power in fluctuations at any scales corresponds to $n_s = 1$. Using current Planck data, the measurement of the scalar spectrum power-law index is $n_s = 0.9603 \pm 0.0073$.

The theory of inflation states that present structures are evolved from small fluctuations of the continuous matter, density, velocity, and temperature fields. The physics at scales smaller than the radius of curvature of the Universe are well described by Eulerian hydrodynamics and Newtonian mechanics. It is made of five equations, the continuity equation, the equation of motion, the Poisson equation, the equation of state and the equation of energy. The application of this set of equations describes the large densities and velocities that occur in the observable Universe. In the linear approximation and under the separation of variables, two solutions describe the time evolution of the over-density field at time t and location r, denoted $\delta(r,t) = \rho(r,t)/\bar{\rho} - 1$, where $\rho(r,t)$ is the density at time t and location r, and $\bar{\rho}$ is the mean density of the Universe. These solutions are a growing mode, denoted $D(t)$, and a decaying mode, denoted $D_{-}(t)$,

$$\delta(r,t) = A(r)D(t) + B(r)D_{-}(t) \tag{1.2}$$

Ignoring the decaying mode that vanishes rather quickly, one can define the dimension less growth rate of the over-density, denoted f,

$$\frac{d\delta(r,t)}{dt} = H\delta\frac{d\log D}{d\log a} = H\delta f \tag{1.3}$$

The growth rate depends on the redshift z, the cosmological model. It is accurately approximated by $f = \Omega_m(z)^{0.6} + \Omega_\Lambda(1+\Omega_m/2)/70$ (Lahav et al., 1991).
Besides the time evolution that applies to the complete Euclidean space, the over-density depends on the space component through $A(r)$. This function is usually described in Fourier space as the comparison of the spatial modes present at $z=\infty$ and $z=0$, denoted $T(k)$,

$$T(k) = \frac{\delta(k, z=0)\delta(k=0, z=\infty)}{\delta(k, z=\infty)\delta(k=0, z=0)} = \widehat{A(r)} \tag{1.4}$$

The analytical expression of $T(k)$ proposed by Eisenstein and Hu (1998) provides results with a precision $\sim 10\%$. This description takes in account different effects: Compton drag, velocity overshoot, baryon infall, adiabatic damping, Silk damping, and cold dark matter (CDM) growth suppression. The main physical scales and orders of magnitude involved in the transfer function are the following.
At the redshift of matter-radiation equality, $z_{eq} = 3391 \pm 60$, the radiation-dominated Universe becomes matter-dominated. The amplitude of perturbations on smaller scales than the particle horizon is damped. At z_{eq} the particle horizon is ~ 93 Mpc.
Then, baryons are freed from the Compton drag of the photons. This release takes place between z_{eq} and the redshift at which baryon-drag optical depth equals unity, denoted $z_d = 1059.25 \pm 0.58$. At this period, two effects dominate. First, the small scale perturbations in the photon-baryon-fluid propagate as acoustic waves at speed $c_s \approx 0.56c$ within the sound horizon. Secondly, the coupling photon-baryon is not perfect. There are diffusion processes, that damp perturbations below scales of about 19 Mpc, referred to as 'Silk damping'.
Finally, the linear matter power spectrum for a standard inflation model reads:

$$P_m(k,z) \propto k^{n_s} |T(k,z)|^2 D^2(z) \ [\mathrm{Mpc}^3] \tag{1.5}$$

The amplitude is not specified, but is obtained with low redshift observables as galaxy clustering or cluster counts.

1.1.2 Galaxy power spectrum

In the standard cosmological model, at large scales, the relationship between the field of over-density of galaxies δ_g and the field of over-density of dark matter δ_m is usually modeled by the following equation:

$$\delta_g(k) = \left(b_g(z) + \frac{d\ln D(z)}{d\ln a(z)}\mu^2\right)\delta_m(k) + \varepsilon \tag{1.6}$$

in which we denoted the linear galaxy bias $b_g(z)$, the linear growth factor $D(z)$, the expansion factor $a(z) = 1/(1+z)$, the cosine of the angle between $\mathbf{k}$ and the observer's line of sight μ, and a spatially uncorrelated noise term ε (random variable, 'white noise') (Kaiser, 1987; McDonald and Roy, 2009). The galaxy bias is a general function, that describes the relation between the galaxies and the dark matter underlying field. One cannot observe directly δ_g. The observable is the power spectrum of galaxies, P_g, that is related to the matter power spectrum, P_m, by

$$P_g(k) = \left(b_g(z) + \frac{d \ln D(z)}{d \ln a(z)} \mu^2 \right)^2 P_m(k) + \bar{n}^{-1} \tag{1.7}$$

As long as galaxies are Poisson distributed around the biased large-scale density field, the white noise power spectrum is given by the inverse of number density of galaxies, $\bar{n}^{-1} = \varepsilon^2$.

To obtain a measurement of the BAO standard ruler to a sub per cent precision with the galaxy power spectrum, one cannot rely only on the linear theory of structure formation, and a detailed the study of $b_g(z)$ is mandatory.

1.2 Non-linear gravitational clustering: distortion of the BAO signal

The non-linear gravitational collapse, the galaxy clustering bias, and the redshift distortions modify the galaxy clustering measurement. It degrades the precision on the measurement of the acoustic scale.

In order to connect the initial perturbations to the non-linear structures we observe today, one needs a description of the non linear collapse of structures. Zel'dovich (1970) described the non-linear collapse of a perturbation of the cosmic density field. He showed the collapse generically occurs first along one direction, producing a sheet-like structure and then filaments. Press and Schechter (1974); Sheth et al. (2001) estimated the mass distribution of collapsed objects in the non linear regime (spherical and ellipsoidal) produced by hierarchical clustering of initial (top-hat) filtered gaussian perturbations. Nowadays, numerical simulations provide the most accurate description of the growth of structure and galaxy clustering in the non-linear regime, assuming Newtonian approximation; see for example the very recent work by Vogelsberger et al. (2012). In the quest of precision cosmology, in particular to address the dark energy puzzle, simulations should be done in a general relativistic framework (Buchert and Ostermann, 2012; Buchert et al., 2013).

The simulations show that non-linear structure formation and, to a lesser extent, redshift distortions erase the higher harmonics of the acoustic oscillations (Meiksin et al., 1999; Seo and Eisenstein, 2005). There are three non-linear gravitational effects that affect the BAO feature. First, the velocity flow that moves matter by $\sim 10 h^{-1}$ Mpc relative to its initial position. It displaces pairs out of the 150 Mpc separation bin, and broadens the peak in configuration space; in Fourier space it corresponds to the damping of the oscillations. Secondly the redshift-space distortions further increase this effect. These two effects come from astrophysical processes occurring at small scales approximatively smaller than 20 Mpc. Thirdly, mode coupling becomes relevant on larger scales as time increases; this effect reduces the contrast of the peak (Meiksin et al., 1999; Jeong and Komatsu, 2006). This raises the values of the non diagonal coefficients of the 2PCF covariance matrix (Schulz and White, 2006) and degrades the Gaussianity of the initial field of perturbation. Some of the information initially contained in the two point correlation function is contained in higher order correlation functions.

It is not obvious to understand what small scale non-linearity could not only broaden the acoustic peak, but shift it, introducing systematic errors, *i.e.* an effect that systematically moves all galaxy pairs separated by 150 Mpc to smaller or larger separations. Regarding this issue, Eisenstein et al. (2007b) showed that non-linear gravitational physics can only introduce shifts of $\sim$0.5% for objects with $b \sim 2$ at z $=0$ (decreasing at higher redshift) Therefore, the characteristic separation of 150 Mpc is not shifted but rather blurred by the action of cluster formation and bulk flows that move matter around; see Fig. 1.1.

Two year later, Padmanabhan and White (2009) characterized the non-linear evolution of the BAO feature as traced by the dark matter and halos, using a combination of perturbation theory and N-body

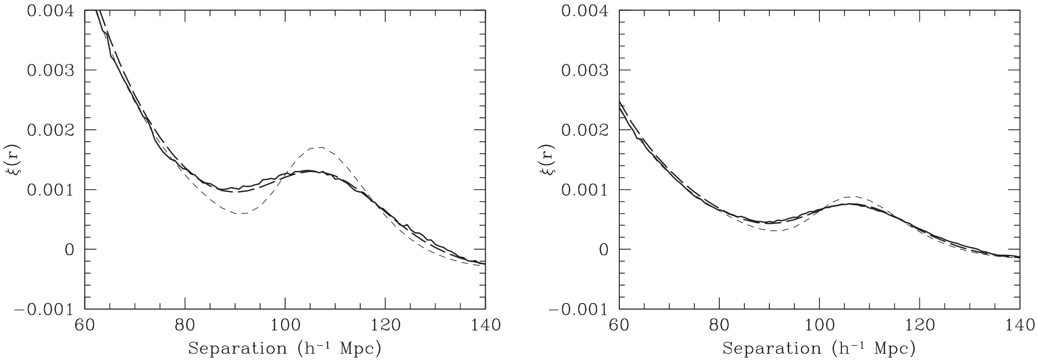


Figure 1.1: Fig. 5 from Eisenstein et al. (2007b). Correlation function from simulations compared with a model with the small-scale linear power restored. Left, $z = 0.3$; right, $z = 1$. In both panels, the simulation data are the solid line, the linear correlation function is the short-dashed thin line, and the model correlation function is the long-dashed line. The model seems to correctly match the smearing of the peak in the correlation function.

simulations. They confirmed that the acoustic peak traced by the dark matter is both broadened and shifted as structure forms, and that this shift is well described by second-order perturbation theory. They demonstrated that these theoretical systematics are smaller than the statistical precision of upcoming surveys and that modeling has the potential to reduce these systematics to below the levels of cosmic variance limited surveys.

Recently, a simulation by Mehta et al. (2011) showed it is possible with current reconstruction algorithms and a knowledge of the galaxy bias at the 10% level to correct from the small systematic shift due to galaxy bias from ∼0.5% to 0.15%; see Fig. 1.2.

In brief, the acoustic scale is a robust standard ruler because it is fundamentally a 150 Mpc clustering signature. We have shown here that the degradations of the signature have benign quasi-linear causes. These can be computed to high accuracy with N-body gravitational simulations. The remaining challenge for the method is the enormous survey volumes required to measure the scale to high precision. In particular to link the theory to the observation, by describing the estimators of the two point correlation function or of the power spectrum.

1.3 Estimating the power spectrum of the galaxy field

The estimation of the two-point correlation function (2PCF) on a galaxy catalog was explained by Fall (1979), then by Davis and Peebles (1983); Landy and Szalay (1993); Hamilton (1993), and lately by (Vargas-Magaña et al., 2012). Under the assumption that the distribution of galaxies to represents a stationary random process, Fall (1979) derived that the isotropic 2PCF, denoted $\xi(r)$, is the joint probability of finding galaxies in the elemental volumes δv_1 and δv_2 separated by a distance r where N is the mean spatial density of galaxies

$$\delta P(r) = N^2[1 + \xi(r)]\delta v_1 \delta v_2 \tag{1.8}$$

where P, is not strictly a true probability as it is not normed. Points distributed in a Poisson or in a uniformly random manner have $\xi = 0$. It means the probability of finding two points at any two points would be independent of their separation r and equal to $N^2\delta v_1 \delta v_2$. $\xi(r)$ is therefore interpreted as an "excess probability" of finding galaxies within a given separation.

Landy and Szalay (1993); Hamilton (1993); Vargas-Magaña et al. (2012) give different estimators of the 2PCF based on the count of pairs of galaxies compared to that of random mock galaxy catalog. In this thesis, all 2PCF estimations are done via these pair-counting estimators. Alternatively, one can directly estimate of the power spectrum of galaxies from a catalog of galaxy positions with the

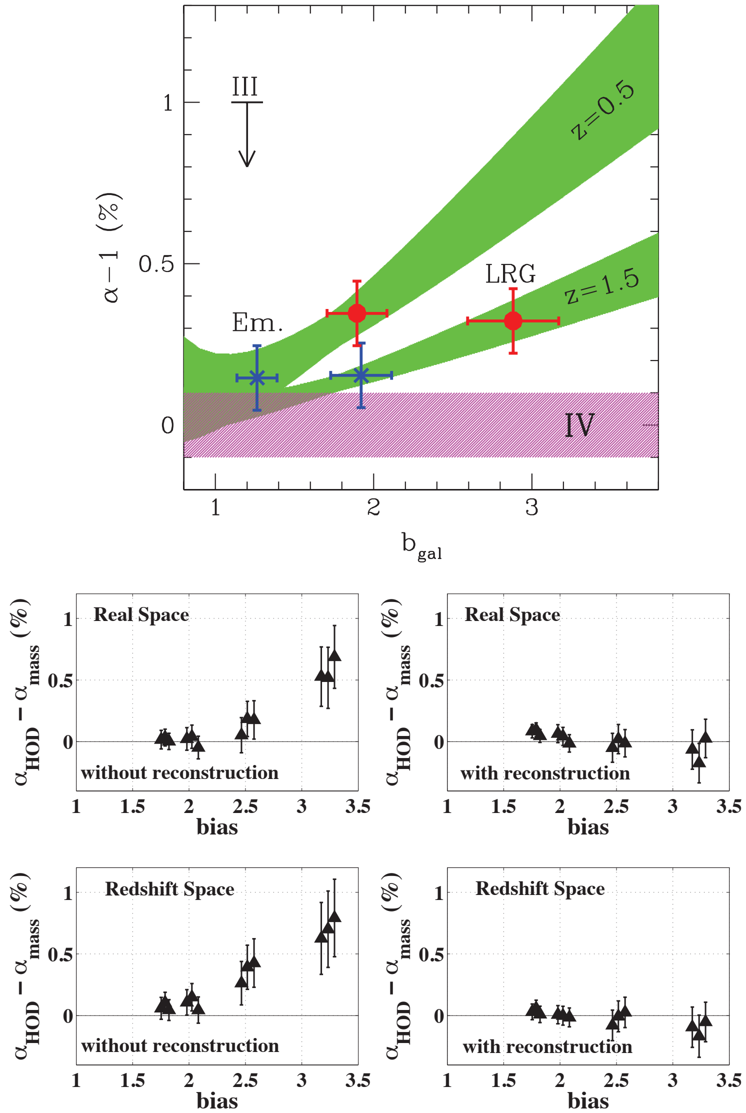


Figure 1.2: α is the shift in the acoustic scale. A good agreement between theory and observations should give $\alpha = 1$. The top panel is taken from Padmanabhan and White (2009). It estimates of the systematic shift in the acoustic scale (if not corrected) as a function of halo bias and redshift for current fiducial ΛCDM cosmology. The width of the colored regions denotes the error estimated from the difference between the Press-Schechter and Sheth-Tormen forms and demonstrates approximately how errors in the theory propagate into a residual shift. Two example populations of halos are shown: a $b(z = 0) = 1$ sample [blue crosses] to represent emission line galaxies, and a $b(z = 0) = 1.5$ sample [red circles] to represent elliptical galaxies. The error bars correspond to a 10% measurement of the bias. Also shown are the nominal distance accuracies for Stage III (currently underway) and Stage IV (future) BAO experiments. The bottom panel is from Mehta et al. (2011). It shows one can correct with reconstruction from the systematic shift due to galaxy bias on α_{HOD} down to precision of 0.15% to recover α_{mass}. There is some evidence for a shift between the most biased HODs and the mass case. However, this effect disappears after we apply the reconstruction algorithm.

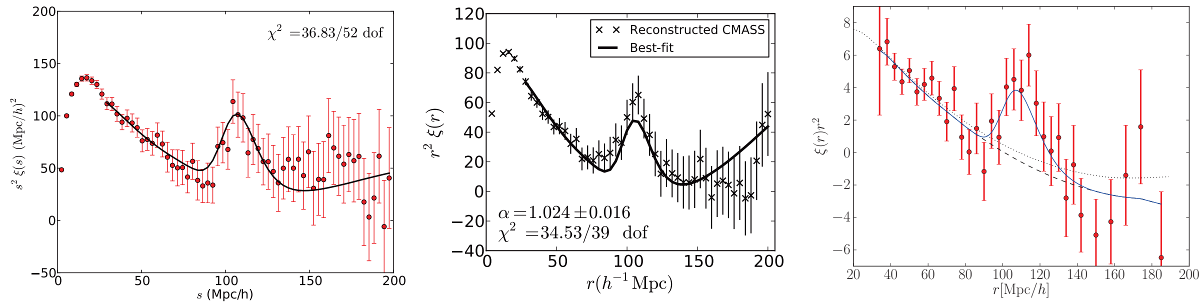


Figure 1.3: Reconstructed two-point correlation function vs. distance. Left, using the LRGs at $z = 0.35$ (Padmanabhan et al., 2012). Middle, using the CMASS at $z = 0.56$ (Anderson et al., 2012). Right, using the Lyα forest (Slosar et al., 2013). The amplitude of the peak varies from tracer to tracer.

'FKP algorithm' that consists in Fourier transforming the galaxy field (Feldman et al., 1994; Hamilton, 2009b). A third set of techniques directly inspired from the CMB analysis, as the Karhunen-Loeve linear compression scheme (Vogeley and Szalay, 1996; Tegmark et al., 1997; Hamilton, 2009a) or the quadratic pre-compression scheme (Tegmark, 1997; Padmanabhan et al., 2001) will be very efficient with future gigantic surveys. On large scale, both classes of estimators agree and are equivalent in terms of computation time.

1.4 Measurements: a short history

The BAO feature was first observed in early 2005 both in the SDSS luminous red galaxy sample (Eisenstein et al., 2005), and the 2dFGRS data (Cole et al., 2005). Other concurrent clustering analysis studies on the SDSS-II LRG sample were carried out and confirmed the first measurement (Padmanabhan et al., 2007; Sánchez et al., 2009; Kazin et al., 2010; Percival et al., 2010). The success of these experiment enabled the start of another generation of observing programs such as 6dfGRS (Beutler et al., 2011), WiggleZ (Blake et al., 2011c) and SDSS-III/BOSS (Anderson et al., 2012). Fig. 1.3 shows the two-point correlation function measured.

The comoving size of the BAO feature in the 2PCF at redshift z in the line-of-sight direction, $r_\parallel$, is related to the observed redshift size Δz by the Hubble parameter $H(z)$:

$$r_\parallel = \frac{cz}{H(z)}. \tag{1.9}$$

The comoving size of the BAO feature in the 2PCF at redshift z in the transverse direction, $r_\perp$ is related to the observed angular size $\Delta\theta$ by the angular diameter $d_A(z)$:

$$r_\perp = (1+z)d_A(z)\Delta\theta. \tag{1.10}$$

Using the isotropic 2PCF, one obtains a measurement of a blended distance, denoted d_V, given equation (1.11), to be compared with r_s (equation 1.1):

$$d_V(z) \equiv \left[r_\perp^2 r_\parallel\right]^{1/3} = \left[(1+z)^2 d_A^2(z)\frac{cz}{H(z)}\right]^{1/3} \tag{1.11}$$

Models of the 2PCF are fitted to determine the so-called shift-parameter α defined as

$$(d_V/r_s) = \alpha(d_V/r_s)_{\rm fiducial} \tag{1.12}$$

This fitting procedure has been extensively tested against incorrect assumptions on the underlying cosmology and galaxy bias yielding fairly good results at the per cent level (Padmanabhan and White, 2009; Xu et al., 2012; Padmanabhan et al., 2012). The current measurement of $d_V/r_s\ (z)$ compared to different cosmological models is shown on Fig. 1.4. The measurements are in agreement with WMAP ΛCDM best-fit. The precision ranges between 2 and 5%.

The use of the monopole of the 2PCF assumes that the enhanced clustering amplitude along the line-of-sight due to redshift space distortions does not alter the relative importance of radial and angular modes when calculating spherically averaged correlation functions. This assumption holds to derive the acoustic scale to a 2% precision. For the next era of precision measurements, this assumption does not hold anymore, and anisotropic 2PCF models are required; see (Reid et al., 2012; Samushia et al., 2013; Sanchez et al., 2013).

Recently a novel BAO technique was developed and applied to Ly_α forest from quasars (QSO). McDonald (2003); McDonald and Eisenstein (2007) proposed to use the Lyman-α forest to measure the angular diameter distance and Hubble parameter at $2 < z < 4$ using the standard ruler provided by the BAO. In order to reach a competitive cosmological measurement, the three-dimensional density field on the scale of the BAO feature should be sampled with a sufficiently high density of order of ~ 15 QSOs deg^{-2}. Techniques using both the photometry and variability of the quasars have been developed to identify QSOs candidates at $z > 2.2$ in the most effective possible way (Palanque-Delabrouille et al., 2011; Ross et al., 2012b). Such a measurement is part of the SDSS-III/BOSS spectroscopic survey where $z > 2.2$ quasars are observed. The first result are stunning, as they clearly show the phase of deceleration of the Universe at $z > 0.8$ in agreement with the ΛCDM model, see Fig. 1.5. The current precision of the measurement of $H(z = 2.3)$ is of about 3.5% (Pâris et al., 2012; Busca et al., 2012; Lee et al., 2013; Kirkby et al., 2013; Slosar et al., 2013), confirming the predictions by simulations (Le Goff et al., 2011).

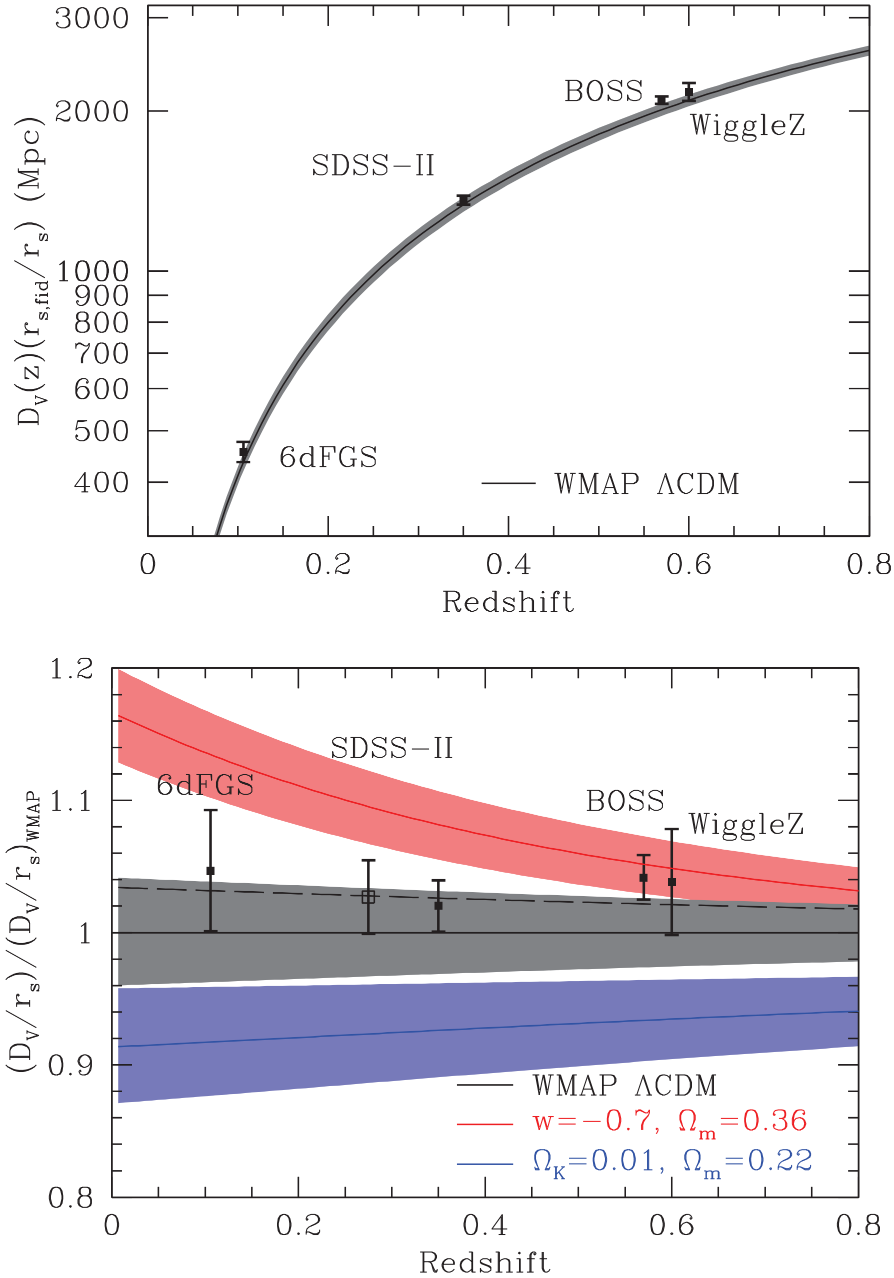


Figure 1.4: Results taken from (Anderson et al., 2012). **Top.** Distance-redshift relation ($d_V(z)/r_s$ times the fiducial r_s to restore a distance) from various BAO measurements from spectroscopic data sets. Included here are the CMASS measurement at $z = 0.57$, the 6dF Galaxy Survey measurement at $z = 0.1$ (Beutler et al., 2011), the SDSS-II LRG measurement at $z = 0.35$ (Padmanabhan et al., 2012; Xu et al., 2012; Mehta et al., 2012), and the WiggleZ measurement at $z = 0.6$ (Blake et al., 2011c). The grey region is the 1σ prediction from WMAP under the assumption of a flat Universe with a cosmological constant (Komatsu et al., 2011). The agreement between the various BAO measurements and this prediction is excellent. **Bottom.** The BAO distance-redshift relation divided by the best-fit flat, ΛCDM prediction from WMAP in grey. The dashed line is the best-fit CMB+LRG+CMASS+flat ΛCDM, which clearly is a good fit to all data sets. Also shown are the predicted regions from varying the spatial curvature to $\Omega_\kappa = 0.01$ (blue band) or varying the equation of state to $w = -0.7$ (red band).

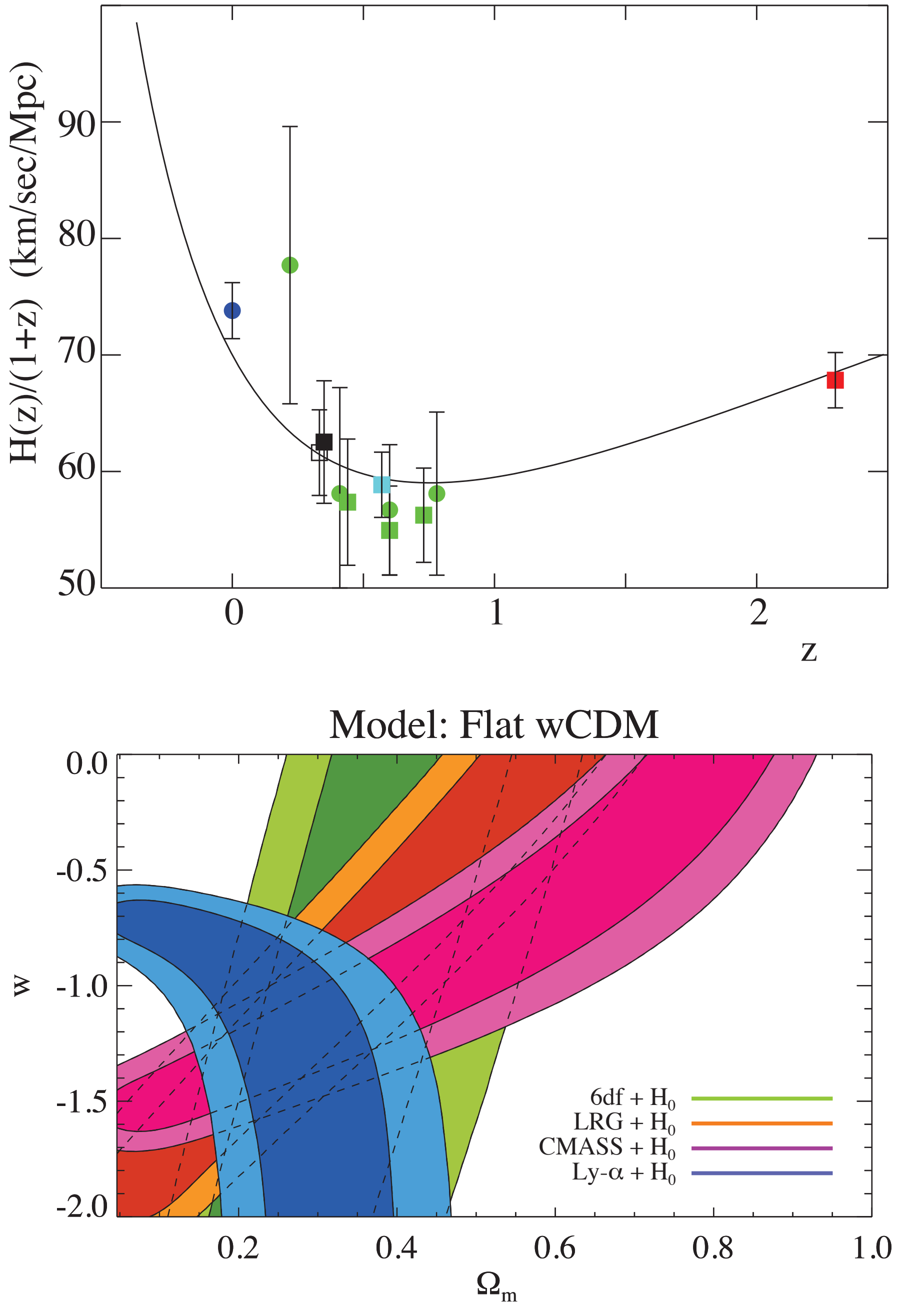


Figure 1.5: Results using Lyα forest QSO taken from (Busca et al., 2012). **Top.** Measurements of $H(z)/(1+z)$ vs. z demonstrating the acceleration of the expansion for $z < 0.8$ and deceleration for $z > 0.8$. It shows the Lyα BAO-based measurements at $z = 2.3$ (red square), the SDSS DR7 BAO measurement at $z = 0.35$ [Xu et al. (2012); filled black square] and [Chuang and Wang (2012); open black square], the BOSS BAO-Alcock-Paczynski measurement at z = 0.57 [Reid et al. (2012), cyan square], and the WiggleZ BAO-Alcock-Paczynski measurement [Blake et al. (2012), green square]. The green circles are from WiggleZ (Blake et al., 2011c) Alcock-Paczynski data combined with supernova data. The blue circle is the H_0 measurement of (Riess et al., 2011). The line is the ΛCDM prediction for $(h, \Omega_m, \Omega_\Lambda) = (0.7, 0.27, 0, 73)$. **Bottom plot.** Constraints on the matter and dark-energy density parameters $(\Omega_m, \Omega_\Lambda)$ assuming a dark-energy pressure-density ratio $w = -1$. The blue regions are the 1 and 2 σ constraints derived from the Lyα-BAO combined with a measurement of H_0 (Riess et al., 2011). Also shown are the contours from lower redshift measurements of d_V/r_s (also combined with H0) at z = 0.11 [6dFGRS: Beutler et al. (2011)], z = 0.35 [LRG: Percival et al. (2010)] and z = 0.57 [CMASS: Anderson et al. (2012)]. All constraints use a WMAP7 (Komatsu et al., 2011) prior on the baryon-to-photon ratio η but do not otherwise incorporate CMB results.

1.5 Uncertainty on the measurement of the acoustic scale

The measurement of the BAO scale using the 2PCF of galaxies at a given redshift is subject to systematic errors of order of a few percent. We can classify them into two categories. First, the systematic errors due to the survey geometry and its design *i.e.* the evaluation of its completeness as a function of the position on the sky and redshift. Secondly, the astrophysical process associated to the galaxies observed, occurring at small scales, underlying the biasing function. Currently both systematic error sources are under control and understood to a lower level than statistical errors.

Completeness, selection function

For surveys with $z < 0.7$, completeness systematics errors are smaller than 1% on the BAO scale measurement, and it exists patches to cope with them. For future surveys, it might be more tricky to handle this systematics to obtain precision of the order of 0.1%, see the investigations I led on this issue in chapter 3.

There are three main concerns about completeness (Ross et al., 2012a). The first kind of systematic error is due to the variations of the photometric calibration across the footprint of the survey and its impact on the galaxy population selected and then observed. The second source of error comes from the relation between the observing conditions on the observed density of galaxies, in fact, the varying observational conditions affect the redshift success rate, that will not be uniform across the survey area. Weights applied on the angular selection function is the standard mean to correct for these two effects. The third systematic error is related to the radial completeness of tracers, and remind to the question about the best way to simulate the radial selection function and the importance of the effects related to galaxy evolution. It is possible to make mock catalogs that are insensitive to systematic shifts in redshift. The observation of these mocks with slightly varying redshift distributions shows the systematic effect is very small compared to the statistical errors of the measurement, in particular on large scales. This result is shown on Fig. 1.6 (Ross et al., 2012a).

Concerning future 'DETF stage IV BAO surveys', the precision aims being at the sub-percent level, it will be necessary to improve the way to handle the systematics due to the target selection, that are currently at about 1% and might be even higher for fainter tracers.

Bulk flows, galaxy bias

To obtain a high precision on the measure of the BAO scale, it is necessary to correct the 2PCF from the dominant non-linear effect of clustering. The bulk flows at a scale of 20 h^{-1} Mpc that form large scale structures smear the BAO peak: it is smoothed by the velocity of pairs. At redshift 1 the rms displacement for biased tracers due to bulk flows is 8.5 h^{-1} Mpc in real space and 17 h^{-1} Mpc in redshift space; see Eisenstein et al. 2007a,b. We can cope with this systematic effect using a reconstruction technique; it consists in correcting this smoothing effect by statistically displacing the galaxies to their original position. The algorithm applied to SDSS-III/CMASS galaxies is described by Padmanabhan et al. (2012); see Fig. 1.7.

The reconstruction algorithm applied on the SDSS-II Data Release 7 (Abazajian et al., 2009) LRG sample sharpens the BAO feature and reduces the errors from 3.5 percent to 1.9 percent. This sample has a density of tracers of 10^{-4} h^3 Mpc^{-3} and the optimum smoothing applied is 15 h^{-1} Mpc (Padmanabhan et al., 2012). Once applied to the DR9-SDSS-III/BOSS spectroscopic sample (different patches cover 3,275 deg^2 on a total of 10,000 deg^2), this reconstruction method sharpens the BAO peak allowing a detection at high significance, but does not significantly improve the precision on the distance measure due to the gaps in the current survey; see Anderson et al. (2012). To allow an optimum reconstruction using a smoothing of 5 h^{-1} Mpc, a dense and contiguous galaxy survey seems a mandatory requirement: the gaps in the survey footprint must be smaller than 1 h^{-1}Mpc, and the sampling density higher than 3×10^{-4} h^3 Mpc^{-3}. This should not be an issue for the final BOSS data set and future surveys.

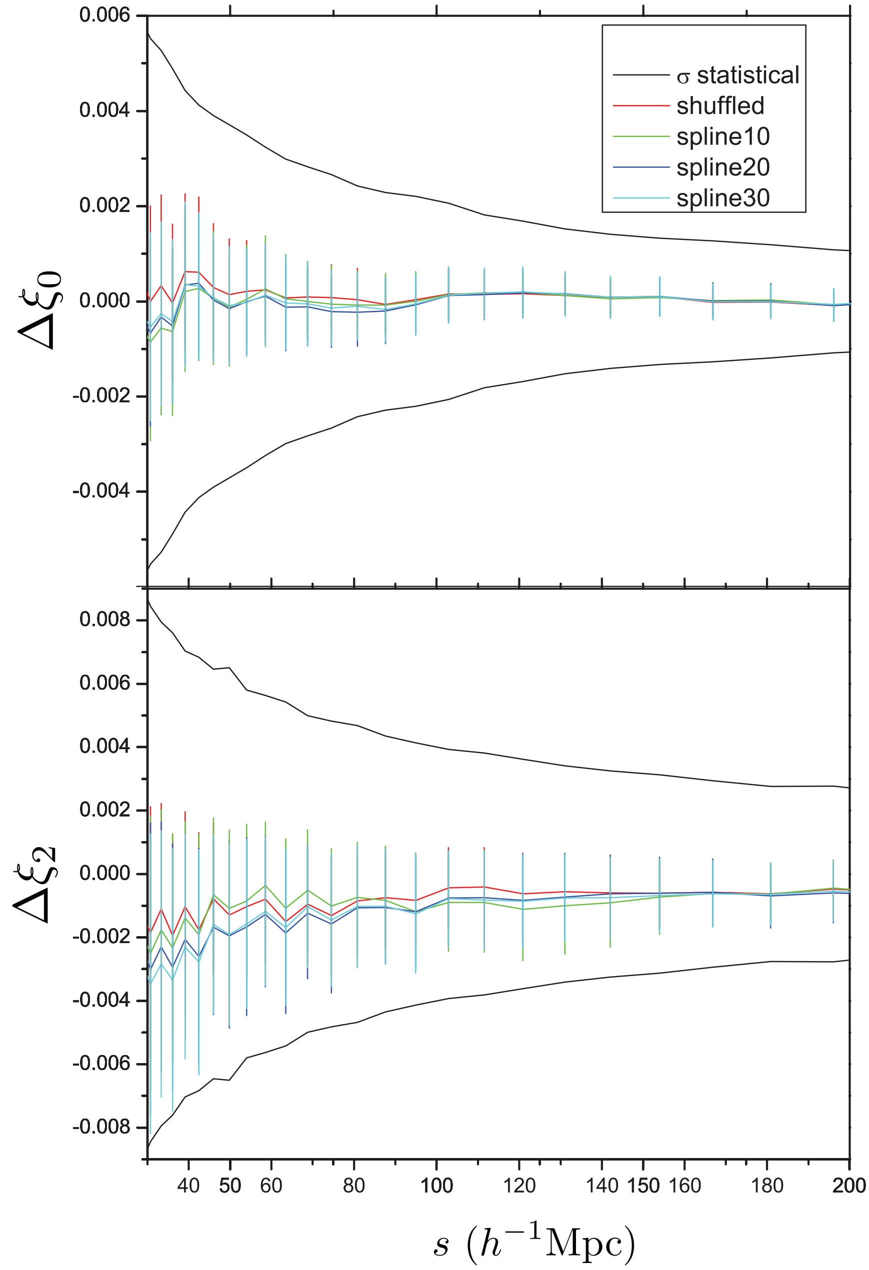


Figure 1.6: Taken from Ross et al. (2012a). ξ_0 and ξ_2 are the monopole and quadrupole projections of the 2PCF. **Top.** Average offsets and 1σ deviations from a true ξ_0 as a function of the separation in redshift space, s, for different methods of assigning redshifts to random catalogues, determined using the 600 mock galaxy catalogues. The solid black line corresponds to the statistical error bars. **Bottom.** as above, for ξ_2.

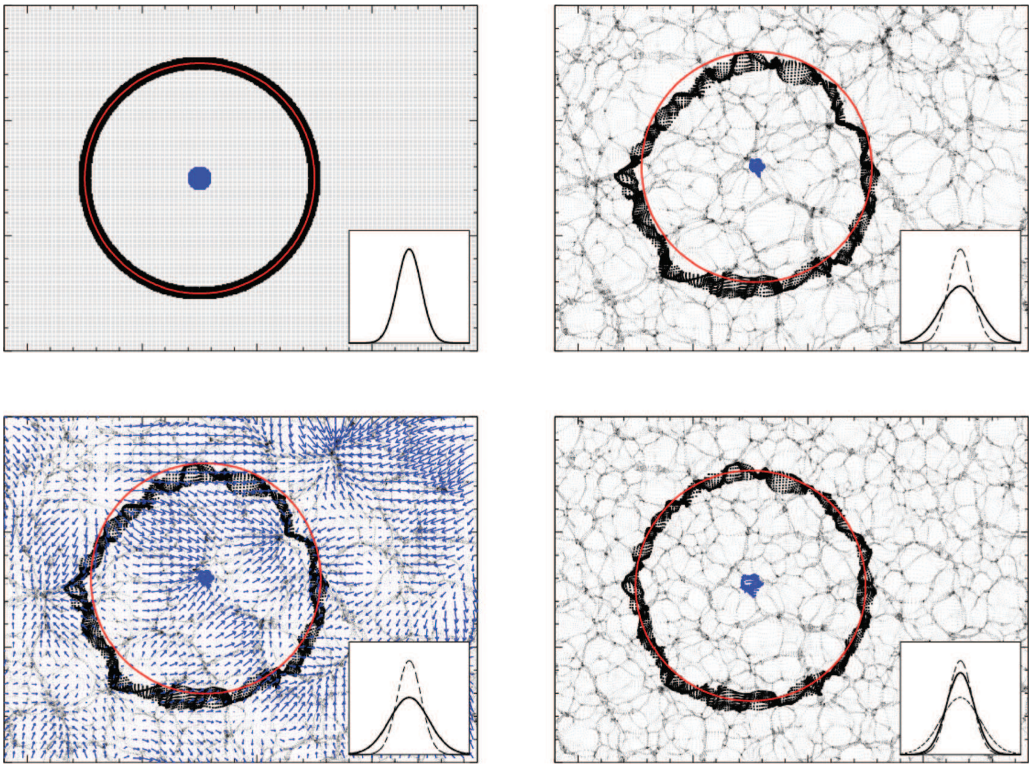


Figure 1.7: Taken from Padmanabhan et al. (2012). A pictorial explanation of how density field reconstruction can improve the measurement of the acoustic scale. In each panel, a thin slice of a simulated cosmological density field is shown. **Top-left.** In the early Universe, the initial densities are very smooth. We mark an ideal acoustic feature as a broad sphere of radius 150 Mpc and Gaussian profile. A Gaussian with the same rms width as the radial distribution of the black points from the centroid of the blue points is shown in the inset. **Top-right.** Particles are displaced according to the Zel'dovich approximation (Zel'dovich, 1970). The red circle shows the initial radius of the ring, centered on the current centroid of the blue points. The large-scale velocity field has caused the black points to spread out; this causes the broadening of the acoustic feature. The inset shows the current rms radius of the black points relative to the centroid of the blue points (solid line) compared to the initial rms (dashed line). **Bottom-left.** As before, but overplotted with the Lagrangian displacement field, smoothed by a 10 h^{-1}Mpc Gaussian filter. The concept of reconstruction is to estimate this displacement field from the final density field and then move the particles back to their initial positions. **Bottom-right.** We displace the present-day position of the particles by the opposite of the displacement field in the previous panel. Because of the smoothing of the displacement field, the result is not uniform. However, the acoustic ring has been moved substantially closer to the red circle. The inset shows the new rms radius of the black points (solid), compared to the initial width (long-dashed) and the uncorrected present-day width (short-dashed). The narrower peak will make it easier to measure the acoustic scale. Note that the algorithm applied to the data is more complex than was just described, but this figure illustrates simply reconstruction.

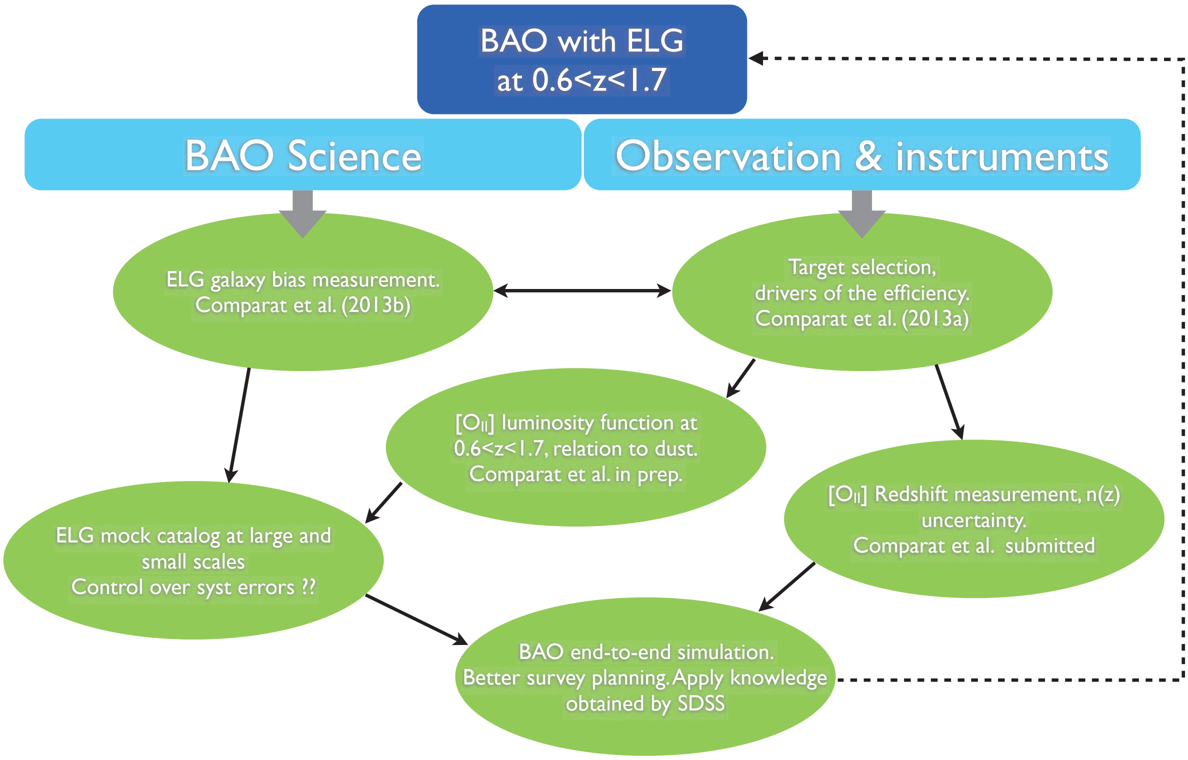


Figure 1.8: Mind map connecting the different parts of my thesis within the project of measuring the BAO feature in the galaxy power spectrum of emission line galaxies at intermediate redshift.

1.6 Plan of the dissertation

During my 3 years of research at the Laboratoire d'astrophysique de Marseille, I investigated the target selection of BAO tracers for $z > 0.6$. This study presented in chapter 2 links the theoretical requirements to observe the BAO to the current photometric surveys to draw the targets from.

Then, in the third and fourth chapters, I focus on the the two main sources of systematic errors on the determination of the BAO scale we expect due to using emission line galaxies as tracers of the matter.

First I show that the incompleteness in the redshift distribution, due to the measurement of the redshift with [OII] , is related to the instrumental resolution. Then, I find there are two interesting regimes. For an observation of the brightest [OII] emitters, a moderate resolution is sufficient, whereas for a fainter survey, the highest the resolution, the best.

Secondly, in chapter 4, I focus on the clustering of these tracers on small and intermediate scales to determine the galaxy bias and the cross-correlation coefficient *i.e.* the relation between the galaxies acting as BAO tracers and the dark matter halos they inhabit. I estimated the linear galaxy bias of the selections discussed in the second chapter and I find they are highly biased. As the time required to observed at a given signal to noise the galaxy power spectrum decreases with the square of the bias, using such tracers it requires less observing time to observe the BAO. Although, it constitutes a new challenge for the creation of mock catalogs.

The mind map characterizing the plan of my thesis is shown on Fig. 1.8. This figure connects all the parts constituting my research and its future development.

Chapter 2

Investigating emission-line galaxy surveys with the Sloan Digital Sky Survey infrastructure

Johan Comparat, Jean-Paul Kneib, Stephanie Escoffier, Julien Zoubian, Anne Ealet, Fabrice Lamareille, N. Mostek, Oliver Steele, Eric Aubourg, Stephen Bailey, Adam S. Bolton, Joel Brownstein, Kyle Dawson, Jian Ge, Olivier Ilbert, Alexie Leauthaud, Claudia Maraston, Will Percival, Nicholas P. Ross, Carlo Schimd, David J. Schlegel, Donald P. Schneider, Daniel Thomas, Jeremy L. Tinker, Benjamin A. Weaver. Published by MNRAS in 2013, 428-498 (arXiv:1207.4321)

Figure 2.1: Vincent Van Gogh. La nuit étoilée. 1889.

2.1 Motivation and contribution

At the start of my PhD in 2010, I proposed with Jean-Paul Kneib the observation of a sample of emission line galaxies (ELG) at the call for ancillary proposal of the SDSS-III experiment. The aim of the proposal was to try and observe high redshift ($0.6 < z < 1.7$) galaxies that will be targeted by the BigBOSS (now MS-DESI) experiment to probe the BAO peak. I put together all the figures and calculations to write this proposal. The proposal was accepted and 2000 fibers were allocated to the observation of ELGs. I provided the list of targets to be observed during this program within the CFHT-LS wide fields number 1, 3 and 4, that are overlapping with the SDSS-III/BOSS footprint. Each spectrum was observed during about 1h with the SDSS-III/BOSS spectrograph.

The analysis of these observations led to the publication mentioned above. The first step was to determine the redshift efficiency of the current SDSS redshift solver. Were the output redshift the right ones ? It came out they were not. A few collaborators and myself therefore inspected all the spectra by eye and with the results coming from three different pipelines, to assign the correct redshift and provide a truth table concerning the actual redshift of the observations.

These inspections enabled me to interact with the BOSS pipeline team and to get acquainted with the issues the redshift solvers have to assign correct redshifts to ELG. Then to obtain the physical properties of the observed targets I interacted with O. Ilbert, F. Lamareille, D. Thomas and O. Steele in order to understand the values output by their pipeline: measurement of emission lines, stacked spectra, redshift uncertainty ... The measurement of the redshift and the derivation of the physical properties of ELG was very interesting as we went down to the best extraction of the information located in this data set.

