

# Clustering properties of $g$ -selected galaxies at $z \sim 0.8$

Ginevra Favole<sup>1,2\*</sup>, Johan Comparat<sup>1,2,5†</sup>, Francisco Prada<sup>3,1,2,4</sup>, Gustavo Yepes<sup>5</sup>, Eric Jullo<sup>6</sup>, Anna Niemiec<sup>6</sup>, Jean-Paul Kneib<sup>7,6</sup>, Sergio A. Rodríguez-Torres<sup>1,2,5‡</sup>, Anatoly Klypin<sup>8</sup>, Ramin A. Skibba<sup>9</sup>, Cameron K. McBride<sup>10</sup>, Daniel J. Eisenstein<sup>10</sup>, David J. Schlegel<sup>3</sup>, Sebastián E. Nuza<sup>11</sup>, Chia-Hsun Chuang<sup>1,2§</sup>, Timothée Delubac<sup>7</sup>, Christophe Yèche<sup>12</sup>, Donald P. Schneider<sup>13,14</sup>

<sup>1</sup>*Instituto de Física Teórica (IFT) UAM/CSIC, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*

<sup>2</sup>*Campus of International Excellence UAM/CSIC, Cantoblanco, E-28049 Madrid, Spain*

<sup>3</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA, 94720, USA*

<sup>4</sup>*Instituto de Astrofísica de Andalucía (CSIC), Granada, E-18008, Spain*

<sup>5</sup>*Departamento de Física Teórica M-8, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*

<sup>6</sup>*Laboratoire d'Astrophysique de Marseille - LAM, Université d'Aix-Marseille & CNRS, UMR7326, F-13388 Marseille, France*

<sup>7</sup>*Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland*

<sup>8</sup>*Astronomy Department, New Mexico State University, MSC 4500, PO Box 30001, Las Cruces, NM 88003-8001, USA*

<sup>9</sup>*Center for Astrophysics and Space Sciences, University of California, 9500 Gilman Drive, San Diego, CA 92093, USA*

<sup>10</sup>*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

<sup>11</sup>*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482, Potsdam, Germany*

<sup>12</sup>*CEA, Centre de Saclay, IRFU, 91191, Gif-sur-Yvette, France*

<sup>13</sup>*Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802*

<sup>14</sup>*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802*

## ABSTRACT

Current and future large redshift surveys, as the Sloan Digital Sky Survey IV extended Baryon Oscillation Spectroscopic Survey (SDSS-IV/eBOSS) or the Dark Energy Spectroscopic Instrument (DESI), will use Emission-Line Galaxies (ELG) to probe cosmological models by mapping the large-scale structure of the Universe in the redshift range  $0.6 < z < 1.7$ . With current data, we explore the halo-galaxy connection by measuring three clustering properties of  $g$ -selected ELGs as matter tracers in the redshift range  $0.6 < z < 1$ : (i) the redshift-space two-point correlation function using spectroscopic redshifts from the BOSS ELG sample and VIPERS; (ii) the angular two-point correlation function on the footprint of the CFHT-LS; (iii) the galaxy-galaxy lensing signal around the ELGs using the CFHTLenS. We interpret these observations by mapping them onto the latest high-resolution MultiDark Planck N-body simulation, using a novel (Sub)Halo-Abundance Matching technique that accounts for the ELG incompleteness. ELGs at  $z \sim 0.8$  live in halos of  $(1 \pm 0.5) \times 10^{12} h^{-1} M_{\odot}$  and  $22.5 \pm 2.5\%$  of them are satellites belonging to a larger halo. The halo occupation distribution of ELGs indicates that we are sampling the galaxies in which stars form in the most efficient way, according to their stellar-to-halo mass ratio.

**Key words:** galaxies: distances and redshifts — galaxies: haloes — galaxies: statistics — cosmology: observations — cosmology: theory — large-scale structure of Universe

## 1 INTRODUCTION

By investigating the properties of galaxy clustering within the cosmic web, it is possible to constrain cosmology and infer the growth of structure and the expansion history of

the Universe (Weinberg et al. 2013). In fact, galaxy clustering measurements using last-generation large-volume redshift surveys, as the Sloan Digital Sky Survey (SDSS; York et al. 2000; Gunn et al. 2006; Smee et al. 2013) and the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Eisenstein et al. 2011; Dawson et al. 2013) provide robust information about both the evolution of galaxies and the cosmological framework in which these complex structures live. In order to interpret such measurements, we need to understand the relation between the theory-predicted dark

\* E-mail: g.favole@csic.es

† Severo Ochoa IFT Fellow

‡ Campus de Excelencia Internacional UAM/CSIC Scholar

§ MultiDark Fellow

matter field and its luminous counterpart i.e., the discrete galaxy map (Cooray & Sheth 2002).

Luminous low-redshift galaxies have already been connected to their dark matter halos in a precise manner, through weak lensing and clustering analysis as a function of galaxy luminosity and stellar mass. Baldry et al. (2004), Zehavi et al. (2011) and Guo et al. (2015) measured the clustering properties of the SDSS “blue cloud” and “red sequence” in the local Universe (SDSS median redshift  $z \sim 0.1$ ; Abazajian et al. 2009), as a function of magnitude and color. Their results show that at a given luminosity, the blue sample has a lower clustering amplitude and a smaller correlation length compared to the red one.