The second part of the work was to evaluate the usefulness of ELGs for BAO measurement. I wrote python libraries to compute the linear power spectrum, cosmological distances and volumes in order to derive all the numbers, figures and tables of the section 2.3. Finally to extrapolate to new selections and decide what photometric depth was desired to have the highest efficiency, I wrote a code to simulate a degradation of a photometry to a worse quality, so that I could determine what is the limit in terms of precision in the parent photometry to have a 70 or a 60 per cent efficiency.

Based on this publication, I presented a poster for the BigBOSS collaboration at the AAS in Austin, USA in january 2012. I also presented this work two summers in a row at the BigBOSS summer workshop in LBNL, Berkeley, USA. Through these travels I have set solid collaborative links with the americans working on this subject.

The results from this publication opened a new opportunity as we demonstrated that we could do an ELG-based BAO survey with the SDSS telescope infrastructure. In part, this work lead to the After-Sloan-III eBOSS proposal lead by my supervisor Jean-Paul Kneib and submitted September 1st, 2011. In this proposal, I was responsible for the writing of the part BAO with ELG. Shortly after in November 2011, the AS3 board selected eBOSS to be one of the projects to be done during the next 6 years, starting in 2014, as the next generation SDSS cosmological survey. This success, has motivated me to continue the analysis of this sample (chapter 4) and try to optimize the instrumentation for the eBOSS survey (chapter 3). Fig. 2.2 illustrates the mind map characterizing this chapter.

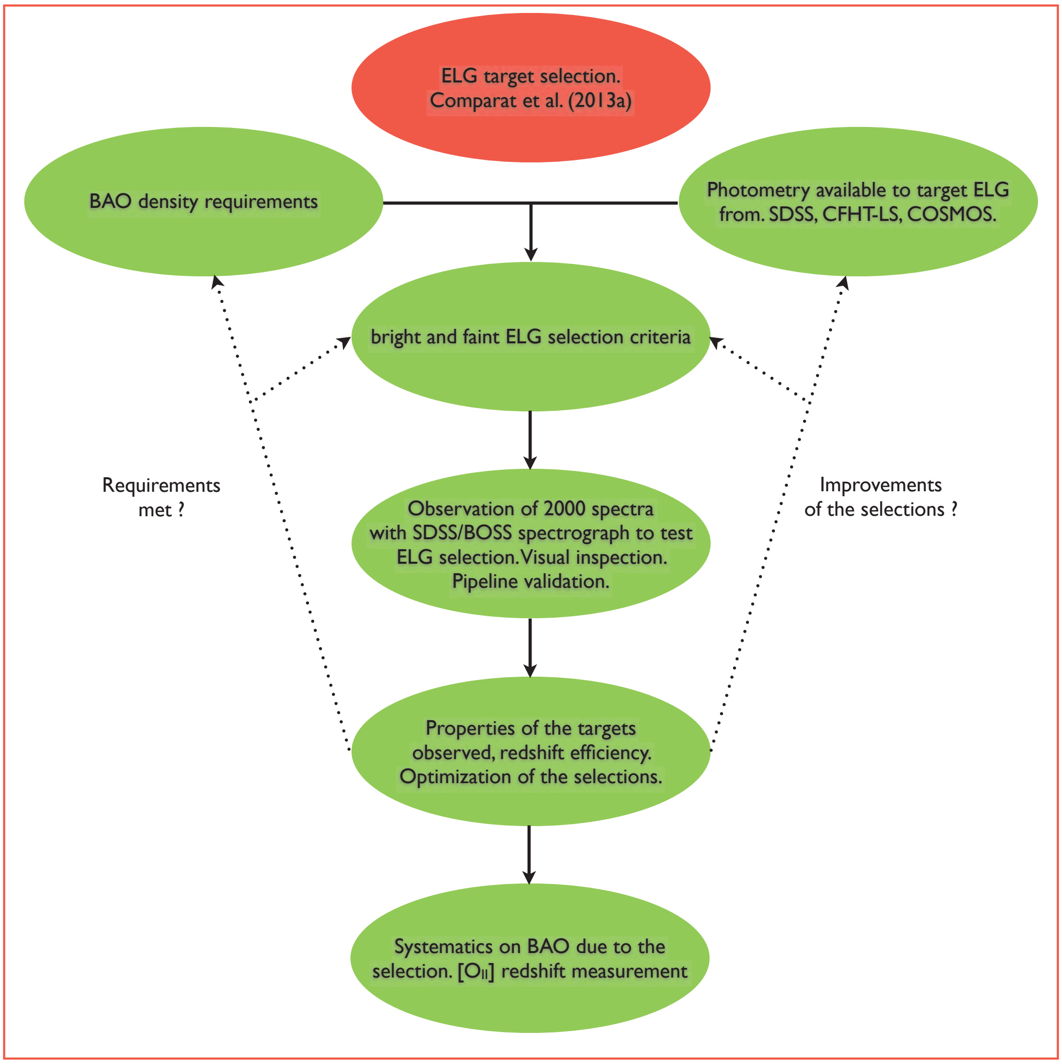


Figure 2.2: Mind map; investigating emission-line galaxy surveys with the Sloan Digital Sky Survey infrastructure.

2.2 Introduction

With the discovery of the acceleration of the expansion of the universe (Riess et al., 1998; Perlmutter et al., 1999), possibly driven by a new form of energy with sufficient negative pressure, recent results have concluded that ~ 96 percent of the energy density of the universe is in a form not conceived by the Standard Model of particle physics and not interacting with the photons, hence dubbed "dark". Lying at the heart of this discovery is the distance-redshift relation mapped by the type Ia supernovae (SnIa) combined with the temperature power spectrum of the cosmic microwave background fluctuations. Since the first detections, a great amount of data has been observed up to redshift $z \sim 1$ (Riess et al. 1998,Perlmutter et al. 1999,Astier et al. 2006,Wood-Vasey et al. 2007, Riess et al. 2004, Riess et al. 2007, Dawson et al. 2009; Riess et al. 2011). The current precision and accuracy required to obtain deeper insight on the cosmological model using SnIa is limited by the systematic errors; therefore to assess a firm picture of the cosmological model, a joint analysis with other cosmological probes is mandatory.

Corresponding to the size of the well-established sound horizon in the primeval baryon-photon plasma before photon decoupling (Peebles and Yu, 1970), the BAO scale provides a standard ruler allowing for geometric probes of the global metric of the universe. In the late-time universe it manifests itself in an excess of galaxies with respect to an unclustered (Poisson) distribution at the comoving scale $r \sim 100h^{-1}$Mpc — corresponding to a fundamental wave mode $k \sim 0.063h$Mpc^{-1}. The value of this scale at higher redshift is accurately measured by the peaks in the CMB power spectrum (*e.g.* Komatsu et al. 2009, 2011). Galaxy clustering and CMB observations therefore allow for a consistent comparison of the same physical scale at different epochs.

The first detection of the 'local' BAO (Cole et al., 2005; Eisenstein et al., 2005) were based on samples at low redshift $z \leq 0.4$. Further analysis on a larger redshift range ($z > 0.5$) and a wider area confirm the first result, reducing the errors by a factor of 2 (Percival et al., 2010; Blake et al., 2011c). Measurements of the BAO feature have thus become an important motivation for large galaxy redshift surveys; the small amplitude of the baryon acoustic peak, and the large value of $r_{\rm BAO}$, require comoving volumes of order of $\sim 1{\rm Gpc}^3h^{-3}$ and at least 10^5 galaxies to ensure a robust detection (*e.g.* Tegmark 1997; Blake and Glazebrook 2003).

BAO studies using luminous red galaxies (LRG) are currently being pushed to $z = 0.7$ by the Baryonic Oscillation Spectroscopic Survey (BOSS) experiment as part of the Sloan Digital Sky Survey III (SDSS-III) survey (Eisenstein et al., 2011). So far, with a third of the spectroscopic data, the BAO feature has been measured at $z = 0.57$ with a 6.7σ significance (Anderson et al., 2012). The final data set, which will be completed by mid-2014, will have a mean galaxy density of about 150 galaxies per square degree over 10,000 deg^2. Recently, the WiggleZ experiment has obtained a significant $\sim 4.9\sigma$ detection of the BAO peak at $z = 0.6$, by combining information from three independent galaxy surveys: the SDSS, the 6-degree Field Galaxy Survey (6dFGS) and the WiggleZ Dark Energy Survey (Blake et al., 2011c). In contrast to SDSS, WiggleZ has mapped the less biased, more abundant emission line galaxies (Drinkwater et al., 2010).

The next generation of cosmological spectroscopic surveys plans to map the high-redshift universe in the redshift range $0.6 \leq z \leq 2$ using the largest possible volume; see BigBOSS (Schlegel et al., 2011), PFS-SuMIRe[1], and EUCLID[2]. To achieve this goal one needs tracers covering this redshift. Above redshift $z \sim 0.6$ the bulk of galaxy population is composed of star forming galaxies (Abraham et al., 1996; Ilbert et al., 2006) and the number density of LRG starts to decrease. it is therefore compelling to build a large sample of such type of galaxies, which allows one to cover a large area and hence a large volume. The main challenges for future BAO surveys is to efficiently select targets for which a secure redshift can be measured within a short exposure time. Contrary to continuum-based LRG surveys, the observational strategy of next generation surveys such as BigBOSS, PFS-SuMIRe, and EUCLID is based on redshift measurements using emission lines, which are a common feature of

[1]http://sumire.ipmu.jp/en/
[2]http://sci.esa.int/euclid

Table 2.1: Requirements for measuring at 1 percent the BAO signal. $\bar{n}$ is the required density to overcome shot noise. The observed density [galaxy deg^{-2}] is the projection of $\bar{n}$ on the sky. The Sample variance area required [deg^2] corresponds to the area necessary to obtain an effective volume of 1 Gpc^3 h^{-3} at $k_1 \simeq 0.063\ h$ Mpc^{-1} and $k_2 \simeq 0.12\ h$ Mpc^{-1} given the value $\bar{n}$. $N_{\rm req}$ is the number of thousands of redshifts required to detect BAO at k_1 and k_2: it is the product of the required observed density multiplied by the area required. 'req.' stands for required.

redshift range	Shot Noise req. $\bar{n}(k_1)$ $10^{-4}h^3$Mpc^{-3}	$\bar{n}(k_2)$	observed density req. [deg^{-2}] for k_1	for k_2	Sample variance area req. [deg^2]	$N_{\rm req}$ [10^3 redshifts] for k_1	for k_2
$[0.3, 0.6]$	1.0	2.1	33	71	6188	204	440
$[0.6, 0.9]$	1.1	2.5	75	162	2585	194	419
$[0.9, 1.2]$	1.3	2.9	121	261	1615	195	421
$[1.2, 1.5]$	1.5	3.2	164	354	1227	201	435
$[1.5, 1.8]$	1.7	3.6	273	589	1041	284	613

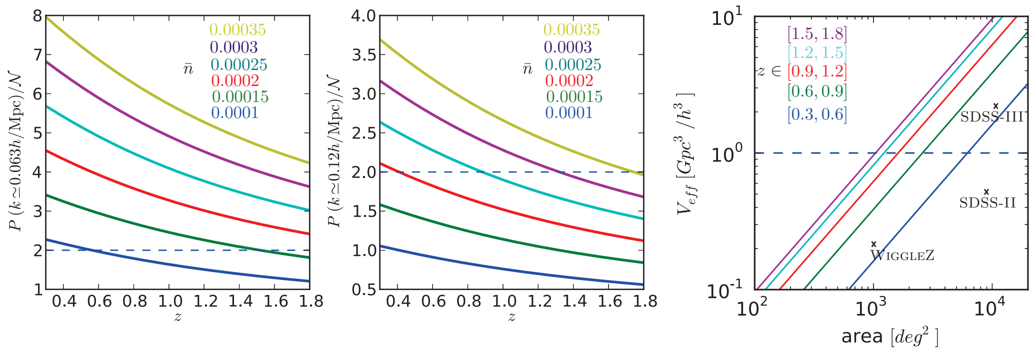


Figure 2.3: $P/\mathcal{N} = \bar{n}P$, ratio of the fiducial power spectrum over the shot noise as a function of redshift at $k \simeq 0.063$ **(left)** and at $k \simeq 0.12$ **(center)** for $\bar{n} \in [1, 3.5]\ 10^{-4}\ h^3$Mpc^{-3}. **(Right)** Effective volume in Gpc^3 h^{-3} sampled as a function of sky coverage, for five different redshift bins from 0.3 to 1.8, the sampling density used for each bin is given in the column $\bar{n}(k_2)$ of Table 2.1. To reach a comoving volume of 1 Gpc^3 h^{-3} over the redshift range $[0.6, 0.9]$, it is necessary to cover about 2500 deg^2.

star-forming galaxies. In this paper we focus on targeting strategies for selecting luminous ELGs at $0.6 < z < 1.7$ using optical photometry, and we test our strategies using the BOSS spectrograph on the SDSS telescope (Gunn et al., 2006).

The plan of the paper is as follow. In section 2.3, we derive the necessary ELG redshift distribution to detect the BAO feature. In section 2.4 we explain how the ELG selection criteria were designed using different photometric catalogs, based on the performances of the BOSS spectrograph. In section 2.5 we compare observed spectra issued from this selection with simulations and we discuss the efficiency of the proposed selection schemes. In section 2.6 we discuss the main physical properties of the ELGs. In section 2.7, we present the redshift distribution of the observed ELGs and how to improve the selection. In appendix 2.9 we display a representative set of the spectra observed.

Throughout this study we assume a flat ΛCDM cosmology characterized by $(\Omega_m, n_s, \sigma_8) = (0.27, 0.96, 0.81)$. Magnitudes are given in the AB system.

2.3 Baryonic acoustic oscillations with ELG

2.3.1 Density and geometry requirements

In order to constrain the distance-redshift relation at $z > 0.6$ using the BAO, we need a galaxy sample that covers the volume of the universe observable at this redshift. In this section we derive the required mean number density of galaxy, $\bar{n}(z)$, and the area to be covered in order to observe the BAO feature at the one percent level.

The statistical errors in the measure of the power spectrum of galaxies $P(k, z)$, evaluated at redshift z and at scale k, arise from sample variance and shot noise (Kaiser, 1986). Denoting the latter as $\mathcal{N}(z) = 1/\bar{n}(z)$, to measure a significant signal the minimal requirement is

$$\bar{n}(z)P(k,z) = \frac{P(k,z)}{\mathcal{N}(z)} \gtrsim 2. \tag{2.1}$$

As the amplitude of the power spectrum decreases with redshift, the required density increases with redshift. *e.g.*, at $z = 0.6$, we need a galaxy density of $\bar{n} = 2.1 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$; at $z = 1.5$, $\bar{n} = 3.2 \times 10^{-4}\ h^3\mathrm{Mpc}^{-3}$. The full trend in redshift bins is given in Table 2.1 and in Figures 2.3 a) and b) which show equation (2.1) as a function of redshift for $k \simeq 0.063\ h\ \mathrm{Mpc}^{-1}$ and $k \simeq 0.12\ h\ \mathrm{Mpc}^{-1}$ (the location of the first and the second harmonics of the BAO peak in the linear power spectrum).

In order to minimize the sample variance, we must sample the largest possible volume (a volume of 1 $\mathrm{Gpc}^3\ h^{-3}$ roughly corresponds to a precision in the BAO scale measurement of 5 percent). To quantify this calculation, we use the effective volume sampled V_{eff}, defined as (Tegmark et al., 1997)

$$V_{\mathrm{eff}}(k) = 4\pi \int dr\, r^2 \left[\frac{\bar{n}(r)b^2(z)P(r,k)}{1+\bar{n}(r)b^2(z)P(r,k)}\right]^2. \tag{2.2}$$

In this calculation, we assume a linear bias according to the DEEP2 study by Coil et al. (2008) that varies according to the redshift as $b(z) = b_0(1 + z)$, with $b(z = 0.8) = 1.3$. The bias could be larger for the more luminous ELGs, that are thought to be the progenitors of massive red galaxies (Cooper et al., 2008). We shall evaluate the bias of ELGs more precisely in a future paper. The corresponding area to be surveyed in order to reach $V_{\mathrm{eff}} \sim 1\mathrm{Gpc}^3\ h^{-3}$ is shown in Table 2.1, setting redshift bins of width $\Delta z = 0.3$ from $z = 0.3$ to $z = 1.8$. The Figure 2.3 c) shows the behavior of V_{eff} as a function of the area for a given slice of redshift with $\bar{n}$ given in the third column of Table 2.1. For the redshift range $[0.6, 0.9]$ the survey area must be $\gtrsim$2,500 deg^2. For the redshift range $[0.9, 1.2]$ the survey area must be $\gtrsim$1,600 deg^2. The observation of $[0.6, 1.7]$ with a single galaxy selection thus needs 2,500 deg^2 to sample the BAO at all redshifts.

Reconstruction of the galaxy field

To obtain a high precision on the measure of the BAO scale, it is necessary to correct the 2-point correlation function from the dominant non-linear effect of clustering. The bulk flows at a scale of 20 h^{-1} Mpc that form large scale structures smear the BAO peak: it is smoothed by the velocity of pairs (At redshift 1 the rms displacement for biased tracers due to bulk flows is 8.5 h^{-1} Mpc in real space and 17 h^{-1} Mpc in redshift space) (Eisenstein et al., 2007a,b).

Reconstruction consists in correcting this smoothing effect. The key quantity that allows reconstruction on a data sample is the smoothing scale used to reconstruct the velocity field and should be as close to 5 h^{-1} Mpc as possible in order to measure the bulk flows without being biased by other non-linear effects that occur on smaller scales.

The reconstruction algorithm applied on the SDSS-II Data Release 7 (Abazajian et al., 2009) LRG sample sharpens the BAO feature and reduces the errors from 3.5 percent to 1.9 percent. This sample has a density of tracers of $10^{-4}\ h^3\ \mathrm{Mpc}^{-3}$ and the optimum smoothing applied is 15 h^{-1} Mpc (Padmanabhan et al., 2012). On the SDSS-III/BOSS data in our study (different patches cover 3,275

deg^2 on a total of 10,000 deg^2), reconstruction sharpens the BAO peak allowing a detection at high significance, but does not significantly improve the precision on the distance measure due to the gaps in the current survey (see Anderson et al. 2012).

To allow an optimum reconstruction using a smoothing three times smaller (5 h^{-1} Mpc) it is necessary to have a dense and contiguous galaxy survey : gaps in the survey footprint smaller than 1 Mpc and a sampling density higher than $3 \times 10^{-4}\ h^3\ \text{Mpc}^{-3}$. This setting should reduce the sample variance error on the acoustic scale by a factor four.

2.3.2 Observational requirements

A mean galaxy density of $3 \times 10^{-4}\ h^3\ \text{Mpc}^{-3}$ can be reached by a projected density of 162 galaxies deg^{-2} with $0.6 < z < 0.9$, 261 deg^{-2} with $0.9 < z < 1.2$, 354 with $1.2 < z < 1.5$, and 589 with $1.5 < z < 1.8$. Considering a simple case where a survey is divided in three depths, the shallow one covering 2,500 deg^2 should contain 419,000 galaxies ; the medium 421,000 galaxies over 1,600 deg^2 ; and the deep 435,000 galaxies over 1,200 deg^2. This represents a survey containing 1,350,000 measured redshifts in the redshift range $[0.6, 1.5]$. It is a challenge to build a selection function that enables the observation of these projected densities.

Given a ground-based large spectroscopic program that measures 1.5×10^6 spectra (it corresponds to about 4 years of dark time operations on SDSS telescope dedicated to ELGs), the challenge is to define a selection criterion that samples galaxies to measure the BAO on the greatest redshift range possible. We define the selection efficiency as the ratio of the number of spectra in the desired redshift range and the number of measured spectra. Under the aforementioned constraint, the galaxy selection must reach an efficiency of ~ 90 percent.

2.3.3 Previous galaxy targets selections

To reach densities of tracers $\gtrsim 10^{-4}\ h^3\ \text{Mpc}^{-3}$ at $z > 0.6$ with a high efficiency, a simple magnitude cut is not enough. Such a selection would be largely dominated by low-redshift galaxies. The use of colour selections is necessary to narrow the redshift range of the target selection for observations.

SDSS-I/II galaxies are selected with visible colours in the red end of the colour distribution of galaxies, resulting in a sample of LRG and not ELGs (Eisenstein et al., 2001). The projected density of LRG is $\sim 120\ \text{deg}^{-2}$ with a peak in the redshift distribution at $z \sim 0.35$. With the SDSS-I/II LRG sample, the distance redshift relation was reconstructed at 2 percent at $z = 0.35$.

BOSS has currently completed about half of its observation plan. The tracers used by BOSS are, as SDSS-I/II LRG, selected in the red end of the color distribution of galaxies, they are called CMASS (it stands for 'constant mass' galaxies) and the selection will is detailed in Dawson et al. (2013). The current BAO detection using the data release 9 (a third of the observation plan) with the CMASS tracers at $z \sim 0.57$ has a 6.7σ significance (Anderson et al. 2012).

WiggleZ blue galaxies are selected using UV and visible colours: they have a density of 240 galaxies deg^{-2} and a peak in the redshift distribution around $z = 0.6$ (Drinkwater et al., 2010). The WiggleZ experiment has obtained a 4.9σ detection of the BAO peak at $z = 0.6$ (Blake et al., 2011c).

At their peak density, both of these BAO surveys reach a galaxy density of $3 \times 10^{-4}\ h^3\ \text{Mpc}^{-3}$, which guarantees a significant detection of the BAO.

Galaxy selections beyond $z = 0.6$ were already performed by surveys such as the VIMOS-VLT Deep Survey[3] (VVDS, see Le Fèvre et al. 2005a), DEEP2[4] (see Davis et al. 2003) or Vimos Public Extragalactic Redshift Survey[5] (VIPERS, see Guzzo et al. 2013), but they are not tuned for a BAO analysis. The DEEP2 Survey selected galaxies using BRI photometry in the redshift range $0.75 - 1.4$ on a few square degrees with a redshift success of 75 percent using the Keck Observatory. It studied the evolution of properties of galaxies and the evolution of the clustering of galaxies compared to samples

[3] http://cesam.oamp.fr/vvdsproject/
[4] http://deep.berkeley.edu/index.html
[5] http://vipers.inaf.it/project.html

at low-redshift. In particular, insights in galaxy clustering to $z = 1$ brings a strong knowledge about the bias of these galaxies (Coil et al., 2008). The VVDS wide survey observed 20 000 redshift on 4 deg^2 limited to $I_{AB} < 22.5$ (Garilli et al., 2008); they studied the properties of the galaxy population to redshift 1.2 and the small scale clustering around $z = 1$. The VIPERS survey maps the large scale distribution of 100 000 galaxies on 24 deg^2 in the redshift range $0.5 - 1.2$ to study mainly clustering and redshift space distortions. Their colour selection, based on *ugri* bands, is described in more detail in the section 2.7.

2.4 Color Selections

Our aim is to explore different colour selections that focus on galaxies located in $0.6 < z < 1.7$ with strong emission lines, so that assigning redshifts to these galaxies is feasible within short exposure times (typically one hour of integration on the 2.5m SDSS telescope). The methodology used here has been first explored and experimented by Davis et al. (2003), Adelberger et al. 2004, Drinkwater et al. 2010. Adelberger et al. (2004) derived different colour selections for faint galaxies (with $23 < R < 25.5$) at redshifts $1 < z < 3$ based on the Great Observatories Origins Deep Survey data (GOODS, see Dickinson et al. 2003). Drinkwater et al. (2010) selected ELGs using UV photometry from the Medium Imaging Survey of the GAlaxy EVolution EXplorer (MIS-GALEX, see Martin et al. 2005) data combined with SDSS, to obtain a final density of 238 ELGs per square degree with $0.2 < z < 0.8$ over ~ 800 square degrees.

Our motivation is to probe much wider surveys than GOODS or GALEX (ultimately a few thousands square degrees) and to concentrate on intrinsically more luminous galaxies (typically with $g < 23.5$) with a redshift distribution extending to redshift 1.7.

The selection criteria studied in this work are designed for a ground-based survey and more specifically for the SDSS telescope, a 2.5m telescope located at Apache Point Observatory (New Mexico, USA), which has a wide field of view to carry out LSS studies (Gunn et al., 2006). The current BOSS spectrographs cover a wavelength range of $3600 - 10200$ Angstrom. Its spectral resolution, defined by the wavelength divided by the resolution element, varies from $R \sim 1,600$ at $3,600$ Angstrom to $R \sim 3,000$ at $10,000$ Angstrom (Eisenstein et al., 2011). The highest redshift detectable with the [OII] emission line doublet ($\lambda 3727, \lambda 3729$) is thus $z_{\rm max} = 1.7$.

To select ELGs in the redshift range $[0.6, 1.7]$ we have explored two different selection schemes: first using *u,g,r* photometry and secondly using *g,r,i* photometry.

2.4.1 Photometric data properties: SDSS, CFHT-LS and COSMOS

The photometric SDSS survey, delivered under the data release 8 (DR8, Aihara et al. 2011), covers 14,555 square degrees in the 5 photometric bands *u, g, r, i, z*. It is the largest volume multi-color extragalactic photometric survey available today. The 3σ magnitude depths are: $u = 22.0$, $g = 22.2$, $r = 22.2$, $i = 21.3$; see Fukugita et al. (1996) for the description of the filters and Gunn et al. (1998) for the characteristics of the camera. The magnitudes we use are corrected from galactic extinction.

The Canada France Hawaii Telescope Legacy Survey[6] (hereafter CFHTLS) covers ~ 155 deg^2 in the *u,g,r,i,z* bands. The transmission curves of the filters differ slightly[7] from SDSS. The data and cataloging methods are described in the T0006 release document[8]. The 3σ magnitude depths are: $u = 25.3$, $g = 25.5$, $r = 24.8$, $i = 24.5$. The CFHT-LS photometry is ten times (in r and i) to thirty times (in u) deeper than SDSS DR8, however the CFHTLS covers a much smaller field of view than SDSS DR8. The magnitudes we use are corrected from galactic extinction. The CFHT-LS wide photometric redshift catalogs are presented in Ilbert et al. (2006), and Coupon et al. (2009) ; the photometric redshift accuracy is estimated to be $\sigma_z < 0.04(1 + z)$ for $g \leq 22.5$. This photometric

[6]http://www.cfht.hawaii.edu/Science/CFHLS/
[7]http://cadcwww.dao.nrc.ca/megapipe/docs/filters.html
[8]http://terapix.iap.fr/cplt/T0006-doc.pdf

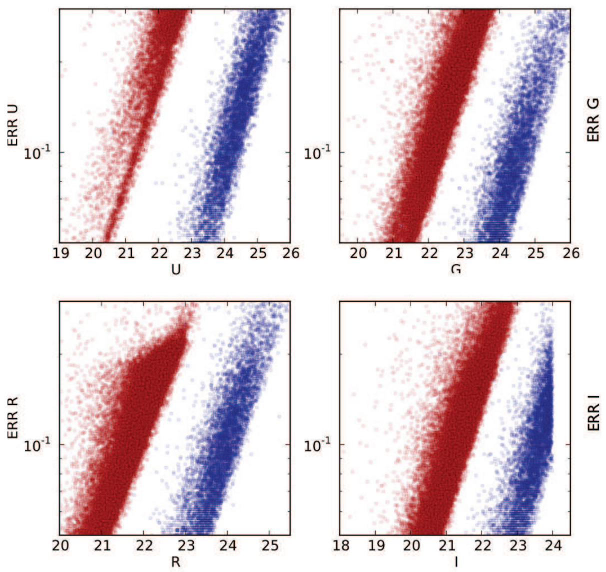


Figure 2.4: The four bands *ugri* and their precision are illustrated; in red for SDSS photometry (ModelMag and its error from SDSS pipeline); in blue for CFHTLS photometry (magnitude from the photo-Z catalog). The u band quality is limiting the precision of the colour selection on SDSS photometry. Note that the photometric redshift CFHTLS catalog is cut at $i = 24$, and the SDSS data is R-selected with $err_R \leq 0.2$.

redshift catalog is cut at $i = 24$, beyond which photometric redshifts are highly unreliable. Fig. 2.4 displays the relative depth between SDSS and CFHT-LS wide surveys in the u,g,r,i-bands.

COSMOS is a deep 2 deg^2 survey that has been observed at more than 30 different wavelengths (Scoville et al., 2007b). The COSMOS photometric catalog is described in Capak et al. (2007) and the photometric redshifts in Ilbert et al. (2009). The COSMOS Mock Catalog, (hereafter CMC; see[9]) is a simulated spectro-photometric catalog based on the COSMOS photometric catalog and its photometric redshift catalog (Jouvel et al. 2009, Zoubian et al. 2012, in preparation). The magnitudes of an object in any filter can be computed using the photometric redshift best-fit spectral templates.

The limiting magnitudes of the CMC in the each band are the same as in the real COSMOS catalog (detection at 5σ in a 3" diameter aperture): $u < 26.4$, $g < 27$, $r < 26.8$, $i < 26.2$. For magnitudes in the range $14 < m < 26$ in the g,r,i bands from the Subaru telescope and in the u band from CFHTLS, the CMC contains about 280,000 galaxies in 2 deg^2 to COSMOS depth. The mock catalog also contains a simulated spectrum for each galaxy. These simulated spectra are generated with the templates used to fit COSMOS photometric redshifts. Emission lines are empirically added using Kennicutt calibration laws (Kennicutt, 1998; Ilbert et al., 2009), and have been calibrated using zCOSMOS (Lilly et al., 2009) as described in Zoubian et al. 2012, in preparation. The strength of [OII] emission lines was confirmed using DEEP2 and VVDS DEEP luminosity functions (Le Fèvre et al., 2005b; Zhu et al., 2009). Finally a host galaxy extinction law is applied to each spectrum. Predicted observed magnitudes take into account the presence of emission lines.

2.4.2 Color selections

Based on the COSMOS and CFHT-LS photometric redshifts, we explore two simple colour selection functions using the *ugr* and *gri* bands. Fig. 2.5 shows the targets available in the *ugr* and *gri* colour planes. We construct a bright and a faint sample based on the photometric depths of SDSS and CFHT-LS.

[9]http://lamwws.oamp.fr/cosmowiki/RealisticSpectroPhotCat

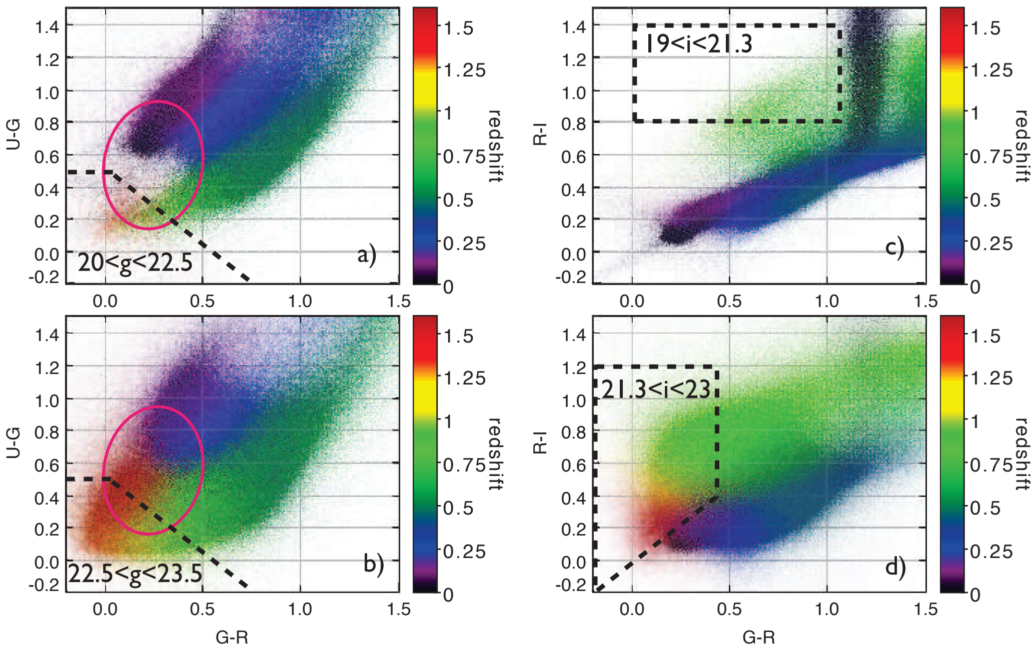


Figure 2.5: Color-color diagrams tinted according to CFHT-LS photometric redshift. Colour selections are indicated with the dashed boxes. Quasars, when overlapping with the colour selection, are located in the pink ellipse. The stellar sequence appears in black. **a)** The bright *ugr* colour selection: u-g vs g-r for $20 < g < 22.5$. In this area the overall density of targets is low. **b)** The faint *ugr* colour selection where the magnitude range is $22.5 < g < 23.5$. The density of available targets increases and the stellar sequence is diminished. **c)** The bright *gri* colour selection: r-i vs g-r diagram for $19 < i < 21.3$. The high-redshift targets are not as visible as in the *ugr* plot. The target density is also small. **d)** The faint *gri* colour selection where the magnitude range is $21.3 < i < 23$. The target density increases and higher redshift targets appear in the blue end of the colour plot while the stellar locus fades. Using SDSS photometry one would obtain similar plots to **a)** and **c)** in terms of target density. The indetermination on the color at $g \sim 22$ is ~ 0.2 for SDSS photometry and ~ 0.02 for the CFHT-LS. At $g \sim 23$, it is ~ 1 for SDSS photometry and ~ 0.1 for the CFHT-LS.

2.4.2.1 *ugr* selection

The *ugr* colour selection is defined by $-1 < u - r < 0.5$ and $-1 < g - r < 1$ that selects galaxies at $z \geq 0.6$ and ensures that these galaxies are strongly star-forming ($u-r$ cut). The cut $-1 < u-g < 0.5$ removes all low-redshift galaxies ($z < 0.3$). Finally the magnitude range is $20 < g < 22.5$ and $g <23.5$ for the bright and the faint samples, resp. Fig. 2.5 a) and b).

2.4.2.2 *gri* selection

The bright *gri* colour selection is defined by the range $19 < i < 21.3$. We select blue galaxies at $z \sim 0.8$ with $0.8 < r - i < 1.4$ and $-0.2 < g - r < 1.1$ (Fig. 2.5 c). In the faint range $21.3 < i < 23$, we tilt the selection to select higher redshifts with $-0.4 < g - r < 0.4$, $-0.2 < r - i < 1.2$ and $g - r < r - i$ (Fig. 2.5 d).

Table 2.2: Properties of the colour-selected samples. ‘b’ stands for bright ($20 < g < 22.5$ for *ugr* or $19 < i < 21.3$ for *gri*) and ‘f’ for faint ($22.5 < g < 23.5$ for *ugr* or $21.3 < i < 23$ for *gri*). We indicate the mean and standard deviation of the redshift distribution for each set. For the CMC sample, [OII] fluxes are expressed in units of $10^{-17}\mathrm{erg\,cm^{-2}\,s^{-1}}$; the median, the first and third quartiles of the flux distribution are indicated. The *ugr* tends to select the highest [OII] fluxes.

selection			$\#\mathrm{deg}^{-2}$	$\bar{u}$	$\bar{g}$	$\bar{r}$	$\bar{i}$	$\bar{z}$	σ_z	$\bar{f}_{[\mathrm{OII}]}$	$Q^1_{f_{[\mathrm{OII}]}}$	$Q^3_{f_{[\mathrm{OII}]}}$
CMC	*ugr*	b	130.0	21.98	21.87	21.69	-	1.25	0.53	61.74	46.47	88.39
		f	1450.8	23.27	23.18	22.98	-	1.19	0.38	16.60	13.06	22.26
	gri	b	257.2	-	22.69	21.87	20.93	0.80	0.21	13.85	8.65	22.21
		f	2170.5	-	23.34	23.09	22.55	0.93	0.31	10.23	6.83	15.99
CFHT-W1	*ugr*	b	193.3	21.95	21.8	21.7	-	1.28	0.38			
		f	1766.8	23.37	23.19	23.07	-	1.29	0.31			
	gri	b	361.4	-	22.62	21.8	20.82	0.81	0.11			
		f	3317.5	-	23.34	23.11	22.55	1.03	0.35			
CFHT-W3	*ugr*	b	232.2	21.89	21.76	21.69	-	1.27	0.37			
		f	1679.1	23.36	23.18	23.06	-	1.28	0.31			
	gri	b	391.6	-	22.62	21.78	20.8	0.82	0.1			
		f	3334.2	-	23.34	23.11	22.54	1.03	0.33			
SDSS	*ugr*	b	166.96	21.76	21.77	21.52	-					
	gri	b	204.96	-	22.57	21.75	20.76					

2.4.3 Predicted properties of the selected samples

The *ugr* colour selection avoids the stellar sequence, but not the quasar sequence. Hence, the contamination of the *ugr* selection by point-source objects is primarily due to quasars; see Fig. 2.5 a) and b). The resulting photometric-redshift distribution as derived from the CFHT-LS photometric redshift catalog has a wide span in redshift, covering $0.6 < z < 2$ as shown in Fig. 2.6. The distribution is centered at $z = 1.3$ for the bright and the faint sample with a scatter of 0.3 (see Table 2.2). The expected [OII] fluxes are computed from the CMC catalog and are shown in Fig. 2.6. For 90 percent of galaxies in the faint sample, the predicted flux is above 10.6×10^{-17}erg cm^{-2} s^{-1}. The bright sample galaxies show strong emission lines.

The *gri* selection avoids both the stellar sequence and the quasar sequence; see Fig. 2.5 c) and d). Thus the contamination from point-sources should be minimal. Fig. 2.6 shows the photometric redshift distribution of the *gri* selection applied to CFHT photometry. The redshifts are centered at $z = 0.8$ for the bright and 1.0 for the faint sample (see Table 2.2). The expected [OII] flux, computed with the CMC catalog, is shown in Fig. 2.6. Emissions are weaker than for the *ugr* selection as expected.

The different selections shown in Fig. 2.5 and Fig 2.6 are summarized in Table 2.2, which contains the number densities available, mean magnitudes, mean redshifts, and mean [OII] fluxes (when available) of the different samples considered. We have lower densities in the CMC than in the CFHT-LS catalog. This is probably due to cosmic variance as the CMC only covers 2 deg^2.

The SDSS colour-selected samples are complete for the bright samples at $g < 22.5$ and $i < 21.3$, not for the faint samples. The CFHTLS-selected samples are complete for both the bright and faint samples; see Fig.2.7, where the total cumulative number counts (solid line) of the *ugr* and *gri* colour-selected samples are plotted as a function of g and i bands respectively. On the bright end of this Figure, although both photometry are complete at the bright limit, we note a discrepancy between the total amount of target selected on CFHT and SDSS that implies selections on CFHT are denser than on SDSS (difference between the red and blue solid lines). This is due to the transposition of the color selection from one photometric system to the other. In fact, we select targets on SDSS with a transposed criterion from CFHT using the calibrations by Regnault et al. (2009). The transposed criterion is as tight as the original. But as the errors on the magnitude are larger in the SDSS system, their colour distributions are more spread. Therefore the SDSS selection is a little less dense than the CFHT selection.

Targeting the bright range is limited by galaxy density, in the best case one can reach 300 targets deg^{-2} and it contains point-sources (stars and quasars) and low-redshift galaxies. In the faint range, the target density is ten times greater, but the exposure time necessary to assign a reliable redshift will be much longer (one magnitude deeper for a continuum-based redshift roughly corresponds to an exposure five times longer). The stellar, quasars and low-redshift contamination is smaller in the faint range. Fig. 2.6 shows the distributions in redshift and in [OII] flux we expect given a magnitude range and a colour criterion within the framework of the CMC simulation. The main trend is that the *ugr* selection identifies strong [OII] emitters out to $z \sim 2$ where the *gri* peaks at redshift 1 and extends to 1.4 with weaker [OII] emitters.

We also used a criterion to split targets in terms of compact and extended sources, which is illustrated in Fig. 2.7. For CFHT-LS we have used the half-light radius (r_2 value, to be compared to the r_2^{limit} value which defines the maximal size of the PSF at the location of the object considered - see Coupon et al 2009 and CFHT-LS T0006 release document) to divide the sample into compact and extended objects. For SDSS we used the "TYPE" flag, which separates compact (TYPE=6) from extended objects (TYPE=3). For the *ugr* colour selection, the number counts are dominated by compact blue objects (quasars) at $g \leq 22.2$. At $g \geq 22.2$ the counts are dominated by extended ELGs. For comparison we show in Fig. 2.7 the cumulative counts of the XDQSO catalog from Bovy et al. (2011) who identified quasars in the SDSS limited to $g < 21.5$. We notice an excellent match with the bright (compact) *ugr* colour-selected objects. For the *gri* colour selection, there is a low contamination by

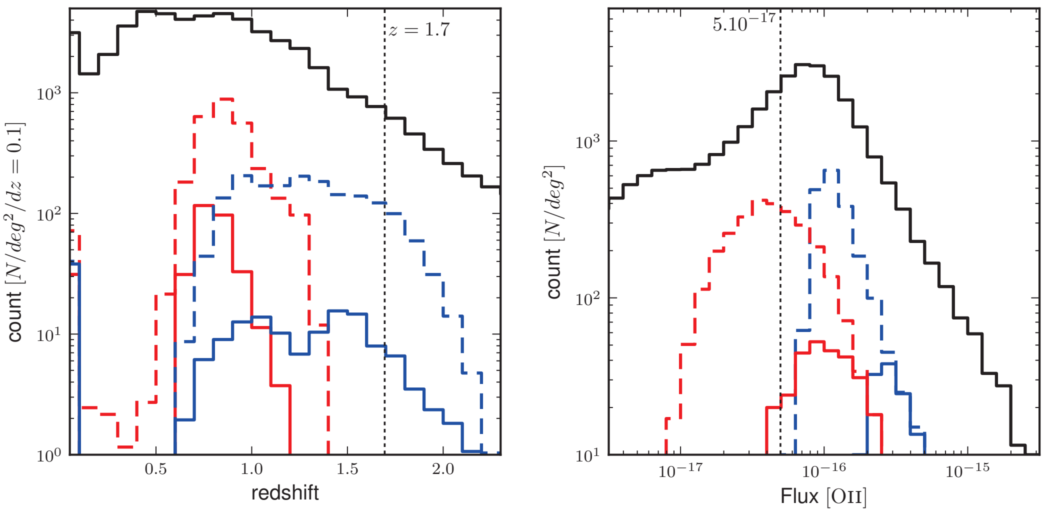


Figure 2.6: Redshift and line flux distribution for various color selections. The blue solid line is the *ugr* selection limited to $20 < g < 22.5$. The blue dashed line is the *ugr* selection limited to $22.5 < g < 23.5$. The red solid line is the *gri* selection limited to $19 < i < 21.3$. The red dashed line is the *gri* selection limited to $21.3 < i < 23.0$. **Left.** Photometric redshift distribution from the CFHT-LS catalog. Black solid line: no colour selection, but limited to $i < 24$. The *ugr* selection is more spread in redshift than the *gri* selection. The vertical dashed line indicates the upper limit of $z = 1.7$ that corresponds to [OII] emission lines at $\sim 1\mu$m. **Right.** Expected [OII] flux distribution from the CMC simulation. The black solid line is the complete CMC catalog. The *ugr* selection identifies stronger line emitters than the *gri* selection scheme. The vertical dashed line indicates the expected mean sensitivity of BOSS in 1h exposure, $5 \times 10^{-17}\mathrm{erg\,s^{-1}\,cm^{-2}}$.

compact objects because the colour box does not overlap with either the stellar or the quasar sequence.

2.5 Observations

To test the reliability of both the bright *ugr* ($g < 22.5$) and the bright *gri* ($i < 21.3$) colour selections, we have conducted a set of dedicated observations, as part of the "Emission Line Galaxy SDSS-III/BOSS ancillary program". The observations were conducted between Autumn 2010 and Spring 2011 using the SDSS telescope with the BOSS spectrograph at Apache Point Observatory. A total of ~2,000 spectra, observed 4 times 15 minutes, were taken in different fields: namely, in the Stripe 82 (using single epoch SDSS photometry for colour selection) and in the CFHT-LS W1, W3 and W4 wide fields (using CFHT-LS photometry). It is not possible to define the effective area of the observations because these were done with spare fibers from the BOSS program that covers this area. These observations were not dedicated to this target class, the BOSS masks/plates were filled with the ELG targets. This data set was released in the SDSS-III Data Release 9[10].

2.5.1 Description of SDSS-III/BOSS spectra

We used the SDSS photometric catalog (Aihara et al., 2011) to select 313 objects according to their *ugr* colours located in the Stripe-82 and 899 objects selected according to their *gri* colours in the CFHT-LS W3 field. In addition we used the CFHT-LS photometry to select 878 *ugr* targets in the CFHT-LS W1 field, and 391 *gri* targets in the CFHT-LS W3 field for observation. The spectra are available in SDSS Data Release 9 and flagged 'ELG'.

[10] http://dr9.sdss3.org/

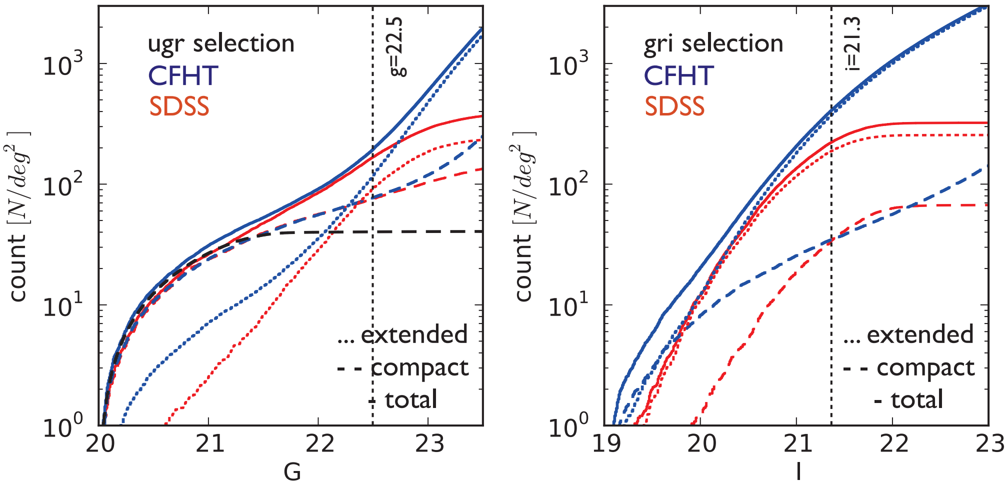


Figure 2.7: Cumulative number counts per square degree of both colour selections using SDSS photometry (in red) and CFHTLS photometry (in blue). The solid line is the whole selection, while the dashed line is for compact (star-like) objects and the dotted line represents extended objects. On the left is the *ugr* selection and on the right is the *gri* selection. The black dashed line represents objects from the XDQSO catalog that have a probability of being a quasar that is greater than 90 percent (Bovy et al., 2011) (the quasar identifications are limited to $g = 21.5$).

All of these spectra were manually inspected to confirm or correct the redshifts produced by two different pipelines (zCode and its modified version that we used to fit the [OII] emission line doublet). As the BOSS pipeline redshift measurement is designed to fit LRG continuum some ELGs with no continuum were assigned wrong redshifts. To classify the observed objects, we have defined seven sub-categories :

Objects with secure redshifts

- 'ELG', Emission-line galaxy (redshift determined with multiple emission lines). Usually these spectra have a weak 'blue' continuum and lack a 'red' continuum. Empirically, using platefit vimos pipeline output, this class corresponds to a spectrum with more than two emission lines with observed equivalent widths $EW \leq -6$ Angstrom ; see examples in Appendix 2.9.

- 'RG', Red Galaxy with continuum in the red part of its spectrum, allowing a secure redshift measurement through multiple absorption lines (*e.g.* Ca K&H, Balmer lines) and the 4000 Angstrom break. Some of these objects have also weak emission (E+A galaxies). Empirically their spectra have a mean $D_n(4000)$ of 1.3 ; where $D_n(4000)$ is the ratio of the continuum level after the break and before the break. These galaxies typically have $i \sim 20$, which is fainter than the CMASS targeted by BOSS.

- 'QSO', Quasars, which are identified through multiple broad lines. Examples are given in Fig. 2.19.

- Stars.

Objects with unreliable redshifts

- 'Single emission line' : the spectra contain only a single emission line which cannot allow a unique redshift determination. For this population, the CFHT T0006 photometric redshifts are compared to the [OII] redshift (assuming the single emission line is [OII]) in Fig. 2.8. The

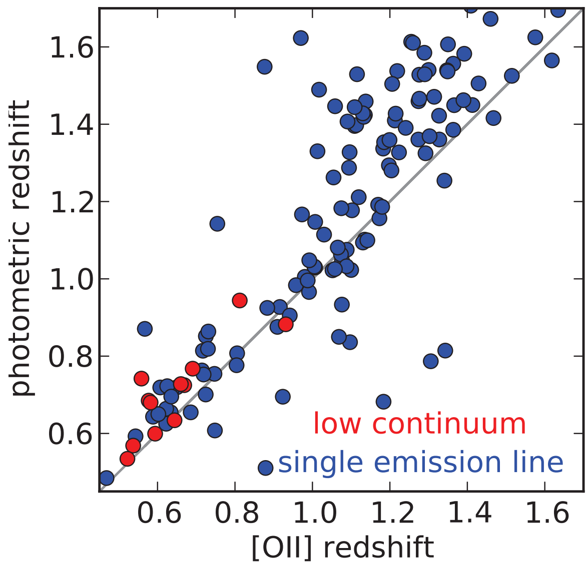


Figure 2.8: T0006 CFHT-LS photometric redshifts of single emission line and low continuum galaxies observed against [OII] redshift. A strong correlation is clearly evident. A slight systematic over-estimation of the photometric redshift is visible above redshift 1.2 (these photometric redshifts were calibrated below 1.2).

two estimates agree very well : 77.7 percent have $(z_{spec} - z_{phot})/(1 + z_{spec}) < 0.1$ for the *gri* selection and 62.7 percent for the *ugr* selection. These galaxies with uncertain redshift tend to have slightly fainter magnitudes with a mean CFHT g magnitude at 22.6 and a scatter of 0.6, whereas for the whole ELGs is 22.4 with a scatter of 0.4.

- 'Low continuum' spectra that show a 4000 Angstrom break too weak for a secure redshift estimate. The agreement between photometric and spectroscopic redshift estimation is excellent : 84.6 percent within 10 percent errors; see Fig. 2.8.

- 'Bad data', the spectrum is either featureless, extremely noisy or both. No redshift could be measured.

The detailed physical properties of the ELGs are discussed in section 2.6 and a number of representative spectra are displayed in Appendix 2.9.

2.5.2 Redshift Identification

The results of the observations are summarized by categories in Table 2.3.

For the targets selected using SDSS photometry and with the *ugr* selection : 32 percent are ELGs at a redshift $z > 0.6$ (100 spectra). The low-redshift ELGs represent 32 percent of the observed targets (101 spectra).The other categories are : 65 'bad data' (20 percent), 30 quasars (10 percent), 10 stars (3.5 percent), and 7 red galaxies with $z < 0.6$ (2.5 percent). With the *gri*-selection, 57 percent of the targets are at $z > 0.6$. However, still 21 percent of the spectra fall into the bad data class.

Using CFHTLS photometry 46 percent ofobserved targets are ELGs at $z > 0.6$ and 14 percent are quasars with the *ugr* selection. With the *gri*-selection, 73 percent are galaxies at $z > 0.6$, five-sixths of which are ELG.

For both selections, targeting with CFHTLS is more efficient than with SDSS. The complete classification of observed targets is in Table 2.3. The redshift distribution of the observed objects is compared to the distributions from the BOSS and WiggleZ current BAO experiments in Fig. 2.9. The

Table 2.3: Observed objects split in categories. The per cents are the ratio between the number of objects in a category and the total amount of observed targets. The difference in efficiency between the use of the CFHT-LS and the SDSS photometry to target ELGs is large.

	gri selection				*ugr* selection			
	SDSS selection		*CFHTLS* selection		*SDSS* selection		*CFHTLS* selection	
Type	Number	%	Number	%	Number	%	Number	%
ELG($z > 0.6$)	450	50	240	61	100	32	402	46
ELG($z < 0.6$)	60	7	3	1	101	32	84	9
RG($z > 0.6$)	73	8	46	12	0	0	0	0
RG($z < 0.6$)	30	3	0	0	7	3	0	0
single emission line	36	4	12	3	0	0	102	12
low continuum	13	1	1	0	0	0	0	0
QSO	8	1	5	1	30	10	126	14
stars	44	5	12	3	10	3	6	1
bad data	185	21	72	18	65	20	158	18
total	899	100	391	100	313	100	878	100

Figure shows that *ugr* and *gri* target selections enable a BAO study at higher redshifts. With a joint selection, we can reach the requirements described in Table 2.1 to detect BAO feature to redshift 1.

2.5.3 Comparison of measured ELGs with the CMC forecasts

To investigate the expected purity of ELG galaxies samples, we created mock catalogs covering redshifts between 0.6 and 1.7. Continuum spectra of ELGs were generated from the Cosmos Mock Catalog and emission lines were added according to the modeling described in Jouvel et al. (2009). Two simulated galaxy catalogs were built, one for each colour selection function (*ugr* and *gri*). Each synthetic spectrum was affected by sky and photon noise as if observed by BOSS spectrographs, by using the SPECSIM1D software. We simulated a set of four exposures of 900 seconds each. The resulting simulated spectra were then analyzed by the ZCODE pipeline (Cannon et al., 2006) to extract the spectroscopic redshift. As our targets are mainly emission line galaxies, we only use the redshift estimate based on fitting discrete emission line templates in Fourier space over all z.

We address the flux measurement of emission lines. This exercise was conducted using the PLATEFIT VIMOS software developed by Lamareille et al. (2009). This software is based on the PLATEFIT software that was developed to analyze SDSS spectra (Tremonti et al., 2004a; Brinchmann et al., 2004). The PLATEFIT VIMOS software was developed to measure the flux of all emission lines after removing the stellar continuum and absorption lines from lower resolution and lower signal-to-noise ratio spectra (Lamareille et al., 2006). The stellar component of each spectrum is fit by a non-negative linear combination of 30 single stellar population templates with different ages (0.005, 0.025, 0.10, 0.29, 0.64, 0.90, 1.4, 2.5, 5 and 11 Gyr) and metallicities (0.2, 1 and 2.5 $Z_\odot$). These templates have been derived using the Bruzual and Charlot (2003) libraries and have been resampled to the velocity dispersion of VVDS spectra. The dust attenuation in the stellar population model is left as a free parameter. Foreground dust attenuation from the Milky Way has been corrected using the Schlegel et al. (1998) maps.

After removal of the stellar component, the emission lines are fit as a single nebular spectrum made of a sum of Gaussians at specified wavelengths. All emission lines are set to have the same width, with the exception of the [OII] λ3727 line which is a doublet of two lines at 3726 and 3729 Angstrom that appear broadened compared to other single lines. Detected emission lines may also be removed from the original spectrum in order to obtain the observed stellar spectrum and measure indices, as well as emission-line equivalent widths. The underlying continuum is obtained by smoothing the stellar spectrum. Equivalent widths are then measured via direct integration over a 5σ bandpass of

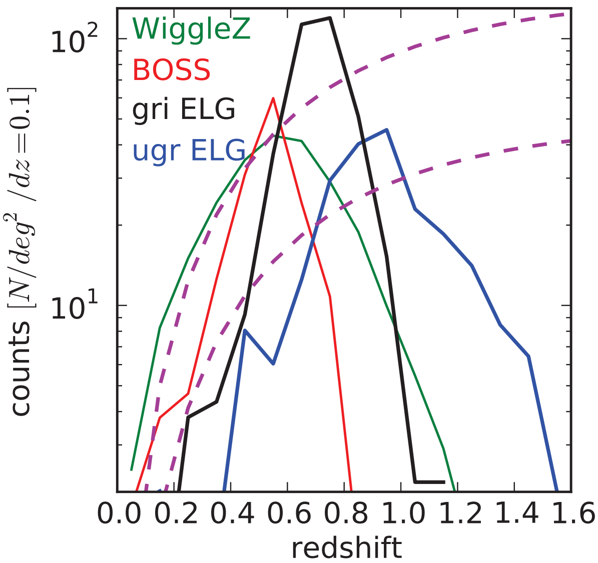


Figure 2.9: Observed redshift distribution for the *ugr* ELGs (blue), the *gri* ELGs (black) compared to the distribution of galaxies from BOSS (red) and WiggleZ (green). Magenta lines represent constant density of galaxies at 1 and 3 $\times 10^{-4}\ h^3\ \mathrm{Mpc}^{-3}$, it constitutes our density goals.

the emission-line Gaussian model divided by the underlying continuum. Then emission lines fluxes are measured for each simulated spectra using the extracted redshift from zCode and the true redshift for cross-checks.

We consider that a redshift has been successfully measured if $\Delta z/(1+z) < 0.001$. We believe that this threshold could be lowered to 10^{-4} in the future by using a more advanced redshift solver. Using the current pipeline, we can distinguish these two regimes.

The first regime is the redshift range $z < 1.0$. Many emission lines ([OII], Hβ, [OIII]) are present in the SDSS spectrum. For $g < 23.5$, 91 percent of the redshift are measured sucessfully. Among the remaining 9 percent, catastrophic failures represent 3.5 percent (the pipeline outputs a redshift between 0 and 1.6 with $\Delta z/(1+z) > 0.01$). Inaccurate redshifts represent 3.9 percent (the pipeline outputs a redshift between 0 and 1.6 with $0.001 < \Delta z/(1+z) < 0.01$) and 1.5 percent are not found by the pipeline ($z = -9$ is output).

The second regime is the redshift range $1.0 \leq z < 1.7$: the redshift determination hinges on the identification of the [OII] doublet. For $g < 23.5$, 66.8 percent of the redshifts are measured sucessfully. 19.1 percent are catastrophic failures and 14.1 percent are inaccurate redshifts. Work is ongoing to improve the redshift measurement efficiency at $z > 1$. In the second regime, the minimum [OII] flux required to compute a reliable redshift depends on the redshift / wavelength, because of the strong OH sky lines in the spectrum. We infer from the observed spectra that to measure a reliable redshift, we require a 5σ detection of the [OII] lines, which means a (blended or not) detection of two peaks in the emission line separated by $2(1+z)$. The detection significance is defined from the 1d spectrum. From the data the faintest 5σ detections are made with a flux of $4 \times 10^{-17}\mathrm{erg\,s^{-1}\,cm^{-2}}$ and the brightest 5σ detections need a flux of $2 \times 10^{-16}\mathrm{erg\,s^{-1}\,cm^{-2}}$ to be on top of sky lines. The simulation shows the same thresholds; see Fig. 2.10. The simulation confirms the detection limit we observe. The bottom plot of Fig. 2.10 raises the issue that the sky time variation has a non-negligible impact on the detection limit of the [OII] emission doublet for redshifts $z > 1.1$. Though this ELG sample is too small to address this issue. In fact the sample was observed during ten different nights and the number of ELG with $z > 1.1$ is less than 60. It is thus not possible to derive a robust trend comparing the [OII] detections to the sky value of each observation. Handling this issue would require a sample

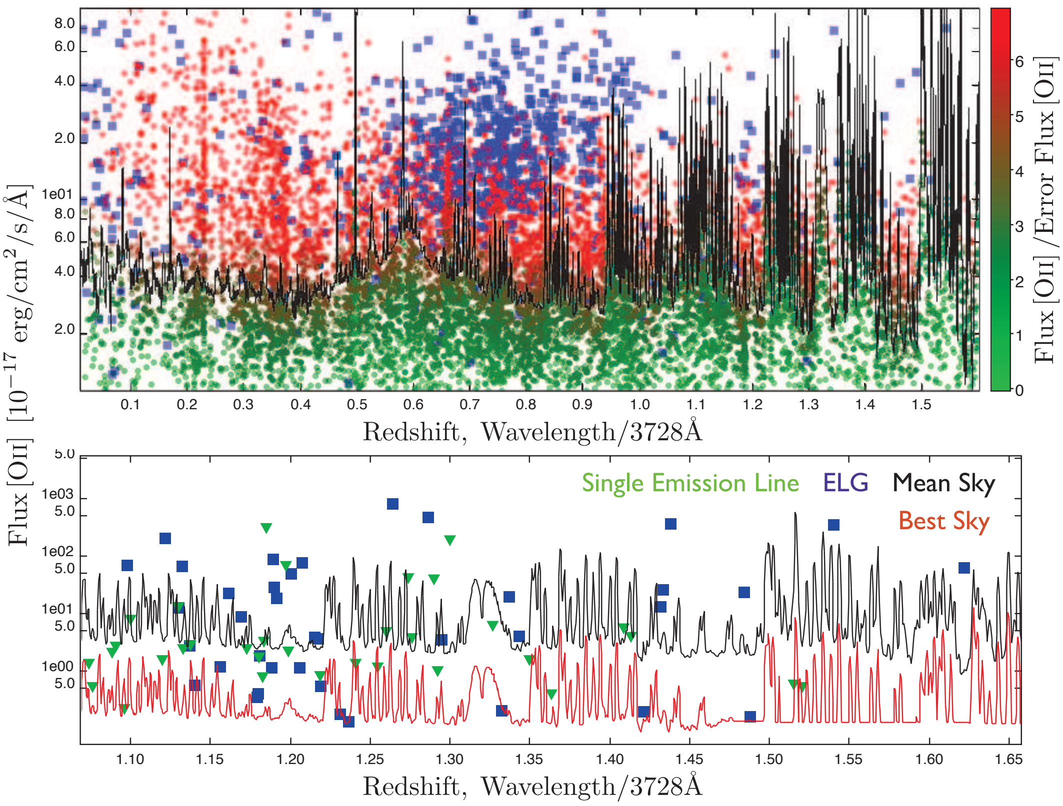


Figure 2.10: Top : Flux detected in the [OII] doublet (green or red filled circle for a simulated spectrum and blue filled square for a real galaxy observation) versus redshift. The simulation is coloured according to the measured line flux divided by the error. The black solid line represents a typical sky as observed by BOSS spectrograph; the 5σ detections of [OII] follow the sky flux. The [OII] detections above 5σ (red circles) follow the tendency of the sky fluctuations : a 5σ detection in a zone with a high sky requires a larger emitted flux. The data are scaled to match a 1 hour exposure on the 2.5m SDSS telescope. The redshifts below the sky flux are assigned using other emission lines or the information in the continuum. **Bottom:** : An expansion of the top panel using $z > 1$ detections. Only measured galaxies are plotted upon the mean sky (black) and the best sky (red) of APO.

Table 2.4: Main properties of the observed galaxies. Fluxes are in unit of $10^{-17}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$. Equivalent Widths 'EW' are in Angstrom.

	gri-selected ELG				*ugr*-selected ELG			
	CFHTLS		SDSS		CFHTLS		SDSS	
	mean	σ	mean	σ	mean	σ	mean	σ
$\mathrm{EW}_{[\mathrm{O\,II}]}$	-14.86	9.01	-16.75	10.13	-50.58	27.24	-30.75	23.04
$\mathrm{Flux}_{[\mathrm{O\,II}]}$	16.85	9.65	18.58	10.37	30.36	30.1	24.23	39.27
EW_{H_β}	-10.28	10.8	-10.72	8.65	-24.27	22.88	-17.18	19.34
Flux_{H_β}	15.44	8.6	14.63	7.72	12.97	15.16	12.57	23.91
$\mathrm{EW}_{[\mathrm{O\,III}]}$	-10.09	10.98	-11.33	10.76	-65.3	91.56	-16.89	30.49
$\mathrm{Flux}_{[\mathrm{O\,III}]}$	17.74	20.15	17.43	21.59	35.13	53.49	13.39	37.79
$12+\log OH$	8.94	0.20	8.92	0.19	8.69	0.21	8.69	0.25
$\mathrm{logSFR}_{[\mathrm{O\,II}]}$	0.97	0.35	0.92	0.45	0.96	1.24	0.76	0.84
$\log(M^*/M_\odot)$	10.85	0.3	10.23	6.87	9.33	0.80	-	-

of ∼500 redshifts in $1.1 < z < 1.6$ observed many times over many nights. With such a sample in hand, we could quantify exactly how to optimize the observational strategy.

2.6 Physical properties of ELGs

All *ugr* and *gri* ELG spectra were analyzed with two different software packages: the PlateFit VIMOS (Lamareille et al., 2009) and the Portsmouth Spectroscopic Pipeline (Thomas et al., 2012). In this section we discuss the following physical properties of the observed ELGs: redshift, star forming rate (SFR), stellar mass, metallicity and classification of the ELG type (Seyfert 2, LINERs, SFG, composite). Observations a larger samples of ELGs are planned to estimate how these quantities vary over time and with their environment, and also to estimate how the clustering depends on these physical quantities. It is key to replace future BAO tracers in the galaxy formation history. With the current sample, we draw simple trends using means and standard deviation of the observed quantities, and we place the ELGs in the galaxy classification made by Lamareille (2010); Marocco et al. (2011).

2.6.1 Main Properties

The main properties of the ELGs are shown in the Table 2.4. The star formation rate was computed using the equation 18 of Argence and Lamareille (2009). The stellar mass was estimated using the CFHTLS *ugriz* photometry. (The errors on the stellar mass using only SDSS photometry were too large to be meaningful, thus the empty cells in the table). The metallicity is estimated using the calibration by Tremonti et al. (2004b). The main trends are :

- The *gri*-selected galaxies of CFHTLS have a higher stellar mass than the *ugr*-selected galaxies.
- The *ugr* selects stronger star-forming galaxies than *gri* (due to the *u*-band selection). There is a factor of two variations in the strength of the measured oxygen lines.
- The *ugr* selects galaxies that have $12+\log[OH] \in [8,9]$ whereas *gri* focuses slightly more on the higher $12+\log[OH] \approx 9$.
- the SFR appears to be independent of the color selection schemes.

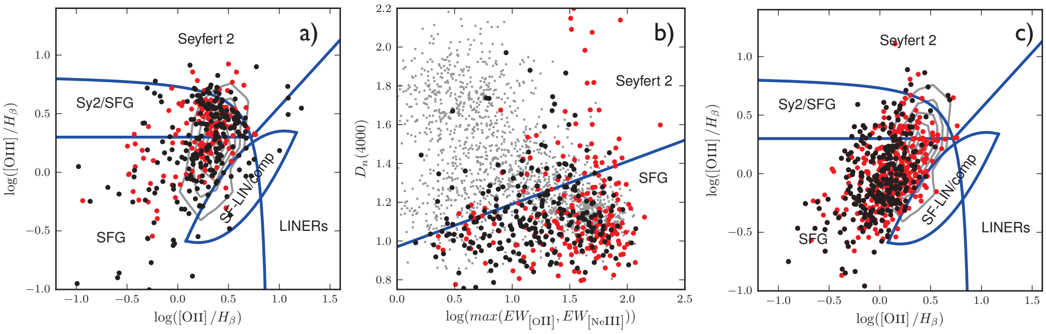


Figure 2.11: Black dots stands for CFHTLS selected ELGs, red dots for SDSS selected ELGs and grey contours or grey pixels for zCOSMOS survey galaxies. **a)** $\log([\mathrm{OIII}]/H_\beta)$ vs. $\log([\mathrm{OII}]/H_\beta)$ for *ugr* ELGs. The *ugr* ELG sample is located in a similar area than zCOSMOS galaxies. **b)** $D_n(4000)$ vs. $\log(max(EW_{[\mathrm{OII}]}, EW_{[\mathrm{NeIII}]}))$ for *ugr* ELGs. Only galaxies from the 'Sy2/SFG' area from the plot a) are represented. The *ugr* ELG is thus mainly composed of 'SFG'. **c)** $\log([\mathrm{OIII}]/H_\beta)$ vs. $\log([\mathrm{OII}]/H_\beta)$ for *gri* ELGs. The *gri* ELG sample is located in the 'SFG' area. The characteristic error is between 0.1 and 0.2 for all the quantities shown.

2.6.2 Classification.

We use a recent classification (Lamareille, 2010; Marocco et al., 2011) for the ELG sample. The classification is made using $\log([\mathrm{OIII}]/H_\beta)$, $\log([\mathrm{OII}]/H_\beta)$, $D_n(4000)$, and $\log(max(EW_{[\mathrm{OII}]}, EW_{[\mathrm{NeIII}]}))$. We compare the ELG sample to zCOSMOS, as zCOSMOS has numerous star forming galaxies in the redshift range we are observing. Fig. 2.11 a) shows that the zCOSMOS and the *ugr* ELG samples are located in three of the five areas delimited by the classification: Seyfert 2 ('Sy2'), Star Forming Galaxies ('SFG'), and a third region where both mix ('Sy2/SFG'). There are a few LINERs and Composite in either sample. Fig. 2.11 b) separates the *ugr* galaxies in the 'Sy2/SFG' area into 'SFG' or 'Sy2', and shows that zCOSMOS galaxies from the 'Sy2/SFG' area are both 'Sy2' and 'SFG' where the *ugr* ELGs in the 'Sy2/SFG' area are mostly 'SFG'. The *gri* observed sample is located in the area of Star Forming Galaxies ('SFG'), whether one considers the one selected on CFHT or on SDSS. Finally, the ELG selected, *ugr* or *gri*, are both in the 'SFG' part of the classification.

2.7 Discussion

2.7.1 Redshift identification rates in *ugr* and *gri*

We summarize the redshift measurement efficiency of the *gri* and *ugr* colour-selected galaxies presented in this paper in Tables 2.3 and 2.5, and we compare the results with those of WiggleZ (Drinkwater et al., 2010), BOSS and VIPERS (the percentages about VIPERS are based on a preliminary subset including only $\sim$ 20 percent of the survey). The original VIPERS selection flag (J. Coupon and O. Ilbert private communication) is defined to have colours compatible with an object at $z > 0.5$ if it has ($r - i \geq 0.7$ and $u - g \geq 1.4$) or ($r - i \geq 0.5(u - g)$ and $u - g < 1.4$) (Guzzo et al., 2013). The efficiencies in the Table 2.5 show that a better photometry and thus more precise colours yield a better efficiency in terms of obtaining objects in the targeted redshift range. It also shows the colour selections proposed in this paper are competitive for building an LSS sample.

To determine the necessary precision on the photometry to stay at the efficiencies observed, we degrade the photometry of the observed ELGs, then reselect them and recompute the efficiencies. Using a photometry less precise than the CFHTLS by a factor of 2.5 in the errors (the ratio of the median values of the mag errors in bins of 0.1 in magnitude equals 2.5) does not significantly change neither the efficiency nor the redshift distribution implied by the colour selection. This change also

Table 2.5: Redshift efficiency in percent. The column 'spectroscopic redshift' quantifies the percentage of spectroscopic redshift obtained for the selection scheme specified in the column 'selection scheme'. The column 'object in z window' is the percentage of spectroscopic redshift located in the range targeted, it is the efficiency of the target selection. The redshift window for the ELG selections is $z > 0.6$. The last column gives the percentage of QSO observed.

selection scheme	spectroscopic redshift	object in z window	quasars
gri SDSS	80	62	1
gri CFHTLS	82	73	1
ugr SDSS	80	32	10
ugr CFHTLS	78	56	13
WiggleZ	60	35	-
BOSS	95	95	-
VIPERS	80	70	-

corresponds to loosening the colour criterion by 0.1 mag. For the *eBOSS* survey a photometry 2.5 times less precise than CFHTLS should be sufficient to maintain a high targeting efficiency (for comparison, at $g = 22$ SDSS has an SNR 10 times smaller than CFHTLS); Fig. 2.12 shows the smearing of the galaxy positions in the colour-colour plane for a degraded photometry.

2.7.2 Measurement of the [OII] doublet, single emission line spectra

For ground-based spectroscopic surveys observing ELGs with $1 < z < 1.7$, the only emission line remaining in the spectrum to assign the spectroscopic redshift is the [OII] doublet. For the redshift to be certain the doublet must be split (*i.e.*, we do not want the target to be classified as 'single emission line' ELG). Fig. 2.13 shows a subsample of the observed bright *ugr* ELGs where [OII] doublets are well resolved.

We can circumvent the 'single emission line' ELG issue (Fig. 2.8) by increasing the resolution of the spectrograph. This modification will enhance a better split of [OII], and will increase the room available to observe the doublet by rendering sky lines narrower. The sky acts as an observational window and prevents some narrow redshift ranges to be sampled by the spectrograph; see Fig. 2.10. An increase in the resolution means the the light is more spread, and if throughput is not maximized, the exposure time has to be increased to reconstruct properly the doublet above the mean sky level.

We performed a simulation of the [OII] emission line fit to quantify by which amount the resolution must be increased to have no 'single emission line' ELGs. We fit one or two Gaussians on a doublet with a total flux of 10^{-16}erg s^{-1} cm^{-2} (lowest 'single emission line' flux observed) contaminated by a noise of 3×10^{-17}erg s^{-1} cm^{-2} (typical BOSS dark sky). The χ^2 of the two fits are equal at low resolution and become disjoint in favor of the fit with 2 Gaussians for a resolution above 3000 at 7454.2 Angstrom (*i.e.* [OII] at redshift 1). Such an increase in resolution could help assigning proper redshifts to 'single emission line' ELGs.

2.7.3 How and why redshift went wrong

The main difference in redshift measurement efficiency between SDSS and CFHT-LS colour selection is mainly due to the difference in photometry depth. Using calibrations made by Regnault et al. (2009), it is possible to translate the colour selection criteria from CFHT-LS magnitudes to SDSS magnitudes. The colour difference can be as large as 1 magnitude as the SDSS magnitude cut is close to the detection limit of the SDSS survey; see Fig. 2.14 where SDSS *gri* colour-selected galaxies are represented with their CFHTLS magnitudes.

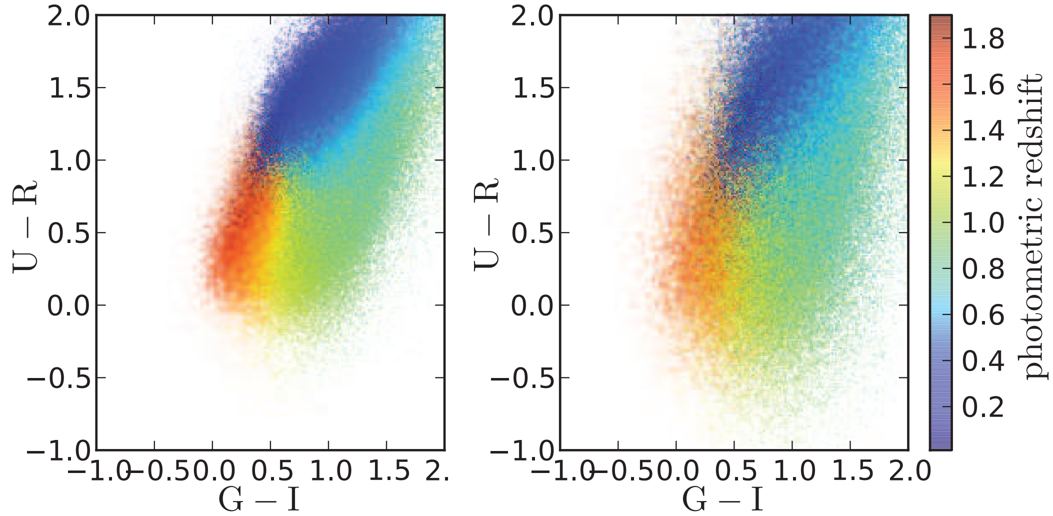


Figure 2.12: u-r vs. g-i colored according to the photometric redshift. On the left CFHT-LS photometry, on the right CFHT-LS photometry degraded by a factor 2.5. This comparison shows how the degradation of the photometry smears the clean separations between galaxy populations in redshift.

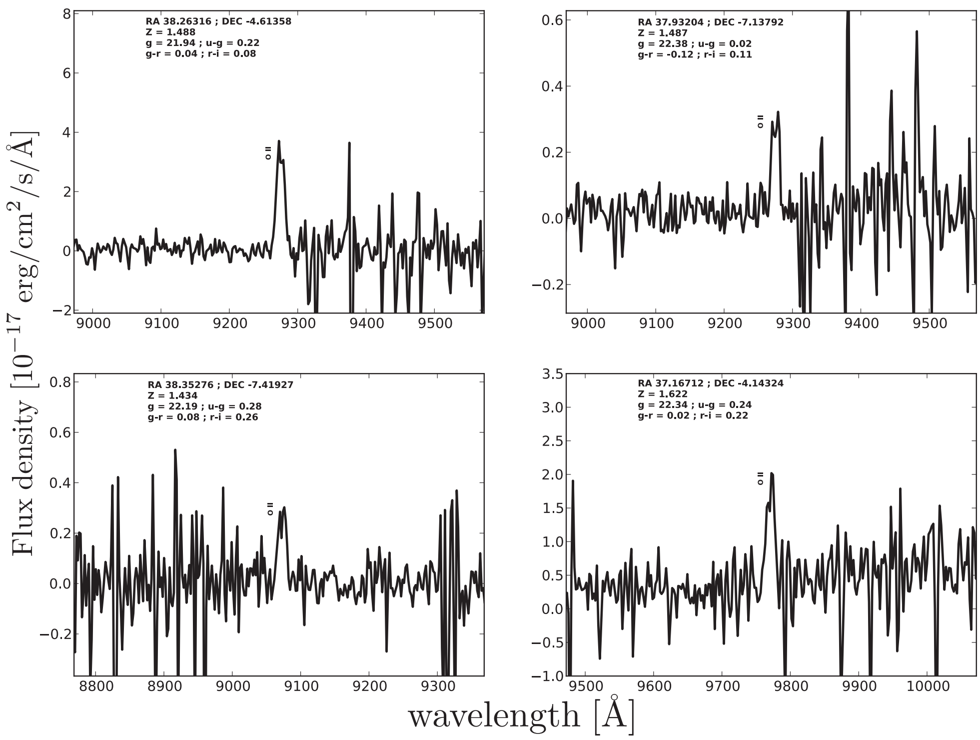


Figure 2.13: Observed Spectra zoomed on the [OII] doublet, extracted from the *ugr* ELG sample. The doublet is enough split (apart from a little blending) to assign the correct spectroscopic redshift without the help of other emission lines.

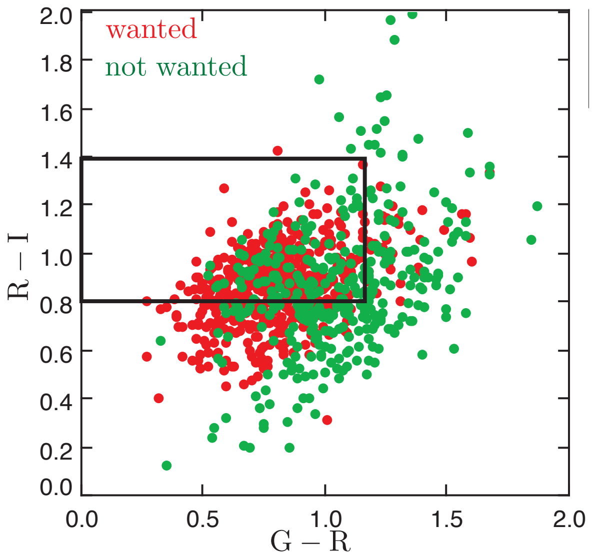


Figure 2.14: *gri* selection based on colours from SDSS (black box) represented by CFHTLS magnitudes. The scatter is quite large: about half the targets would not have been selected if we used CFHTLS photometry. The 'wanted' objects are galaxies at $z > 0.6$ or quasars and 'unwanted' objects are the rest.

2.7.4 Improving the ELG selection for future surveys

We suggest a few ways to increase the redshift measurement efficiency and reach the requirements set in the second section.

For the *ugr* selection : lowering *u*-*g* cut to 0.3 diminishes the contamination by low-redshift galaxies. Additional low-redshift galaxies can be removed from the selection through an inspection of the images. Some of the low-redshift galaxies are quite extended, and one could mistake a high-redshift merger for an extended low-redshift galaxy. Visual inspection reduces the low-redshift share from 9 percent to 4 percent. The compact and extended selection on the CFHT data is very efficient at identifying quasars. There is also room for improving the spectroscopic redshift determination and thus re-classify 'single emission line' galaxies : they represent a 12 percent share, among which 10 percent are at $z > 0.6$. It seems reasonable to assume an efficiency improvement from 46 percent ELG($z > 0.6$) to 61 percent. Thus a total efficiency of $\sim 75\%$ (61 percent ELG($z > 0.6$) + 14 percent quasar)

For the *gri* selection : improving the spectroscopic redshift determination pipeline can gain up to 5 percent efficiency thus increasing from 73 to 78 percent of ELG($z > 0.6$).

We have also optimized target selections for BAO sampling density using the four bands *ugri*. We find that the optimum selections have a redshift distribution close to the smooth combination of the *gri* and *ugr* selections discussed here; see Fig. 2.15.

2.8 Conclusion

We present an emission-line galaxy selection that can provide a sample from which one can measure the BAO feature in the 2-point correlation function at $z > 0.6$. The efficiencies range from ~ 60 to 80%. With the photometry available today we can plan for a BAO measure to redshift 1 with the BOSS spectrograph.

A representative set of photometric surveys that might be available for target selection in a near future on more than 2,000 square degrees are :

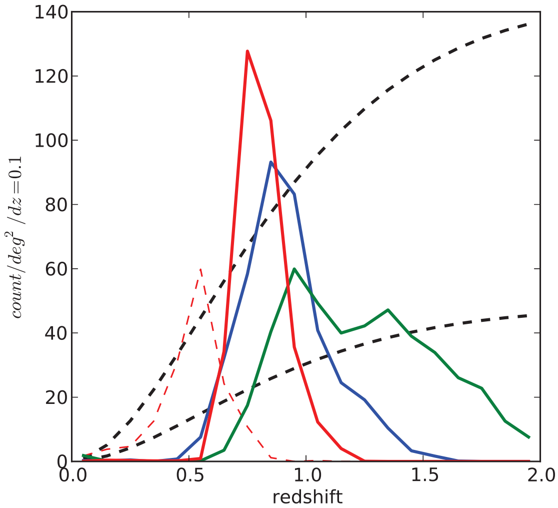


Figure 2.15: Photometric redshift distributions obtained using the *ugri* bands. The dashed black lines are the low and high density goals mentioned in Section 2.3, $\bar{n} = 10^{-4}$ and 3×10^{-4} h^3Mpc^{-3}. The dashed red line is the BOSS CMASS sample. The solid blue line is the distribution enhanced by the *ugri* selection, it has a projected sky density of ~ 340 deg^{-2}. The solid red line is the *gri* selection (projected sky density ~ 350 deg^{-2}). The solid green is the *ugr* selection (projected sky density ~ 400 deg^{-2}). It shows the possibility of making a selection able to sample $[0.6, 1.2]$ for a BAO experiment.

- The Kilo Degree Survey (KIDS)[11] aims at observing 1500 square degrees in the *ugri* bands with 3σ depth of 24.8, 25.4, 25.2, 24.2 using the VST.
- the South Galactic Cap U-band Sky Survey[12] (SCUSS) aims a 5σ limiting magnitude of 23.0
- the Dark Energy Survey (DES) aims at observing 5,000 square degrees in *griz* bands with 10 σ depth of 24.6, 24.1, 24.3, 23.9. This survey does not include the u band (The Dark Energy Survey Collaboration, 2005; Banerji et al., 2008).
- the Large Synoptic Survey Telescope (LSST) (Ivezic et al., 2008) plans to observe 20,000 square degrees in *ugrizy* bands with 5 σ depth of 26.1, 27.4, 27.5, 26.8, 26.1, 24.9.

Using such deeper photometric surveys and improved pipelines, it should be possible to probe BAO to redshift $z = 1.2$ in the next 6 years, *e.g.* by the *eBOSS* experiment, and to $z = 1.7$ in the next 10 years, *e.g.* by PFS-SuMIRE or *BigBOSS* experiment.

2.9 Table of a subsample of observed galaxies at $z > 0.6$

[11] http://kids.strw.leidenuniv.nl/
[12] http://batc.bao.ac.cn/Uband/

Table 2.6: 30 ELGs from the *gri* selection and their observed properties. The error on the redshift determined by the software is not yet fully reliable and thus not displayed. The complete ELG dataset will be included in SDSS Data Release 9 in July 2012.

RA (J2000)	DEC (J2000)	redshift	g	r	i	r_2	r_2^{limit}	$logM_*$	$EW_{[OII]}^{3727}$	$flux_{[OII]}^{3727}$	$EW_{[OIII]}^{5007}$	$flux_{[OIII]}^{5007}$	D_{4n}
213.6918335	52.7059669	0.6055	22.79	22.13	21.28	3.46	2.4	10.5	-15.0	14.15	-3.98	4.57	1.12
214.8627014	51.775528	0.6057	22.96	21.87	20.85	3.16	3.04	10.55	-4.1	6.5	-3.06	5.37	1.21
214.234848	53.0964775	0.6074	23.0	22.29	21.18	4.16	2.69	10.35	-13.97	12.06	-5.67	4.43	1.45
213.7070465	52.2989693	0.6169	22.2	21.52	20.65	5.11	2.62	10.52	-6.95	7.76	-11.71	18.04	1.27
216.1340179	53.6467667	0.6216	22.62	21.72	20.79	4.03	2.63	10.67	-14.3	20.84	-8.74	9.13	1.11
213.3887024	53.5097351	0.6233	21.97	21.3	20.28	5.98	2.16	10.98	-17.69	19.41	-1.75	3.02	1.15
215.3585205	52.8954582	0.6319	22.71	22.06	21.24	4.2	2.85	10.41	-6.48	5.94	-5.89	5.49	1.05
213.268158	53.4914665	0.6325	22.79	21.73	20.71	3.06	2.16	10.83	-3.63	3.33	-2.96	5.85	1.21
216.0396729	53.8974152	0.6453	22.17	21.39	20.52	3.75	2.63	10.72	-13.94	15.13	-10.69	19.87	1.25
213.7992401	53.8096962	0.6486	22.78	21.81	20.78	3.96	2.69	10.62	-10.3	10.82	-9.41	12.07	1.26
213.3646545	52.2972374	0.6516	22.55	21.97	20.8	5.33	2.4	10.65	-6.27	8.9	-8.67	9.83	1.05
214.332962	52.1850281	0.6713	22.92	22.1	21.13	3.39	3.04	10.58	-5.77	6.99	-11.25	19.21	1.35
215.8718262	52.804966	0.6728	22.86	21.97	21.06	4.51	2.85	10.39	-22.57	18.81	-14.13	14.35	1.38
216.0170441	53.2218399	0.6745	22.73	21.77	20.74	4.43	2.63	10.67	-6.25	6.94	-9.26	13.23	1.32
213.0988007	53.3722191	0.7115	22.2	21.58	20.63	3.56	2.16	10.83	-6.5	10.66	-5.93	10.26	1.44
213.3162994	52.2124023	0.7172	22.38	22.12	21.16	4.59	2.4	10.09	-28.72	23.85	-17.58	14.36	1.31
216.207016	53.3441353	0.7214	22.68	21.65	20.75	6.29	2.63	10.91	-5.67	7.19	-12.6	21.03	1.33
215.2261353	54.4270706	0.7232	22.51	21.75	20.84	4.74	2.77	10.91	-15.3	23.72	-15.91	29.68	1.33
214.0466309	54.2588539	0.7332	22.38	21.85	20.88	3.56	2.94	10.34	-15.25	12.34	-14.08	18.21	1.05
216.872467	53.8220406	0.7384	22.52	21.76	20.84	4.64	2.69	11.05	-8.17	8.28	-6.1	11.15	1.27
215.8093567	52.568676	0.7435	22.07	21.24	20.22	3.54	2.85	11.07	-8.75	12.48	-10.71	25.15	1.29
214.6755829	52.3814735	0.7474	22.72	21.87	20.91	3.66	2.62	10.78	-37.32	41.68	-4.67	11.07	1.19
214.9938507	53.0233078	0.7536	22.37	21.57	20.54	3.85	2.62	11.15	-10.54	10.98	-3.2	5.18	1.28
216.4916077	52.4496536	0.7655	22.24	21.59	20.64	4.69	2.85	11.01	-5.49	8.42	-3.18	5.86	1.38
215.0812225	54.1586342	0.7663	22.58	22.17	21.1	3.44	2.94	10.56	-11.64	16.8	-6.87	9.62	1.13
213.9476013	52.7977905	0.7667	22.56	21.56	20.34	3.31	2.62	10.84	-1.5	1.94	-2.29	4.6	1.38
214.9521484	53.7672997	0.7691	22.91	21.95	20.86	3.44	2.69	10.7	-8.77	11.81	-2.11	3.95	1.36
215.9055939	53.82481	0.7727	21.84	20.94	19.96	7.69	2.63	10.75	-22.57	16.87	-6.3	8.81	1.18
213.0868225	53.8581505	0.7814	22.47	21.93	21.09	5.33	2.16	10.65	-26.32	18.35	-20.79	12.91	1.22
214.3099976	53.5430412	0.7861	22.19	21.75	20.93	5.76	2.69	10.54	-10.87	9.3	-5.97	9.67	1.1
215.021637	53.6823692	0.7931	22.99	22.1	21.27	4.73	2.69	10.6	-16.75	16.98	-8.86	11.09	1.65

Figure 2.16: *ugr*-selected ELGs at $z < 0.5$. The spectra are represented in f_λ convention. The coloured image on the left is courtesy of CFHT cutout service. The black and white image in the center is the g-band. The observed object is the one in the centre of the images.

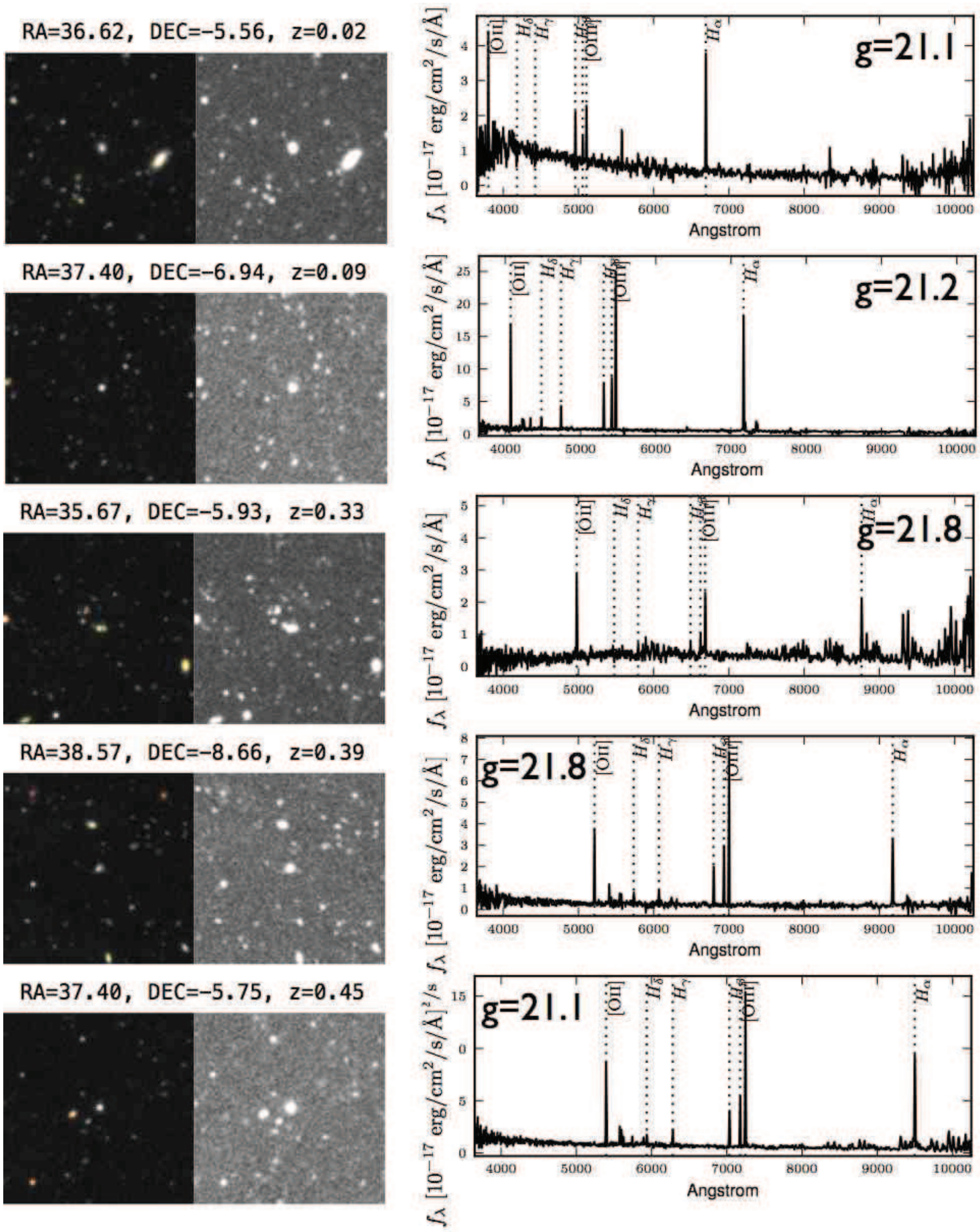


Figure 2.17: *ugr*-selected ELGs at $0.5 < z < 1.0$. The spectra are represented in f_λ convention. The coloured image on the left is courtesy of CFHT cutout service. The black and white image in the center is the *g*-band. The observed object is the one in the centre of the images.

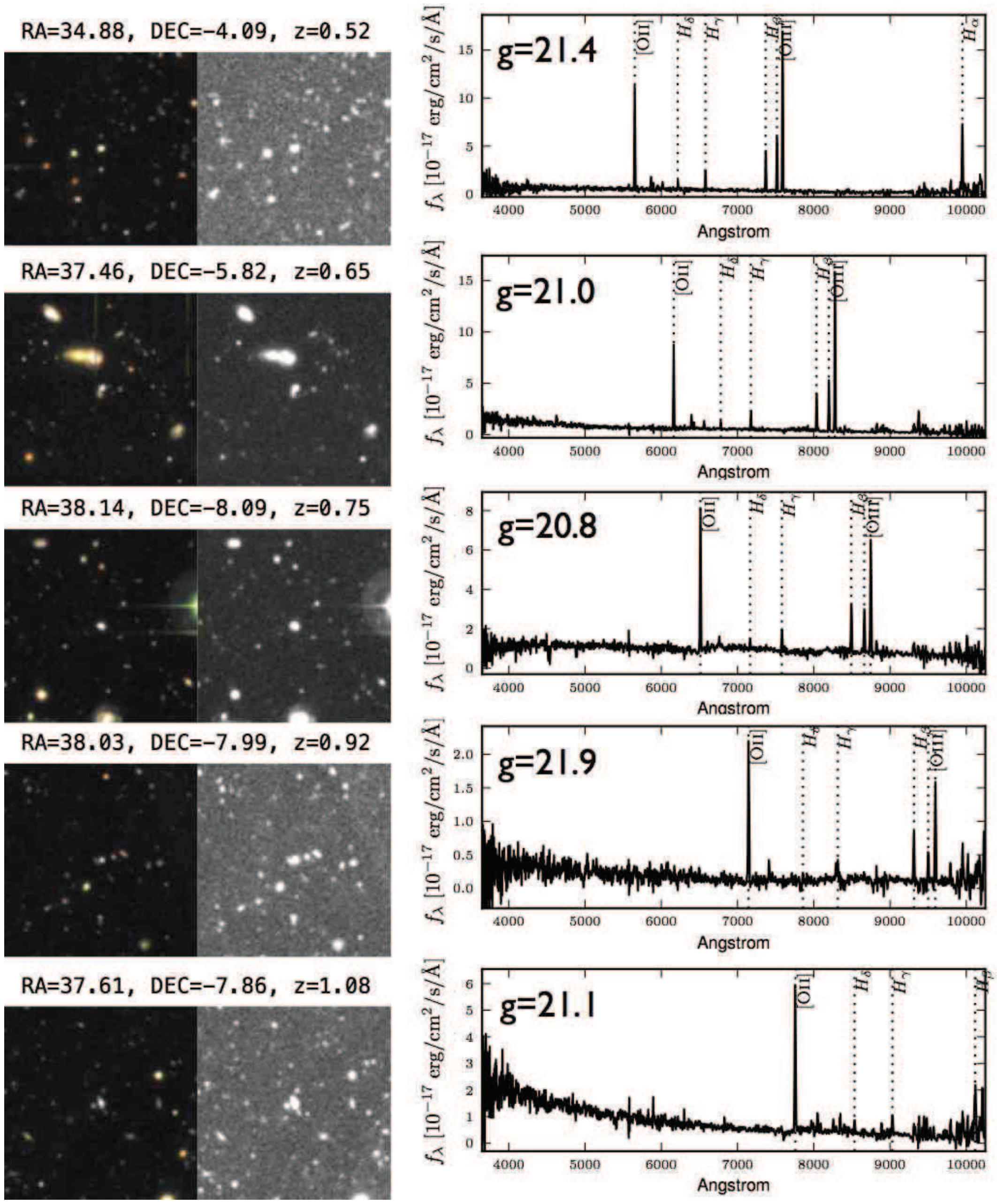


Figure 2.18: *ugr*-selected ELGs at $1.0 < z < 1.6$. The spectra are represented in f_λ convention. The coloured image on the left is courtesy of CFHT cutout service. The black and white image in the center is the g-band. The observed object is the one in the centre of the images.

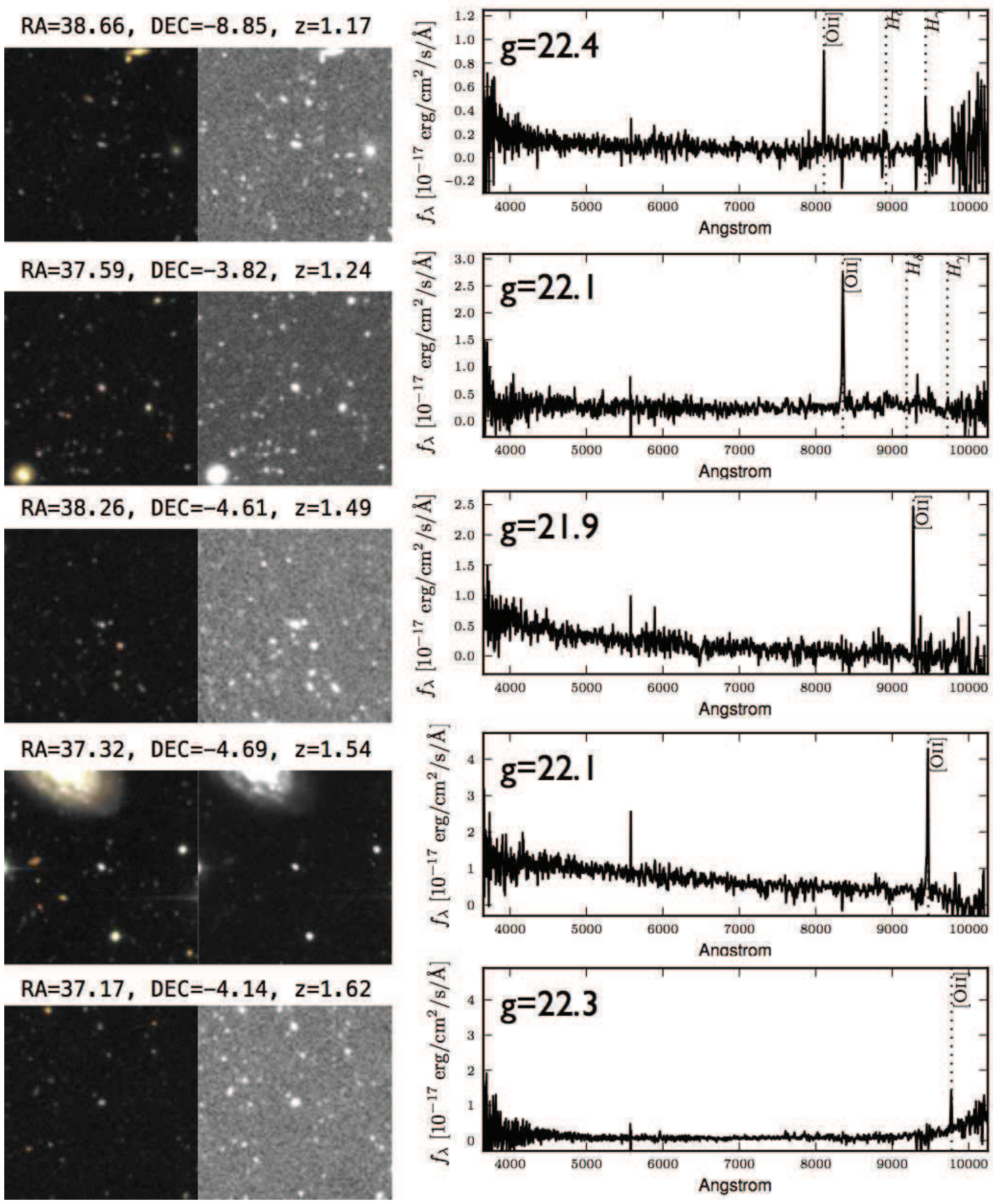
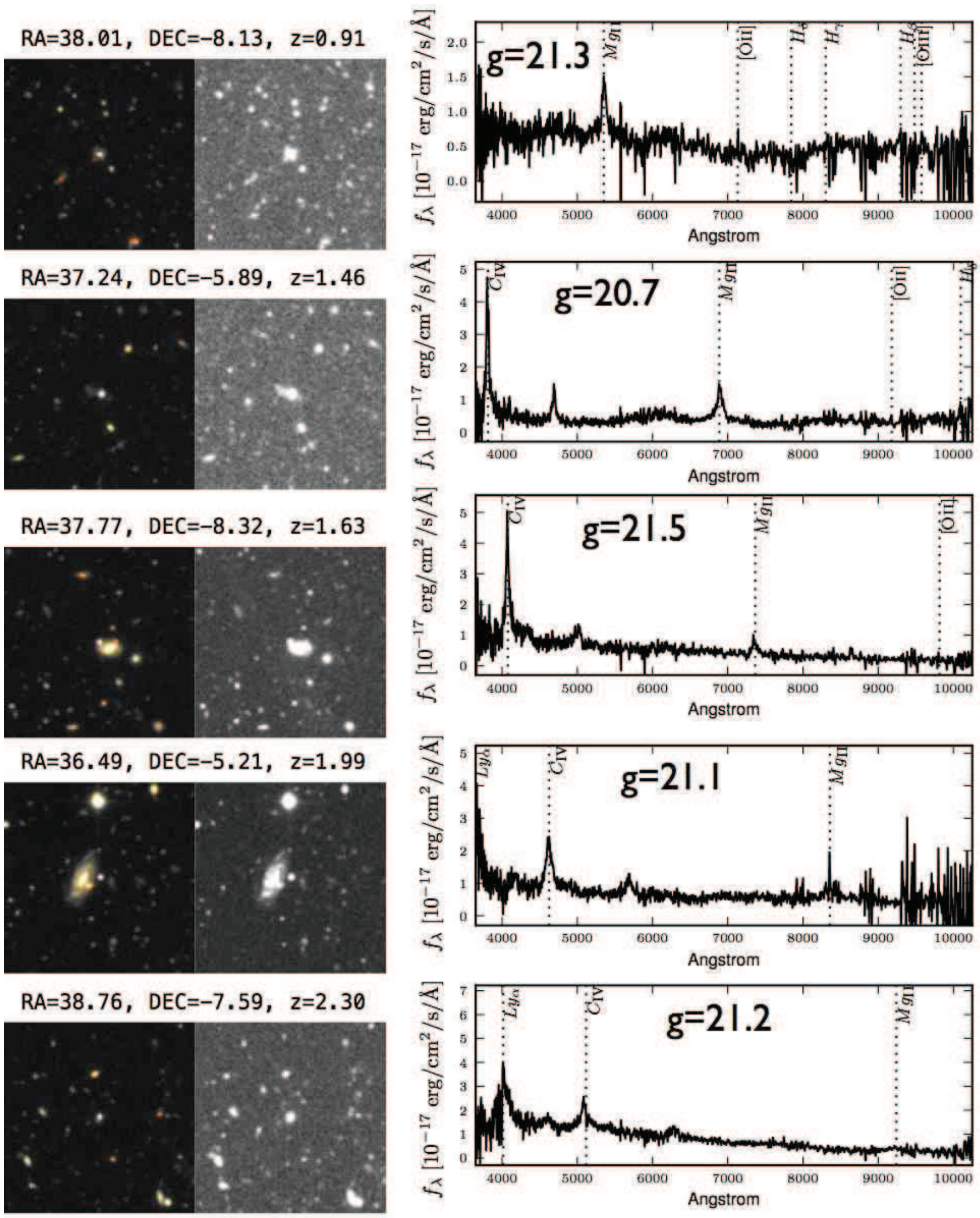


Figure 2.19: *ugr*-selected quasars at $1.0 < z < 2.3$. The spectra are represented in f_λ convention. The coloured image on the left is courtesy of CFHT cutout service. The black and white image in the center is the *g*-band. The observed object is the one in the centre of the images.



Chapter 3

Measuring galaxy [OII] emission line doublet with future ground-based wide-field spectroscopic surveys

J. Comparat, J.-P. Kneib, R. Bacon, N. Mostek, J. Newman, David J. Schlegel C. Yeche. Submitted in may 2013.

3.1 Motivation and contribution

The eBOSS project being accepted, a large effort in the community was made to optimize the survey plan with the previous knowledge of SDSS surveys. The most important issue for a BAO survey is to have reliable redshifts *i.e.* to make sure that a spectrum observed with a significant signal to noise gets a certain redshift. This issue is particularly important beyond redshift 1, where the [OII] emission line is the only strong line present in the spectrum; see Fig. 3.1 the mind map of this chapter.

To answer this question properly, I conducted the simulation described in this chapter in a collaboration with the DEEP2 team, that are experts in high resolution spectroscopy at $z = 1$, and with collaborators from the CRAL in Lyon, France, that have built the MUSE instrument for the VLT and know the instrumental constraints to build spectrographs. The question being, should we build a higher resolution spectrograph for eBOSS ? If yes what is the optimum resolution and is this feasible ? The answer is clear and funding was asked for building an extra spectrograph for the SDSS telescope, which will enable the multiplexing to reach 1500 fibers by adding a 500 fibers red-arm spectrograph.

The simulation package I developed and calibrated on observations gave me a deeper insight on the spectra observed in the ELG ancillary program (creation of the emission line, addition of noise, fit of the emission lines, comparison of the fits, and dependance on the redshift). In particular I found multiple ways to improve the current BOSS redshift solver in order to obtain reliable redshifts for ELG. This enabled me to get really involved in the pipeline team of BOSS. I am now performing myself tests on the pipeline to commit the changes proposed in a robust manner.

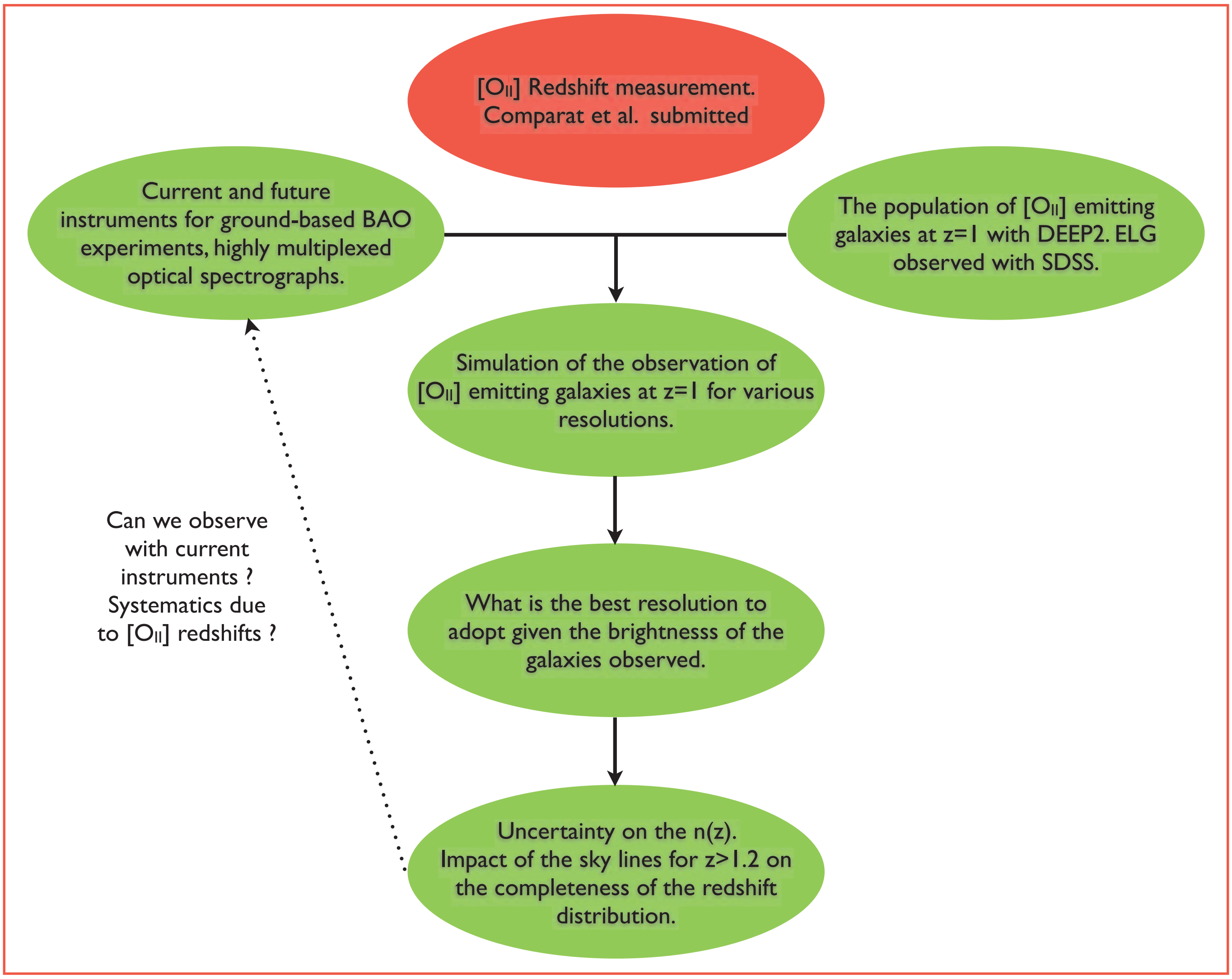


Figure 3.1: Mind map; measuring galaxy [OII] emission line doublet with future ground-based wide-field spectroscopic surveys.

Abstract

The next generation of wide-field spectroscopic redshift surveys will map the large-scale galaxy distribution in the redshift range $0.7 \leq z \leq 2$ to measure baryonic acoustic oscillations (BAO). The primary optical signature used in this redshift range comes from the [OII] emission line doublet, which provides a unique redshift identification that can minimize confusion with other single emission lines. To derive the required spectrograph resolution for these redshift surveys, we simulate observations of the [OII] doublet and measure $\Delta\chi^2$ values for various instrument resolutions, and line velocities. We foresee two strategies about the choice of the resolution for future spectrographs for BAO surveys.
For bright [OII] emitter surveys (like SDSS-IV/eBOSS), a resolution of $R \sim 3300$ is sufficient to obtain 90% of doublets in order to train the pipeline to recover all the remaining 10%. The impact of the sky lines on the completeness in redshift is below 6%.
For faint [OII] emitter surveys (like MS-DESI), we recommend to push the resolution to the highest, the limit being in the end given by the number of pixels (4k by 4k) on the detector and the number of spectroscopic channels (2 or 3).

3.2 Introduction

Following the successful baryonic acoustic oscillation (BAO) measurement in the galaxy clustering from SDSS (Eisenstein et al., 2005), 2dF (Cole et al., 2005), Wiggle-Z (Blake et al., 2011a), and the Baryonic Oscillation Spectroscopic Survey (BOSS) (Anderson et al., 2012), there is a strong motivation in the community to plan the next generation of spectroscopic redshift surveys for BAO. In particular, these future ground-based surveys plan to map the galaxy distribution in the universe and use the galaxy power spectrum to precisely measure the BAO signature and strongly constrain cosmological parameters.

Examples of this new paradigm are the following projects : MS-DESI (previously BigBOSS and DESpec), PFS-SuMiRe, Weave, 4Most.

The BigBOSS project which planned to map 14,000 deg^2 of sky using 5000 motorized fibers over a 7 deg^2 field of view and measure 22 million galaxy redshifts (see Schlegel et al. (2011) for a global survey description and Mostek et al. (2012a) for the current survey parameters).

DESpec aims to build a 4000 fiber spectrograph for the Blanco 4-meter telescope at Cerro Tololo Inter-American Observatory (CTIO) in Chile. The spectrograph would be interchangeable with the CCD imager of the existing Dark Energy Camera (DECam), thus accessing a field of view of 3.8 square degrees in a single exposure. In a survey of 350 nights, the DESpec collaboration proposes to obtain spectroscopic redshifts for ∼8 million galaxies over 5000 deg^2 (Abdalla et al., 2012).

The Subaru Prime Focus Spectrograph (PFS) is a 2400 fiber-fed optical and near-infrared spectrograph ($380 < \lambda < 1300$nm), for the 8m Subaru telescope. The Subaru Strategic Program of about 300 nights will dedicate about 100 nights to constrain the nature of dark energy via a survey of emission line galaxies spanning a comoving volume of $9.3\,h^{-3}$Gpc^3 for $0.8 < z < 2.4$. (Ellis et al., 2012)

The WEAVE galaxy evolution and cosmology survey on the William Herschel Telescope will provide emission-line redshifts for all star-forming galaxies at $z < 1.3$ using H_α and [OII] emission and at $z > 2$ using Ly_α emission for $\sim 10^7$ objects over 10,000 deg^2 (Dalton et al., 2012).

The 4MOST consortium aims to develop a wide-field (> 3 square degree), high-multiplex (> 1500 fibres) spectroscopic survey facility for the ESO VISTA 4m telescope. 4MOST will run permanently on the telescope to perform a 5 year public survey yielding more than 20 million spectra at resolution $R \sim 5000$ ($\lambda = 390 - 1000$nm) and more than 2 million spectra at $R \sim 20,000$ ($395 - 456.5$ nm & $587 - 673$ nm) (de Jong et al., 2012).

Redshifts for the large-scale galaxy distribution will be mostly determined from the emission line features of star-forming galaxies between $0.7 \leq z \leq 2$. Table 3.1 lists the primary emission lines that are available at optical and NIR wavelengths within this redshift range. Of these lines, the [OII] doublet at ($\lambda 3727, \lambda 3729$) will provide the most consistently available feature. In order to avoid confusion with

Table 3.1: Emission lines available at optical and near-infrared wavelengths. Taken from Atomic Line List[a]. $\lambda_{\rm vac}$ [] is the wavelength emitted in vacuum in , the orbital transition is given under the column 'term', 'J-J' is the spin state. The last column gives the energy transition that occurs in electron-Volt.

line name	$\lambda_{\rm vac}$ []	term	J-J	energy level [eV]
[OII]	3727.092	E2 4So-2Do	3/2 - 3/2	0 - 3.326568
[OII]	3729.875	E2 4So-2Do	3/2 - 5/2	0 - 3.324086
H_β	4862.683	E1 2-4	* - *	10.198834 - 12.748544
[OIII]	4960.295	M1 3P-1D	1-2	0.014032 - 2.513566
[OIII]	5008.240	M1 3P-1D	2 - 2	0.037961 - 2.513566
H_α	6564.61	E1 2-3	* - *	10.198834 - 12.087510

[a]http://www.pa.uky.edu/ peter/atomic/

other prominent emission lines (H$_\alpha$, H$_\beta$, [OIII]), the [OII] doublet should be resolved over the instrumented wavelength range where no other lines are available to make an unambiguous identification; such is the case in the MS-DESI and PFS-SuMIRe spectrographs (see Jelinsky et al. (2012) for the current MS-DESI spectrograph design, and Vives et al. (2012) for PFS-SuMIRe spectrograph design).

Previous emission line redshift surveys have had different strategies concerning the use of emission lines for measuring the redshift. Wiggle-Z with a spectral resolution of 1300 obtained 60% of reliable redshifts (18% based on the detection of the [OII] doublet *i.e.* the doublet is resolved or partially resolved), and 40% of unreliable redshifts; (Drinkwater et al., 2010). DEEP2 survey with a resolution of 6000 obtained 71% of reliable redshifts (14.8% based on the detection of the [OII] doublet *i.e.* the doublet is resolved or partially resolved), 10% between reliable and unreliable, and 19% of unreliable redshifts (Newman et al., 2012). The difference between these redshift efficiencies is related to the resolution and the wavelength coverage. Indeed, if the [OII] emission is the only one available in the spectrum, at high resolution the doublet is split and the redshift is reliable. Whereas, at lower resolution the doublet is not always split and might therefore be confused as another emission line.

In Section 3.3, we derive the minimum resolution necessary to resolve the doublet in the case of an observation without noise. In Section 3.4, we describe our simulation of [OII] doublet detections based on DEEP2 spectral observations. We discuss the results of our simulation in Section 3.5.

3.3 Instrumental requirements

First, let us define our notation. R is the resolution of the spectrograph, $\lambda_a = 3727.092$, $\lambda_b = 3729.875$ are the individual [OII] doublet emission wavelengths and $\lambda_{[O_{II}]} = (\lambda_a * 3.326568 + \lambda_b * 3.324086)/(3.326568 + 3.324086) = 3728.483$ is the energy-weighted mean [OII] wavelength. The observed wavelength separation between the emission lines depends on the redshift; we denote in Eq. 3.1 this separation as $\delta_{[O_{II}]}(z)$.

$$\delta_{[O_{II}]}(z) = (\lambda_b - \lambda_a)(1+z) = 2.783(1+z) \tag{3.1}$$

We can thus define the resolution, $R_{[O_{II}]}$, as the minimal resolution required to properly sample a theoretical [OII] doublet (with zero intrinsic width) without loss of information by Eq. 3.2 (Nyquist-Shannon sampling theorem, Shannon and Weaver 1975). Note that $R_{[O_{II}]}$ is independent of redshift.

$$R_{[O_{II}]} = 2\frac{(1+z)\lambda_{[O_{II}]}}{\delta_{[O_{II}]}(z)} = 2679. \tag{3.2}$$

However, a real galaxy has an intrinsic velocity dispersion, Δv, that broadens the emission lines from a theoretical δ-function profile. Assuming the line profile is dominated by thermal Doppler broadening

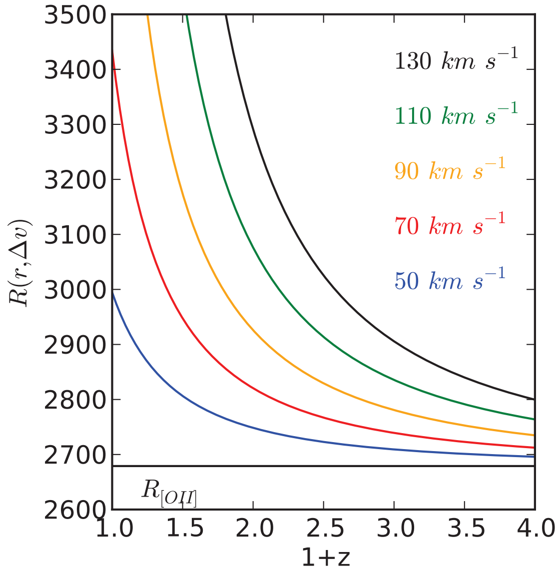


Figure 3.2: $R(z, \Delta v)$. The bottom blue line is for a velocity dispersion of $\delta v = 50\,\mathrm{km\,s^{-1}}$, every line has its velocity dispersion increased by $20\,\mathrm{km\,s^{-1}}$ to finish with the top red line representing $v = 130\,\mathrm{km\,s^{-1}}$. The horizontal line is at the Nyquist minimal resolution R_{Ny} to sample the doublet in the case of a zero width emission line.

in the host galaxy interstellar medium, the observed wavelength width of the broadened [OII] line profiles is given by $\delta\lambda_v$ defined in Eq. 3.3 where c is the speed of light.

$$\delta\lambda_v = \lambda_{[OII]}\frac{\Delta v}{c} \tag{3.3}$$

In this simplified case, the intrinsic velocity dispersion is equivalent to the standard deviation in a Gaussian profile. For example, a galaxy at $z = 1$ with $\Delta v = 50\,\mathrm{km\,s^{-1}}$ has a line width of $\delta\lambda_v \sim 0.6$, which represents $\sim 10\%$ of the wavelength separation between the doublet peaks.

Furthermore, the spectral resolution of the instrument also broadens the width of the [OII] lines. By defining $R = \lambda/\Delta\lambda$, the change in line width due to resolution is given by $\delta\lambda_R(z)$ defined in Eq. 3.4. Note that the broadening due to instrumental resolution depends on the redshift because the position of [OII] changes with redshift while $\Delta\lambda$ is constant in wavelength (for grim spectrograph):

$$\delta\lambda_R(z) = (1+z)\frac{\lambda_{[OII]}}{R} \tag{3.4}$$

By performing a quadrature sum of the components in Eq. 3.3 and 3.4, we obtain the observed width, denoted $w_{[OII]}(z)$, of an individual line in the [OII] doublet, see Eq. 3.5.

$$w_{[OII]}(z) = \lambda_{[OII]}\sqrt{\frac{(1+z)^2}{R^2} + \frac{\Delta v^2}{c^2}} \tag{3.5}$$

In order to Nyquist sample the observed [OII] doublet at redshift z, the individual line width of the doublet has to be at least twice the doublet separation, or $w_{[OII]}(z) = 2\delta_{[OII]}(z)$. Rewriting Eq. 3.5 in terms of this minimum sampling requirement (Eq. 3.2), we derive the relation 3.6. Fig. 3.2 plots the function $R(z, \Delta v)$ and shows that the resolution required to split the doublet decreases with redshift and increases with the velocity dispersion. It converges asymptotically towards $R_{[OII]}$. The colored lines show the minimum resolution required to resolve the [OII] doublet with Δv at redshift z.

$$R(z, \Delta v) = \left[\frac{1}{R^2_{[OII]}} - \frac{\Delta v^2}{(1+z)^2 c^2}\right]^{-1/2}. \tag{3.6}$$

For a galaxy at $z = 1$ ([OII] is observed at $\lambda \sim 7456$) with $\Delta v = 100\,\mathrm{km\,s^{-1}}$, the minimum resolution required is $R_{min} = 3000$. For a galaxy at $z = 1$ with $\Delta v = 70\,\mathrm{km\,s^{-1}}$, it is $R_{min} = 2800$. The spectrograph currently used by SDSS-III/BOSS reaches $R \sim 2700 > R_{[OII]}$ at 9320, and therefore it theoretically splits the [OII] doublet for galaxies with $\Delta v < 50\,\mathrm{km\,s^{-1}}$ at $z \geq 1.5$. With this spectrograph, the observation of the [OII] doublet of galaxies with $\Delta v = 100\,\mathrm{km\,s^{-1}}$ will be highly-blended. We show a variety of examples of [OII] doublets observed with SDSS telescope around redshift 1. It shows at this resolution, both emission lines are always blended, but that sometimes, we see the two lines; see Fig 3.3. Even if the lines are blended, we can still assign a reliable redshift if we see the two peaks.

At low resolution, it is possible to actually see [OII] doublets when the lines peaks and valley falls exactly right opposite to the pixels. In the following, when we state 'the doublet is resolved', it is true wherever the emission line lands on the detector.

In classical spectrographs, the resolution element $\Delta\lambda$ is roughly constant with wavelength, and therefore the spectral resolution R is a linear function of the observed wavelength λ. R also depends on $(1+z)$ assuming the restframe wavelength of a given emission line is at the observed wavelength (see Eq. 3.4). We must therefore define the minimum resolution requirement to be at the lowest redshift limit where [OII] becomes the only emission line available in the spectrum. The resolution requirement will automatically be satisfied for all higher redshifts.

As an example, let us consider a spectrograph with a spectral range extending to $1\mu m$. As the observed galaxy redshift increases beyond $z = 1$, the H_β and [OIII] emission lines are redshifted beyond the wavelength limit of the spectrograph and we are left with only the [OII] doublet. This is the case of many spectrographs using deep-depletion CCD detectors such as SDSS-III BOSS and VLT-FORS2. Spectrographs with a near-infrared channel cutting at $1.3\mu m$, such as the foreseen SUMIRE/PFS experiment (Vives et al., 2012), will have an upper redshift limit of $z = 1.6$ for H_β and [OIII] , beyond which [OII] will again be the only strong emission line. Fig. 3.4 shows prominent emission lines and the two spectral cutoffs.

In this study, we consider the more common case where the spectral range is limited to $< 1\mu m$. Fig. 3.5 shows an example of two galaxy spectra that are just below and above $z = 1$, these were observed with the SDSS-BOSS spectrograph that has a range of ~ 0.36 to $\sim 1\mu m$; see the sections 2.6, 2.5 and 2.9, that describe the observations, their physical properties, and displays more spectra, or Comparat et al. (2013).

3.4 Simulation

To confirm the theoretical considerations of Section 2, we simulate observations of [OII] doublets in the presence of Poisson noise. Future massive spectroscopic redshift surveys are primarily focused on obtaining redshifts with only emission lines, which is less demanding in terms of exposure time than requiring the detection of the continuum. For these applications, a Gaussian profile is sufficient to simulate the gross resolution effects.

Because spectrograph resolution increases with wavelength, the minimal resolution requirement is determined at the shortest wavelength where the [OII] doublet becomes the only major emission line feature in the spectrum. Assuming an instrumental wavelength limit of 1μm, the resolution requirement for [OII] is therefore defined at $\lambda_{obs}([\mathrm{OII}], z = 1) \sim 7450$Å.

3.4.1 Parameters of the simulation

Of interest for this work, DEEP2 has obtained a complete spectroscopic sample with [OII] emitter redshift measurements at $z = 1$. The magnitude depth is $r = 24.1$ and the [OII] flux limit is $5 \times 10^{-17}\mathrm{erg\,cm^{-2}\,s^{-1}}$ (see Newman et al. (2012) for the latest data release and the survey overview). These limits are deeper than the target selection limits for BAO surveys currently under development.

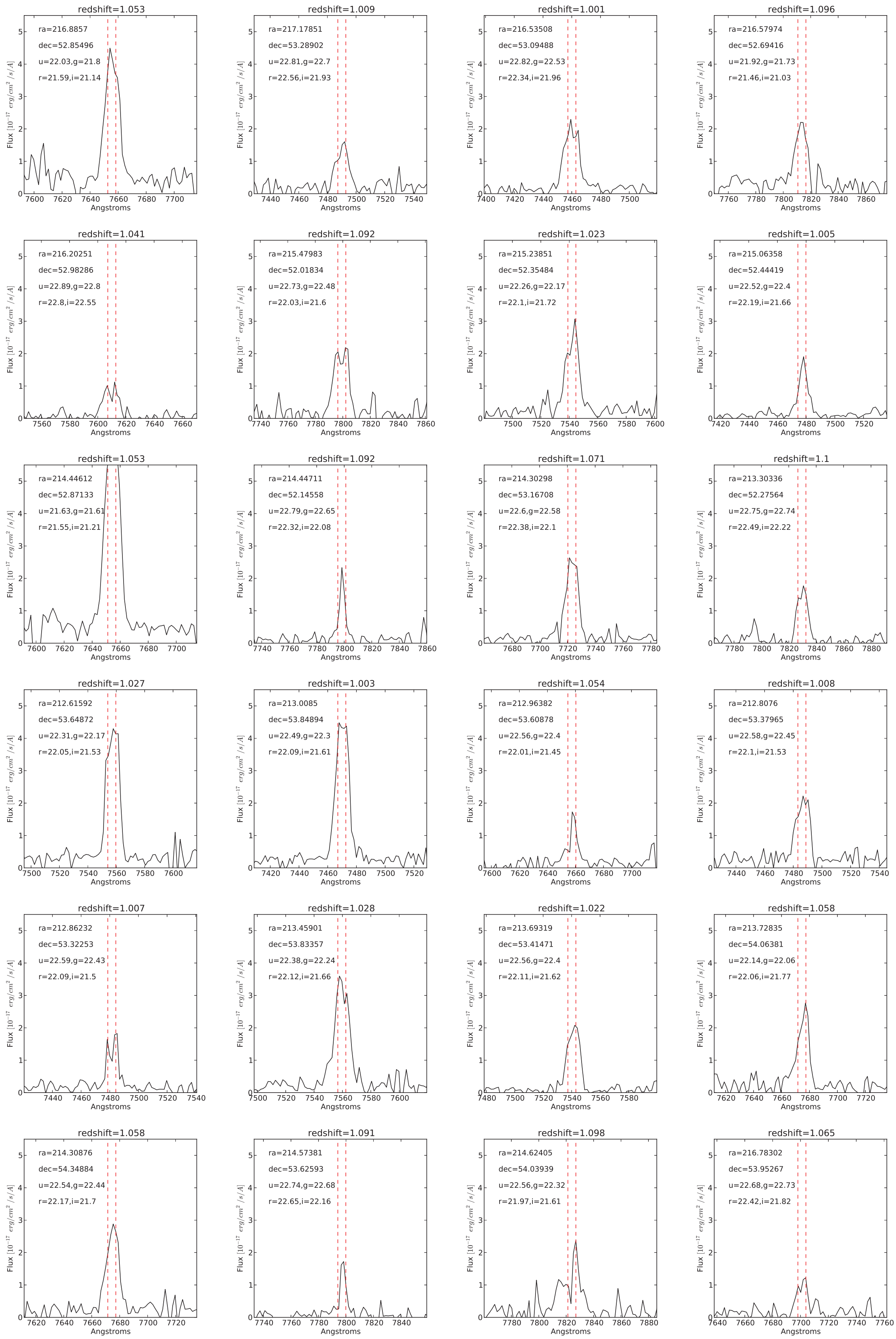


Figure 3.3: Examples of [OII] observations at $R \sim 2000$ for $1 < z < 1.1$. The spectrum is in black, the position of the emission lines are given by the red vertical dashed lines. Both emission lines are always blended, but sometimes, we see the two lines.

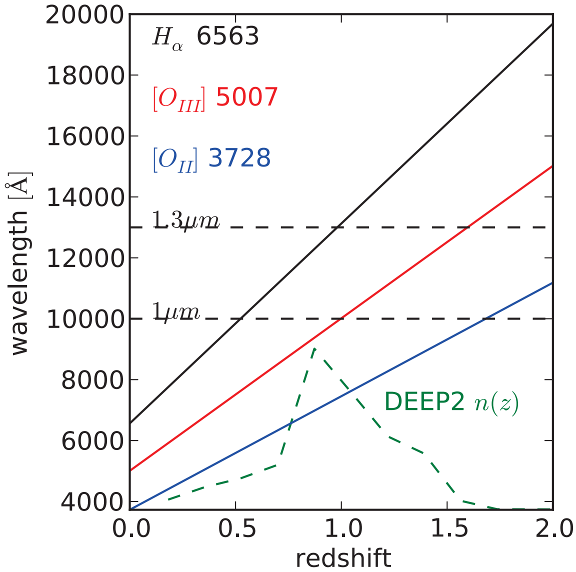


Figure 3.4: Wavelength vs. redshift. Beyond $z = 1$, the only emission line remaining on a spectrum reaching $1\mu m$ is [OII] 3727. For a spectrograph covering out to $1.3\mu m$, [OIII] 5007 is present until $z = 1.6$. The rescaled redshift distribution of DEEP 2 DR4 galaxies with redshift quality better than 3 and i brighter than 24 is drawn (green dashes).

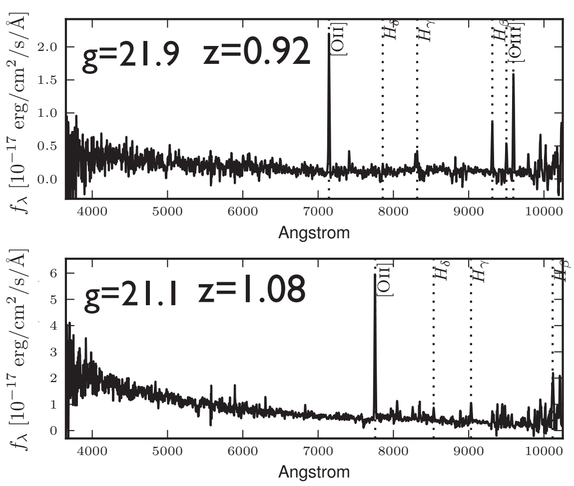


Figure 3.5: Two emission line galaxy spectra observed by the *Sloan* telescope at redshift 0.92 and 1.08, with observed g-band magnitudes of 21.9 and 21.1 respectively. On the top one we see the redshift is assessed by many emission lines and on the bottom one there is only the strong [OII] emission line available. H_β and H_γ emission lines are in the noise.

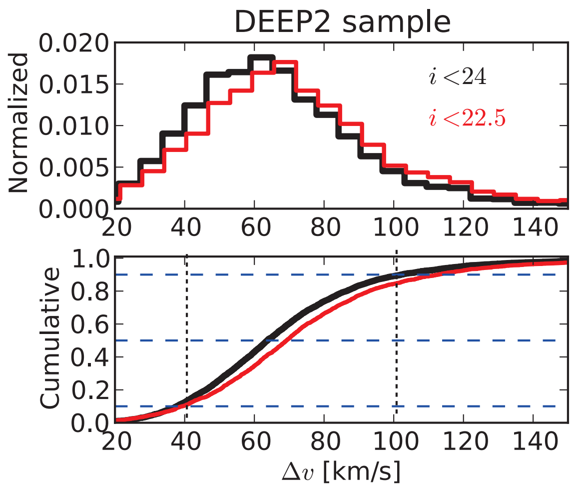


Figure 3.6: Normed and cumulative distributions of the velocity dispersion measured in the DEEP2 DR4 sample. The horizontal dashed lines are at 10 and 90%. The brighter sample, in red, is slightly shifted towards the higher velocity dispersions compared to the fainter sample in black.

Furthermore, DEEP2 used the DEIMOS spectrograph at Keck with a resolution R=6000 using a grim (Faber et al., 2003) and was limited by the galaxy continuum signal-to-noise.

The range of velocity dispersions used in our simulation is empirically determined by observations of $z \sim 1$ [OII] emitters within the DEEP2 redshift survey. In fact the resolution of DEEP2 allows an accurate measurement of the distribution of the velocity dispersion; see Fig. 3.6 that shows the velocity dispersion of galaxies with $0.9 < z < 1.1$ and $\sigma_{\Delta v} < 0.1\Delta v$ for a bright sample with $i < 22.5$ and a faint sample $i < 24$. In order to concentrate on the bulk of the galaxy distribution, we set the lower (upper) limit of the investigated range at $\Delta v = 20\,\mathrm{km\,s^{-1}}$ ($120\,\mathrm{km\,s^{-1}}$) that constitute the first and ninth deciles of the distribution.

In terms of instrumental resolution, we explore the range of $2500 < R < 6000$ sampled by steps of $\delta_R = 3$ in resolution. To avoid aliasing problems, for each doublets we add a random number smaller than 3 to the resolution, in order to sample correctly the complete resolution range. We use a sampling of 3 pixels per resolution element. Our results will span a meaningful range of resolutions for numerous spectrographs at λ_{obs}([OII]) $\sim$ 7500Å, including the current SDSS-III/BOSS spectrograph ($R \sim 2500$, Smee et al. (2012)) and future spectrographs such as PFS-SUMIRE with $R \sim 3000$ (Vives et al., 2012)) or MS-DESI with $R \sim 4000$ (Jelinsky et al., 2012).

We use a Gaussian function, to model the [OII] doublet, given by Eq. 3.7. This function produces an emission line centered at λ_0 and normalized to a total flux F_0. The profile width σ_g is linked to the velocity dispersion by $\sigma_g = \lambda_{[O_{II}]}\Delta v/c$. The Gaussian profile has a set exponential drop off from the emission peak value, and therefore it may not represent systematic effects like scattered light within the spectrograph. A Moffat profile can recover information about the wings of the emission line when β is allowed to vary. However, the Moffat model is only attractive if the data has a high spectral resolution and high signal to noise ratio. Otherwise, the information in the wings will have low significance due to measurement noise.

$$f_{\mathrm{gaussian}}(\lambda, \lambda_0, \sigma_g, F_0) = \frac{F_0}{\sqrt{2\pi}\sigma_g} Exp\left[\frac{(\lambda-\lambda_0)^2}{2\sigma_g^2}\right]. \tag{3.7}$$

We calibrate the flux f and the sky level to a recent emission line galaxy observational study performed at the SDSS Telescope; see Table 2.4. They showed the nominal observed total line flux is $\sim 30 \times 10^{-17}\mathrm{erg\,cm^{-2}\,s^{-1}}$ and the nominal sky brightness is $\sim 3 \times 10^{-18}\mathrm{erg\,cm^{-2}\,s^{-1}\,{}^{-1}arcsec^{-2}}$ at $\sim$ 7400. This noise level corresponds to detections with a SNR above 7 which should be typical

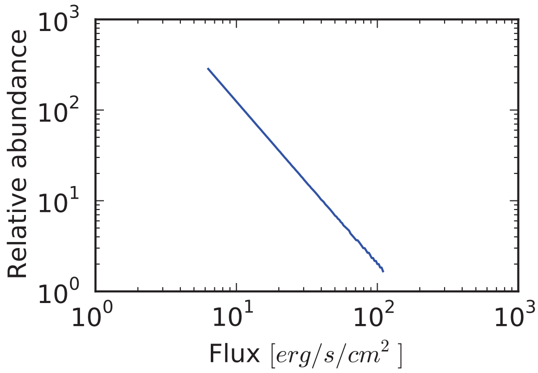


Figure 3.7: Relative abundance of emission lines simulated vs. [OII] flux. We determine the relative abundance of emission lines at a given flux with the [OII] luminosity function at $z \sim 1$ measured by Zhu et al. (2009) on DEEP2.

of observations in future BAO survey. In the simulation, we use fluxes f from a broader range, $6 < f < 100 \times 10^{-17}$erg cm^{-2} s^{-1}. We determine the relative abundance of emission lines at a given flux with the [OII] luminosity function at $z \sim 1$ measured by Zhu et al. (2009) on DEEP2; see Fig. 3.7.

Steps of the simulation

First, we first make a Gaussian doublet at λ_{obs}([OII]) $\sim$ 7450 for a given resolution R, velocity dispersion Δv , and flux f. Next, we sample the doublet spectrum at resolution R with 3 pixel per resolution element. We add Poisson sky noise on each pixel (this is the dominating contribution of the observed noise). This creates a mock observation of the [OII] emission doublet for the Gaussian profile. Finally, we fit two models to the simulated doublet: a single Gaussian profile, and a double Gaussian profile. From each fit, we compute the SNR and the χ^2 to compare the detections. χ^2 is defined as the usual 'reduced chi-square statistics' by Eq. 3.8 where n_{dof} is the number of degrees of freedom, O is the array of observed values, M^1 is the model with one line, M^2 is the model with 2 lines, and N is the noise. The number of degrees of freedom vary from 35 to 94 (depending on the spectral resolution used).

$$\chi_{i=1\,or\,2} = \frac{1}{n_{dof}} \sum_{k\,\in\,\text{pixels}} \frac{(O_k - M_k^i)^2}{N_k^2} \tag{3.8}$$

The SNR is calculated with a Fisher matrix of the form given in Eq. 3.9.

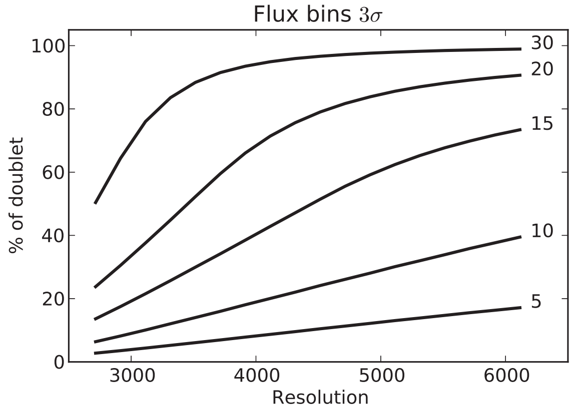
$$SNR = 1/\sqrt{\begin{pmatrix} \sum_{k\,\in\,\text{pixels}} \left(\frac{M_k}{N_k}\right)^2 & \sum_{k\,\in\,\text{pixels}} \frac{M_k}{N_k^2} \\ \sum_{k\,\in\,\text{pixels}} \frac{M_k}{N_k^2} & \sum_{k\,\in\,\text{pixels}} \left(\frac{1}{N_k}\right)^2 \end{pmatrix}^{-1} [0,0]} \tag{3.9}$$

3.5 Results

3.5.1 Doublet detection and resolution at $z \sim 1$

The simulation contains $\sim 15 \times 10^6$ simulated [OII] lines sampling the velocity dispersion, resolution, and flux range set in the above.

To statistically differentiate whether an observation of [OII] is identified as a doublet or a single emission line (SEL), given that the numbers of degrees of freedom is high ($35 < n_{dof} < 94$), we use



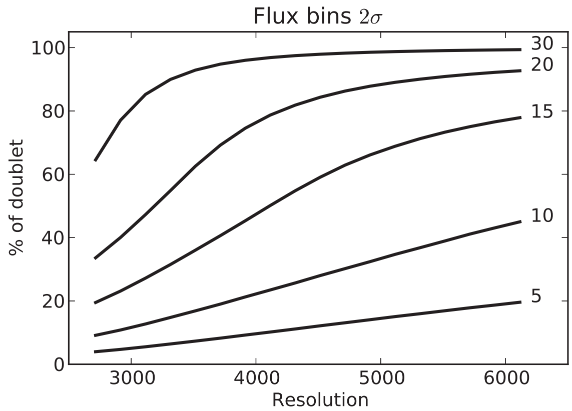


Figure 3.8: Share of doublets at the 3 and 2σ (confidence level of 99%, 95% from top to bottom) vs. resolution for $i < 24$ doublets at $z = 1$ for different flux bins. Each line corresponds to a survey with a the flux detection limit given on the right end of each line in units of $10^{-17}\mathrm{erg\,cm^{-2}\,s^{-1}}$. eBOSS corresponds to the line 30 and MS-DESI to the line 10.

$\Delta\chi^2 = \chi_1/n_{dof1} - \chi_2/n_{dof2}$. A $\Delta\chi^2 = 9$ means the single line emission model is ruled out at 3σ or with a 99.7% confidence level. We compute the share of emission line with $i < 24$ (convolved by the velocity dispersion distribution of Fig. 3.6 in black) detected as a doublet at the 2 and 3 σ confidence levels at redshift 1 as function of the resolution for different [OII] flux detection limit, see Fig. 3.8.

The main trend is that the percentage of doublets increases as a function of the resolution. In the regime of low [OII] fluxes (below the line 12), the gain is linear. It indicates we should push for the highest resolution possible. For higher [OII] fluxes, the marginal increase of the doublet share is large for low resolutions and diminishes for higher resolution. This result advocates two strategies. For a survey aiming only to observe the brightest [OII] emitters (on Fig. 3.8), it is not necessary to aim for the highest resolution. $R = 3300$ is sufficient to obtain 90% of doublets. And for $R > 3300$, the marginal cost of an extra percent of doublets decreases. For a survey aiming to observe all [OII] emitters (MS-DESI line 10 on Fig. 3.8), it is necessary to push the resolution to its highest.

The DEEP 2 survey dealt with SEL using a neural network (Kirby et al., 2007). They showed that given a fair spectroscopic sample of an observed population with reliable redshifts, it is possible to infer correct redshifts to nearly 100% of the [OII] SEL. The H_α, H_β, and [OIII] SEL cases are not as well handled by the neural network with efficiencies of $\sim 90\%$, $\sim 60\%$, and $\sim 60\%$ respectively.

The combination of the two latter points shows it will be possible to derive robust [OII] redshifts where [OII] is the only emission line available in the spectrograph, even if the fraction of 3σ doublet detections is small.

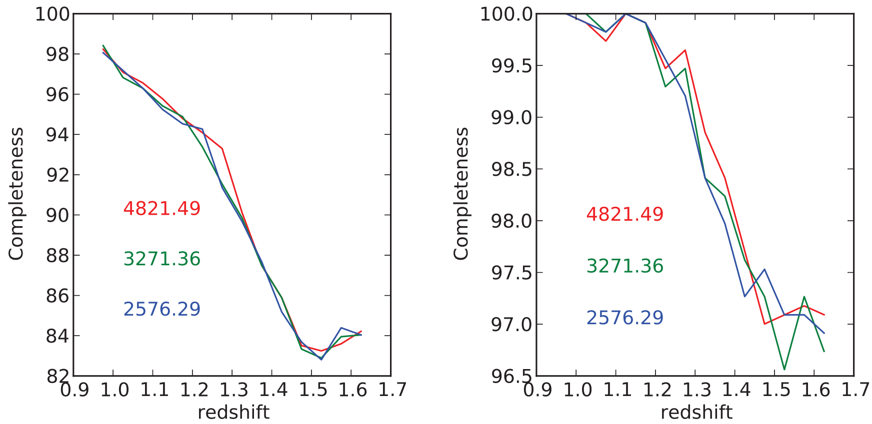


Figure 3.9: Completeness vs. redshift for three resolutions. The left panel is in the case of an [OII] emission line with total flux of 10^{-16}erg cm^{-2} s^{-1} (MS-DESI-like) and the right with 3×10^{-16}erg cm^{-2} s^{-1} (eBOSS-like). The sky brightness is 10% of that observed in a fiber of 2 arcsec diameter fiber (we assume a sky subtraction efficient at 90%).

3.5.2 Higher redshift, sky lines, completeness

The sky lines have an observed width of one resolution element, therefore their width varies with the resolution. In the case of a single sky line located on a doublet, it is not a problem to subtract the sky line and obtain an accurate redshift. In the case of many contiguous sky lines, it can cover completely a doublet and prevent from getting any redshift in the zone or at a higher flux limit. To quantify the impact of the sky lines obstruction as a function of redshift, we simulate at various resolutions the observation of a sky spectrum. The sky spectrum is taken from Hanuschik (2003).

At a given resolution, we convert the wavelength array of the sky into a redshift array corresponding to the [OII] redshift. We scan the redshift array by steps of 0.0005 (it corresponds to the desired precision of a spectroscopic redshift). At each step, we compare the median value of the sky to the flux measured in the middle of an [OII] doublet with (where it is the lowest). If the median value of the sky is greater than the value of the doublet, we consider we cannot fit a redshift. From this, we compute the percentage of the redshift range where we can fit spectroscopic redshifts.

We run this simulation for two settings. A bright survey with [OII] fluxes $\sim 30\times10^{-17}$erg cm^{-2} s^{-1} and fibers of 2" diameter (eBOSS-like). A faint surveys with [OII] fluxes $\sim 10\times10^{-17}$erg cm^{-2} s^{-1} and fibers of 1.5" diameter (MS-DESI-like). The completeness as a function of the resolution is displayed on Fig. 3.9. It shows the impact of the sky lines on the completeness depends weakly on the resolution. For the brighter scenario the completeness is greater than 95% for the redshift range. For the faint survey case, the completeness is greater than 80% and the discrepancy between different resolutions remains under a few percents.

This demonstrates how the sky lines impact the completeness of the redshift distribution of a spectroscopic survey. It shows it is necessary to have the smaller fiber possible to diminish the impact of the sky, but the increase in resolution is not useful to cope with this problem. Finally the the completeness in redshift is not driven by the resolution at the first order, but by the robustness of sky subtraction and the strength of the [OII] flux. To obtain a precise estimate of the impact of sky lines on the redshift distribution completeness, a full end-to-end simulation is needed.

3.5.3 Integrated velocity profile

The integrated velocity profile of each galaxy within a fiber is assumed to be Gaussian in this study, although galaxy rotation may make this more complicated. There is not sufficient data to explore this particular complication: Nearby galaxies are not representative of the properties of these high-redshift galaxies, and Epinat et al. (2012) is limited to a sample of only 50 galaxies in the redshift range $0.6 < z < 1.6$.

3.6 Conclusion

Large spectroscopic redshift surveys are being designed to measure galaxy redshifts using the [OII] emission line doublet and trace the large-scale matter distribution. We have shown how observations of the doublet evolves with the instrumental resolution and the line velocity dispersion. We quantified the impact of sky lines on the completeness in redshift.

In light of the numbers obtained, we foresee two strategies about the choice of the resolution for future spectrographs for BAO surveys. For bright [OII] emitter surveys (like SDSS-IV/eBOSS), a resolution of $R \sim 2500$ (current SDSS spectrograph) is sufficient to obtain a fair sample of doublets (60%) in order to train the pipeline to recover all the [OII] redshifts. Increasing the resolution to 3300 allows to get 90% of doublets. For a small increase in resolution, the redshift determination efficiency doubles. The impact of the sky lines on the completeness in redshift is below 6%. For faint [OII] emitter surveys (like MS-DESI), we recommend to push the resolution to the highest. Knowing there is a limited number of pixels on the detector (4k), and that the highest resolution possible on a three channel spectrograph is $R \sim 4500$ at 7500, to go beyond, it is necessary to use a four channel spectrograph. Practically with a a resolution of 4500, one would obtain 25% of doublets, which is enough to train the pipeline to assign correct redshift. Ongoing preliminary developments on redshift solving softwares show this rather conservative study is a lower limit to the % of doublets we will obtain. On current programs were the redshift is fitted on the complete spectrum, efficiency is higher. The next step for future surveys is now to develop complete end-to-end simulation tools, with a particular focus on the redshift solver.

Chapter 4

Stochastic bias of color-selected BAO tracers by clustering – weak-lensing analysis

Johan Comparat, Eric Jullo, Jean-Paul Kneib, Carlo Schimd, HuanYuan Shan, T. Erben, Olivier Ilbert, Joel Brownstein, A. Ealet, S. Escoffier, B. Moraes, N. Mostek, J. A. Newman, M. E. S. Pereira, F. Prada, D. Schlegel, Donald P. Schneider, Carlos H. Brandt.
Published by MNRAS on june 3rd, 2013 (arXiV: 1302.5655)

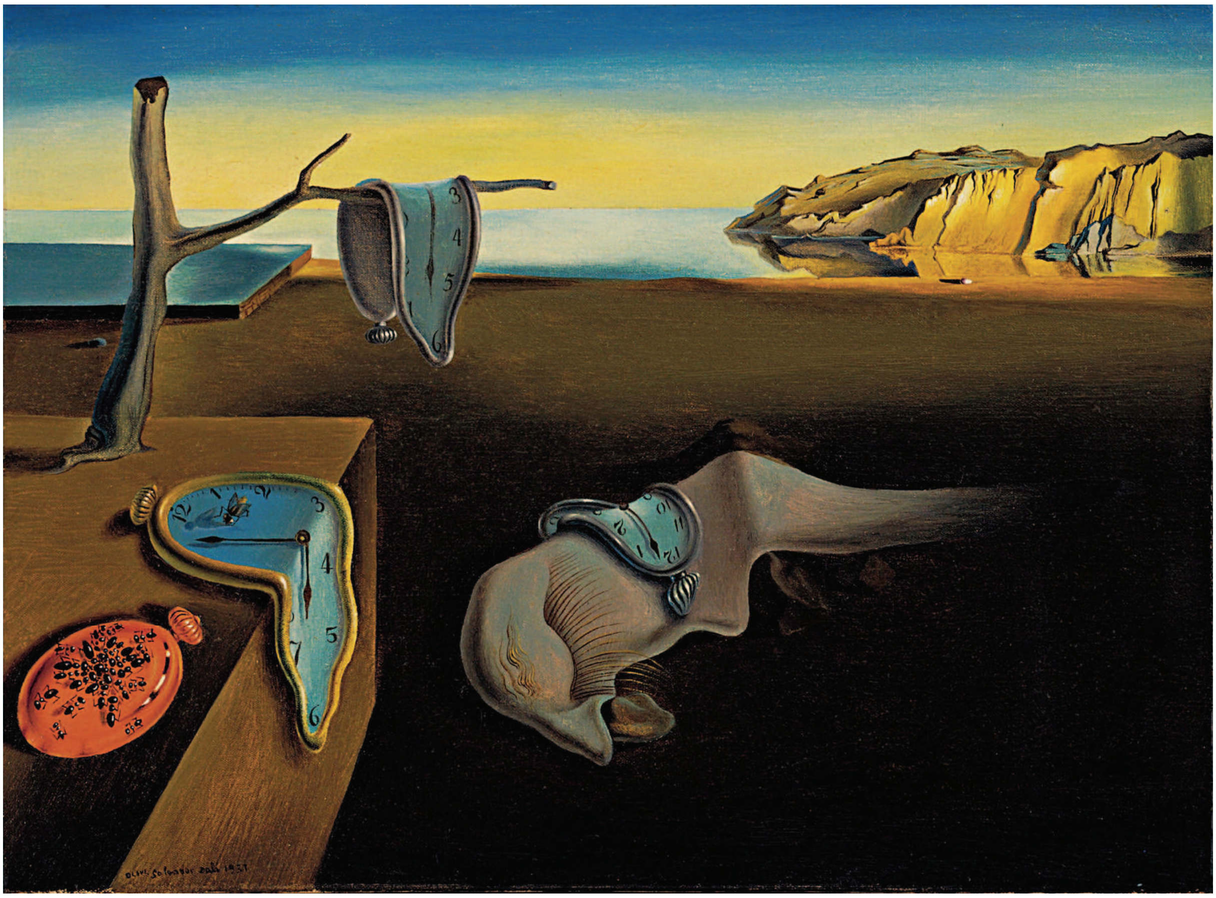


Figure 4.1: Salvador Dali. La persistence de la mémoire. 1931.

4.1 Motivation and contribution

In order to optimize the survey strategy of the eBOSS project, that will start on September 2014, it is important to better understand the clustering properties of the possible BAO tracing targets in order to start the modeling of the targets. The aim being to try and evaluate the systematic uncertainties before observing the targets.

The bulk flows of biased tracers smears the BAO peak and can induce a systematic error on the determination of the size of the standard ruler. This is the reason why I started investigating the galaxy bias of BAO tracers in a broad manner. However for the tracers to be used in future surveys, the lack of spectroscopic redshift to derive clean 3D clustering measurements is a first hurdle. To circumvent this issue, we used deeper photometric fields such as the CFHT-LS and SDSS Stripe-82 deep co-add that have accurate photometric redshifts. Using these surveys, I computed angular correlation functions of BAO tracers. But from this measurement only, it was difficult to obtain a value of the galaxy bias I could firmly believe in. See the mind map of this chapter on Fig. 4.2.

Luckily, in 2010 I participated in a proposal accepted by the CFHT to re-observe the Stripe 82 under better seeing conditions in order to perform a weak lensing survey (known as the CS82 survey). In this proposal, I defined the tiling of the footprint of about 170 square degrees to avoid the bright stars. Once the observations were finished I participated in the data reduction by inspecting the observations, making the star-masks, and other default masks. It was a difficult and long task to accomplish, but very important to understand the difference between a photometry with a seeing of 0.6 arc seconds and that of SDSS co-add.

At that period, E. Jullo arrived at the CNRS/LAM and his expertise in the determination of the galaxy bias came right on time as the shear catalogs were almost finished. I learned a lot from him about using gravitational lensing and the galaxy bias. A bit later I met F. Prada that convinced me to perform a sub halo abundance match with the Multidark simulation to connect for the first time these BAO tracers with a dark matter simulation. This project was very interesting as I learned the physics of weak lensing, and the use of large N-body simulations.

In this context, I wrote the code to derive angular correlation functions, and the fitting procedure for the HOD and SHAM models. I embedded these codes in the ones I wrote about cosmology, distance estimators and power spectrum determination. I build a package 'pyLSS' containing all these classes, with its own documentation, that I distribute to my collaborators (I recently added to this package, the 3D correlation function estimation and the derivation of luminosity functions). When I obtained the first results of galaxy bias for BAO tracers, the packages written by M. Kilbinger 'CosmoPMC' were released, and I could perform the same calculation ten times faster. I therefore switched to the PMC sampling to obtain the final results presented in this work.

This work led me to encounter C. McBride and D. Eisenstein in a BOSS collaboration meeting. They invited me over at Harvard CFA to collaborate with them to work on creating mock catalogs for the ELGs (a two months stay in February and March 2013). This collaboration based on my previous work confirmed what we already found and enabled us to go one step forward in terms of the modeling of the galaxy population targeted by ELG-based BAO surveys. Finally, this collaboration made me chose a post-doc position where I could access and work on N-body simulations to provide SDSS-IV collaboration with mock catalogs to best understand the measurements to be expected from eBOSS and MS-DESi.

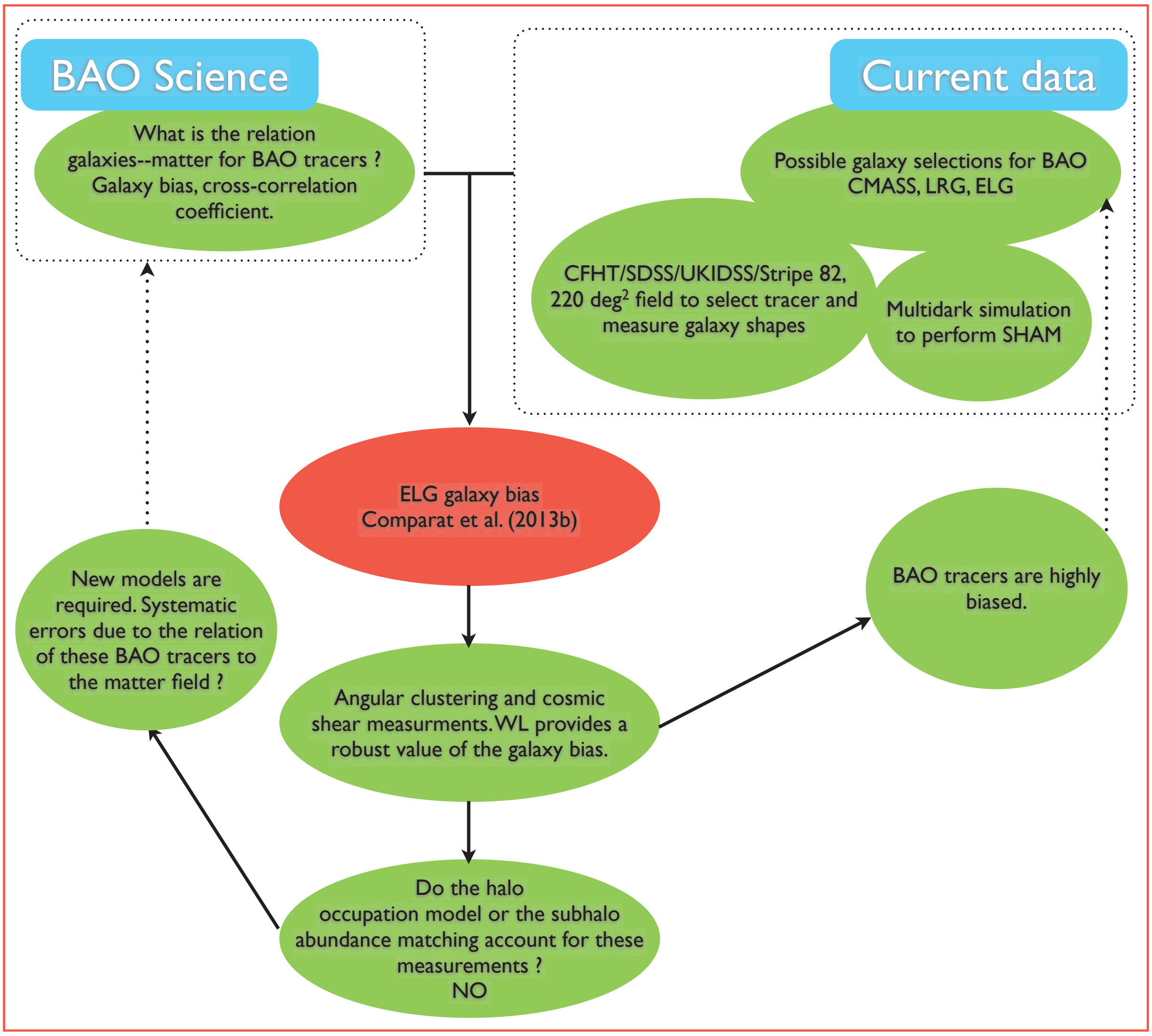


Figure 4.2: Mind map of the chapter about galaxy bias.

4.2 Introduction

Working within the standard cosmological framework described by the Friedmann model, recent analyses of cosmic microwave background (CMB) and type Ia supernovae (SnIa) data confirm the picture of our Universe as almost spatially flat, with an energy content known at the percent level and dominated by dark components, and undergoing a phase of accelerated expansion since the last 7 Gyr (*e.g.* Komatsu et al. 2011; Suzuki et al. 2012). A confirmation of these results is provided by the use of the baryonic acoustic oscillations (BAO) signature in the clustering of galaxies (Frieman et al., 2008), which offers a standard ruler for the measurement of the angular distance-redshift relation. Current measurements of BAO have precision larger than 2%, due to statistical errors. Beutler et al. (2011) provides the latest measurement at $z = 0.106$ on the 6dF Galaxy Survey (Jones et al., 2009) with a precision of 4.5%. Blake et al. (2011b) and Anderson et al. (2012) made the latest measurements at $z \simeq 0.6$ using, the WiggleZ Dark Energy Survey (Drinkwater et al., 2010) and the Sloan Digital Sky Survey III Baryonic Oscillation Spectroscopic Survey (SDSS-III/BOSS hereafter) (Eisenstein et al., 2011; Dawson et al., 2013), respectively. At larger redshift, BAO using the Ly_α forest of quasars have a current precision of the measurement of $H(z = 2.3)$ of about $\sim 3.5\%$; see Pâris et al. (2012); Busca et al. (2012); Lee et al. (2013); Kirkby et al. (2013); Slosar et al. (2013). Future surveys plan detections of the BAO at the sub-percent level in the redshift range, $0.7 \lesssim z \lesssim 2$, to bridge the gap between the aforementioned results. It constitutes the next step toward the investigation of the accelerated expansion of the Universe.

The measurement of BAO in the power spectrum of galaxies relies on the assumption that 'galaxies are tracers of the underlying matter distribution' (Kaiser, 1984). The validity and the limits of this assumption are not yet completely understood theoretically or observationally. To constrain cosmology with the measurement of the BAO feature in the galaxy clustering, it is necessary to understand the relation between the galaxy distribution and the dark matter field. Currently, to bridge the gap between dark matter haloes and galaxies, there are different methods.

The first approach is to use a model that provides how the galaxies are distributed in haloes (Halo Occupation Distribution model; hereafter HOD); see Dekel and Lahav (1999); Ma and Fry (2000); Seljak (2000); Cooray and Sheth (2002). The HOD model is fitted on the measure of the galaxy clustering to predict the properties of the haloes that the galaxies inhabit. In particular, HOD predicts the galaxy bias; see White et al. (2011); Zehavi et al. (2011); Parkinson et al. (2012) for the HOD applied to galaxies used in BAO surveys.

The alternative method is based on weak lensing (WL). With recent surveys, it has become possible to combine on small areas (a few $\deg^2$) different statistical measures of the galaxy distribution: the galaxy clustering, the cosmic shear, and the galaxy galaxy lensing. Using COSMOS data, (Capak et al., 2007; Scoville et al., 2007b; Ilbert et al., 2009) Jullo et al. (2012); Leauthaud et al. (2012) led the way in combining these measures to constrain the galaxy-halo relation and the evolution of the galaxy bias. Improved HOD models are being developed to interpret the combination of both measures (Leauthaud et al., 2011; Cacciato et al., 2012; More et al., 2012; van den Bosch et al., 2012).

The application of such analysis on thousands of square degrees by combining weak lensing with galaxy clustering is however currently limited by the depth and the point spread function (PSF) characterization of the photometry (Huff et al., 2011).

Another possibility to determine the galaxy bias is the direct comparison with large simulations of dark matter using subhalo abundance matching technique (SHAM), (Kravtsov et al., 2004; Vale and Ostriker, 2004; Conroy et al., 2006; Trujillo-Gomez et al., 2011). Applied to the analysis of WiggleZ and SDSS-III/BOSS observations, this technique shows similar results as the HOD; see Prada et al. (2012); Nuza et al. (2012); Parkinson et al. (2012).

In this paper, we investigate the clustering amplitude of galaxies selected to be the BAO tracers of future large spectroscopic surveys as SDSS-IV/eBOSS, BigBOSS, DESspec, PFS-SuMiRe, Euclid (Comparat et al., 2013; Schlegel et al., 2011; Laureijs et al., 2011). The clustering amplitude being proportional to $\propto (b\sigma_8)^2$, we consider a constant $\sigma_8 = 0.81$ (value from CMB analysis WMAP7

Komatsu et al. (2011)) in order to derive the value of the bias b. For this analysis, we combine an HOD approach with a weak lensing analysis. We obtain rather quantitative rough estimates of the galaxy bias, that should not be taken at face value. We make the point for a high clustering amplitude (or a high bias), justifying target selection scheme of future BAO surveys and a more in depth investigation to make forecasts for these surveys accounting for biases at these levels. In fact signal-to-noise for BAO studies scales as the density of tracers times the clustering amplitude of these tracers, and hence BAO studies profit from highly clustered tracers.

In section 4.3 we examine the galaxy bias and the two different methods used to measure it: the halo occupation distribution model and the weak lensing aperture statistics method. We also discuss the limits of each measurement. In section 4.4 we describe the data on which our measurement is based. In section 4.5 we discuss the bias measurement from clustering and weak lensing for each tracer and the associated errors. Finally, in section 4.6 we discuss the feasibility of future BAO studies using the described tracers.

Throughout this paper we assume a flat ΛCDM cosmology characterized by WMAP7 parameters $H_0 = 100$ km s^{-1} Mpc^{-1}, with h=0.71, $\Omega_{\rm CDM} = 0.226$, $\Omega_b = 0.0455$, $\Omega_\Lambda = 0.7285$, $\sigma_8 = 0.81$, $n_{\rm initial} = 0.966$ (Komatsu et al., 2011). Magnitudes are given in the AB system (Oke and Gunn, 1983).

4.3 Galaxy clustering – weak-lensing analysis

Galaxy clustering and gravitational lensing probe, respectively, the galaxy overdensity field $\delta_g(\mathbf{r}, z)$, which is a discrete random variable function of position $\mathbf{r}$ and redshift z, and the matter overdensity field, $\delta_m(\mathbf{r}, z)$, which is a discrete random variable related to the latter by some functional form generically called "bias", eventually allowing for stochasticity (Tegmark and Peebles, 1998; Dekel and Lahav, 1999; Tegmark and Bromley, 1999).[1] As random variables, only the N-point correlation functions (or functions thereof) are meaningful and can be measured; by combining galaxy clustering and gravitational lensing one can measure the auto-correlation functions $\langle\delta_g^2\rangle$ and $\langle\delta_m^2\rangle$ and the cross-correlation $\langle\delta_g\delta_m\rangle$ and eventually define the (linear) bias, b_g, and stochasticity, r, parameters, after Hoekstra et al. (2002)

$$b_g^2 = \frac{\langle\delta_g^2\rangle}{\langle\delta_m^2\rangle}\ ,\quad r = \frac{\langle\delta_g\delta_m\rangle}{\sqrt{\langle\delta_g^2\rangle\langle\delta_m^2\rangle}}. \tag{4.1}$$

A positive (negative) value for r indicates correlation (anticorrelation) between the galaxy and the matter fields, while $r = 1$ ($r = -1$) indicates total correlation (total anticorrelation).

4.3.1 b_g by clustering, an HOD approach

To derive the galaxy bias, we need the auto-correlation of matter and of galaxies. The galaxy auto-correlation function $\langle\delta_g^2\rangle$ is estimated with Landy and Szalay (1993) 'minimum variance estimator'. Then we use a HOD to interpret the measure of auto-correlation, as it provides an insight into the relation between the galaxies and the haloes they inhabit.

The main ingredients of the model are the following. The halo mass function, $n(M, z)$, gives the density of haloes of mass M present at redshift z; see Sheth and Tormen (1999). The large scale halo bias, $b_h(M, z)$, that is taken from Sheth et al. (2001) with the parameters and scale dependence described in Tinker et al. (2005). The auto-correlation of the matter we use is that from Smith et al. (2003). The parameter of the model is a function that gives the number of galaxies present in a halo of mass M, designated $N(M)$.

$$N(M) = \frac{1}{2}\left[1 + erf\left(\frac{\log M - \log M_{\rm min}}{\sigma_{\log M}}\right)\right]\left[1 + \left(\frac{M - M_0}{M_1}\right)^{\alpha}\right] \tag{4.2}$$

[1] For clarity, we drop in the following the space and time dependence from δs.

For a given $N(M)$ the HOD model outputs the corresponding auto-correlation function for galaxies. A fit on the galaxy auto-correlation function is performed to derive the analytical form of $N(M)$. During the fit, the cosmological parameters are fixed at their fiducial value (WMAP7). The large scale galaxy bias of a volume-limited sample of galaxies described by $N(M)$ is obtained with equation (4.3).

$$\langle b_g \rangle = \frac{\int n(M,z) b_h(M,z) N(M) dM}{\int n(M,z) N(M) dM}. \tag{4.3}$$

For more details about the implementation of the HOD we use in this study, see Cooray and Sheth (2002); Tinker et al. (2005); Coupon et al. (2012). We compute correlation functions with the software ATHENA[2] and the HOD fits are performed using CosmoPMC[3]; see Kilbinger et al. (2011, 2009, 2010); Coupon et al. (2012).

The main limit of the model is the input auto-correlation of matter. This information is taken from Smith et al. (2003) and has errors of $\sim 8\%$ at small scales and $< 3\%$ at large scales $k < 10h\,\mathrm{Mpc}^{-1}$ in the ΛCDM paradigm. Also the HOD model is not valid in different cosmological paradigms and can therefore not be used to rule out cosmologies: it can only be used to replace galaxy samples in the galaxy clustering history, which is the aim of this paper.

4.3.2 b_g and r by weak-lensing

Using weak-lensing analysis, we derive both b_g and r. For this purpose, we need the three auto and cross-correlations mentioned before. The WL measurement is based on the analysis of the distribution of shapes of the galaxies located behind the BAO tracers. Their intrinsic shapes are distorted by the foreground mass around the BAO tracers. The auto-correlation of the ellipticities of the background galaxies contains the information of the matter clustering between the background galaxies and us, $\langle \delta_m^2 \rangle$. The cross-correlation of the ellipticities with the position of the foreground galaxies contains the information of the cross-correlation $\langle \delta_g \delta_m \rangle$. The angular auto-correlation $\langle \delta_g^2 \rangle$ of galaxies is the same as for the HOD.

To be able to compare the auto- and cross-correlation results, the measurements are convolved with Bessel kernels to transform the correlations in aperture correlations. This formalism is described in Schneider et al. (1998) and applied on RCS data in Hoekstra et al. (2002), GaBoDS data in Simon et al. (2007), and on COSMOS in Jullo et al. (2012). We use the same routines as in Jullo et al. (2012) to measure the aperture auto-correlation of galaxies, denoted $\langle \mathcal{N}_{ap}^2(\theta) \rangle$, that of matter, denoted $\langle M_{ap}^2(\theta) \rangle$, and the cross-correlation, denoted $\langle \mathcal{N}_{ap}(\theta) M_{ap}(\theta) \rangle$ defined from the power spectrum of the galaxies P_n, of the matter P_κ, and their cross-power spectrum $P_{n\kappa}$.

$$\langle \mathcal{N}_{ap}^2(\theta) \rangle = 2\pi \int_0^\infty dl\, l P_n(l) \left[\frac{12 J_4(l\theta)}{\pi (l\theta)^2} \right]^2 \tag{4.4}$$

$$\langle M_{ap}^2(\theta) \rangle = 2\pi \int_0^\infty dl\, l P_\kappa(l) \left[\frac{12 J_4(l\theta)}{\pi (l\theta)^2} \right]^2 \tag{4.5}$$

$$\langle \mathcal{N}_{ap}(\theta) M_{ap}(\theta) \rangle = 2\pi \int_0^\infty dl\, l P_{n\kappa}(l) \left[\frac{12 J_4(l\theta)}{\pi (l\theta)^2} \right]^2 \tag{4.6}$$

where J_4 is the Bessel function of order 4, see Schneider et al. (1998); Simon et al. (2007).
The first output is the galaxy bias in an aperture of an angular radius θ; see equation (4.7). The second measurement is the cross-correlation coefficient r, that measures the randomness; see equation (4.8).

$$b(\theta) = f_1(\theta, \Omega_m, \Omega_\Lambda) \sqrt{\frac{\langle \mathcal{N}_{ap}^2(\theta) \rangle}{\langle M_{ap}^2(\theta) \rangle}}, \tag{4.7}$$

[2] http://www2.iap.fr/users/kilbinge/athena/
[3] http://www2.iap.fr/users/kilbinge/CosmoPMC/

$$r(\theta) = f_2(\theta, \Omega_m, \Omega_\Lambda) \frac{\langle \mathcal{N}_{ap}(\theta) M_{ap}(\theta) \rangle}{\sqrt{\langle \mathcal{N}^2_{ap}(\theta) \rangle \langle M^2_{ap}(\theta) \rangle}}. \tag{4.8}$$

The functions f_1 and f_2 are normalizations. They correct for the fact that different cosmological volumes are probed by the different statistics. The functions are computed assuming a Smith et al. (2003) non-linear power spectrum with unbiased foreground galaxies ($b = r = 1$), that uses the Eisenstein and Hu (1998) transfer function, constrained by the WMAP7 cosmological parameters (Komatsu et al., 2011).

The WL also uses the Smith et al. (2003) non-linear power spectrum and has therefore the same precision limitations as the HOD. Moreover, an additional constraint is that, to be able to extract the WL signal, the sources must be behind the lenses. Because of this point, we are not able to have a significant measure of the matter galaxy cross-correlation by WL for the faintest galaxy samples (see section 4.4).

4.4 Data

In this section, we present the data sets that are used for this analysis. First, we explain the catalogs used to target galaxies tracing BAO, and we detail the selection functions for the BAO tracers. Then we describe the dark matter halo simulation we used. Finally, we describe the catalogs containing the galaxy shape measurements.

4.4.1 SDSS-Stripe 82 galaxy catalog

The Sloan Digital Sky Survey (York et al., 2000; Eisenstein et al., 2011), delivered under the data release 9 (DR9, Ahn et al. 2012), covers 14,555 square degrees in the 5 photometric bands u, g, r, i, z. It is the largest volume multi-color extragalactic photometric survey available today. The 3σ magnitude depths are: $u = 22.0$, $g = 22.2$, $r = 22.2$, $i = 21.3$; see Fukugita et al. (1996) for the description of the filters, Gunn et al. (1998) for the characteristics of the camera, and Gunn et al. (2006) for the description of the 2.5m telescope located at Apache Point Observatory (New Mexico, USA).

The data set used here for the selection of the different BAO tracers is the Sloan Digital Sky Survey deep Stripe 82 coadd (Annis et al., 2011). It covers an area of 225 deg^2 in one continuous stripe $-1.25° < \delta_{\rm J2000} < 1.25°$, $-34° < \alpha_{\rm J2000} < 46°$ with five photometric bands u, g, r, i, z. We also use the infrared counterpart of this area observed by WISE at $3.4\mu m$ (Wright et al., 2010). Table 4.1 gives an overview of the tracers selected in this analysis, see section 4.4.2.

The photometric redshift distribution of the BAO tracers are obtained using the CFHT-LS Wide W4 field photometric redshift catalog T0007. The photometric redshift accuracy is estimated to be $\sigma_z < 0.08(1 + z)$ for $i \leq 24$. The data and cataloging methods are described in Ilbert et al. (2006); Coupon et al. (2009), the T0007 release document[4]. The CFHT-LS W4 field overlaps the stripe 82 on a few square degrees. Both datasets have similar depths in magnitude. From a position match of the targets selected on the stripe 82, we obtain a sub-catalog of the targets with photometric redshifts, from which we infer the redshift distribution of the selections. The redshift distributions are shown on Fig. 4.3.

With current data, it is not feasible to obtain a reliable photometric redshift on the complete stripe 82. Therefore, we cannot construct clean volume limited samples, as required by the HOD analysis. We can only make magnitude limited samples. The impact of the galaxies located in the tails of redshift distribution on the angular clustering is discussed in paragraph 4.5.1.

[4] http://www.cfht.hawaii.edu/Science/CFHTLS/

Table 4.1: Samples of BAO tracers used in this analysis, 'b' stands for bright and 'f' for faint. $\bar{z}$ is the mean redshift of the sample, σ_z is the dispersion of the redshift distribution, N the total amount of tracers in the sample, N_S the density of tracers per deg^2, '$\perp$ scale' is the transverse scale in h^{-1} Mpc corresponding to one degree at $\bar{z}$.

sample	$\bar{z}$	σ_z	N	N_S deg^{-2}	$\perp$ scale h^{-1} Mpc deg^{-1}
CMASS	0.53	0.10	22k	98	16.0
LRG-WISE b	0.58	0.12	15k	70	17.8
LRG-WISE f	0.69	0.13	59k	275	18.5
ELG gri b	0.80	0.08	119k	530	19.2
ELG gri f	0.81	0.15	226k	1007	19.3
ELG ugri b	0.95	0.24	156k	693	20.3
ELG ugri f	0.94	0.22	436k	1938	20.2
ELG ugr b	1.28	0.36	77k	341	21.5
ELG ugr f	1.26	0.34	237k	1052	21.5

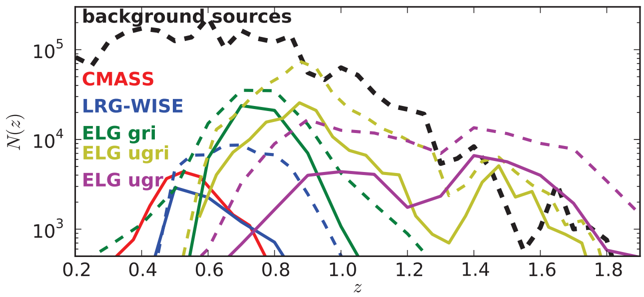


Figure 4.3: Redshift distributions of the samples. For each sample, solid (dashed) lines refer to the bright (faint) sample. 22 000 CMASS (red), 15 000 (59 000) LRG-WISE (blue), 119 000 (226 000) *gri* ELG (green), 156 000 (436 000) *ugri* (yellow), 77 000 (237 000) *ugr* ELG (magenta). The dashed black line represents the redshift distribution of the background sources used in the weak-lensing analysis.

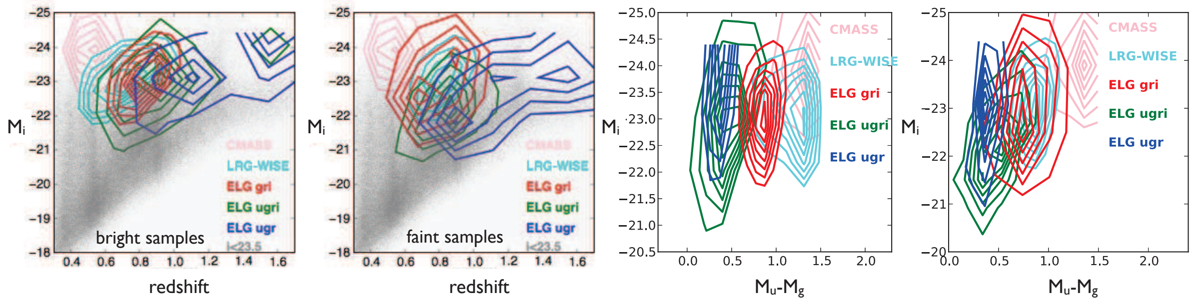


Figure 4.4: Absolute magnitude M_i vs redshift and rest-frame M_u-M_g. All galaxies with $i < 23.5$ (grey), CMASS (pink), LRG-WISE (light blue), *gri* ELG (red), *ugri* ELG (green), *ugr* ELG (blue). First panel shows the bright selections are quite luminous with $M_i \sim -23$. The second panel shows the faint selections are less luminous than the bright selections with $M_i \sim -22$. The third panel shows the bright selections span the complete rest-frame color range from *ugr* ELG (blue) to the CMASS (pink). The fourth panel shows the faint selections. Their rest frame colors are similar to that of the bright selections, except for the LRG-WISE faint galaxies that are bluer than the bright sample. The *ugri* ELG overlap with the *gri* and *ugr* ELG selections, which is expected as the selection encompasses galaxies from both other selections. This figure was made with the same CFHT W4 matched sample as for the redshift distributions.

4.4.2 BAO tracer selections

The latest BAO tracers used at $z > 0.5$ are SDSS-III Constant Mass galaxies (CMASS) at $z \simeq 0.53$ (Dawson et al., 2013) and WiggleZ Emission Line Galaxies at $z \simeq 0.6$ (Drinkwater et al., 2010). The next stage BAO experiments plan to use Luminous Red Galaxies (LRG) and Emission Line Galaxies (ELG) at $z > 0.7$ (Schlegel et al., 2011). The target selection of ELG tracers is extensively discussed in chapter 2 taken from Comparat et al. (2013).

The selection of the tracers used in this analysis is only based on photometric criteria, depth and color. We do not use the photometric redshifts for the selection as their quality varies strongly across the stripe. The use of photometric redshift to target ELGs would make the selection function very complicated. The selection criteria target the most luminous galaxies in incremental redshift intervals to $z = 2$; see Fig. 4.4 which compares all samples in absolute magnitude M_i versus redshift or rest-frame M_u-M_g. The cleanest way to select BAO tracers would be to select by stellar mass, but with current data this is not possible.

4.4.2.1 CMASS at $z \simeq 0.5$

The CMASS tracer is the SDSS-III/BOSS experiment BAO tracer (Dawson et al., 2013). The current BAO detection using the data release 9 (Ahn et al., 2012) (a third of the observation plan) with the CMASS tracers has a 6.7σ significance (Anderson et al., 2012). Its selection is described in detail in Dawson et al. (2013). The parent catalog of CMASS selection on stripe 82 contains 22034 tracers *i.e.* ~ 98 deg^{-2}. We use the complete CMASS selection, not only the galaxies confirmed by spectroscopy, in order to avoid fiber collision issues. The redshift distribution is shown by the red solid line in Fig 4.3. The mean redshift is 0.53 with a dispersion of 0.1. We use snapshot number 62 of the multidark simulation. The density of the tracer at $\bar{z}$ is $2.64 \times 10^{-4} h^3$Mpc^{-3}, which corresponds to a circular velocity cut of $V_{\rm cir} = 384.1$ km s^{-1}. This value of V_{cir} is consistent with the study done by Nuza et al. (2012). The halo mass distribution peaks at $\log_{10} M_{200}/M_\odot = 13.17^{+0.34}_{-0.19}$. The minimum satellite fraction expected is 5%.

4.4.2.2 LRG-WISE at $z \simeq 0.65$

This selection is the extension in redshift of the red CMASS selection. We name this sample the 'Luminous Red Galaxies selected with WISE' (LRG-WISE); see Schlegel et al. (2011), Prakash et al. (2013), in preparation. On stripe 82, it selects 15643 galaxies *i.e.*, a density of 70 $\deg^{-2}$ at a mean redshift of 0.58 for the bright sample with $i < 20$ and ~ 59000 galaxies at a mean redshift of 0.69 for the faint sample with $i < 21$ *i.e.*, a density of 262 $\deg^{-2}$. The selection is as follows :

- we use dereddened r,i,z and $3.4\mu m$ (WISE 1) magnitudes
- SDSS z-band flag must be true, and WISE channel 1 flag must be 0
- Objects with extreme astrometric uncertainties were rejected
- bright selection : ($i < 20$ or $z < 20$) *and* $r - i > 1$ *and* $r - W1 > 2(r - i) - 0.5$ *and* $r > 16$ *and* $r < 22.5$
- faint selection : ($i < 21$ or $z < 21$) *and* $r - i > 1$ *and* $r - W1 > 2(r - i) - 0.5$ *and* $r > 16$ *and* $r < 22.5$

The redshift distribution of the bright (faint) samples are shown with a solid (dashed) blue line on Fig. 4.3. Their densities are 1.57 and $5.06 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, respectively. We use the multidark snapshot 58 for the bright and the faint samples SHAM. The samples correspond to cuts at $V_{\mathrm{cir}} = 431.5$ km s^{-1} and 323.6 km s^{-1}, respectively. The halo mass distributions peak at $\log_{10} M_{200}/M_\odot = 13.28^{+0.31}_{-0.17}$, and $\log_{10} M_{200}/M_\odot = 12.92^{+0.36}_{-0.2}$, respectively. They constitute an upper limit for the halo mass these tracer inhabit. The minimum satellite fractions expected are 4% and 7%, respectively.

4.4.2.3 ELG *gri* at $z \simeq 0.8$

The *gri* ELG selection, on the top-right panel of Fig. 2.5, is an intermediate selection between the reddest galaxies selected in SDSS-III and the bluest ones selected by the *ugr* and *ugri* schemes (see below), identifies emission-line galaxies with emission lines a small Balmer break. On Stripe 82, it selects $\approx 119\,000$ and $226\,000$ galaxies for the bright and faint samples, which correspond to a projected density of about 530 and 1000 $\deg^{-2}$ at mean redshift $z \simeq 0.8$ and 0.81, respectively. The bright sample is limited by $19 < i < 21.5$, and the faint sample by $19 < i < 22$. The densities are respectively 8.16 and 13.37 $\times 10^{-4}\, h^3 \mathrm{Mpc}^{-3}$. We use multidark snapshot number 56 to perform the SHAM. The samples match with a circular velocity cut of 292.2, 250.7 km s^{-1}. The halo mass distributions peak at $\log_{10} M_{200}/M_\odot = 12.77^{+0.37}_{-0.21}$, and $\log_{10} M_{200}/M_\odot = 12.34^{+0.44}_{-0.25}$, respectively. They constitute an upper limit for the halo mass these tracer inhabit. The minimum satellite fractions expected are 8% and 10%, respectively.

4.4.2.4 ELG *ugr* at $z \simeq 1.2$

The *ugr* ELG selection, on the top-left panel of Fig. 2.5, selects the blue tail of the color distribution aiming for galaxies with strong emission lines at higher redshift ($z \simeq 1.2$). The *ugr* redshift distribution is flatter and extends to higher redshifts. The bright sample is limited by $20 < g < 22.5$, and the faint sample by $20 < g < 23$. For the bright and faint sample it selects, 77 000 and 237 000 galaxies, corresponding to a density of 341 and 1052 $\deg^{-2}$, respectively. Their mean redshift are at 1.28 and 1.26. We use multidark snapshot number 48. The bright sample density, $1.45 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, corresponds to a cut at 434.5 km s^{-1}, which gives an upper limit of $\log_{10} M_{200}/M_\odot \lesssim 13.21^{+0.29}_{-0.16}$. The faint sample density, $5.76 \times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, corresponds to a cut at 315.0 km s^{-1}, which gives an upper limit of $\log_{10} M_{200}/M_\odot \lesssim 12.81^{+0.34}_{-0.19}$. The minimum satellite fractions expected are 3% and 7%, respectively.

4.4.2.5 ELG *ugri* at $z \simeq 0.95$: building SDSS-IV/eBOSS

The *ugri* ELG selection is designed to be the BAO tracer sample for SDSS-IV/eBOSS project (eBOSS collaboration 2013 in preparation). It is a mix of the two previous *gri* and *ugr* ELG selections. The selection criteria are

- $20 < g < 22.5$ for the bright; $20 < g < 23$ for the faint
- $-0.5 < u - r < 0.7(g - i) + 0.1$ and $g - i > -0.5$

The mean redshifts are 0.95 and 0.94. The bright sample contains 156 000 galaxies and the faint sample 436 000 (*i.e.* densities of 693 and 1938 deg^{-2}). We use multidark snapshot number 52 for the SHAM. The densities are respectively 5.44 and 16.45 $\times 10^{-4} h^3 \mathrm{Mpc}^{-3}$, corresponding to mass upper limits of $\log_{10} M_{200}/M_\odot \lesssim 12.84^{+0.35}_{-0.19}$ and $\log_{10} M_{200}/M_\odot \lesssim 12.48^{+0.4}_{-0.23}$. The minimum satellite fractions expected are 7% and 11%, respectively.

4.4.3 Multidark-Stripe 82 halo catalog

We use the Multidark dark matter simulation to characterize qualitatively our samples to the first order using sub-halo abundance matching (SHAM). To perform the match, we use snapshots of the Multidark simulation corresponding to the mean redshift of our tracers. Every snapshot occupies a volume of 1 $h^{-3}\mathrm{Gpc}^3$ filled with haloes obtained with bound density maximum algorithm with a cut-off at 200 times the critical density ($\log_{10} M_{200}/M_\odot$). These data sets are taken from the database Multidark[5] (Riebe et al., 2011).

In each snapshot, we select the dark matter halo population that corresponds to the number density of galaxies with $\bar{z} - \sigma_z < z < \bar{z} + \sigma_z$ selected on the Stripe 82. This links each galaxy sample to a halo sample. The density, velocity cuts and characteristic masses and satellite fractions are given for each sample in Table 4.2.

The SHAM brings two components to this analysis. Under the hypothesis that the galaxy sample is complete in mass. The abundance matched subhalo sample has a mass distribution characterized by the three quartiles of $\log_{10} M_{200}/M_\odot$ and a satellite fraction f_{sat}. If the galaxy sample is not complete in mass, the mass distribution of the subhaloes the galaxies live in will be shifted towards lower masses and the satellite fraction will increase. As we cannot demonstrate that the BAO tracers selected are complete in mass and that they are tracing the most massive structure at their redshift, we consider the subhalo mass scales obtained with the SHAM procedure are an upper boundary, and that the satellite fraction is a lower limit to the true values. Secondly, it provides the mass distribution of the subhaloes. Given that we measure clustering over ~ 200 deg^2 and weak lensing on ~ 150 deg^2 and not the full sky, we check the mean halo mass deduced from the maximum likelihood HOD model is between the first and third quartile of the sub halo mass distribution given by SHAM.

From the qualitative numbers obtained by the SHAM, the LRG-WISE faint sample appeasr to be located in massive subhalos of $\sim 10^{13} M_\odot$ with a low satellite fraction, similarly to the CMASS tracers. Concerning the ELG, the satellite fraction are higher and the subhalo masses are lower: $\lesssim 10^{12.8} M_\odot$ for the bright samples and $\lesssim 10^{12.6} M_\odot$ for the faint samples.

4.4.4 CFHT-Stripe 82 weak-lensing catalog

The existing comprehensive data sets available on the SDSS Stripe-82 area were complemented with high-quality i'-band observations at CFHT. The main purpose of this one-band Megaprime@CFHT (see Boulade et al. 2003) follow-up survey is to enrich already existing multi-colour data with deep optical observations suitable for weak lensing studies. The CFHT Stripe-82 program (CFHT/Stripe82) is a large collaborative effort between the Canadian/French and Brazilian communities[6]. The observations

[5] http://www.multidark.org/

[6] The survey CFHT identification numbers are 10BB09, 10BC22, and 10BF23

Table 4.2: Abundance matching of the samples of BAO tracers used in this analysis, 'b' stands for bright and 'f' for faint. The density of tracers is in units of $10^{-4}\,h^3\mathrm{Mpc}^{-3}$. $V_{\rm cir}$ or $V_{\rm max}$ is the velocity cut applied to the multidark simulation snapshot to match the density d. The obtained sample is characterized by the three quartiles of its distributions given in the columns $\log_{10} M_{200}/M_\odot$. f_{sat} is the fraction of satellites in the halo catalog.

sample	density d	$V_{\rm cir}$ km s^{-1}	$\log_{10} M_{200}/M_\odot$ M_d	f_{sat}
CMASS	2.64	384	$13.17^{+0.34}_{-0.19}$	5%
LRG-WISE b	1.57	431	$13.28^{+0.31}_{-0.17}$	4%
LRG-WISE f	5.06	324	$12.92^{+0.36}_{-0.2}$	7%
ELG gri b	8.16	292	$12.77^{+0.37}_{-0.21}$	8%
ELG gri f	13.37	251	$12.59^{+0.41}_{-0.23}$	10%
ELG ugri b	5.44	317	$12.84^{+0.35}_{-0.19}$	7%
ELG ugri f	16.45	238	$12.48^{+0.4}_{-0.23}$	11%
ELG ugr b	1.45	435	$13.21^{+0.29}_{-0.16}$	3%
ELG ugr f	5.76	315	$12.81^{+0.34}_{-0.19}$	7%

encompass 169 MegaPrime new pointings resulting in an effective survey area of about 150 deg^2. This area comprises 169-4(due to fringing)=165 pointings CFHT/Stripe82 and 8 CFHT-LS Wide pointings *i.e.* a total of 173 pointings. All the data were obtained under excellent seeing conditions, and the final co-added images show an image quality between 0.4 and 0.8 arcsec with a median of 0.59 arcsec. Each pointing was obtained in four dithered observations with an exposure time of 410s, each resulting in a 5-σ limiting magnitude in a 2 arcsec diameter aperture of about 24.

The data processing, which starts with the Elixir preprocessed images obtained from CADC and ends with co-added images and all necessary quantities to perform weak gravitational lensing studies, follows closely the descriptions given in Erben et al. (2012).

We exploit the opportunity that the CFHTLS includes the HST COSMOS survey field, and verify the calibration of our shear measurement pipeline for ground-based data against an independent analysis of much higher resolution space-based data. We stack the subsets of the CFHTLS DEEP2 imaging to the same depth as the CFHT Stripe 82 survey and analyze it using KSB method (Kaiser et al., 1995). The KSB method is used to measure the shapes of all the selected galaxies. Our implementation of KSB is based on the KSBf90[7] pipeline (Heymans et al., 2006b). This has been generically tested on simulated images containing a known shear signal as part of the Shear TEsting Programme (STEP: Heymans et al. (2006a); Massey et al. (2007)) and the GRavitational lEnsing Accuracy Testing challenge (GREAT08: Bridle et al. (2010); GREAT10: Kitching et al. (2012)). The systematic errors of this pipeline are well controlled in all cases. This pipeline has also been used on the shear measurement of CFHTLS-Wide W1 fields (Shan et al., 2012). Finally, we match galaxies in the CFHT/Stripe82 catalog to those of the HST COSMOS catalog (Leauthaud et al., 2010) to calibrate the measurements.

Here is the description of the calibration tests we made on the lensing catalog. The multiplicative calibration component is $(m_1, m_2) = (0.13, 0.14)$ and the additive component is $(c_1, c_2) = (0.0003, 0.0015)$. These coefficients are obtained from simulations and connect through the linear relation between the observed galaxy ellipticities and the true ones ($e_{\rm obs} = (1+m)[\gamma + e_{\rm true}] + c$). They mainly depend on the algorithm that measures the shapes.

If the PSF anisotropy is small, the shear γ can be recovered to first-order the observed ellipticity

[7] `http://www.roe.ac.uk/~heymans/KSBf90/Home.html`

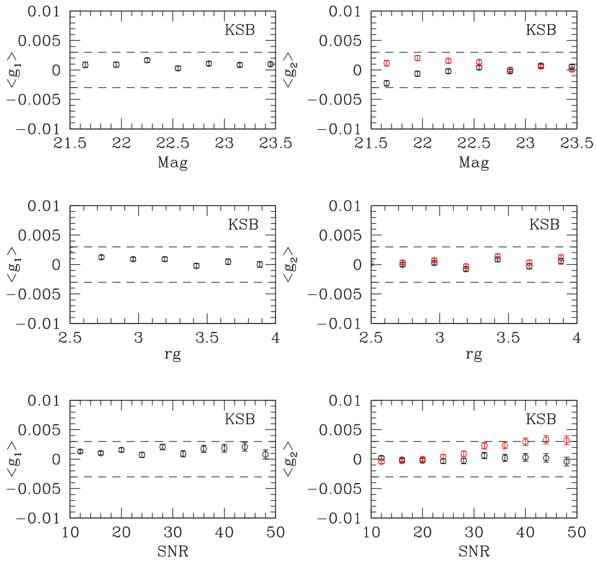


Figure 4.5: Average shear measurements $\langle g_1 \rangle$ (left) and $\langle g_2 \rangle$ (right) of KSB from galaxies in i'-band observations of the entire CFHT/Stripe82 field. In the absence of additive systematics, these should be consistent with zero. In practice, they always remain within the dashed lines that indicate an order of magnitude lower than the 1–10% shear signal around clusters. Top, middle and bottom panels show trends as a function of galaxy magnitude, size and signal-to-noise ratio. The black points denote the average shear from all the galaxies in the two shear catalogs. The red points denote $\langle g_2 \rangle$ before the correction described in (i).

$e^{\rm obs}$ of each galaxies via

$$\gamma = P_\gamma^{-1}\left(e^{\rm obs} - \frac{P^{\rm sm}}{P^{\rm sm*}}e^*\right), \tag{4.9}$$

where asterisks indicate quantities that should be measured from the PSF model interpolated to the position of the galaxy, $P^{\rm sm}$ is the smear polarisability, and P_γ is the correction to the shear polarisability that includes the smearing with the isotropic component of the PSF. The ellipticities are constructed from a combination of an object weighted quadrupole moments, and the other quantities involve higher order shape moments. All definitions are taken from Luppino and Kaiser (1997). With the assumption $\kappa \sim 0$, the reduced shear $g = \frac{\gamma}{1-\kappa} \sim \gamma$.

In order to check for significant additive systematics we performed four tests.

1. We test the average shear $\langle \gamma \rangle$ across all 173 pointings of the CFHT/Stripe82 field. Fig. 4.5 demonstrates that the average shear is almost consistent with zero as expected, for galaxies of all sizes, magnitudes and signal-to-noise ratios (SNR). As in Heymans et al. (2012), $\langle g_2 \rangle$ is clearly dependent on the SNR (red points). We quantify the additive component of the calibration correction with a simple relation $\langle g_2 \rangle = a + b\,SNR$. We find a best fit reduced $\chi^2 = 0.54$ with $a = -0.002, b = 0.000118$. The corrected average shear is almost consistent with zero as expected, for galaxies of all sizes, magnitudes and signal-to-noise ratios (SNR). Furthermore, the average shear of each pointing has also been tested, see Fig. 4.6. We find that the mean shear $\langle g_1 \rangle$ of pointing S82p24p is around -0.0126 ± 0.0017, which is far from zero compared to other observation pointings. Systematic error residual signal may exist in this pointing. We remove the related galaxies in the final catalogue.

2. Star-galaxy cross correlation (residual systematics due to imperfect PSF). For the total 173 pointings, we find our PSF correction is well within requirements for our analysis. Fig. 4.7 shows the correlation $\xi_{\rm sys}$ between the corrected shapes of galaxies and the uncorrected shapes

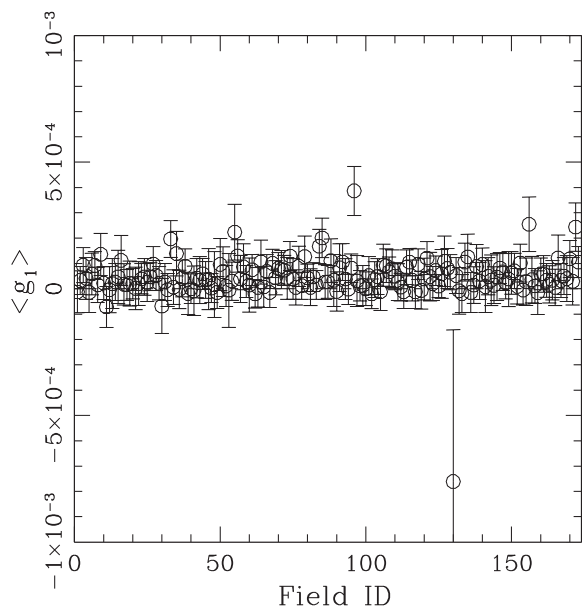


Figure 4.6: Average shear $\langle g_1 \rangle$ of 173 pointings, in which S82p24p field is biased from zero.

of the stars. We normalize the star-galaxy ellipticity correlation by the uncorrected star-galaxy ellipticity correlation to assess its impact on shear measurements (Bacon et al., 2003); see Fig. 4.8. Though, we find ξ_{sys} is obviously higher than ξ_+ in one pointing S82p35p at scale $\theta >$ 3arcmin, which reveals the presence of a systematic error residual signal in this pointing. We remove the related galaxies in the final catalogue (Only the field with $\delta_\xi < 0$ is excluded).

3. Two-point shear correlation functions. We calculate the two-point shear correlation functions of the shear catalogue and compare them to an analytical prediction. The comparison between the analytical prediction of halo model and the measured ξ_+ is good, see Fig. 4.9. Note the cross-correlation $\xi_{\rm tx}$ between the tangential and rotated shear should be zero.

4. E, B-mode signals. We test for the (nonphysical) B-mode signal. The B-mode signal corresponds to the imaginary component of $P_\kappa(\ell)$. We rotate all galaxy shears through 45° and remeasure the E-mode signal. We measure that the B-mode signals are around zero except for small scales; see Fig. 4.10.

For the lensing catalog, we only select galaxies fainter than the foreground samples by 1 magnitude in i-band, so that we have no overlap between the two samples. The complete redshift distribution of the sources used for the weak lensing analysis is displayed on Fig. 4.3 with a thick dashed black line labeled 'background sources'.

4.5 Clustering and lensing measurements

In this section, we present the measurements made through clustering and weak lensing analysis of the data. The correlation function measurements are displayed on Figs. 4.11-4.15. Figs. 4.16 and 4.17 show the galaxy bias measurement. Table 4.3 summarizes the details of the measurements.

4.5.1 Angular clustering measure and HOD models

We measure the angular clustering, denoted $w(\theta)$, for angles in the range 0.001° to 1.5°. In this manner we avoid the signal due to blended sources. In the case of LRG-WISE, the clustering analysis starts

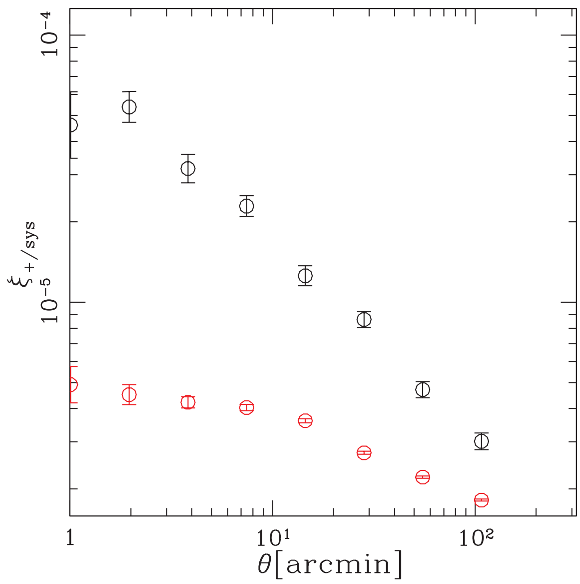


Figure 4.7: The cross-correlation between shear measurements and stellar ellipticities as a function of the separation between galaxies and stars, averaged throughout the CFHTLS Stripe 82 field. The black and red points are shear correlation function and star-galaxy cross correlation, respectively. If all residual influence of the observational PSF has been successfully removed from the galaxy shape measurements, the red points should be consistent with zero.

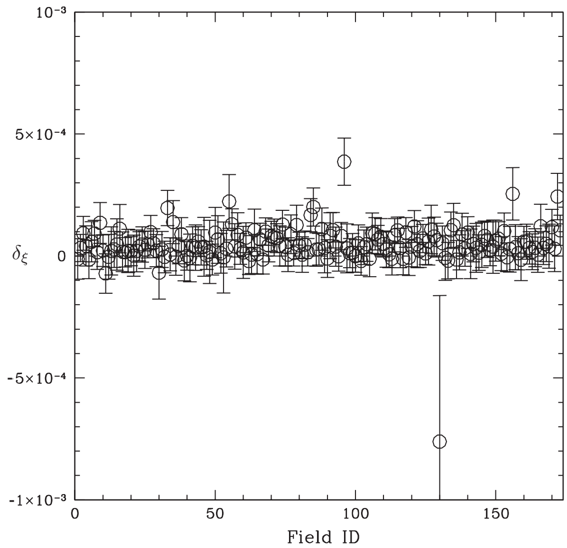


Figure 4.8: Difference δ_ξ between shear auto-correlation ξ_+ and star-galaxy cross correlation $\xi_{\rm sys}$ at scale $\theta = 3$arcmin. The difference delta of the S82p35p pointing is higher than others.

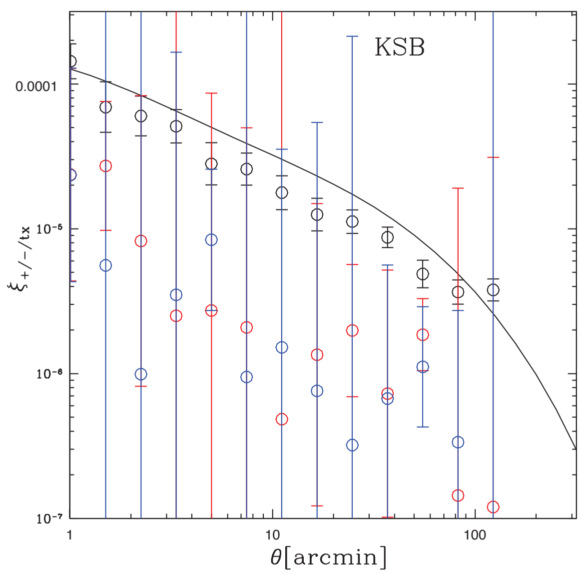


Figure 4.9: The two-point shear correlation functions of the shear catalogue: ξ_+ (black points), ξ_- (red points), and $\xi_{\rm tx}$ (blue points). Solid lines show the analytical predictions of ξ_+. Note the cross-correlation $\xi_{\rm tx}$ between the tangential and rotated shear should be zero. The solid line is the theoretical prediction for a WMAP7-cosmology.

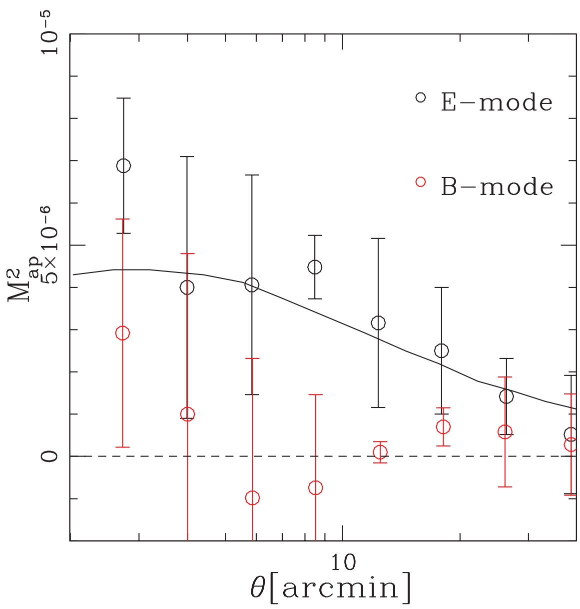


Figure 4.10: The aperture mass from the entire CFHT/Stripe82 field. Red points show the B-mode, black points show the E-mode. The error bars of the E-mode include statistical noise added in quadrature to the non-Gaussian cosmic variance. Only statistical uncertainty contributes to the error budget for the B-mode. The errorbars are 1σ confidence level. Within 2σ errors the B-modes are consistent with zero. The solid line is the theoretical prediction for a WMAP7-cosmology.

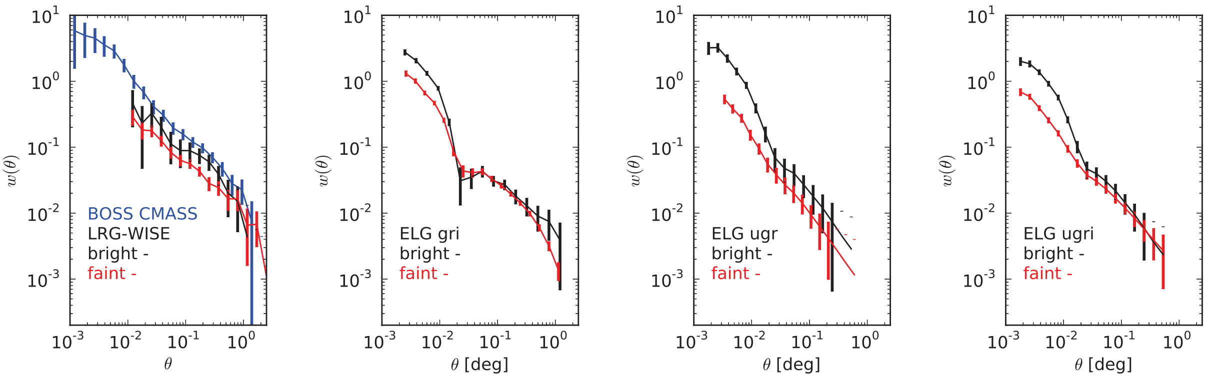


Figure 4.11: Angular clustering of the selected tracers. The wide angular span of the clustering allows the HOD analysis, except for the LRG-WISE, that are limited to $\theta > 0.01°$. On smaller scales than 0.002, the number of pairs available to evaluate the 2PCF drops and the error bars go up; this is cut out of these plots.

at 0.01°. We use between 15 and 20 angular bins regularly log-spaced. The errors and the covariance matrix are obtained from 112 bootstrap realizations of the angular clustering measure. Fig. 4.11 shows a comparison of the angular clustering of all previously defined samples.

The most precise estimations of the angular correlation functions are between 0.01° and 0.1°, due to a balance between the density of BAO tracers and the geometry of the survey. For the sparse tracers (the bright selections at low redshift), below 0.01°, there are only a few pairs of tracers. For the densest tracers, the correlation function is precise down to 0.001°. For all tracers above 0.1°, given the geometry of the survey (a long stripe), the correlation function measurements are covariant. We correct from the integral constraint as in Wall and Jenkins (2012).

The main trends in the angular clustering measurements follow expectations. The lower redshift samples are more clustered than higher redshift samples. The bright samples are more clustered than the faint samples in particular at small scales.

To interpret the angular clustering, we use the HOD parametrization given in (4.2), where $erf(x) = \frac{2}{\pi}\int_0^x e^{-t^2}dt$ is the 'error function'. From $N(M)$ we deduce the angle averaged 3D correlation function, that we project using the Limber equation (Limber, 1954; Simon, 2007) to obtain $w_{\rm HOD}$. The χ^2 accounting for the accuracy of the fit is defined by equation (4.10).

$$\chi^2 = (w_{\rm observed} - w_{\rm HOD})^T.C^{-1}.(w_{\rm observed} - w_{\rm HOD}) + \frac{(N^{\rm gal}_{\rm mes} - N^{\rm gal}_{\rm mes})^2}{\sigma^2_{\rm Ngal}} \tag{4.10}$$

The main source of error is the uncertainty on the number of galaxies representing the redshift distribution of each population. This uncertainty lies between 5 and 20% for the bright tracers (CMASS and ELG bright). For the faint tracers it is between 30 and 40% (error denoted $\sigma_{\rm Ngal}$). The second source of errors is the statistical errors on the measurement of $w(\theta)$, which is accounted for by the covariance matrix C. Another source of errors is the cosmic variance. The fractional error on the measured density due to cosmic variance is $\sigma_{cv} < 0.1$; see Newman and Davis (2002). We account for this when computing the error on the HOD derived parameters. We add to the HOD error the dispersion on the halo mass obtained with the SHAM $\sigma^2_{\rm SHAM}(M_{\rm halo}) + \sigma^2_{\rm HOD}(M_{\rm halo}) = \sigma^2_{\rm total}(M_{\rm halo})$. In practice, $\sigma^2_{\rm SHAM}(M_{\rm halo}) >> \sigma^2_{\rm HOD}(M_{\rm halo})$ we therefore only use $\sigma^2_{\rm SHAM}(M_{\rm halo})$; see Table 4.3.

We perform additional tests to search for systematic errors. We test if the angular clustering measured is representative of the redshift distribution. We compute the angular clustering on only the tracers that have a reliable photoZ. We find a discrepancy on the angular clustering smaller by a factor of three than the error on the angular clustering measurement. We therefore assume the angular clustering measurement is representative of the redshift distribution.

The validity of the use of an HOD model can be discussed. $N(M)$ is valid to predict the clustering when applied to a volume limited sample. Current data does not allow us to construct a clean volume limited sample, as we do not have good photometric redshifts for all the tracers. We cannot cut a clean sample with $z \in \bar{z} \pm \sigma_z$. We quantify the impact of the wide redshift distribution on the angular clustering on large scale, where it matters for the determination of the galaxy bias, using the Multidark simulation. We compare the angular clustering of haloes SHAM selected with a top-hat redshift distribution (1 if $z \in \bar{z} \pm \sigma_z$, else 0) and the observed redshift distribution. The large scale angular clustering $\theta > 0.08°$ is unchanged. The small scale angular clustering does change. Therefore we assume the large scale angular clustering measured on the full sample is representative of the sample located in $\bar{z} \pm \sigma_z$, and that the bias, within the error bars will be correct. The values of the galaxy bias derived by the HOD method are given in Table 4.3 and plotted in Figs. 4.16 and 4.17. This shows the major flaw of this HOD parametrization, that cannot account for a galaxy population with a wide redshift distribution. A hint to try and model this would be to introduce a redshift dependence for the parameters describing $N(M)$. We further discuss this issue in section 4.6.

Table 4.3: Results from the clustering and the weak lensing analysis. The number of degrees of freedom is ~ 18 for the HOD fits. The values of the χ^2 show current HOD models fail in explaining angular clustering of the ELGs. The column $\lesssim M_d$ states the relation $M_{halo} - -M_d$. $b_{\text{clustering}}$ is the galaxy bias obtained with the HOD, $\langle b_{\text{WL}} \rangle$ with the weak-lensing, and $\langle r_{\text{WL}} \rangle$ is cross-correlation coefficient. The column satellite fraction compares the results from the SHAM and the HOD.

sample	maximum likelihood HOD parameter					χ^2/dof	halo mass		f_{sat}		$b_{w(\theta)}$	$\langle b_{\text{WL}} \rangle$	$\langle r_{\text{WL}} \rangle$
	$log(\frac{M_{min}}{M_\odot})$	$log(\frac{M_1}{M_\odot})$	$log(\frac{M_0}{M_\odot})$	σ_{logM}	α		$\langle M_{halo} \rangle$	$\leftrightarrow M_d$	HOD	SHAM			
CMASS	13.43	14.57	10.73	0.24	1.79	0.301	13.77	$>$	5	5	2.5 ± 0.5	2.1 ± 0.3	1.2 ± 0.4
LRG-WISE b	-	-	-	-	-	-	-	-	-	4	$2.3^{+0.5}_{-0.7}$	1.8 ± 0.3	-
LRG-WISE f	-	-	-	-	-	-	-	-	-	7	2.0 ± 0.2	1.5 ± 0.2	-
ELG gri b	13.51	14.51	11.07	0.89	0.81	16.05	13.22	$=$	6	8	2.0 ± 0.4	2.1 ± 0.4	-
ELG gri f	13.25	14.93	11.99	0.26	0.69	3.37	13.55	$>$	9	10	$2.5^{+0.7}_{-0.3}$	1.6 ± 0.4	-
ELG ugri b	12.41	12.04	12.42	0.58	0.64	2.88	13.15	$=$	52	7	$2.0^{+0.31}_{-0.17}$	1.7 ± 0.3	-
ELG ugri f	13.04	14.44	13.08	0.69	0.54	6.95	13.08	$>$	5	11	$2.1^{+0.5}_{-0.2}$	1.6 ± 0.3	-
ELG ugr b	13.17	13.40	13.07	0.63	0.62	1.54	13.24	$<$	16	3	$2.7^{+0.3}_{-0.2}$	-	-
ELG ugr f	13.10	14.38	12.92	0.59	0.77	6.53	13.12	$=$	4	7	$2.5^{+0.5}_{-0.2}$	-	-

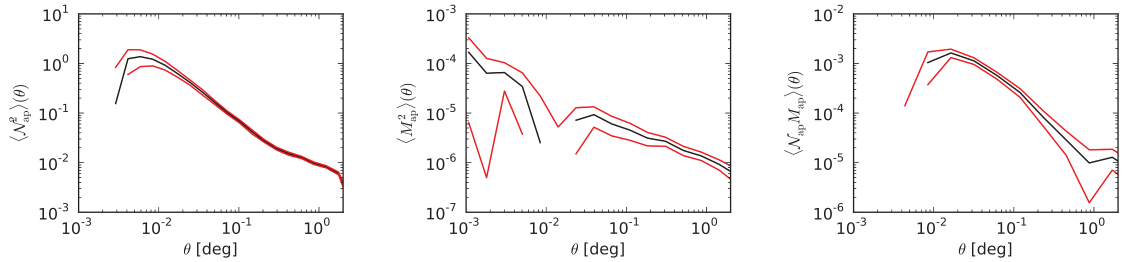


Figure 4.12: WL measurements for the BOSS-CMASS sample. From the left to the right the galaxy autocorrelation $\mathcal{N}^2(\theta)$, the matter autocorrelation $M_{ap}^2(\theta)$, and the galaxy-matter cross-correlation $\mathcal{N}(\theta)M_{ap}(\theta)$. The black line represent the mean value and the red lines the 1σ error contours. Measurements of $\mathcal{N}^2(\theta)$ and $M_{ap}^2(\theta)$ are clean between 0.05° and 1°. $\mathcal{N}(\theta)M_{ap}(\theta)$ is clean between 0.02° and 0.2°. We thus can deduce b_g and r.

4.5.2 Weak lensing measurements

We measure $\mathcal{N}(\theta)M_{ap}(\theta)$, $\mathcal{N}^2(\theta)$, $M_{ap}^2(\theta)$ with the same routines as in Jullo et al. (2012). The signal-to-noise ratio obtained on the bias is ~ 3 for tracers with $z < 1$ and of ~ 1.5 for $z > 1$. Due to this low signal to noise measurement, we exclude the *ugr* ELG sample at $z \simeq 1.2$ from the WL analysis. The measurements are shown on Figs. 4.12 (CMASS), 4.13 (LRG-WISE), 4.14 (*gri* ELG), 4.15 (*ugri* ELG). We derive the galaxy bias for all these samples; see Table 4.3. The cross-correlation $\mathcal{N}(\theta)M_{ap}(\theta)$ is measured with precision only for the lower redshift CMASS sample. Therefore r is computed only for this sample.

The main source of error arises from the error on the shape measurement; see the data description section. The uncertainty on the redshift distribution can also potentially introduce some systematic errors. All statistical errors are propagated with bootstrap as described in Jullo et al. (2012). For each foreground and background galaxy catalog, 100 samples are randomly drawn and processed to compute the aperture auto and cross-correlation functions as well as the bias and correlation coefficient. The background source catalogs are cut at 1 magnitude fainter than the foreground. Systematic error on $M_{ap}^2(\theta)$ on large scale due to the lack of information at small scale is below 1% (Kilbinger et al., 2006).

4.5.3 Cross-correlation coefficient and galaxy bias

Measurements of the galaxy clustering as a function of the luminosity were performed in the local universe by Coil et al. (2008); Skibba and Sheth (2009); Zehavi et al. (2011); Guo et al. (2012). They showed the clustering amplitude is greater for more luminous samples. Thus we expect to measure high values of the galaxy bias for all samples. There is no straightforward existing relation that can predict the expected galaxy bias.

4.5.3.1 Cross-correlation coefficient

We measure the cross-correlation coefficient only for the CMASS sample, $r = 1.16 \pm 0.35$. This measurement is a first, and implies the CMASS galaxies are fully correlated with the matter field. This strongly supports current BAO analysis made on this sample (Anderson et al., 2012).

With current data, it is not possible to constrain the cross-correlation coefficient of the other galaxy samples.

4.5.3.2 Galaxy bias

For the CMASS, LRG-WISE bright, *gri* ELG bright, and *ugri* ELG bright samples, the measurements of the two methods are in agreement; see Fig. 4.16. We find the bias is above 1.5 for the all ELG

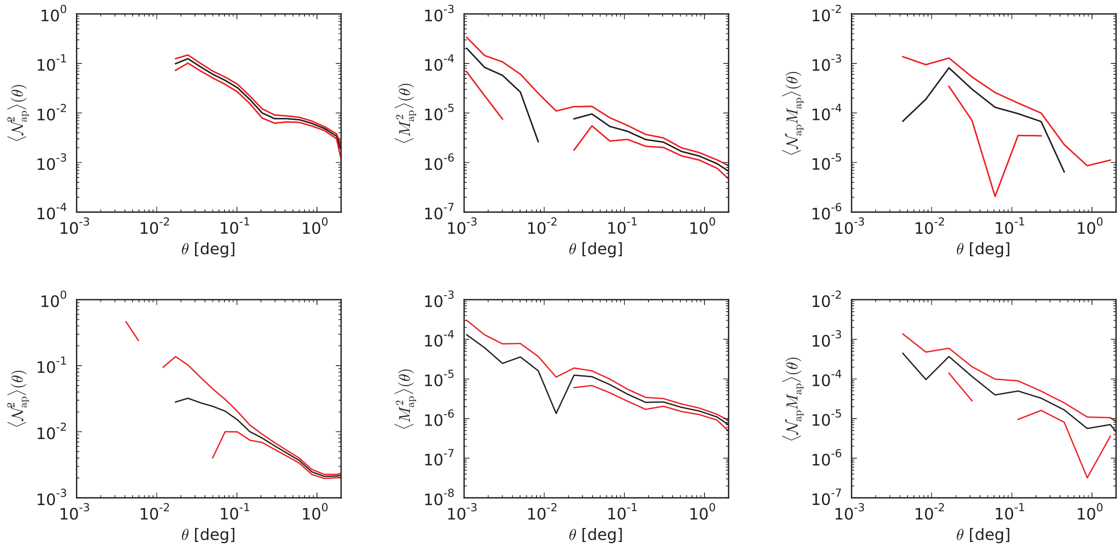


Figure 4.13: WL measurements for the LRG-WISE bright (top row) faint (bottom row) sample. From the left to the right the galaxy auto-correlation $\mathcal{N}^2(\theta)$, the matter autocorrelation $M^2_{ap}(\theta)$, and the galaxy-matter cross-correlation $\mathcal{N}(\theta)M_{ap}(\theta)$. The black line represent the mean value and the red lines the 1σ error contours. Measurements of $M^2_{ap}(\theta)$ and $\mathcal{N}^2(\theta)$ are clean for $\theta > 0.06°$ for both samples. $\mathcal{N}(\theta)M_{ap}(\theta)$ is not.

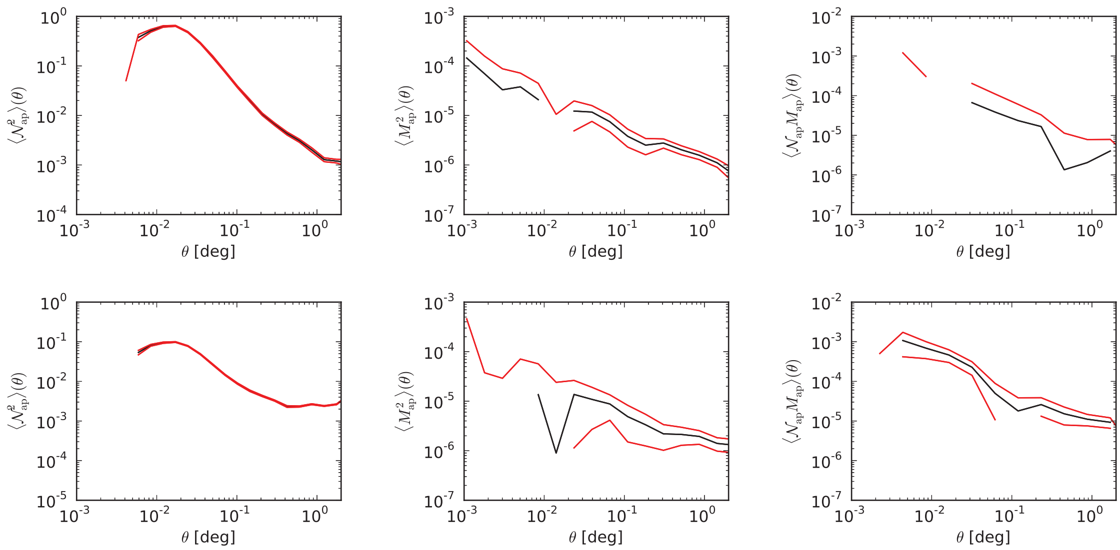


Figure 4.14: WL measurements for the *gri* ELG bright (top row) faint (bottom row) sample. From the left to the right the galaxy auto-correlation $\mathcal{N}^2(\theta)$, the matter autocorrelation $M^2_{ap}(\theta)$, and the galaxy-matter cross-correlation $\mathcal{N}(\theta)M_{ap}(\theta)$. The black line represent the mean value and the red lines the 1σ error contours. Measurements are clean between $0.1°$ and $1°$ for the $M^2_{ap}(\theta)$ and $\mathcal{N}^2(\theta)$ of the bright sample and $\mathcal{N}^2(\theta)$ of the faint sample. The other lensing measurements are not robust.

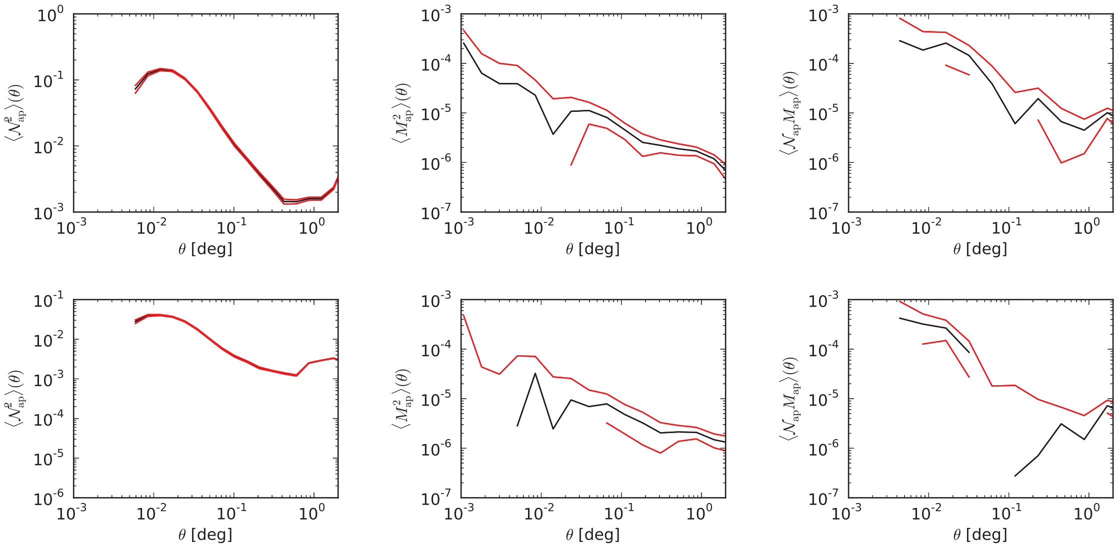


Figure 4.15: WL measurements for the *ugri* ELG bright (top row) faint (bottom row) sample. From the left to the right the galaxy auto-correlation $\mathcal{N}^2(\theta)$, the matter autocorrelation $M^2_{ap}(\theta)$, and the galaxy-matter cross-correlation $\mathcal{N}(\theta)M_{ap}(\theta)$. The black line represent the mean value and the red lines the 1σ error contours. Measurements are clean between $0.05°$ and $1°$ for the $M^2_{ap}(\theta)$ and $\mathcal{N}^2(\theta)$ of the bright sample and $\mathcal{N}^2(\theta)$ of the faint sample. The other lensing measurements are not robust.

selections, which is consistent with the measurements made by Coupon et al. (2012) on its brightest samples. We consider both measurements give a trend about the size of dark matter haloes these galaxies inhabit; see Fig. 4.17. Comparing the values of the bias with the curves of the bias for a constant halo mass after Tinker et al. (2005) and with the mass distributions obtained via SHAM shows all these samples are tracing the most massive haloes at their respective redshifts. From Fig. 4.16 we note that the bias from the HOD seems systematically higher than that of WL. The bias of the HOD corresponds to the bias of a sample that is complete in mass (represented by $N(M)$). Though the assumption of completeness in mass does not really hold for color-selected samples. The bias output by the WL does not need this assumption, and is thus lower than the value obtained via the HOD and probably closer to the reality for the bright samples where signal to noise is sufficient. We think this is the primary reason to this systematic discrepancy. With current data, it is not possible to estimate precisely stellar masses and understand this discrepancy in greater details.

For the LRG-WISE faint, *gri* ELG faint and *ugri* ELG faint samples, there is a systematic disagreement between the methods : the clustering analysis outputs a higher bias than the weak lensing analysis; see Fig. 4.16. The values are consistent at 1σ. Thus for the LRG-WISE faint, *gri* ELG faint and *ugri* ELG faint samples, we find the combination of current surveys "CFHT/Stripe82 + CFHT-LS + SDSS-S82 deep coadd" is too shallow to derive a reliable value of the bias. We need a fair spectroscopic sample of such tracers to improve the photometric redshift distribution and reduce the error on the redshift distribution and a deeper survey for the shape of background galaxies.

For the *ugr* ELG bright (and faint), with current data, the WL analysis is not feasible. The galaxy bias derived using only the clustering information is although of prime interest. It is consistent with the fact that the *ugr* ELG bright sample occupies the most massive halos at $z \sim 1.2$.

4.5.3.3 Satellite fraction, halo mass and HOD

The shape of the angular clustering for the ELG samples suggests a high satellite fraction. Indeed there is a strong bump for $\theta < 0.02°$. The HOD models fail to reproduce trustworthy satellite fractions

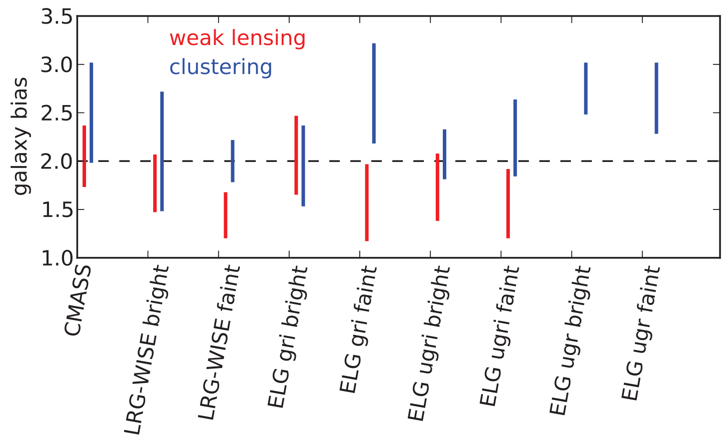


Figure 4.16: Large scale galaxy bias of the different BAO tracers ordered by mean redshift. The bright samples are more biased than the faint samples. The HOD and WL bias agree within the error bars for the bright samples. For the faint samples, the HOD bias is not in agreement and tends to be greater than the weak lensing estimation.

for the ELG samples, see Table 4.3. For the CMASS sample, that is approximately complete in mass and a volume limited sample, HOD and SHAM methods are in agreement and obtain 5% of satellites. For the *gri* ELG sample, which is the reddest of the ELG samples, the HOD and SHAM satellite fractions are in agreement, but remain below 10%. Given the observed small scale bump, we expected satellite fraction of the order of 15 to 25%. For the *ugri* and the *ugr* ELG samples, the satellite fractions do not really make sense. They are not consistent between the bright and the faint sample, which is puzzling. It shows the HOD models have a high degeneracy when they fit the small angular scales of a sample with a wide redshift distribution.

A similar effect is seen for the mean halo mass predicted by the HOD; it is always 0.5 dex higher than the expectations set by the SHAM. The clustering curves of the brightest galaxies in each redshift bin therefore constitute a new challenge to halo occupation models.

4.6 Discussion

The measurements presented in this paper are meaningful only for the CMASS, LRG-WISE bright, *gri* ELG bright, *ugri* ELG bright, and *ugr* ELG bright (clustering only) samples. The following discussion is therefore restricted to these tracers unless otherwise stated (we set aside the ELG faint samples).

4.6.1 Are these selections suitable for future BAO studies ?

The signal to noise in BAO goes as the number density ($\bar{n}$) times the power spectrum, which amplitude scales as $\propto b^2$ for a constant σ_8. Thus $SNR_{BAO} \propto \bar{n}P \propto \bar{n}b^2$. Thus the higher the bias (or the amplitude, it is equivalent), the smaller is the number density required to a reach a given signal to noise, and thus the faster the survey. For the same signal to noise ratio, the density required using tracers with $b = 1.5$ (resp. $b = 2$) is 2.25 (resp. 4) times smaller than using tracers with a bias $b = 1$. A BAO survey using these tracers has thus a real advantage compared to selecting galaxies with lower absolute luminosity that have a lower clustering amplitude. These numbers constitute a zeroth order value of the galaxy bias that can be used in simulation to predict the accuracy of the detection and anticipate the requirements on geometry and density for surveys as SDSS-IV/eBOSS, DESspec, BigBOSS, PFS-SuMiRe, and Euclid. A check could be performed with n-Body simulations

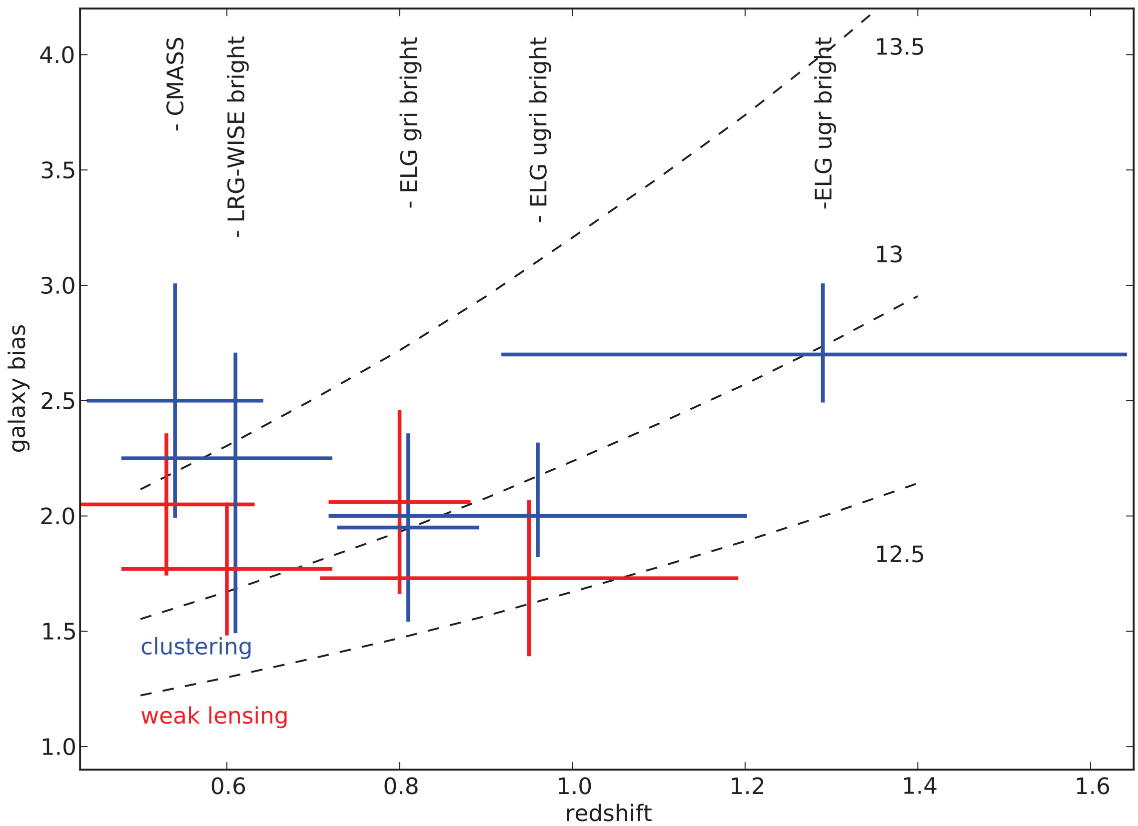


Figure 4.17: Large scale galaxy bias of the different BAO tracers vs. redshift. The interquartile are shown as error bars (25% - 75% intervals). The dashed black lines correspond to the bias of haloes of mass $\log_{10} M_{200}/M_{\odot}$ of 12.5, 13 and 13.5 from bottom to the top, after Tinker et al. (2005). The figure shows the ELGs inhabit haloes of mass between $10^{12.5}$ and $10^{13} M_{\odot}$. The comparison with the SHAM halo mass values (Table 4.2) shows these selections are tracing the most massive halos at their mean redshift.

by trying to reproduce the angular clustering measurements from this paper.

The primary role of the target selection, to select highly-biased galaxies, is here fulfilled. There is, however, a trade-off. The galaxy bias is the source of the dominating systematic on the BAO measurement; see Mehta et al. (2011) and references therein. With current reconstruction algorithms, and a knowledge of the bias better than 10%, simulations show it is possible to reduce the systematic errors down to 0.15%. This study enables to start calibrating simulations on observed data to understand how to cope with this systematic effect.

4.6.2 Sensitivity to input cosmological parameters

In this study, we use WMAP 7 fiducial values for cosmological parameters, yet the latest parameters obtained by Planck differ slightly, see Planck Collaboration et al. (2013). The mean change of the angular correlation function of matter on linear scales (between 0.05 and 0.5°) at $z \sim 0.8$ with the redshift distribution of the bright *gri* ELG is of $\sim 7\%$. The WMAP parameters predict a matter clustering stronger by $\sim 7\%$ than that of Planck. Therefore with Planck parameters, the galaxy bias would increase by $\sim 3.5\%$ compared to the values we measure here.

4.6.3 Complementarity of the lower redshift selections

The redshift distribution of the CMASS, LRG-WISE and *gri* ELG are overlapping in $0.6 < z < 0.7$; see Fig. 4.3. The match between the catalogs shows that 20% of the tracers are common between the LRG-WISE bright sample and the CMASS. Therer is no common tracers between the *gri* ELG and the LRG-WISE bright or the CMASS. These three samples are therefore complementary to map the large scale structures in this redshift zone, and particularly to perform cross-correlation analysis.

4.6.4 ELG ugri selection, an intermediate selection

The *ugri* ELG color selection is an hybrid between the *ugr* and the *gri* selections. The clustering analysis shows the $w(\theta)$ of the *ugri* is located between that of the *ugr* and the *gri* $w(\theta)$. It is thus consistent.

4.6.5 Hints on luminosity, stellar mass, halo mass, HOD modeling and N-body simulations

Wake et al. (2011) showed that stellar mass and halo mass are correlated for $1 < z < 2$. Li et al. (2006) reported that on large scales the galaxy clustering amplitude dependence on stellar mass is positively correlated to the luminosity. Our selections target the most luminous galaxies (in their redshift range). We therefore expect to track high stellar masses and high halo masses. The combination of the clustering analysis from this work and of the estimates of stellar masses from Table 2.4 demonstrate that these galaxies trace haloes with large masses ($M_{halo} \sim 10^{13} M_\odot$) *i.e.* we comply with the general picture.

This analysis is complementary to similar analyses based on faint and deep 'pencil beam' surveys, that investigate the clustering properties of average luminosity galaxies, see Leauthaud et al. (2012); Mostek et al. (2012b). In fact our target selections aim for galaxies that are not in large numbers in such surveys. Given the trends in these papers, our results are in good agreement with what galaxy bias is expected for the brightest galaxies at redshift $z = 1$.

These clustering measurements are the first of its kind and are of prime interest to improve our knowledge of the halo occupation distribution of such galaxy populations. In this work, we assumed the angular clustering measured was representative of the clustering of a volume limited sample at the mean redshift of the sample in order to derive its large scale bias. This assumption is not simple to test with current data, in particular with current photometric redshifts. Though this might be one of the reasons why the HOD model fails to describe in details the observed angular clustering at small scales, in particular the satellite fractions. In fact the properties of the selected galaxies evolve with redshift *i.e.* the lower redshift ones are less luminous than the higher redshift ones. The HOD model used here does not have parameters that evolve with redshift, it is a fixed halo population convolved with the redshift distribution. These considerations are oriented towards the construction of a mock catalog.

4.7 Conclusion

We measured, for the first time, the angular clustering and the cosmic shear of the future samples selected for BAO studies. We have validated the use of this method on the CMASS sample as our measurement agrees with previous robust results. Moreover, we show that the CMASS sample is well correlated to the dark matter field.

We demonstrate that the bright ELG samples are highly biased and therefore suited to detect BAO. The galaxies to be targeted for the near future SDSS-IV/eBOSS BAO experiment will provide a high signal-to-noise ratio in the galaxy power spectrum.

This work provides a strong basis to develop simulations and mock catalogs of these galaxy populations in order to understand the details of the relation ELG - dark matter haloes.

When pushing the method further in depth (ELG faint samples), we are limited by the knowledge of the redshift distribution of the tracers and by the depth of galaxy shape catalogs for weak-lensing. Both methods disagree when the redshift distribution is poorly known, *i.e.*, with an uncertainty of about 40%. Within the next few years, deeper observations in the infra-red and improved photometric redshifts should provide the necessary data to describe with precision the relation between galaxies and matter probed by the faint target selection schemes.

Chapter 5

Conclusion

Figure 5.1: Paul Gauguin. D'où venons-nous ? Que sommes-nous ? Où allons-nous ? 1898.

If we trust General Relativity and the Cosmological Principle, there is an overwhelming evidence that the Universe is spatially flat and undergoing an accelerated expansion since half of its life after Big Bang. Cosmological probes based on geometric measurements at different cosmic epochs are expected to provide a firm insight into the mechanism responsible for such acceleration – dubbed dark energy or dark gravity.

The Baryon Acoustic Oscillation (BAO) feature in the two-point correlation functions of galaxies provide a standard ruler to probe the late-time accelerated expansion of the Universe through the measurement of the redshift-distance relation.

As discussed in the first chapter, its measure is going to be very accurate ($< 1\%$) at epochs surveyed by ongoing and future galaxy observational projects, covering thousands of square degrees. The main sources of systematic errors on the BAO measurement come from the uncertainty on the redshift distribution of the tracers used and from the spatial homogeneity of the photometry setting the target selection.

In the second chapter, I studied several galaxy selection schemes aiming at building an ELG sample in the redshift range $0.6 < z < 1.7$, that would be suitable for future BAO studies using the Baryonic Oscillation Spectroscopic Survey (BOSS) spectrograph on the Sloan Digital Sky Survey (SDSS) telescope. I explored two different colour selections using both the SDSS and the Canada France Hawai Telescope Legacy Survey (CFHT-LS) photometry in the *u, g, r*, and *i* bands and evaluated their performance selecting luminous ELG for spectroscopic observation on the Sloan 2.5m telescope.. From about 2,000 ELGs observed, I identified two selection schemes, dubbed *ugr* and *gri*, that have a redshift measurement efficiency above 70 percent.

The *gri* selection efficiency can be tuned by improving the automated determination of the spectroscopic redshift, this can increase the efficiency by 5 per cent raising the number of galaxies at $z > 0.6$ from 73 to 78 per cent.

The *ugr* selection can be improved to reach a 70% efficiency by lowering the *u-g* color selection to 0.3 in order to diminish the contamination by low-redshift galaxies. Additional low-redshift galaxies can be removed from the selection through an inspection of the images: some of the low-redshift galaxies are quite extended, and one could mistake a high-redshift merger for an extended low-redshift galaxy. A visual inspection reduces the low-redshift share from 9 percent to 4 percent. Moreover, the compact and extended selection on the CFHT data is very efficient at identifying quasars. It is also possible to improve the spectroscopic redshift determination and thus re-classify the 12% of 'single emission line' galaxies (10 percent are at $z > 0.6$).

Based on these color-selections, I have implemented target selections for BAO using the four bands *ugri* finding a resulting selection with a redshift distribution close to the smooth combination of the *gri* and *ugr* selections. Current observations, performed by another ancillary program of which I am the principal investigator, show this new selection is very efficient at selecting galaxies with strong emission lines. These result support the feasibility of massive ELG surveys using the BOSS spectrograph on the SDSS telescope for a BAO detection at redshift $z \sim 1$. In the light of the future the SDSS-IV/*eBOSS* experiment, conceived to combine the use of the BAO ruler with redshift space distortions, the *ugri* selection scheme will be applied to select emission line galaxies in the redshift $0.6 < z < 1.2$.

Chapters 3 presents the systematic errors of a BAO survey based on ELG due to the use of the [OII] signature in the galaxy spectrum as a redshift proxy. The [OII] emission line doublet is an optical signature, which provides a unique redshift identification that can minimize confusion with other single emission lines. Will the use of [OII] introduce a bias in the observed redshift distribution? From the simulated observations of the [OII] doublet for various instrument resolutions and line velocities, I derived the required spectrograph resolution for redshift surveys pinpointing the redshift range $0.6 < z < 1.7$. Two strategies about the choice of the resolution for future spectrographs for BAO surveys emerged:

- for bright [OII] emitter surveys (like SDSS-IV/eBOSS), a resolution of $R \sim 2500$ (as in the current SDSS spectrograph) is sufficient to obtain a fair sample of doublets (60%) in order to train the pipeline to recover all the [OII] redshifts. Increasing the resolution to 3300 allows one to get 90% of doublets. For a small increase in resolution (2500 to 3300), the redshift determination efficiency gains 50%. The impact of the sky lines on the completeness in redshift is below 6%;

- for faint [OII] emitter surveys (like MS-DESI), this study recommends the highest resolution. Knowing there is a limited number of pixels on the detector (about 4000), and that the highest resolution possible on a three channel spectrograph is $R \sim 4500$ at 7500 Angstrom, to go beyond, a four channel spectrograph becomes mandatory. Although, with a resolution of $R \sim$4500, one would obtain 25% of doublets, which is enough to train the pipeline to assign correct redshift.

Besides, these simulations showed that the impact of the sky lines on the redshift completeness is independent on the resolution and only depends on the sky subtraction and on the flux of the observed emission lines.

Preliminary developments on redshift pipelines show this study is conservative and constitutes a lower limit to the % of doublets to be obtained. The next step for future surveys is now to develop complete end-to-end simulation tools, with a particular focus on the redshift solver. In this aim I am involved in a collaboration with scientists of the Lawrence Berkeley National Laboratory and the University of Utah (USA). This effort is clearly identified in the SDSS collaboration.

The observation of the BAO based on highly biased LRG or the low bias Lya forest showed the relation between the luminous matter tracer and its dark counterpart plays a key role in understating the details of the measurement of the BAO standard ruler. The use of the ELG to track the matter field is a new challenge: how are the ELG related to the underlying matter field? Is a linear bias model sufficient? What is the amount of stochasticity in this relation?

The fourth chapter is the first effort made in this direction. Applying the aforementioned galaxy selection schemes, I analyzed the angular clustering at the redshifts 0.5, 0.7, 0.8, 1, and 1.2 and fitted to a halo occupation distribution (HOD) model to derive the properties of the haloes these galaxies inhabit, in particular the galaxy bias on large scales. Jointly, we have also performed a weak lensing analysis (aperture statistics) to extract the galaxy bias and the cross-correlation coefficient to be further compared the HOD prediction. To this purpose, we considered a data set composed of the photometry of the deep co-addition on SDSS Stripe 82 (225 deg^2), of CFHT/Stripe82 deep i band and of the WISE infrared photometric band W1: a unique combination to perform this analysis.

The analysis on the SDSS-III/CMASS selection at $z = 0.5$ is in agreement with previous studies on the same tracer, and its cross-correlation coefficient is $r = 1.16 \pm 0.35$.

For the higher redshift bins, the trends that the brightest galaxy populations selected are strongly biased ($b > 1.5$) is confirmed: unless a precise estimation is still not permitted by the current data sets, the high bias of bright ELGs is demonstrated, providing a well-suited tracer for the detection of the BAO. Emission line galaxies will therefore represent a privileged target for the near future SDSS-IV/eBOSS experiment.

Pushing the method with ELG faint samples, showed that the current knowledge of the redshift distribution of the tracers and the current depth of galaxy shape catalogs for weak-lensing measurements is limited with present data sets. Within the next few years, deeper observations and improved photometric redshifts should provide the necessary data to describe with precision the relation between galaxies and matter probed by the faint target selection schemes.

This work helped me establish strong collaboration with the German, Brazilian and Canadian cosmology communities when reducing the CFHT/Stripe82 data.

This work supports future ground-based spectroscopic cosmology projects:

- eBOSS experiment, starting in august 2014 will observe about 270 000 ELGs at redshift $0.6 < z < 1.2$ with SDSS telescope infrastructure.

- MS-DESI (the new name of BigBOSS/DES-Spec) obtained the CD-0 from the american department of energy, and aims at observing 22 000 000 ELG at $0.6 < z < 1.7$ with a 4-m class telescope, see Schlegel et al. (2011); Mostek et al. (2012a)
- The Subaru Prime Focus Spectrograph Strategic Program of about 300 nights will survey ELG at $0.8 < z < 2.4$, see (Ellis et al., 2012).
- The WEAVE galaxy evolution and cosmology survey on the William Herschel Telescope will provide emission-line redshifts for all star-forming galaxies at $z < 1.3$ using H_α and [OII] emission and at $z > 2$ using Ly_α emission for $\sim 10^7$ objects over 10,000 deg^2 (Dalton et al., 2012).
- The 4MOST spectroscopic survey will run permanently on the ESO/VISTA 4m-class telescope to perform a 5 year public survey yielding more than 20 million spectra at resolution $R \sim 5000$ ($\lambda = 390 - 1000$nm) and more than 2 million spectra at $R \sim 20,000$ ($395 - 456.5$ nm and $587 - 673$ nm) (de Jong et al., 2012).

These projects will investigate both the cosmology and the formation and evolution of galaxies at the epoch of the transition between the dark-matter era and the current dark-energy era.

In the next twenty years, the huge increment of data from cosmology, astronomy, and particle physics is expected to provide a deep insight on fundamental physics. Future data are expected to allow a sufficient precision to test if the current acceleration of the geometry is due to a dark energy component such as a dynamical scalar field (Peebles and Ratra (1988) ; see Copeland et al. (2006) for a review), a new theory of gravitational interactions beyond general relativity (Clifton et al., 2012), or non-linear dynamical effects due to matter inhomogeneities intrinsic to general relativity and related to the averaging problem (Buchert and Räsänen, 2012). Besides, the very precise cosmological measurement of BAO, combined with other probes, very likely will provide a dramatic understanding of the dark matter sector, and in particular of massive neutrinos.

Bibliography

Abazajian K. N., Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., Allende Prieto C., An D. et al. The Seventh Data Release of the Sloan Digital Sky Survey. *ApJS*, 182:543–558, june 2009.

Abdalla F., Annis J., Bacon D., Bridle S., Castander F., Colless M. et al. The Dark Energy Spectrometer (DESpec): A Multi-Fiber Spectroscopic Upgrade of the Dark Energy Camera and Survey for the Blanco Telescope. *ArXiv e-prints*, september 2012.

Abraham R. G., Tanvir N. R., Santiago B. X., Ellis R. S., Glazebrook K., van den Bergh S. Galaxy morphology to I=25 mag in the Hubble Deep Field. *MNRAS*, 279:L47–L52, april 1996.

Adelberger K. L., Steidel C. C., Shapley A. E., Hunt M. P., Erb D. K., Reddy N. A. et al. Optical Selection of Star-forming Galaxies at Redshifts $1 < z < 3$. *ApJ*, 607:226–240, may 2004.

Ahn C. P., Alexandroff R., Allende Prieto C., Anderson S. F., Anderton T., Andrews B. H. et al. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. *ApJS*, 203:21, december 2012.

Aihara H., Allende Prieto C., An D., Anderson S. F., Aubourg É., Balbinot E. et al. The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III. *ApJS*, 193:29, april 2011.

Albrecht A., Bernstein G., Cahn R., Freedman W. L., Hewitt J., Hu W. et al. Report of the Dark Energy Task Force. *ArXiv Astrophysics e-prints*, september 2006.

Anderson L., Aubourg E., Bailey S., Bizyaev D., Blanton M., Bolton A. S. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. *MNRAS*, 427:3435–3467, december 2012.

Annis J., Soares-Santos M., Strauss M. A., Becker A. C., Dodelson S., Fan X. et al. The SDSS Coadd: 275 deg^2 of Deep SDSS Imaging on Stripe 82. *ArXiv e-prints*, november 2011.

Argence B., Lamareille F. Emission-lines calibrations of the star formation rate from the Sloan Digital Sky Survey. *A&A*, 495:759–773, march 2009.

Astier P., Guy J., Regnault N., Pain R., Aubourg E., Balam D. et al. The Supernova Legacy Survey: measurement of Ω_M, $\Omega_?$ and w from the first year data set. *A&A*, 447:31–48, february 2006.

Bacon D. J., Massey R. J., Refregier A. R., Ellis R. S. Joint cosmic shear measurements with the Keck and William Herschel Telescopes. *MNRAS*, 344:673–685, september 2003.

Banerji M., Abdalla F. B., Lahav O., Lin H. Photometric redshifts for the Dark Energy Survey and VISTA and implications for large-scale structure. *MNRAS*, 386:1219–1233, may 2008.

Bennett C. L., Larson D., Weiland J. L., Jarosik N., Hinshaw G., Odegard N. et al. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *ArXiv e-prints*, december 2012.

Bennett C. L., Smoot G. F., Hinshaw G., Wright E. L., Kogut A., de Amici G. et al. Preliminary separation of galactic and cosmic microwave emission for the COBE Differential Microwave Radiometer. *ApJ*, 396:L7–L12, september 1992.

Beutler F., Blake C., Colless M., Jones D. H., Staveley-Smith L., Campbell L. et al. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *MNRAS*, 416:3017–3032, october 2011.

Blake C., Bridle S. Cosmology with photometric redshift surveys. *MNRAS*, 363:1329–1348, november 2005.

Blake C., Brough S., Colless M., Contreras C., Couch W., Croom S. et al. The WiggleZ Dark Energy Survey: joint measurements of the expansion and growth history at z<1. *MNRAS*, 425:405–414, september 2012.

Blake C., Brough S., Colless M., Contreras C., Couch W., Croom S. et al. The WiggleZ Dark Energy Survey: the growth rate of cosmic structure since redshift z=0.9. *MNRAS*, 415:2876–2891, august 2011a.

Blake C., Davis T., Poole G. B., Parkinson D., Brough S., Colless M. et al. The WiggleZ Dark Energy Survey: testing the cosmological model with baryon acoustic oscillations at z= 0.6. *MNRAS*, 415: 2892–2909, august 2011b.

Blake C., Glazebrook K. Probing Dark Energy Using Baryonic Oscillations in the Galaxy Power Spectrum as a Cosmological Ruler. *ApJ*, 594:665–673, september 2003.

Blake C., Kazin E. A., Beutler F., Davis T. M., Parkinson D., Brough S. et al. The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations. *MNRAS*, 418:1707–1724, december 2011c.

Boggess N. W., Mather J. C., Weiss R., Bennett C. L., Cheng E. S., Dwek E. et al. The COBE mission - Its design and performance two years after launch. *ApJ*, 397:420–429, october 1992.

Boulade O., Charlot X., Abbon P., Aune S., Borgeaud P., Carton P.-H. et al. MegaCam: the new Canada-France-Hawaii Telescope wide-field imaging camera. In Iye M., Moorwood A. F. M., éditeurs, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4841 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 72–81, march 2003.

Bovy J., Hennawi J. F., Hogg D. W., Myers A. D., Kirkpatrick J. A., Schlegel D. J. et al. Think Outside the Color Box: Probabilistic Target Selection and the SDSS-XDQSO Quasar Targeting Catalog. *ApJ*, 729:141–+, march 2011.

Branch D., Tammann G. A. Type IA supernovae as standard candles. *ARA&A*, 30:359–389, 1992.

Bridle S., Balan S. T., Bethge M., Gentile M., Harmeling S., Heymans C. et al. Results of the GREAT08 Challenge: an image analysis competition for cosmological lensing. *MNRAS*, 405:2044–2061, july 2010.

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T. et al. The physical properties of star-forming galaxies in the low-redshift Universe. *MNRAS*, 351:1151–1179, july 2004.

Bruzual G., Charlot S. Stellar population synthesis at the resolution of 2003. *MNRAS*, 344:1000–1028, october 2003.

Buchert T., Nayet C., Wiegand A. Lagrangian theory of structure formation in relativistic cosmology II: average properties of a generic evolution model. *ArXiv e-prints*, march 2013.

Buchert T., Ostermann M. Lagrangian theory of structure formation in relativistic cosmology: Lagrangian framework and definition of a nonperturbative approximation. *Phys. Rev. D*, 86(2):023520, july 2012.

Buchert T., Räsänen S. Backreaction in Late-Time Cosmology. *Annual Review of Nuclear and Particle Science*, 62:57–79, november 2012.

Busca N. G., Delubac T., Rich J., Bailey S., Font-Ribera A., Kirkby D. et al. Baryon Acoustic Oscillations in the Ly-\alpha\ forest of BOSS quasars. *ArXiv e-prints*, november 2012.

Cacciato M., Lahav O., van den Bosch F. C., Hoekstra H., Dekel A. On combining galaxy clustering and weak lensing to unveil galaxy biasing via the halo model. *MNRAS*, 426:566–587, october 2012.

Cannon R., Drinkwater M., Edge A., Eisenstein D., Nichol R., Outram P. et al. The 2dF-SDSS LRG and QSO (2SLAQ) Luminous Red Galaxy Survey. *MNRAS*, 372:425–442, october 2006.

Capak P., Aussel H., Ajiki M., McCracken H. J., Mobasher B., Scoville N. et al. The First Release COSMOS Optical and Near-IR Data and Catalog. *ApJS*, 172:99–116, september 2007.

Chuang C.-H., Wang Y. Measurements of H(z) and D_A(z) from the two-dimensional two-point correlation function of Sloan Digital Sky Survey luminous red galaxies. *MNRAS*, 426:226–236, october 2012.

Clifton T., Ferreira P. G., Padilla A., Skordis C. Modified gravity and cosmology. *Phys. Rep.*, 513: 1–189, march 2012.

Coil A. L., Newman J. A., Croton D., Cooper M. C., Davis M., Faber S. M. et al. The DEEP2 Galaxy Redshift Survey: Color and Luminosity Dependence of Galaxy Clustering at z˜1. *ApJ*, 672: 153–176, january 2008.

Cole S., Percival W. J., Peacock J. A., Norberg P., Baugh C. M., Frenk C. S. et al. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362:505–534, september 2005.

Comparat J., Kneib J.-P., Escoffier S., Zoubian J., Ealet A., Lamareille F. et al. Investigating emission-line galaxy surveys with the Sloan Digital Sky Survey infrastructure. *MNRAS*, 428:1498–1517, january 2013.

Conroy C., Wechsler R. H., Kravtsov A. V. Modeling Luminosity-dependent Galaxy Clustering through Cosmic Time. *ApJ*, 647:201–214, august 2006.

Cooper M. C., Newman J. A., Weiner B. J., Yan R., Willmer C. N. A., Bundy K. et al. The DEEP2 Galaxy Redshift Survey: the role of galaxy environment in the cosmic star formation history. *MNRAS*, 383:1058–1078, january 2008.

Cooray A., Sheth R. Halo models of large scale structure. *Phys. Rep.*, 372:1–129, december 2002.

Copeland E. J., Sami M., Tsujikawa S. Dynamics of Dark Energy. *International Journal of Modern Physics D*, 15:1753–1935, 2006.

Coupon J., Ilbert O., Kilbinger M., McCracken H. J., Mellier Y., Arnouts S. et al. Photometric redshifts for the CFHTLS T0004 deep and wide fields. *A&A*, 500:981–998, june 2009.

Coupon J., Kilbinger M., McCracken H. J., Ilbert O., Arnouts S., Mellier Y. et al. Galaxy clustering in the CFHTLS-Wide: the changing relationship between galaxies and haloes since$z \sim 1.2\star$. *A&A*, 542:A5, june 2012.

Cunha C., Huterer D., Doré O. Primordial non-Gaussianity from the covariance of galaxy cluster counts. *Phys. Rev. D*, 82(2):023004, july 2010.

Dalton G., Trager S. C., Abrams D. C., Carter D., Bonifacio P., Aguerri J. A. L. et al. WEAVE: the next generation wide-field spectroscopy facility for the William Herschel Telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8446 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, september 2012.

Davis M., Faber S. M., Newman J., Phillips A. C., Ellis R. S., Steidel C. C. et al. Science Objectives and Early Results of the DEEP2 Redshift Survey. In P. Guhathakurta, éditeur, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4834 de *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 161–172, february 2003.

Davis M., Peebles P. J. E. A survey of galaxy redshifts. V - The two-point position and velocity correlations. *ApJ*, 267:465–482, april 1983.

Dawson K. S., Aldering G., Amanullah R., Barbary K., Barrientos L. F., Brodwin M. et al. An Intensive Hubble Space Telescope Survey for $z > 1$ Type Ia Supernovae by Targeting Galaxy Clusters. *AJ*, 138:1271–1283, november 2009.

Dawson K. S., Schlegel D. J., Ahn C. P., Anderson S. F., Aubourg É., Bailey S. et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *AJ*, 145:10, january 2013.

de Jong R. S., Bellido-Tirado O., Chiappini C., Depagne É., Haynes R., Johl D. et al. 4MOST: 4-metre multi-object spectroscopic telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8446 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, september 2012.

Dekel A., Lahav O. Stochastic Nonlinear Galaxy Biasing. *ApJ*, 520:24–34, july 1999.

Dickinson M., Papovich C., Ferguson H. C., Budavári T. The Evolution of the Global Stellar Mass Density at $0 < z < 3$. *ApJ*, 587:25–40, april 2003.

Dolney D., Jain B., Takada M. Baryon oscillations and dark-energy constraints from imaging surveys. *MNRAS*, 366:884–898, march 2006.

Drinkwater M. J., Jurek R. J., Blake C., Woods D., Pimbblet K. A., Glazebrook K. et al. The WiggleZ Dark Energy Survey: survey design and first data release. *MNRAS*, 401:1429–1452, january 2010.

Dünner R., Hasselfield M., Marriage T. A., Sievers J., Acquaviva V., Addison G. E. et al. The Atacama Cosmology Telescope: Data Characterization and Mapmaking. *ApJ*, 762:10, january 2013.

Einstein A. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354:769–822, 1916.

Eisenstein D. J., Annis J., Gunn J. E., Szalay A. S., Connolly A. J., Nichol R. C. et al. Spectroscopic Target Selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample. *AJ*, 122: 2267–2280, november 2001.

Eisenstein D. J., Hu W. Baryonic Features in the Matter Transfer Function. *ApJ*, 496:605, march 1998.

Eisenstein D. J., Seo H.-J., Sirko E., Spergel D. N. Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak. *ApJ*, 664:675–679, august 2007a.

Eisenstein D. J., Seo H.-J., White M. On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter. *ApJ*, 664:660–674, august 2007b.

Eisenstein D. J., Weinberg D. H., Agol E., Aihara H., Allende Prieto C., Anderson S. F. et al. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. *AJ*, 142:72, september 2011.

Eisenstein D. J., Zehavi I., Hogg D. W., Scoccimarro R., Blanton M. R., Nichol R. C. et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633:560–574, november 2005.

Ellis R., Takada M., Aihara H., Arimoto N., Bundy K., Chiba M. et al. Extragalactic Science and Cosmology with the Subaru Prime Focus Spectrograph (PFS). *ArXiv e-prints*, june 2012.

Epinat B., Tasca L., Amram P., Contini T., Le Fèvre O., Queyrel J. et al. MASSIV: Mass Assembly Survey with SINFONI in VVDS. II. Kinematics and close environment classification. *A&A*, 539:A92, march 2012.

Erben T., Hildebrandt H., Miller L., van Waerbeke L., Heymans C., Hoekstra H. et al. CFHTLenS: The Canada-France-Hawaii Telescope Lensing Survey - Imaging Data and Catalogue Products. *ArXiv e-prints*, october 2012.

Faber S. M., Phillips A. C., Kibrick R. I., Alcott B., Allen S. L., Burrous J. et al. The DEIMOS spectrograph for the Keck II Telescope: integration and testing. In Iye M., Moorwood A. F. M., éditeurs, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4841 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1657–1669, march 2003.

Fall S. M. Galaxy correlations and cosmology. *Reviews of Modern Physics*, 51:21–43, january 1979.

Feast M. W., Catchpole R. M. The Cepheid period-luminosity zero-point from HIPPARCOS trigonometrical parallaxes. *MNRAS*, 286:L1–L5, march 1997.

Feldman H. A., Kaiser N., Peacock J. A. Power-spectrum analysis of three-dimensional redshift surveys. *ApJ*, 426:23–37, may 1994.

Fowler J. W., Acquaviva V., Ade P. A. R., Aguirre P., Amiri M., Appel J. W. et al. The Atacama Cosmology Telescope: A Measurement of the $600 < l < 8000$ Cosmic Microwave Background Power Spectrum at 148 GHz. *ApJ*, 722:1148–1161, october 2010.

Freedman W. L., Madore B. F., Scowcroft V., Burns C., Monson A., Persson S. E. et al. Carnegie Hubble Program: A Mid-infrared Calibration of the Hubble Constant. *ApJ*, 758:24, october 2012.

Frieman J. A., Turner M. S., Huterer D. Dark energy and the accelerating universe. *Annual Review of Astronomy and Astrophysics*, 46(1):385–432, 2008. URL http://www.annualreviews.org/doi/abs/10.1146/annurev.astro.46.060407.145243.

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P. The Sloan Digital Sky Survey Photometric System. *AJ*, 111:1748, april 1996.

Garilli B., Le Fèvre O., Guzzo L., Maccagni D., Le Brun V., de la Torre S. et al. The Vimos VLT deep survey. Global properties of 20,000 galaxies in the $I_{AB} < 22.5$ WIDE survey. *A&A*, 486:683–695, august 2008.

Gunn J. E., Carr M., Rockosi C., Sekiguchi M., Berry K., Elms B. et al. The Sloan Digital Sky Survey Photometric Camera. *AJ*, 116:3040–3081, december 1998.

Gunn J. E., Siegmund W. A., Mannery E. J., Owen R. E., Hull C. L., Leger R. F. et al. The 2.5 m Telescope of the Sloan Digital Sky Survey. *AJ*, 131:2332–2359, april 2006.

Guo H., Zehavi I., Zheng Z., Weinberg D. H., Berlind A. A., Blanton M. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Luminosity and Color Dependence and Redshift Evolution. *ArXiv e-prints*, december 2012.

Guzzo L., Scodeggio M., Garilli B., Granett B. R., Abbas U., Adami C. et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS). An unprecedented view of galaxies and large-scale structure at $0.5 < z < 1.2$. *ArXiv e-prints*, march 2013.

Hamilton A. J. S. Toward Better Ways to Measure the Galaxy Correlation Function. *ApJ*, 417:19, november 1993.

Hamilton A. J. S. Power Spectrum Estiamtion II. Linear Maximum Likelihood. In Martínez V. J., Saar E., Martínez-González E., Pons-Bordería M.-J., éditeurs, *Data Analysis in Cosmology*, volume 665 de *Lecture Notes in Physics, Berlin Springer Verlag*, pages 433–456, 2009a.

Hamilton A. J. S. Power Spectrum Estimation. I. Basics. In Martínez V. J., Saar E., Martínez-González E., Pons-Bordería M.-J., éditeurs, *Data Analysis in Cosmology*, volume 665 de *Lecture Notes in Physics, Berlin Springer Verlag*, pages 415–431, 2009b.

Hanuschik R. W. A flux-calibrated, high-resolution atlas of optical sky emission from UVES. *A&A*, 407:1157–1164, september 2003.

Heymans C., Van Waerbeke L., Bacon D., Berge J., Bernstein G., Bertin E. et al. The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *MNRAS*, 368:1323–1339, may 2006a.

Heymans C., Van Waerbeke L., Miller L., Erben T., Hildebrandt H., Hoekstra H. et al. CFHTLenS: The Canada-France-Hawaii Telescope Lensing Survey. *ArXiv e-prints*, september 2012.

Heymans C., White M., Heavens A., Vale C., van Waerbeke L. Potential sources of contamination to weak lensing measurements: constraints from N-body simulations. *MNRAS*, 371:750–760, september 2006b.

Hinshaw G., Larson D., Komatsu E., Spergel D. N., Bennett C. L., Dunkley J. et al. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ArXiv e-prints*, december 2012.

Hoekstra H., van Waerbeke L., Gladders M. D., Mellier Y., Yee H. K. C. Weak Lensing Study of Galaxy Biasing. *ApJ*, 577:604–614, october 2002.

Holtzman J. A. Microwave background anisotropies and large-scale structure in universes with cold dark matter, baryons, radiation, and massive and massless neutrinos. *ApJS*, 71:1–24, september 1989.

Hu W., Sugiyama N. Small-Scale Cosmological Perturbations: an Analytic Approach. *ApJ*, 471:542, november 1996.

Huff E. M., Hirata C. M., Mandelbaum R., Schlegel D., Seljak U., Lupton R. H. Seeing in the dark – I. Multi-epoch alchemy. *ArXiv e-prints*, november 2011.

Ilbert O., Arnouts S., McCracken H. J., Bolzonella M., Bertin E., Le Fèvre O. et al. Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey. *A&A*, 457:841–856, october 2006.

Ilbert O., Capak P., Salvato M., Aussel H., McCracken H. J., Sanders D. B. et al. Cosmos Photometric Redshifts with 30-Bands for $2 - deg^2$. *ApJ*, 690:1236–1249, january 2009.

Ivezic Z., Tyson J. A., Acosta E., Allsman R., Anderson S. F., Andrew J. et al. LSST: from Science Drivers to Reference Design and Anticipated Data Products. *ArXiv e-prints: 0805.2366*, may 2008.

Jarosik N., Bennett C. L., Dunkley J., Gold B., Greason M. R., Halpern M. et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results. *ApJS*, 192:14, february 2011.

Jelinsky P., Bebek C., Besuner R., Carton P.-H., Edelstein J., Lampton M. et al. The BigBOSS spectrograph. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8446 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, september 2012.

Jeong D., Komatsu E. Perturbation Theory Reloaded: Analytical Calculation of Nonlinearity in Baryonic Oscillations in the Real-Space Matter Power Spectrum. *ApJ*, 651:619–626, november 2006.

Jones D. H., Read M. A., Saunders W., Colless M., Jarrett T., Parker Q. A. et al. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. *MNRAS*, 399:683–698, october 2009.

Jouvel S., Kneib J., Ilbert O., Bernstein G., Arnouts S., Dahlen T. et al. Designing future dark energy space missions. I. Building realistic galaxy spectro-photometric catalogs and their first applications. *A&A*, 504:359–371, september 2009.

Jullo E., Rhodes J., Kiessling A., Taylor J. E., Massey R., Berge J. et al. COSMOS: Stochastic Bias from Measurements of Weak Lensing and Galaxy Clustering. *ApJ*, 750:37, may 2012.

Kaiser N. On the spatial correlations of Abell clusters. *ApJ*, 284:L9–L12, september 1984.

Kaiser N. A sparse-sampling strategy for the estimation of large-scale clustering from redshift surveys. *MNRAS*, 219:785–790, april 1986.

Kaiser N. Clustering in real space and in redshift space. *MNRAS*, 227:1–21, july 1987.

Kaiser N., Squires G., Broadhurst T. A Method for Weak Lensing Observations. *ApJ*, 449:460, august 1995.

Kazin E. A., Blanton M. R., Scoccimarro R., McBride C. K., Berlind A. A., Bahcall N. A. et al. The Baryonic Acoustic Feature and Large-Scale Clustering in the Sloan Digital Sky Survey Luminous Red Galaxy Sample. *ApJ*, 710:1444–1461, february 2010.

Keisler R., Reichardt C. L., Aird K. A., Benson B. A., Bleem L. E., Carlstrom J. E. et al. A Measurement of the Damping Tail of the Cosmic Microwave Background Power Spectrum with the South Pole Telescope. *ApJ*, 743:28, december 2011.

Kennicutt, Jr. R. C. Star Formation in Galaxies Along the Hubble Sequence. *ARA&A*, 36:189–232, 1998.

Kilbinger M., Benabed K., Cappe O., Cardoso J.-F., Coupon J., Fort G. et al. CosmoPMC: Cosmology Population Monte Carlo. *ArXiv e-prints*, january 2011.

Kilbinger M., Benabed K., Guy J., Astier P., Tereno I., Fu L. et al. Dark-energy constraints and correlations with systematics from CFHTLS weak lensing, SNLS supernovae Ia and WMAP5. *A&A*, 497:677–688, april 2009.

Kilbinger M., Schneider P., Eifler T. E- and B-mode mixing from incomplete knowledge of the shear correlation. *A&A*, 457:15–19, october 2006.

Kilbinger M., Wraith D., Robert C. P., Benabed K., Cappé O., Cardoso J.-F. et al. Bayesian model comparison in cosmology with Population Monte Carlo. *MNRAS*, 405:2381–2390, july 2010.

Kirby E. N., Guhathakurta P., Faber S. M., Koo D. C., Weiner B. J., Cooper M. C. The DEEP2 Galaxy Redshift Survey: Redshift Identification of Single-Line Emission Galaxies. *ApJ*, 660:62–71, may 2007.

Kirkby D., Margala D., Slosar A., Bailey S., Busca N. G., Delubac T. et al. Fitting Methods for Baryon Acoustic Oscillations in the Lyman-{\alpha} Forest Fluctuations in BOSS Data Release 9. *ArXiv e-prints*, january 2013.

Kitayama T., Suto Y. Semianalytic Predictions for Statistical Properties of X-Ray Clusters of Galaxies in Cold Dark Matter Universes. *ApJ*, 469:480, october 1996.

Kitching T. D., Balan S. T., Bridle S., Cantale N., Courbin F., Eifler T. et al. Image analysis for cosmology: results from the GREAT10 Galaxy Challenge. *MNRAS*, 423:3163–3208, july 2012.

Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Gold B., Hinshaw G. et al. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *ApJS*, 180:330–376, february 2009.

Komatsu E., Smith K. M., Dunkley J., Bennett C. L., Gold B., Hinshaw G. et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18–+, february 2011.

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B. et al. The Dark Side of the Halo Occupation Distribution. *ApJ*, 609:35–49, july 2004.

Lahav O., Lilje P. B., Primack J. R., Rees M. J. Dynamical effects of the cosmological constant. *MNRAS*, 251:128–136, july 1991.

Lamareille F. Spectral classification of emission-line galaxies from the Sloan Digital Sky Survey. I. An improved classification for high-redshift galaxies. *A&A*, 509:A53+, january 2010.

Lamareille F., Brinchmann J., Contini T., Walcher C. J., Charlot S., Pérez-Montero E. et al. Physical properties of galaxies and their evolution in the VIMOS VLT Deep Survey. I. The evolution of the mass-metallicity relation up to z ˜ 0.9. *A&A*, 495:53–72, february 2009.

Lamareille F., Contini T., Brinchmann J., Le Borgne J.-F., Charlot S., Richard J. Spectrophotometric properties of galaxies at intermediate redshifts (z ˜ 0.2-1.0). II. The Luminosity - Metallicity relation. *A&A*, 448:907–919, march 2006.

Landy S. D., Szalay A. S. Bias and variance of angular correlation functions. *ApJ*, 412:64–71, july 1993.

Laureijs R., Amiaux J., Arduini S., Augueères J. ., Brinchmann J., Cole R. et al. Euclid Definition Study Report. *ArXiv e-prints*, october 2011.

Le Fèvre O., Paltani S., Arnouts S., Charlot S., Foucaud S., Ilbert O. et al. A large population of galaxies 9 to 12 billion years back in the history of the Universe. *Nature*, 437:519–521, september 2005a.

Le Fèvre O., Vettolani G., Garilli B., Tresse L., Bottini D., Le Brun V. et al. The VIMOS VLT deep survey. First epoch VVDS-deep survey: 11 564 spectra with $17.5 < I_{AB} < 24$, and the redshift distribution over $0 < z < 5$. *A&A*, 439:845–862, september 2005b.

Le Goff J. M., Magneville C., Rollinde E., Peirani S., Petitjean P., Pichon C. et al. Simulations of BAO reconstruction with a quasar Ly-α survey. *A&A*, 534:A135, october 2011.

Leauthaud A., Finoguenov A., Kneib J.-P., Taylor J. E., Massey R., Rhodes J. et al. A Weak Lensing Study of X-ray Groups in the Cosmos Survey: Form and Evolution of the Mass-Luminosity Relation. *ApJ*, 709:97–114, january 2010.

Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H. A Theoretical Framework for Combining Techniques that Probe the Link Between Galaxies and Dark Matter. *ApJ*, 738:45, september 2011.

Leauthaud A., Tinker J., Bundy K., Behroozi P. S., Massey R., Rhodes J. et al. New Constraints on the Evolution of the Stellar-to-dark Matter Connection: A Combined Analysis of Galaxy-Galaxy Lensing, Clustering, and Stellar Mass Functions from z = 0.2 to z =1. *ApJ*, 744:159, january 2012.

Lee K.-G., Bailey S., Bartsch L. E., Carithers W., Dawson K. S., Kirkby D. et al. The BOSS Lyα Forest Sample from SDSS Data Release 9. *AJ*, 145:69, march 2013.

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z. The dependence of clustering on galaxy properties. *MNRAS*, 368:21–36, may 2006.

Lightman A. P., Schechter P. L. The Omega dependence of peculiar velocities induced by spherical density perturbations. *ApJS*, 74:831, december 1990.

Lilly S. J., Le Brun V., Maier C., Mainieri V., Mignoli M., Scodeggio M. et al. The zCOSMOS 10k-Bright Spectroscopic Sample. *ApJS*, 184:218–229, october 2009.

Limber D. N. The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II. *ApJ*, 119:655, may 1954.

Luppino G. A., Kaiser N. Detection of Weak Lensing by a Cluster of Galaxies at Z = 0.83. *ApJ*, 475: 20, january 1997.

Ma C.-P., Fry J. N. Deriving the Nonlinear Cosmological Power Spectrum and Bispectrum from Analytic Dark Matter Halo Profiles and Mass Functions. *ApJ*, 543:503–513, november 2000.

Madore B. F., Freedman W. L. The Cepheid distance scale. *PASP*, 103:933–957, september 1991.

Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J. Galaxy halo masses and satellite fractions from galaxy-galaxy lensing in the Sloan Digital Sky Survey: stellar mass, luminosity, morphology and environment dependencies. *MNRAS*, 368:715–731, may 2006.

Mantz A., Allen S. W., Rapetti D., Ebeling H. The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints. *MNRAS*, 406:1759–1772, august 2010.

Marocco J., Hache E., Lamareille F. Spectral classification of emission-line galaxies from the Sloan Digital Sky Survey. II. A supplementary diagnostic for AGNs using the $D_n(4000)$ index. *A&A*, 531: A71, july 2011.

Martin D. C., Fanson J., Schiminovich D., Morrissey P., Friedman P. G., Barlow T. A. et al. The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission. *ApJ*, 619:L1–L6, january 2005.

Massey R., Heymans C., Bergé J., Bernstein G., Bridle S., Clowe D. et al. The Shear Testing Programme 2: Factors affecting high-precision weak-lensing analyses. *MNRAS*, 376:13–38, march 2007.

Matsubara T. Correlation Function in Deep Redshift Space as a Cosmological Probe. *ApJ*, 615: 573–585, november 2004.

McDonald P. Toward a Measurement of the Cosmological Geometry at z ˜ 2: Predicting Lyα Forest Correlation in Three Dimensions and the Potential of Future Data Sets. *ApJ*, 585:34–51, march 2003.

McDonald P., Eisenstein D. J. Dark energy and curvature from a future baryonic acoustic oscillation survey using the Lyman-α forest. *Phys. Rev. D*, 76(6):063009, september 2007.

McDonald P., Roy A. Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS. *J. Cosmology Astropart. Phys.*, 8:20, august 2009.

Mehta K. T., Cuesta A. J., Xu X., Eisenstein D. J., Padmanabhan N. A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - III. Cosmological measurements and interpretation. *MNRAS*, 427:2168–2179, december 2012.

Mehta K. T., Seo H.-J., Eckel J., Eisenstein D. J., Metchnik M., Pinto P. et al. Galaxy Bias and Its Effects on the Baryon Acoustic Oscillation Measurements. *ApJ*, 734:94, june 2011.

Meiksin A., White M., Peacock J. A. Baryonic signatures in large-scale structure. *MNRAS*, 304: 851–864, april 1999.

More S., van den Bosch F., Cacciato M., More A., Mo H., Yang X. Cosmological Constraints from a Combination of Galaxy Clustering and Lensing – II. Fisher Matrix Analysis. *ArXiv e-prints*, june 2012.

Mostek N., Barbary K., Bebek C. J., Dey A. T., Edelstein J., Jelinsky P. et al. Mapping the universe with BigBOSS. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8446 de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, september 2012a.

Mostek N., Coil A. L., Cooper M. C., Davis M., Newman J. A., Weiner B. The DEEP2 Galaxy Redshift Survey: Clustering Dependence on Galaxy Stellar Mass and Star Formation Rate at z˜1. *ArXiv e-prints*, october 2012b.

Newman J. A., Cooper M. C., Davis M., Faber S. M., Coil A. L., Guhathakurta P. et al. The DEEP2 Galaxy Redshift Survey: Design, Observations, Data Reduction, and Redshifts. *ArXiv e-prints*, march 2012.

Newman J. A., Davis M. Measuring the Cosmic Equation of State with Counts of Galaxies. II. Error Budget for the DEEP2 Redshift Survey. *ApJ*, 564:567–575, january 2002.

Nolta M. R., Dunkley J., Hill R. S., Hinshaw G., Komatsu E., Larson D. et al. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Angular Power Spectra. *ApJS*, 180:296–305, february 2009.

Nuza S. E., Sanchez A. G., Prada F., Klypin A., Schlegel D. J., Gottloeber S. et al. The clustering of galaxies at z˜0.5 in the SDSS-III Data Release 9 BOSS-CMASS sample: a test for the LCDM cosmology. *ArXiv e-prints*, february 2012.

Oke J. B., Gunn J. E. Secondary standard stars for absolute spectrophotometry. *ApJ*, 266:713–717, march 1983.

Padmanabhan N., Schlegel D. J., Seljak U., Makarov A., Bahcall N. A., Blanton M. R. et al. The clustering of luminous red galaxies in the Sloan Digital Sky Survey imaging data. *MNRAS*, 378: 852–872, july 2007.

Padmanabhan N., Tegmark M., Hamilton A. J. S. The Power Spectrum of the CFA/SSRS UZC Galaxy Redshift Survey. *ApJ*, 550:52–64, march 2001.

Padmanabhan N., White M. Calibrating the baryon oscillation ruler for matter and halos. *Phys. Rev. D*, 80(6):063508, september 2009.

Padmanabhan N., Xu X., Eisenstein D. J., Scalzo R., Cuesta A. J., Mehta K. T. et al. A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey. *MNRAS*, 427:2132–2145, december 2012.

Palanque-Delabrouille N., Yeche C., Myers A. D., Petitjean P., Ross N. P., Sheldon E. et al. Variability selected high-redshift quasars on SDSS Stripe 82. *A&A*, 530:A122, june 2011.

Pâris I., Petitjean P., Aubourg É., Bailey S., Ross N. P., Myers A. D. et al. The Sloan Digital Sky Survey quasar catalog: ninth data release. *A&A*, 548:A66, december 2012.

Parkinson D., Riemer-Sørensen S., Blake C., Poole G. B., Davis T. M., Brough S. et al. The WiggleZ Dark Energy Survey: Final data release and cosmological results. *Phys. Rev. D*, 86(10):103518, november 2012.

Peebles P. J. E. *Principles of Physical Cosmology.* 1993.

Peebles P. J. E., Ratra B. Cosmology with a time-variable cosmological 'constant'. *ApJ*, 325:L17–L20, february 1988.

Peebles P. J. E., Yu J. T. Primeval Adiabatic Perturbation in an Expanding Universe. *ApJ*, 162:815, december 1970.

Penzias A. A., Wilson R. W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ*, 142:419–421, july 1965.

Percival W. J., Reid B. A., Eisenstein D. J., Bahcall N. A., Budavari T., Frieman J. A. et al. Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample. *MNRAS*, 401: 2148–2168, february 2010.

Perlmutter S., Aldering G., Goldhaber G., Knop R. A., Nugent P., Castro P. G. et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *ApJ*, 517:565–586, june 1999.

Planck, AMI Collaborations, :, Aghanim N., Arnaud M., Ashdown M. et al. Planck Intermediate Results II: Comparison of Sunyaev-Zeldovich measurements from Planck and from the Arcminute Microkelvin Imager for 11 galaxy clusters. *ArXiv e-prints*, april 2012.

Planck Collaboration, Ade P. A. R., Aghanim N., Armitage-Caplan C., Arnaud M., Ashdown M. et al. Planck 2013 results. XVI. Cosmological parameters. *ArXiv e-prints*, march 2013.

Planck Collaboration, Ade P. A. R., Aghanim N., Arnaud M., Ashdown M., Atrio-Barandela F. et al. Planck intermediate results. III. The relation between galaxy cluster mass and Sunyaev-Zeldovich signal. *ArXiv e-prints*, april 2012.

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J. Halo concentrations in the standard Λ cold dark matter cosmology. *MNRAS*, 423:3018–3030, july 2012.

Press W. H., Schechter P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, february 1974.

Rapetti D., Allen S. W., Mantz A., Ebeling H. The observed growth of massive galaxy clusters - III. Testing general relativity on cosmological scales. *MNRAS*, 406:1796–1804, august 2010.

Regnault N., Conley A., Guy J., Sullivan M., Cuillandre J., Astier P. et al. Photometric calibration of the Supernova Legacy Survey fields. *A&A*, 506:999–1042, november 2009.

Reichardt C. L., Shaw L., Zahn O., Aird K. A., Benson B. A., Bleem L. E. et al. A Measurement of Secondary Cosmic Microwave Background Anisotropies with Two Years of South Pole Telescope Observations. *ApJ*, 755:70, august 2012.

Reid B. A., Samushia L., White M., Percival W. J., Manera M., Padmanabhan N. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measurements of the growth of structure and expansion rate at z = 0.57 from anisotropic clustering. *MNRAS*, 426:2719–2737, november 2012.

Reyes R., Mandelbaum R., Seljak U., Baldauf T., Gunn J. E., Lombriser L. et al. Confirmation of general relativity on large scales from weak lensing and galaxy velocities. *Nature*, 464:256–258, march 2010.

Riebe K., Partl A. M., Enke H., Forero-Romero J., Gottloeber S., Klypin A. et al. The MultiDark Database: Release of the Bolshoi and MultiDark Cosmological Simulations. *ArXiv e-prints*, august 2011.

Riess A. G., Filippenko A. V., Challis P., Clocchiatti A., Diercks A., Garnavich P. M. et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038, september 1998.

Riess A. G., Macri L., Casertano S., Lampeitl H., Ferguson H. C., Filippenko A. V. et al. A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3. *ApJ*, 730:119, april 2011.

Riess A. G., Strolger L.-G., Casertano S., Ferguson H. C., Mobasher B., Gold B. et al. New Hubble Space Telescope Discoveries of Type Ia Supernovae at redshift 1: Narrowing Constraints on the Early Behavior of Dark Energy. *ApJ*, 659:98–121, april 2007.

Riess A. G., Strolger L.-G., Tonry J., Casertano S., Ferguson H. C., Mobasher B. et al. Type Ia Supernova Discoveries at redshift 1 from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *ApJ*, 607:665–687, june 2004.

Ross A. J., Percival W. J., Sánchez A. G., Samushia L., Ho S., Kazin E. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics. *MNRAS*, 424:564–590, july 2012a.

Ross N. P., Myers A. D., Sheldon E. S., Yèche C., Strauss M. A., Bovy J. et al. The SDSS-III Baryon Oscillation Spectroscopic Survey: Quasar Target Selection for Data Release Nine. *ApJS*, 199: 3, march 2012b.

Rozo E., Wechsler R. H., Rykoff E. S., Annis J. T., Becker M. R., Evrard A. E. et al. Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog. *ApJ*, 708:645–660, january 2010.

Sachs R. K., Wolfe A. M. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *ApJ*, 147:73, january 1967.

Samushia L., Reid B. A., White M., Percival W. J., Cuesta A. J., Lombriser L. et al. The clustering of galaxies in the SDSS-III DR9 Baryon Oscillation Spectroscopic Survey: testing deviations from Λ and general relativity using anisotropic clustering of galaxies. *MNRAS*, 429:1514–1528, february 2013.

Sánchez A. G., Crocce M., Cabré A., Baugh C. M., Gaztañaga E. Cosmological parameter constraints from SDSS luminous red galaxies: a new treatment of large-scale clustering. *MNRAS*, 400:1643–1664, december 2009.

Sanchez A. G., Kazin E. A., Beutler F., Chuang C.-H., Cuesta A. J., Eisenstein D. J. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological constraints from the full shape of the clustering wedges. *ArXiv e-prints*, march 2013.

Schlegel D., Abdalla F., Abraham T., Ahn C., Allende Prieto C., Annis J. et al. The BigBOSS Experiment. *ArXiv e-prints*, june 2011.

Schlegel D. J., Finkbeiner D. P., Davis M. Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds. *ApJ*, 500:525, june 1998.

Schmidt F. Cosmological simulations of normal-branch braneworld gravity. *Phys. Rev. D*, 80(12): 123003, december 2009.

Schneider P., van Waerbeke L., Mellier Y., Jain B., Seitz S., Fort B. Detection of shear due to weak lensing by large-scale structure. *A&A*, 333:767–778, may 1998.

Schulz A. E., White M. Scale-dependent bias and the halo model. *Astroparticle Physics*, 25:172–177, march 2006.

Scoville N., Abraham R. G., Aussel H., Barnes J. E., Benson A., Blain A. W. et al. COSMOS: Hubble Space Telescope Observations. *The Astrophysical Journal Supplement Series*, 172:38–45, september 2007a.

Scoville N., Aussel H., Brusa M., Capak P., Carollo C. M., Elvis M. et al. The Cosmic Evolution Survey (COSMOS): Overview. *ApJS*, 172:1–8, september 2007b.

Seljak U. Analytic model for galaxy and dark matter clustering. *MNRAS*, 318:203–213, october 2000.

Seljak U., Zaldarriaga M. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *ApJ*, 469:437, october 1996.

Seo H.-J., Eisenstein D. J. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. *ApJ*, 598:720–740, december 2003.

Seo H.-J., Eisenstein D. J. Baryonic Acoustic Oscillations in Simulated Galaxy Redshift Surveys. *ApJ*, 633:575–588, november 2005.

Shan H., Kneib J.-P., Tao C., Fan Z., Jauzac M., Limousin M. et al. Weak Lensing Measurement of Galaxy Clusters in the CFHTLS-Wide Survey. *ApJ*, 748:56, march 2012.

Shannon C. E., Weaver W. *The mathematical theory of communication*. 1975.

Sheldon E. S., Johnston D. E., Masjedi M., McKay T. A., Blanton M. R., Scranton R. et al. Cross-correlation Weak Lensing of SDSS Galaxy Clusters. III. Mass-to-Light Ratios. *ApJ*, 703:2232–2248, october 2009.

Sheth R. K., Mo H. J., Tormen G. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12, may 2001.

Sheth R. K., Tormen G. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, september 1999.

Simon P. How accurate is Limber's equation? *A&A*, 473:711–714, october 2007.

Simon P., Hetterscheidt M., Schirmer M., Erben T., Schneider P., Wolf C. et al. GaBoDS: The Garching-Bonn Deep Survey. VI. Probing galaxy bias using weak gravitational lensing. *A&A*, 461: 861–879, january 2007.

Skibba R. A., Sheth R. K. A halo model of galaxy colours and clustering in the Sloan Digital Sky Survey. *MNRAS*, 392:1080–1091, january 2009.

Slosar A., Iršič V., Kirkby D., Bailey S., Busca N. G., Delubac T. et al. Measurement of baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9. *J. Cosmology Astropart. Phys.*, 4:26, april 2013.

Smee S., Gunn J. E., Uomoto A., Roe N., Schlegel D., Rockosi C. M. et al. The Multi-Object, Fiber-Fed Spectrographs for SDSS and the Baryon Oscillation Spectroscopic Survey. *ArXiv e-prints*, august 2012.

Smith R. E., Peacock J. A., Jenkins A., White S. D. M., Frenk C. S., Pearce F. R. et al. Stable clustering, the halo model and non-linear cosmological power spectra. *MNRAS*, 341:1311–1332, june 2003.

Smoot G. F., Bennett C. L., Kogut A., Wright E. L., Aymon J., Boggess N. W. et al. Structure in the COBE differential microwave radiometer first-year maps. *ApJ*, 396:L1–L5, september 1992.

Spergel D. N., Verde L., Peiris H. V., Komatsu E., Nolta M. R., Bennett C. L. et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *ApJS*, 148:175–194, september 2003.

Sunyaev R. A., Zeldovich Y. B. Small-Scale Fluctuations of Relic Radiation. *Ap&SS*, 7:3–19, april 1970.

Sunyaev R. A., Zeldovich Y. B. Formation of Clusters of Galaxies; Protocluster Fragmentation and Intergalactic Gas Heating. *A&A*, 20:189, august 1972.

Suzuki N., Rubin D., Lidman C., Aldering G., Amanullah R., Barbary K. et al. The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above $z > 1$ and Building an Early-type-hosted Supernova Sample. *ApJ*, 746:85, february 2012.

Tegmark M. Measuring Cosmological Parameters with Galaxy Surveys. *Physical Review Letters*, 79: 3806–3809, november 1997.

Tegmark M., Bromley B. C. Observational Evidence for Stochastic Biasing. *ApJ*, 518:L69–L72, june 1999.

Tegmark M., Peebles P. J. E. The Time Evolution of Bias. *ApJ*, 500:L79, june 1998.

Tegmark M., Taylor A. N., Heavens A. F. Karhunen-Loeve Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets? *ApJ*, 480:22, may 1997.

The Dark Energy Survey Collaboration. The Dark Energy Survey. *ArXiv e-prints: 0510.346*, october 2005.

Thomas D., Steele O., Maraston C., Johansson J., Beifiori A., Pforr J. et al. Stellar velocity dispersions and emission line properties of SDSS-III/BOSS galaxies. *ArXiv e-prints*, july 2012.

Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I. On the Mass-to-Light Ratio of Large-Scale Structure. *ApJ*, 631:41–58, september 2005.

Tremonti C. A., Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., White S. D. M. et al. The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *ApJ*, 613:898–913, october 2004a.

Tremonti C. A., Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., White S. D. M. et al. The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *ApJ*, 613:898–913, october 2004b.

Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J. Galaxies in ΛCDM with Halo Abundance Matching: Luminosity-Velocity Relation, Baryonic Mass-Velocity Relation, Velocity Function, and Clustering. *ApJ*, 742:16, november 2011.

Vale A., Ostriker J. P. Linking halo mass to galaxy luminosity. *MNRAS*, 353:189–200, september 2004.

van den Bosch F., More S., Cacciato M., Mo H., Yang X. Cosmological Constraints from a Combination of Galaxy Clustering and Lensing – I. Theoretical Framework. *ArXiv e-prints*, june 2012.

Vargas-Magaña M., Bautista J. E., Hamilton J.-C., Busca N. G., Aubourg É., Labatie A. et al. An optimized correlation function estimator for galaxy surveys. *ArXiv e-prints*, november 2012.

Vikhlinin A., Kravtsov A. V., Burenin R. A., Ebeling H., Forman W. R., Hornstrup A. et al. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *ApJ*, 692:1060–1074, february 2009.

Vives S., Le Mignant D., Madec F., Jaquet M., Prieto E., Martin L. et al. A spectrograph instrument concept for the Prime Focus Spectrograph (PFS) on Subaru Telescope. *ArXiv e-prints*, october 2012.

Vogeley M. S., Szalay A. S. Eigenmode Analysis of Galaxy Redshift Surveys. I. Theory and Methods. *ApJ*, 465:34, july 1996.

Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L. Moving mesh cosmology: numerical techniques and global statistics. *MNRAS*, 425:3024–3057, october 2012.

Wake D. A., Whitaker K. E., Labbé I., van Dokkum P. G., Franx M., Quadri R. et al. Galaxy Clustering in the NEWFIRM Medium Band Survey: The Relationship Between Stellar Mass and Dark Matter Halo Mass at $1 < z < 2$. *ApJ*, 728:46, february 2011.

Wall J. V., Jenkins C. R. *Practical Statistics for Astronomers*. april 2012.

Weinberg N. N., Kamionkowski M. Constraining dark energy from the abundance of weak gravitational lenses. *MNRAS*, 341:251–262, may 2003.

White M., Blanton M., Bolton A., Schlegel D., Tinker J., Berlind A. et al. The Clustering of Massive Galaxies at z ~ 0.5 from the First Semester of BOSS Data. *ApJ*, 728:126, february 2011.

Wilson M. J., Sherwin B. D., Hill J. C., Addison G., Battaglia N., Bond J. R. et al. Atacama Cosmology Telescope: A measurement of the thermal Sunyaev-Zel'dovich effect using the skewness of the CMB temperature distribution. *Phys. Rev. D*, 86(12):122005, december 2012.

Wood-Vasey W. M., Miknaitis G., Stubbs C. W., Jha S., Riess A. G., Garnavich P. M. et al. Observational Constraints on the Nature of Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey. *ApJ*, 666:694–715, september 2007.

Wright E. L., Eisenhardt P. R. M., Mainzer A. K., Ressler M. E., Cutri R. M., Jarrett T. et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *AJ*, 140:1868–1881, december 2010.

Wright E. L., Meyer S. S., Bennett C. L., Boggess N. W., Cheng E. S., Hauser M. G. et al. Interpretation of the cosmic microwave background radiation anisotropy detected by the COBE Differential Microwave Radiometer. *ApJ*, 396:L13–L18, september 1992.

Xu X., Padmanabhan N., Eisenstein D. J., Mehta K. T., Cuesta A. J. A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - II. Fitting techniques. *MNRAS*, 427:2146–2167, december 2012.

York D. G., Adelman J., Anderson, Jr. J. E., Anderson S. F., Annis J., Bahcall N. A. et al. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120:1579–1587, september 2000.

Zehavi I., Zheng Z., Weinberg D. H., Blanton M. R., Bahcall N. A., Berlind A. A. et al. Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity. *ApJ*, 736:59, july 2011.

Zel'dovich Y. B. Gravitational instability: An approximate theory for large density perturbations. *A&A*, 5:84–89, march 1970.

Zhang P., Liguori M., Bean R., Dodelson S. Probing Gravity at Cosmological Scales by Measurements which Test the Relationship between Gravitational Lensing and Matter Overdensity. *Physical Review Letters*, 99(14):141302, october 2007.

Zhu G., Moustakas J., Blanton M. R. The [O II] λ3727 Luminosity Function at z ˜ 1. *ApJ*, 701: 86–93, august 2009.

Chapter 6

Résumé en français

La force dominante à l'échelle du cosmos est la gravité. Ainsi une description des la dynamique et de la structure de l'Univers sera basée sur une théorie de la gravité. Aujourd'hui, c'est la théorie de la relativité générale d'Einstein qui est utilisée (Einstein, 1916). C'est une description de l'espace temps par une métrique déterminée par le contenu en matière et en énergie de l'Univers. Les équations d'Einstein relient le tenseur d'Einstein, qui décrit la géométrie de l'espace temps et le tenseur énergie-impulsion qui prend en compte la matière et l'énergie. Nous ne nous épancherons pas sur la topologie de l'espace-temps, ce travail de thèse est effectué dans le modèle standard (Peebles, 1993).

Pour un univers homogène rempli de matière ou de rayonnement, la relativité générale prédit que l'expansion cosmique va ralentir au fil du temps, en accord avec l'intuition newtonienne. Cependant, des études de supernova lointaines ont constaté que l'expansion de l'univers s'accélère au cours des cinq derniers milliards d'années (Riess et al., 1998; Perlmutter et al., 1999). La compréhension de cette accélération est un vrai défi, plus connu sous le nom de la question de l'énergie noire. Cinq expériences sont en cours (ou prévues) pour mieux comprendre la nature de l'énergie noire: des observations plus précises de supernovae, et du fond diffus cosmologique, l'oscillation acoustique baryonique dans le regroupement de galaxies (BAO), le lentillage gravitationnel faible, et l'étude des amas de galaxies.

6.1 Cosmologie

Les sondes cosmologiques sont à l'intersection entre les observables prédites par la théorie et les objets astrophysiques que nous observons, je donne ici un bref panorama des sondes cosmologiques actuelles. Je traite les oscillations acoustiques des baryons séparément plus en détails, dans le section suivante.

6.1.1 Fond diffus cosmologique

Le fond diffus cosmologique (Cosmic Microwave Background, abrégé CMB) est constitué de photons émis au moment du découplage entre la matière et les photons à un décalage vers le rouge $z \sim 1100$. Il a été prédit en 1940 par Ralph Alpher, Robert Herman et George Gamow, et observé par Penzias and Wilson (1965). Il représente le spectre du corps noir de l'Univers.

Le spectre de puissance des fluctuations de température contient quatre catégories d'information. Le dipôle permet de mesurer la vitesse de l'observateur (celle du groupe local). Aux grandes échelles, l'ampleur des fluctuations du spectre de puissance de la température fournit des informations sur le voyage du photon issu de la surface de dernière diffusion du CMB jusqu'à l'observateur et donc sur les perturbations de la métrique rencontrées par le photon sur le chemin, qui entraînent des changements de fréquence dans les deux directions rouge et bleu (monter un potentiel provoque décalage vers le rouge, et la dilatation du temps lors de l'ascension provoque un décalage vers le bleu) (Sachs and Wolfe, 1967). Les échelles intermédiaires sondent les perturbations de vitesse sur la surface de dernière diffusion et les oscillations acoustiques à l'intérieur de l'horizon (Seljak and Zaldarriaga, 1996). Les petites échelles angulaires dépendent de la teneur en matière noire et les baryons.

Anisotropie extrinsèque, le déficit de photons du CMB

Lorsque les photons du CMB entrent dans un plasma chaud relativiste, ils sont envoyés selon le processus inverse de Compton à des énergies plus élevées. Les photons repartent dans une direction aléatoire et sont ainsi perdus, le CMB semble donc plus froid. Cet effet, appelé effet Sunyaev-Zel'dovich, est mesuré dans les amas de galaxies contenant des gaz à une température de 10^8 K; (Sunyaev and Zeldovich, 1972). Planck a mesuré cet effet[1] (Planck et al., 2012; Planck Collaboration et al., 2012).

Observations et résultats

De nombreuses expériences au sol et trois satellites ont mesuré avec une grande précision le fond diffus cosmologique. Le premier satellite de la NASA Cosmic Background Explorer (COBE) a mesuré la température du fond diffus cosmologique $T_{\rm CMB} = 2.728 \pm 0.002$K et aussi son anisotropie avec une résolution angulaire de l'ordre de 7 degrés de l'ordre de $\Delta T/T \sim 2 \times 10^{-5}$ (Boggess et al., 1992; Smoot et al., 1992; Wright et al., 1992; Bennett et al., 1992).

Le deuxième satellite de la NASA Wilkinson Microwave Anisotropy Probe (WMAP) (Spergel et al., 2003) avec une résolution angulaire de $\sim 0.2°$ a amélioré la mesure du CMB et aussi la compréhension de la cosmologie en calculant le spectre de puissance des fluctuations de l'anisotropie de température. En outre, grâce à la résolution angulaire faible, il a été possible de mesurer l'amplitude des fluctuations de recombinaison, qui est liée aux fluctuations initiales du champ de la matière, voir (Bennett et al., 2012; Hinshaw et al., 2012).

Au même moment, des observations des fluctuations du CMB aux petites échelles ont été effectuées à partir du sol par le télescope du pôle Sud (South Pole Telescope, SPT) (Keisler et al., 2011; Reichardt et al., 2012) et par l'Atacama Cosmology Telescope (ACT) (Fowler et al., 2010; Wilson et al., 2012; Dünner et al., 2013). Ces expériences mesurent le spectre de puissance jusqu'à une échelle d'une seconde d'arc.

Le troisième satellite de l'ESA Planck vient de compléter sa première année d'observation, et d'excellents résultats ont été publiés dans un communiqué, voir Planck Collaboration et al. (2013) à propos de la cosmologie. Des résultats plus complets sont attendus dans quelques années.

Aujourd'hui l'observation du CMB est la sonde la plus efficace pour éliminer les modèles cosmologiques.

6.1.2 Céphéides

Les étoiles variables appelées céphéides voient leurs magnitudes apparentes changer régulièrement, avec des périodes allant de 2 à 150 jours. Cette période est étroitement corrélée à leur luminosité. Une fois la relation période-luminosité calibrée (voir Madore and Freedman (1991); Feast and Catchpole (1997) pour l'étalonnage de la relation), elle peut être appliquée à des céphéides lointaines (dont les distances ne peuvent pas être obtenus à partir de parallaxe) afin d'obtenir leurs distances directement à partir des mesures de leurs périodes de variation. Etant donné que les céphéides sont relativement lumineuses, des télescopes avec une résolution spatiale suffisamment élevée, tels que le télescope spatial Hubble (HST), permettent de déterminer les distances de céphéides éloignées jusqu'à $\sim$ 10 Mpc. La constante de Hubble H_0 peut être obtenue en mesurant le décalage vers le rouge et la distance d'un objet (en ignorant la vitesse propre de l'objet) par l'intermédiaire de la loi expansion de Hubble $cz = d.H_0$. Les dernières estimations de la constante de Hubble avec cette méthode donnent $H_0 = 74.3 \pm 3.6 {\rm km\,s^{-1}\,Mpc^{-1}}$ (Freedman et al., 2012).

6.1.3 Supernova

Une supernova de type 1a (SN1a) est l'explosion d'une étoile qui a un profil de luminosité bien calibré (Branch and Tammann, 1992), qui peut atteindre des luminosités très élevées et donc observables sur des distances cosmologiques de plusieurs milliers de mégaparsecs. Ainsi, on peut obtenir les distances par rapport aux supernova de type 1a en mesurant leurs courbes de lumière. Les distances absolues sont obtenue après une calibration de luminosité sur les supernova de type 1a proches, elles-mêmes calibrées sur les céphéides par exemple.

[1] http://www.sciops.esa.int/index.php?project=PLANCK&page=Planck_Published_Papers

Ensuite, la comparaison de la distance des supernova à leur décalage vers le rouge permet de mesurer la loi de Hubble. L'observation des supernova est à l'origine de la constatation que l'univers est en expansion accélérée (Riess et al., 1998; Perlmutter et al., 1999). Aujourd'hui, elle constitue une preuve pour une constante cosmologique. Depuis les premières détections, il y a eu une augmentation énorme du nombre de données à décalage vers le rouge $z \sim 1$ (Riess et al., 1998; Perlmutter et al., 1999; Astier et al., 2006; Wood-Vasey et al., 2007; Riess et al., 2004, 2007; Dawson et al., 2009; Riess et al., 2011). La précision nécessaire pour obtenir des analyses approfondies du modèle cosmologique en utilisant les SN1a est limitée par les erreurs systématiques de cette sonde, par conséquent, une analyse statistique jointe avec d'autres sondes est obligatoire pour obtenir une image consolidée du modèle cosmologique.

6.1.4 Lentillage gravitationnel faible

Le lentillage gravitationnel faible (Weak Lensing, WL) est l'étude de la distorsion de la forme des galaxies observées en raison de la présence de matière entre la galaxie déformée et l'observateur. Bien que son utilisation est techniquement difficile, c'est un outil très puissant pour mesurer la quantité de matière présente dans l'Univers.
Les mesures du sondage Sloan, SDSS[2], ont fourni des résultats cruciaux concernant la matière noire autour des galaxies (Mandelbaum et al., 2006) et autour des amas de galaxies (Sheldon et al., 2009) par corrélation croisée entre les galaxies identifiées par spectroscopie et le champ de cisaillement des galaxies d'arrière-plan (analyse par 'stack' du très faible signal autour des galaxies sélectionnées ou des amas). Ces mesures sont fondamentales pour la compréhension de la relation entre les halos de matière noire et les galaxies (Leauthaud et al., 2010, 2012) et pour l'utilisation des amas de galaxies comme sondes cosmologiques. En outre, la mesure de cisaillement gravitationnel sur un grand échantillon de galaxies en arrière-plan de l'échantillon du sondage spectroscopique permet également des tests de la relativité générale, à savoir si les deux potentiels scalaires gravitationnels dans la jauge newtonienne sont égaux (Zhang et al., 2007; Reyes et al., 2010). Dans cette thèse, j'utilise les lentilles gravitationnelles afin d'estimer le biais des galaxies qui serviront à observer les oscillations acoustiques des baryons. Les astronomes ont appliqué ces techniques sont le sondage COSMOS (Scoville et al., 2007a; Leauthaud et al., 2012; Jullo et al., 2012) et le sondage CFHT-Lens(Heymans et al., 2012; Erben et al., 2012).

6.1.5 Amas de galaxies

L'abondance des amas de galaxies est une sonde cosmologique reconnue (Vikhlinin et al., 2009; Mantz et al., 2010; Rozo et al., 2010). Les contraintes actuelles sont $\Omega_{\text{matter}} = 0.255 \pm 0.043$. Cette sonde est sensible, non seulement à la géométrie sous-jacente de l'univers, mais aussi à la croissance des structures à grande échelle, ce qui en fait une sonde pour mesurer une éventuelle modification de la gravité (Schmidt, 2009; Rapetti et al., 2010). Les prochains sondages, comme le Dark Energy Survey[3], seront en mesure de détecter environ 20 amas de galaxies par degré carré jusqu'à $z = 1$, ce qui se traduira par une amélioration dans les contraintes cosmologiques à partir des amas de galaxies.

6.2 Oscillation acoustique des baryons

L'empreinte des oscillations acoustiques des baryons de l'Univers primordial dans le regroupement de la matière observée aujourd'hui est apparue comme un moyen efficace pour mesurer l'échelle des distances et l'histoire de l'expansion de l'univers (Peebles and Yu, 1970; Sunyaev and Zeldovich, 1970). Cette méthode est appelée l'oscillation acoustique baryonique (Baryonic Acoustic Oscillations, BAO) et repose sur la mesure de l'empreinte laissée par les ondes sonores dans l'univers primitif sur le regroupement de la matière et des galaxies dans notre Univers proche. Il s'agit de mesurer le pic (ou

[2]http://www.sdss.org/
[3]www.darkenergysurvey.org

l'échelle) acoustique dans la fonction de corrélation à deux points en trois dimensions du champs de galaxies constitue une règle étalon robuste pour sonder la géométrie de l'univers à différents décalages vers le rouge.La distribution des galaxies dans l'univers est donc une sonde directe de l'histoire de la formation des structures dans l'univers.

Nous définissons l'échelle acoustique, r_s, comme la distance comobile que les ondes sonores peuvent se parcourir depuis le Big Bang jusqu'à la recombinaison $z \sim 1100$ (équation 1.1); Holtzman (1989); Hu and Sugiyama (1996); Eisenstein and Hu (1998). L'analyse actuelle des données de WMAP donne une mesure de l'échelle acoustique $r_s = 153.19$ Mpc avec une précision de 1%. Cela correspond sur le fond de ciel micro-ondes à un angle $\theta_A = r_s(z_{\rm rec})/d_A \simeq 0.6\,\rm deg$ (Jarosik et al., 2011). La dernière série de données Planck devrait permettre d'améliorer la précision d'un facteur 4. La théorie des perturbations associée aux simulations numériques montrent que l'imprécision de l'échelle acoustique à travers le temps cosmique est inférieure à 1%, ce qui en fait une excellente règle étalon standard (Eisenstein et al., 2007b). La mesure de la signature des BAO est d'un intérêt primordial pour la cosmologie car il fournit une information géométrique presque dépourvue d'erreur systématique et elle est complémentaire à d'autres sondes cosmologiques tels que le fond diffus cosmologique ou les effets de lentillage gravitationel faible. Cette règle standard ouvre une voie solide à l'étude de l'énergie noire. Pour interprétation la mesure du BAO à partir du champs de galaxies, une théorie concernant la formation des structures à grande échelle est nécessaire.

6.2.1 Regroupement gravitationnel de la matière dans le régime linéaire, la mise en place des oscillations acoustiques des baryons

La distance sur laquelle les ondes acoustiques peuvent se propager dans les premiers millions d'années de l'univers devient une échelle caractéristique, mesurable, non seulement dans les anisotropies CMB (Bennett et al., 2012; Hinshaw et al., 2012), mais aussi dans le regroupement des galaxies (Beutler et al., 2011; Blake et al., 2011c; Padmanabhan et al., 2012; Anderson et al., 2012). Cette échelle peut être calculée avec la théorie de perturbation linéaire une fois que l'on précise les rapports d'abondance baryon-photon et matière-rayonnement. Ceux-ci étant mesurés à partir des détails sur les pics acoustiques du CMB. Cela fait, on peut mesurer la distance angulaire et le paramètre de Hubble en fonction du décalage vers le rouge en utilisant de grands sondages spectroscopiques de galaxies.

Spectre de puissance de la matière

Nous supposons un spectre de puissance des fluctuations initiales de la matière $P_i(k) \propto k^{n_s}$, ce qui signifie que la gravitation n'a pas d'échelle préférée. Harrison et Zeldovich ont montré qu'une équirépartition de l'énergie à toutes les échelles correspond à $n_s = 1$. La théorie de l'inflation dit que les structures actuelles ont évolué à partir des petites fluctuations initiales des champs de la matière, de la densité, de la vitesse et de la température. La physique à des échelles plus petites que le rayon de courbure de l'Univers est bien décrite par l'hydrodynamique d'Euler complétée par la mécanique Newtonienne. C'est résumé par cinq équations, l'équation de continuité, l'équation du mouvement, l'équation de Poisson, l'équation d'état et l'équation de l'énergie. L'application de l'ensemble des cinq équations décrit les zones de grande densité et leur vitesse que l'on retrouve dans l'Univers observable. Dans l'approximation linéaire (densités et des vitesses très petites), deux solutions décrivent l'évolution du champ de densité $\delta(r,t) = \rho(r,t)/\bar{\rho} - 1$. Un mode croissant (noté $D_+(t)$) et un mode décroissant (noté $D_-(t)$); équation (1.2).

En ignorant le mode décroissant qui disparaît assez rapidement, on peut définir le taux de croissance de la densité, f; équation (1.3).
Ce taux de croissance dépend du décalage vers le rouge et du modèle cosmologique et peut être approximé d'une manière assez simple et précise par $f = \Omega_m(z)^{0.6} + \Omega_\Lambda(1+\Omega_m/2)/70$ (Lahav et al., 1991). Dans un espace euclidien, chaque mode a une évolution différente, en plus de l'évolution temporelle qui s'applique à tous les modes, celle-ci est codée dans la fonction de transfert. Cette fonction de transfert prend en compte des effets différents: compton drag, velocity overshoot, baryon

infall, adiabatic damping, Silk damping, and CDM growth suppression (Eisenstein and Hu, 1998). Elle est définie comme la comparaison entre les modes présents maintenant et ceux présents au début; équation (1.4)

Les principales échelles physiques et ordres de grandeurs impliqués dans la fonction de transfert dans le modèle ΛCDM sont les suivants (les applications numériques sont faites dans le modèle de concordance WMAP9 ΛCDM). A $z_{eq} \approx 3526$ l'Univers dominé par le rayonnement devient dominé par la matière. L'horizon de particules à z_{eq} est de 93 Mpc, et toute perturbation sur une échelle plus petite est supprimée. Puis entre z_{eq} et $z_d \approx 970$, lorsque les baryons sont libérés du Compton drag des photons, deux effets dominent. Tout d'abord, les perturbations à petite échelle dans le fluide photons-baryons se propagent sous forme d'ondes acoustiques à une vitesse $c_s \approx 0.56c$ dans l'horizon du son, qui est à z_d d'environ 167 Mpc. Deuxièmement, le couplage entre les photons et les baryons n'est pas parfait. Il y a des processus de diffusion, qui diminuent l'importance des perturbations à des échelles inférieures à 19 Mpc (Silk damping). La forme analytique de la fonction de transfert de Eisenstein and Hu (1998) atteint une précision de 10% pour la plupart des modèles d'Univers.

A partir de la fonction de transfert, du taux de croissance des structures et du spectre de puissance initiale (issu de l'inflation), on obtient le spectre de puissance linéaire à partir de l'équation (1.5). Pour décrire complètement le spectre de puissance, nous avons besoin de savoir son amplitude. La norme usuelle de l'amplitude du spectre de puissance dans un univers ΛCDM est donnée par une convolution du spectre de puissance avec une fonction fenêtre de 8 Mpc d'ouverture. Elle est reliée à la racine carré de la moyenne des fluctuations du spectre de puissance (σ_8) dans un rayon $R = 8$Mpc et à la densité critique nécessaire à l'effondrement sphérique sous gravité à $z = 0$ (δ_c). Les calculs théoriques précis de δ_c et l'observation de σ_8 sont donnés dans Lightman and Schechter (1990); Lahav et al. (1991); Kitayama and Suto (1996); Weinberg and Kamionkowski (2003).

Spectre de puissance des galaxies

Dans le modèle cosmologique standard, à grande échelle, la relation entre le champ de surdensité de galaxies δ_g et le champ de surdensité de la matière noire δ_m est décrite par le modèle de Kaiser (1987); McDonald and Roy (2009); équation (1.6). Dans ce modèle, le spectre de puissance des galaxies observables, P_g, est liée au spectre de puissance de la matière, P_m (équation 1.7). Comme les galaxies sont, dans une bonne approximation, distribuées de manière Poissonienne autour du champ de densité à grande échelle, le spectre de bruit blanc est donné par l'inverse de la densité du champs de galaxies $\bar{n}^{-1}$.

A l'ordre zéro, pour extraire l'information cosmologique contenu dans l'emplacement du pic BAO dans la galaxie fonction d'autocorrélation, il n'est pas nécessaire de comprendre la relation entre les galaxies utilisées et les halos de matière noire sous-jacents. Cela correspondrait à une précision de l'échelle de mesure BAO de l'ordre de 5%. Pour effectuer une analyse cent fois plus précise de la fonction BAO (objectif pour les expériences classées catégorie IV de la Dark Energy Task Force), il est nécessaire de mieux comprendre la relation entre les galaxies utilisées et de la matière noire. C'est-à-dire d'étudier le biais des galaxies $b_g(z)$, qui contient une mine d'informations.

6.2.2 Regroupement gravitationnel non linéaire: la distorsion du signal BAO

Le regroupement des galaxies ne correspond pas exactement au modèle prédit par la théorie linéaire du regroupement de la matière. L'effondrement gravitationnel non linéaire, le biais des galaxies, et les distorsions dues au décalage vers le rouge modifient la mesure du regroupement des galaxies. Et ainsi dégradent la précision avec laquelle on mesure l'échelle standard.

Afin de relier les perturbations initiales aux structures non-linéaires que nous voyons aujourd'hui, nous avons besoin de décrire les modèles actuels d'effondrement non linéaire. Zel'dovich (1970) a décrit le premier l'évolution non-linéaire d'une perturbation cohérente du champ de densité cosmique de façon Lagrangienne au premier ordre. Ce modèle montre que l'effondrement se produit d'abord dans une seule direction, produisant une structure en forme de feuille. En appliquant ce modèle à un champ de densité initiale gaussien, Press and Schechter (1974); Sheth et al. (2001) ont estimé la fonction de masse des objets effondrés par un regroupement hiérarchique. De nos jours, les simulations numériques, basées sur ces modèles poussées au second ordre de la théorie des perturbations Lagrangienne, nous permettent de regarder la formation des structures non-linéaires et leur évolution, voir (Vogelsberger et al., 2012). Des simulations ont montré que la formation de la structure non-linéaire et, dans une moindre mesure, les distorsions décalage vers le rouge effacent les harmoniques supérieures des oscillations acoustiques (Meiksin et al., 1999; Seo and Eisenstein, 2005), et aussi que le biais de regroupement galaxie est une erreur potentielle systématique sur la détermination de l'échelle acoustique (Eisenstein et al., 2007a,b).

Passons dans l'espace réel (le conjugué de l'espace de Fourier), où la signature acoustique est un pic à une séparation de ~ 150 Mpc dans la fonction de corrélation à deux points des galaxies (notée 2PCF, conjuguée de Fourier du spectre de puissance); pour expliquer l'impact des effets non-linéaires. Les effets non linéaires sont les suivants. D'abord le flux dans la direction du groupe dominant en terme de gravitation, qui déplace les galaxies de ~ 10 Mpc comobile par rapport à leurs positions initiales. Il déplace ainsi les paires séparées de 150 Mpc, et élargit le pic (dans l'espace de Fourier cela correspond à l'enveloppe exponentielle des oscillations qui diminue). Deuxièmement, les distorsions pour passer de l'espace réel à l'espace des décalages vers le rouge accroît encore cet effet, et élargit le pic. Ces effets proviennent de processus astrophysiques qui se produisent à petite échelle, < 20 Mpc. Troisièmement, avec le temps, le regroupement couple les modes, ce qui réduit le contraste du pic (Meiksin et al., 1999; Jeong and Komatsu, 2006). Enfin, l'augmentation du bruit due aux petites non-linéarités fera augmenter les valeurs des coefficients non diagonaux de la matrice de covariance de la 2PCF (Schulz and White, 2006), cette notion est résumée sous le titre de biais du halo dépendant de l'échelle ('scale dependant halo bias').

A priori, ces non-linéarités pourraient non seulement élargir le pic acoustique, mais aussi le déplacer *i.e.* induire des erreurs systématiques. Une erreur systématique signifie qu'il y a un effet qui déplace systématiquement toutes les paires à 150 Mpc à des séparations plus ou moins grandes. En ce qui concerne cette question, Eisenstein et al. (2007b) a montré que non linéarités dues à la gravitation ne peuvent introduire des changements que de $\sim$0.5% pour des objets avec un biais $b \sim 2$ à $z = 0$. Par conséquent, la séparation caractéristique de 150 Mpc n'est pas déplacé, mais est plutôt floutée sous l'action de formation de groupes de galaxies qui se déplacent autour de la matière de manière aléatoire (voir la figure 1.1).

L'effet d'ordre suivant est ainsi induite par le biais des galaxies. Padmanabhan and White (2009) a caractérisé l'évolution non linéaire du pic des BAO en utilisant une combinaison de la théorie des perturbations et de simulations N-corps. Ils ont confirmé que le pic acoustique tracé par la matière noire est à la fois élargi et déplacé lorsque les structures se forment, et que ce changement est bien décrite par la théorie des perturbations du second ordre. Ils ont démontré que les erreurs systématique théoriques sont plus petites que la précision statistique des sondages à venir. Une modélisation simple a le potentiel de réduire ces systématique en dessous de la variance cosmique. Récemment, Mehta et al. (2011) ont effectué une simulation pour quantifier si nous pouvons faire face aux effets du biais des galaxies en utilisant des algorithmes de reconstruction améliorés. Ils montrent qu'il est possible de corriger avec la reconstruction du décalage systématique du à un biais jusqu'à une précision de 0.15% voir figure 1.2. Par conséquent, pour atteindre une grande précision dans les sondages futurs, il est nécessaire de connaître le biais des traceurs utilisés avec une précision $\lesssim 10\%$.

En bref, l'échelle acoustique est une règle standard robuste, car elle est fondamentalement une signature à 150 Mpc du regroupement des galaxies. Nous avons montré ici que les dégradations issues des processus quasi-linéaires sur la signature du BAO sont bénignes. Ces effets peuvent être calculés avec une grande précision par des simulations N-corps. Le défi de la méthode reste, car les volumes sondés seront énormes. En particulier pour lier la théorie à l'observation, en décrivant les estimateurs de la fonction de corrélation à deux points ou du spectre de puissance.

6.2.3 Estimer le $P(k)$ dans un champs de galaxies

Il reste à estimer la 2PCF depuis les données, *i.e.* à partir d'un catalogue de positions de galaxies. Cette question a été traitée parFall (1979), puis par Landy and Szalay (1993); Hamilton (1993), et dernièrement par (Vargas-Magaña et al., 2012). Fall (1979) a dérivé que la 2PCF est la probabilité conjointe de trouver des galaxies dans les volumes élémentaires séparés par une distance r. On note la 2PCF par ξ, et elle est définie par l'équation (1.8). Dans un langage mathématique, nous supposons que la distribution des galaxies représente un processus aléatoire stationnaire.

Des points répartis de manière Poissonienne ou d'une manière uniformément aléatoire ont une 2PCF nulle. ξ est interprétée comme l'excès de probabilité de trouver des galaxies à une séparation donnée.

A partir de ces considérations, Landy and Szalay (1993); Hamilton (1993); Vargas-Magaña et al. (2012) donnent des estimateurs différents de 2PCF basé sur le nombre de paires de galaxies observées. Alternativement, on peut estimer directement du spectre de puissance des galaxies à partir d'un catalogue de positions de galaxies avec l'algorithme 'FKP qui consiste à faire une transformation de Fourier du champ galaxie (Feldman et al., 1994; Hamilton, 2009b). Un troisième ensemble de techniques directement inspirées de l'analyse du CMB sera très efficace avec les futures sondages gigantesques (Vogeley and Szalay, 1996; Tegmark et al., 1997; Hamilton, 2009a; Tegmark, 1997; Padmanabhan et al., 2001).

6.2.4 Historique des mesures du BAO

Le pic du BAO a été observé pour la première fois au début de 2005 avec l'échantillon SDSS-II constitué de galaxies rouges lumineuses (LRG) (Eisenstein et al., 2005), et avec les données 2dFGRS (Cole et al., 2005). D'autres études sur l'échantillon SDSS-II LRG ont été effectuées et ont confirmé la première mesure (Padmanabhan et al., 2007; Sánchez et al., 2009; Kazin et al., 2010; Percival et al., 2010). Le succès de ces expériences a permis le début d'une autre génération d'expériences telles 6dfGRS, WiggleZ et SDSS-III/BOSS. Anderson et al. (2012) résume les contraintes BAO obtenus avec les galaxies à $z < 1$. Cette publication intègre les résultats des expériences 6dFGRS (Beutler et al., 2011), SDSS-II (Percival et al., 2010; Padmanabhan et al., 2012; Xu et al., 2012), SDSS-III/BOSS (Anderson et al., 2012), et WiggleZ (Blake et al., 2011c); figure 1.4.

La taille comobile du BAO à un décalage vers le rouge z dans la direction de la ligne de visée, $r_{\parallel}$, (direction transversale, $r_{\perp}$) est liée à la taille de décalage vers le rouge observé Δz (taille angulaire $\Delta\theta$) par le paramètre Hubble $H(z)$ (le diamètre angulaire $d_A(z)$), équation 1.9. En utilisant le monopole de la 2PCF (sa moyenne angulaire), nous obtenons une mesure d'une distance, notée d_V, équation (1.11), à comparer avec r_s (équation 1.1). Les modèles de fonction de corrélation sont ensuite ajustés pour déterminer le paramètre α défini dans l'équation 1.12. Cette procédure d'ajustement a été largement décrit et testé, en particulier, d'être impartial, même en présence d'hypothèses erronées au sujet de la cosmologie et des galaxies (Padmanabhan and White, 2009; Xu et al., 2012; Padmanabhan et al., 2012). La mesure actuelle de $d_V/r_s\ (z)$ comparée à différents modèles cosmologiques est montrée sur la figure 1.4. Les mesures sont en accord avec le modèle WMAP ΛCDM. La précision est comprise entre 2 et 5%.

Lors de l'utilisation du monopole de la 2PCF, nous supposons que l'augmentation de son l'amplitude le long de la ligne de visée en raison de distorsions spatiales par le décalage vers le rouge est peu importante. Cette hypothèse marche pour obtenir l'échelle acoustique à une précision de $\sim 2\%$. Dans le régime de précision des prochaines expériences, cette hypothèse ne tient plus, et des modèles anisotropes de 2PCF sont nécessaires, voir (Reid et al., 2012; Samushia et al., 2013; Sanchez et al., 2013).

Récemment, une nouvelle technique pour mesurer le BAO à partir des forêts Ly_α des quasars a été utilisée. Deux auteurs ont proposé d'utiliser la forêt Ly_α pour mesurer la distance angulaire et le paramètre de Hubble à $2 < z < 4$ en utilisant la règle standard fournie par le BAO (McDonald, 2003; McDonald and Eisenstein, 2007). Les premiers résultats sont extraordinaires, car ils montrent clairement la phase de décélération de l'univers à $z > 0.8$ et cela réconforte le modèle ΛCDM, voir figure 1.5. La précision de la mesure actuelle sur $H(z = 2, 3)$ est d'environ $\sim 3.5\%$, voir Pâris et al. (2012); Busca et al. (2012); Lee et al. (2013); Kirkby et al. (2013); Slosar et al. (2013).

Erreur sur la mesure de l'échelle acoustique

La mesure de l'échelle BAO à un décalage vers le rouge donné est soumis à des erreurs systématiques de l'ordre de quelques pour cent. On peut les diviser en deux catégories. Tout d'abord, la systématique en raison de la géométrie du sondage et de son design *i.e.* l'évaluation de l'exhaustivité en fonction de la position sur le ciel et du décalage vers le rouge des galaxies observées. La deuxième catégorie de systématique, est due au processus astrophysique attaché aux galaxies étudiées, se produisant à petite échelle. Actuellement, les deux sources d'erreurs systématiques qui sont sous contrôle et à un niveau inférieur à celui des erreurs statistiques.

6.3 Défis actuels pour la cosmologie

Aujourd'hui, le principal défi pour la cosmologie est d'expliquer la nature de l'énergie noire et de la matière noire. Le rapport de la Dark Energy Task Force (DETF; Albrecht et al. 2006) a souligné qu'à ce jour, il n'y a pas d'explications théoriques convaincantes de l'énergie noire. Pour cette raison, la communauté doit focaliser ses efforts sur l'observation pour comprendre de ce que l'Univers est fait.

Dans cette thèse, je discute de la façon de construire, d'un point de vue observationnel, les expériences BAO futures, afin d'augmenter la précision d'observation du spectre de puissance des galaxies et de la mesure du BAO.

6.4 Etude des sondages avec des galaxies à raies d'émission avec l'infrastructure du Sloan Digital Sky Survey

Les sondages en cours portant sur un volume comobile suffisant pour mesurer l'échelle BAO sont limités à un décalage vers le rouge $z \lesssim 0.7$. J'ai étudié dans le chapitre 2 plusieurs schémas de sélection de galaxies visant à l'obtention de galaxies à raie d'émission (Emission Line Galaxy, ELG) dans la gamme de décalage vers le rouge de $0.6 < z < 1.7$ qui est appropriée pour les futures études BAO utilisant les spectrographes présents sur le télescope SDSS. J'ai exploré deux sélections en couleurs différentes en utilisant les sondages photométriques SDSS et CFHT-LS dans les bandes u, g, r, et i pour évaluer leur performance dans la sélection d'ELG lumineuses.

A partir de l'observation de 2000 ELGs, nous avons identifié un schéma de sélection qui a une efficacité de 75 pour cent de mesure du décalage vers le rouge. Ces résultats sont détaillés dans le chapitre 2 de la thèse. Ce résultat confirme la faisabilité des sondages massifs d'ELG en utilisant le spectrographe du télescope SDSS-III/BOSS pour une détection du BAO au redshift $z \sim 1$. En particulier, cela prouve la faisabilité du projet SDSS-IV/eBOSS, expérience, qui prévoit d'utiliser le télescope SDSS pour combiner l'utilisation du BAO avec les distorsions spatiales dans le regroupement des galaxies à raie d'émission pour mieux comprendre l'énergie noire.

6.5 Mesure du doublet [OII] des galaxies à raies d'émission avec les futurs sondages spectroscopiques basés au sol

La prochaine génération de sondages spectroscopiques à champ large permettra de cartographier la distribution des galaxies à grande échelle dans la gamme des décalages vers le rouge $0.7 \leq z \leq 2$ pour mesurer les BAO. La signature optique primaire utilisée dans cette gamme de décalage vers le rouge vient de l'émission du doublet [OII] , qui fournit une identification unique, ce qui minimise la confusion avec d'autres raies d'émission. Pour calculer la résolution du spectrographe requis pour ces sondages, j'ai simulé l'observations du doublet [OII] et mesuré les valeurs de $\Delta\chi^2$ pour différentes résolutions instrumentales, et différentes vitesses de dispersion dans les galaxies. Ces résultats sont détaillés dans le chapitre 3 de la thèse.

Nous préconisons deux stratégies sur le choix de la résolution des spectrographes pour les futurs sondages BAO. Pour un sondage observant des émetteurs [OII] très lumineux (comme SDSS-IV/eBOSS), une résolution de R $\sim$ 3500 est suffisante pour obtenir un bon échantillon de doublets pour que le pipeline puisse récupérer tous les décalages vers le rouge [OII] . L'impact de leur profil sur l'exhaustivité de décalage vers le rouge est inférieure à 6%. Pour les sondages à émetteur [OII] faible (comme BigBoss), nous recommandons de pousser pour une résolution la plus élevée possible.

6.6 Biais stochastique des traceurs BAO séléctionnés en couleur par calcul d'aggrégation et de lentille faible.

Nous avons étudié plusieurs schémas de sélection de galaxies, visant à la construction d'un échantillon d'ELG dans la gamme de décalage vers le rouge de $0.6 < z < 1.7$, qui serait approprié pour les futures études BAO. Le mieux étant de fournir un échantillon de galaxie très biaisé. Nous analysons le regroupement angulaire des galaxies à décalage vers le rouge de 0.5, 0.7, 0.8, 1 et 1.2 et nous combinons cette analyse avec un modèle d'occupation de halo de matière noire (HOD) pour déterminer les propriétés des halos que ces galaxies peuplent, en particulier leur biais aux grandes échelles. Nous effectuons également une analyse de lentillage gravitationnel faible (via la statistiques d'ouverture) pour extraire le biais des galaxies ainsi que leur coefficient de corrélation avec la matière noire.

Nous appliquons cette analyse sur un ensemble de données composé de la photométrie de SDSS Stripe 82 (225 deg^2) et du sondage en bande i CFHT/Stripe82, et de l'infrarouge tiré de WISE (bande photométrique W1). Ces résultats sont détaillés dans le chapitre 4 de la thèse.

L'analyse de la sélection SDSS-III/CMASS à $z = 0.5$ est en accord avec les études précédentes sur le traceur, d'ailleurs nous mesurons son coefficient de corrélation $r = 1.16 \pm 0.35$, pour la première fois. Pour les décalage vers le rouge supérieurs, nous confirmons les tendances que les populations de galaxies les plus brillantes sont fortement biaisées ($b > 1.5$). Cependant, nous sommes limités par la profondeur des données pour obtenir des valeurs précises. Une future étude des BAO utilisant ces traceurs du champ de masse aura une détection efficace du BAO.

6.7 Conclusion

Si nous faisons confiance à la relativité générale et au principe cosmologique, alors il y a une preuve claire que l'Univers est Euclidien et qu'il subit une expansion accélérée depuis la moitié de sa vie après Big Bang. Les sondes cosmologiques basées sur des mesures géométriques à différentes époques cosmiques sont censées expliquer le mécanisme responsable de cette accélération - nommé l'énergie noire.

Le pic dans la fonctions de corrélation à deux points des galaxies dû aux oscillations acoustiques des baryons dans l'Univers primordial fournit une règle standard pour sonder l'expansion accélérée de l'Univers et la mesure de la relation distance – redshift.

Comme indiqué dans le premier chapitre, sa mesure par les projets d'observation galaxies actuels et futurs, couvrant des milliers de degrés carrés, va devenir de plus en plus précise ($< 0.1\%$). Les principales sources d'erreurs systématiques sur la mesure du BAO proviennent de l'incertitude sur la distribution en redshift des galaxies utilisées et de l'homogénéité spatiale de la photométrie qui permet de sélectionner les galaxies dans la tranche de redshift voulue.

Dans le deuxième chapitre, j'ai étudié plusieurs schémas de sélection de galaxies visant à la construction d'un échantillon dans la gamme redshift $0.6 < z < 1.7$, qui seraient appropriés pour les futures études BAO utilisant le spectrographe SDSS-III/BOSS. J'ai exploré deux sélections en couleurs différentes en utilisant à la fois la photométrie du SDSS et du CFHT-LS dans les bandes u, g, r et i et évalué leurs performances à sélectioner des ELG lumineuses pour confirmation spectroscopique vais le télescope SDSS de 2.5m. A partir de 2,000 ELGs observées, j'ai identifié deux schémas de sélection, nommé *ugr* et *gri*, qui ont une efficacité de mesure redshift supérieure à 70 pour cent.

L'efficacité de la sélection *gri* peut être augmentée de 5 pour cent en améliorant la détermination automatique du redshift spectroscopique, pour une efficacité finale de 73 à 78 pour cent de galaxies à $z > 0.6$.

La sélection *ugr* peut être améliorée pour atteindre un rendement de 70 % en abaissant la coupure en couleur $u - g$ de 0.3 afin de réduire la contamination par les galaxies à faible redshift. D'autres galaxies à faible redshift peuvent être retirées de la sélection par une inspection des images: quelques-unes des galaxies à faible redshift sont en effet assez étendues et facile à retirer. Une inspection visuelle permet de réduire la part faible décalage vers le rouge de 9 pour cent à 4 pour cent. En outre, la sélection compact - étendu sur les données CFHT est très efficace pour identifier les quasars. Il est également possible d'améliorer la détermination du redshift spectroscopique et donc reclasser les 12 % de galaxies à 'raie d'émission unique' (10 pour cent sont à $z > 0.6$).

Sur la base de ces sélections en couleur, j'ai mis en place une sélections pour le BAO en utilisant les quatre bandes *ugri* obtenue par la combinaison des sélections *gri* et *ugr*. Les observations actuelles, effectuées par un autre programme connexe dont je suis le instigateur principal, montrent que cette nouvelle sélection est très efficace pour sélectionner les galaxies avec de très fortes raies d'émission. Ces résultats soutiennent la faisabilité de relevés d'ELG utilisant le spectrographe de BOSS sur le télescope SDSS pour une détection du BAO au redshift z $\sim$ 1. A l'avenir, l'expérience SDSS-IV/eBOSS, conçue pour combiner l'utilisation de la règle BAO avec des distorsions de l'espace de redshift, utilisera le schéma de sélection *ugri* pour sélectionner les galaxies raies d'émission dans le redshift de $0.6 < z < 1.2$.

L'étude du chapitre 3 détermine la résolution instrumentale nécéssaire pour obtenir des redshift avec la raie d'émission [OII] assez précis pour un sondage BAO. L'émission du doublet [OII] est une signature optique qui fournit une identification de décalage spectral unique qui peut réduire la confusion avec d'autres raies d'émission. Est-ce que l'utilisation de [OII] peut introduire un biais dans la distribution de redshift observée ? D'après les simulations effectuées dans mon travail de thèse deux stratégies sur le choix de la résolution des spectrographes des futures sondages BAO ont émergées:

- pour les enquêtes basées sur des émetteurs [OII] lumineux (comme SDSS-IV/eBOSS), une résolution de $R \sim 2500$ (comme le spectrographe BOSS) est suffisante pour obtenir un bon échantillon de doublets (60 %) afin de récupérer tous les décalages vers le rouge. L'augmentation de la résolution à ~3500 permet d'obtenir 90 % des doublets. Pour une petite augmentation de la résolution, l'efficacité de détermination redshift gagne 30 %. L'impact des raies du ciel sur l'exhaustivité de la distribution en redshift est inférieure à 6%;

- pour les enquêtes basées sur des émetteurs [OII] faibles (comme MS-DESI), cette étude recommande que la résolution soit la plus haute possible. Sachant qu'il y a un nombre limité de pixels sur le capteur (environ 4000), et que la résolution la plus élevée possible sur un spectrographe à trois canaux est de $R \sim 4500$ à 7500 Angstrom, pour aller au-delà, un spectrographe à quatre canaux devient obligatoire. Avec une résolution de $R \sim 4500$, on peut obtenir 25 % des doublets, ce qui est suffisant.

En outre, ces simulations ont montré que l'impact des raies du ciel sur la compluétude de la distribution en redshift est indépendant de la résolution et ne dépend que de la soustraction du ciel.

Des développements préliminaires sur les pipelines de redshift montrent que cette étude est conservatrice et constitue une limite inférieure au % des doublets obtenus. La prochaine étape pour les futures sondages est de développer des outils complets de simulation, avec un accent particulier sur la faç de trouver le bon redshift. Dans ce but, je suis impliqué dans une collaboration avec des scientifiques du Laboratoire national Lawrence Berkeley et de l'Université de l'Utah (USA). Cet effort est clairement identifiée dans la collaboration SDSS.

L'observation des BAO basée sur des LRG fortement biaisé ou sur la forêt Lyα a montré la relation entre le traceur de matière lumineuse et la matière noirs joue un rôle clé dans la mesure du BAO.

L'utilisation des ELG pour suivre le champ de la matière est un nouveau défi: comment sont les ELG liées à la matière noire sous-jacente? Est-ce qu'un modèle de biais linéaire est suffisant? Quelle est la stochasticité de cette relation?

Le quatrième chapitre est le premier effort fait dans ce sens et essaie de rṕondre à ces questions. J'ai analysé le regroupement des galaxies issues des schémas de sélection de galaxies sus-mentionnés. Les ELG les plus brillantes sont fortement biaisées ($b > 1.5$), ces galaxies sont donc un bon traceur pour le BAO. L'application de cette méthode avec les ELG moins brillantes a montré que les connaissances actuelles sur la répartition en redshift des traceurs et que la profondeur actuelle des catalogues de forme sont trop limitées pour effectuer une mesure robuste. Ce travail m'a aidé à établir une collaboration étroite avec les communautés de cosmologie allemandes, brésiliennes et canadiennes lors de la réduction des données CFHT/Stripe82.

Dans les vingt prochaines années, on s'attend à une énorme augmentation du nombre de données prises pour la cosmologie, l'astronomie et la physique des particules. Tout cela dans le but de fournir un aperçu en profondeur sur la physique fondamentale. On s'attend à une précision suffisante pour tester si l'accélération actuelle de la géométrie est due à une composante de l'énergie noire qui serait un champ scalaire dynamique (Peebles and Ratra (1988); voir Copeland et al. (2006) pour une revue), ou à trouver une nouvelle théorie des interactions gravitationnelles au-delà de la relativité générale (Clifton et al., 2012), ou à mettre en évidence des effets dynamiques non-linéaires dus aux inhomogénéités intrinsèques de la matière connexes au problème de la moyenne (Buchert and Räsänen, 2012). Par ailleurs, la mesure cosmologique très précise de BAO, combinée avec d'autres sondes, pourra très probablement fournir une compréhension considérable du secteur de la matière noire, et en particulier des neutrinos massifs.

Chapter 7

Acknowledgements, Remerciements

A tous ceux qui m'ont aidés lors, de ma thèse, je vous remercie en ordre analphabétique par ce petit jeu (la relation est seulement injective) :

Jeannine Comparat	○	○	mais si ca va marcher
Aurelie Penin	○	○	partage d'idee
Emmanuel Hugot	○	○	collaboration
L'equipe administrative du LAM	○	○	recommandations
Jean-Paul Kneib	○	○	relecture
Olivier Ilbert	○	○	gentillesse
Laurence Tresse	○	○	ELG is good for you
Thomas Erben	○	○	barre de traction
Marie Laslandes	○	○	des
Maria Pereira	○	○	relecture une dexieme fois
Delphine Comparat	○	○	volume maximal
Salvador Dali	○	○	soutien quotidien
Daniel Thomas	○	○	galaxy bias
Carlos Brandt	○	○	soutien administratif
Thibault Moutard	○	○	ecoute
Carlo Schimd	○	○	films au bureau
Frederic Chopin	○	○	validation
Fransicso Prada	○	○	bigboss
Bruno Moraes	○	○	cmc
Romain Thomas	○	○	ping pong
Pol Molitor	○	○	anglais
Sebastien Vives	○	○	bibliographie
Thierry Fusco	○	○	blague a deux francs
Jacques Comparat	○	○	lettres
Nicholas Ross	○	○	grimpeur du vendredi
Violeta Gonzalez-Perez	○	○	pipeline automatique qui marche
Anne Ealet	○	○	cla brou
Pin-Wei Wang	○	○	chercher plus vite
France Comparat	○	○	grande mere infatiguable
Joel Brownstein	○	○	nocturne
Pierre Verry	○	○	chemise a carreau
Benjamin Weaver	○	○	boss
Jean-Gabriel Cuby	○	○	partage du mate
Damien Maurin	○	○	sequels
Gabriel Besson	○	○	energie noire
Julien Zoubian	○	○	repassage
Jeremy Tinker	○	○	soleil

Zalpha Challita	○	○	soutien moral
David Schlegel	○	○	explications poussees
Marie Comparat	○	○	strong lensing
Eric Aubourg	○	○	grimpeur du dimanche
Daniel Eisenstein	○	○	manger du bon fromage
Eric Jullo	○	○	peintures
Sebastian Nuza	○	○	simulations
Roland Bacon	○	○	cookies et gateaux a la pause
Marseille	○	○	demande de vulgarisation
python	○	○	logiciel genial
Oliver Steele	○	○	soutien affectif
Donald Schneider	○	○	alpiniste occasionel
Kyle Dawson	○	○	completude
Mathilde Jauzac	○	○	interet pour mes recherches sans faille
Adam Bolton	○	○	bri brou
Sebastien Fontanel	○	○	avis orthogonal
Franz Liszt	○	○	visite de musee a NYC
Christophe Yeche	○	○	miam le bon melon
Alexie Leauthaud	○	○	pousser plus loin
Jose Garcia	○	○	conjonction
topcat	○	○	grimpeur enerve
Nick Mostek	○	○	sonate
Marc Peyre	○	○	mode d'emploi du cmc
Marie-Helene Peyre	○	○	commentaire constructifs
Solene Comparat	○	○	motivation sans faille
Johan Richard	○	○	half-dome
Claudia Maraston	○	○	references
Fabrice Lamareille	○	○	critique acerbe
Raphael Carroy	○	○	eboss
HuanYuan Shan	○	○	bucheron gay
Sylvain Casini	○	○	weak lensing
Jian Ge	○	○	les sous
Will Percival	○	○	grande clope
Olivier Le Fevre	○	○	oh des jolis poissons
Stephanie Escoffier	○	○	affichage de poster a l'ED
Marc Ferrari	○	○	langage facile
Stephen Bailey	○	○	jazz
Jeffrey Newman	○	○	vin et fromage
Christine Da Sylva	○	○	petite pause
Cameron McBride	○	○	opera

Merci mille fois à vous tous.