Guo et al. (2014) investigated the clustering luminosity and colour dependence of BOSS CMASS DR10 (Anderson et al. 2014), and found that more luminous galaxies are more clustered and hosted by more massive halos. For luminous red galaxies (LRGs), these masses are  $\sim 10^{13} - 10^{14} h^{-1} M_{\odot}$ , at fixed luminosity, progressively redder galaxies are more strongly clustered on small scales, which can be explained by having a larger fraction of these galaxies in the form of satellites in massive haloes. Favole et al. (2015) measured galaxy clustering in the BOSS CMASS DR11 (Anderson et al. 2014) sample at  $z > 0.55$  as a function of color, and proposed a new statistic to extract robust information about small-scale redshift-space distortions and large-scale galaxy bias. Consistent with many previous results (e.g., Wang et al. 2007; Zehavi et al. 2005; Swanson et al. 2008), they found that, compared to the blue population, red galaxies reside in more massive halos, show a higher clustering amplitude, large-scale bias and peculiar velocities.

This type of clustering analysis has recently been extended to higher redshifts thanks to the VIMOS Public Extragalactic Survey (VIPERS; Guzzo et al. 2014; Garilli et al. 2014) and DEEP2 survey (Newman et al. 2013). Compared to DEEP2, VIPERS has a much larger volume but has a lower redshift limit however, the signal-to-noise ratio in its spectroscopic measurements is higher. Using VIPERS data, Marulli et al. (2013) measured the clustering properties of galaxies at redshift  $z = 0.8$  as a function of their luminosity and stellar mass, and found that the clustering amplitude and the correlation length increase with these two quantities; see also the PRISM Multi-object Survey (PRIMUS) results by Skibba et al. (2015) and Bray et al. (2015). Mostek et al. (2013) measured the clustering of the red sequence and the blue cloud at  $z = 0.9$ , as a function of their stellar mass and star formation history, using DEEP2 data. They argued that blue galaxies are more clustered in the local Universe than at  $z = 0.9$ , and red galaxies are much more clustered locally than at high redshift. They also suggested that the clustering trend observed with star formation rate (SFR) can be explained mostly by the correlation between stellar mass and clustering amplitude for blue galaxies. Coil et al. (2008) studied the DEEP2 clustering dependence on color and luminosity, and found that the dependence on color is much stronger than with luminosity, and is as strong with color at  $z \sim 1$  as locally. They claimed no dependence of the clustering amplitude on color for galaxies in the red sequence, but a significant dependence for galaxies within the blue cloud. Cooper et al. (2008) investigated the connection between star formation (SF) and environment in DEEP2 data at  $z \sim 0.1$ , and  $z \sim 1$ . Their results indicate that, locally,

galaxies in regions of higher overdensity have lower star formation rates (SFRs), and their stars form more slowly than in their counterparts in lower density regions. At  $z \sim 1$ , this SFR-overdensity relation is inverted; this is in part due to a population of bright, blue galaxies in dense environments, which lacks a counterpart in the local Universe, and is thought to evolve into members of the red sequence from redshift 1 to 0.

The combination of clustering with weak galaxy-galaxy lensing (see e.g., Bartelmann 1999) allows one to gain insight on the large-scale structure formation, and directly probe the stellar-to-halo mass relation (SHMR; Leauthaud et al. 2011). The galaxy-halo connection has been measured at  $z < 1$  by Leauthaud et al. (2012), Shan et al. (2015), and Coupon et al. (2015), using three different weak lensing surveys (COSMOS: Scoville et al. (2007); CFHT-Stripe82 and CFHTLenS<sup>1</sup>: Heymans et al. (2012a); Erben et al. (2013a)); all obtained consistent results. Leauthaud et al. (2012) performed the first joint analysis of galaxy-galaxy weak lensing, galaxy clustering, and galaxy number densities using COSMOS data, and provided robust constraints on the shape and redshift evolution of the SHAM relation in the redshift range  $0.2 < z < 1$ . At low stellar mass, the halo mass scales proportionally to  $M_{\star}^{0.46}$ ; this scaling does not evolve significantly with redshift. At  $M_{\star} > 5 \times 10^{10} M_{\odot}$ , the SHMR rises sharply, causing the stellar mass of a central galaxy to become a poor tracer of its parent halo mass. Combining observations in the CFHT-LenS/VIPERS field from the near-UV to the near-IR, Coupon et al. (2015) found that the SHMR for the central galaxies peaks at  $M_{h,peak} = (1.9_{-0.1}^{+0.2} \times 10^{12} M_{\odot})$ , and its amplitude decreases as the halo mass increases. Hearin et al. (2014) presented new measurements of the galaxy two-point correlation function and the galaxy-galaxy lensing signal from SDSS, as a function of color and stellar mass, and demonstrated that the age-matching model (Hearin & Watson 2013), which states that older halos tend to host galaxies with older stellar populations, exhibits remarkable agreement with these and other statistics of low-redshift galaxies.

Current (Sub)Halo-Abundance Matching (SHAM; Conroy et al. 2006; Trujillo-Gomez et al. 2011; Klypin et al. 2013; Nuza et al. 2013) and Halo Occupation Distribution (HOD; Berlind & Weinberg 2002; Kravtsov et al. 2004; Zheng et al. 2005, 2007) models correctly reproduce the clustering measurement mentioned above. SHAM maps observed galaxies onto dark matter halos directly from N-body cosmological simulations, according to a precise monotonic correspondence between halo and galaxy number densities. The HOD method is an analytical prescription to populate simulated halos with galaxies, in which the assignment is performed by interpolating the halo occupation distribution at the values of the desired halo masses. In this sense, the SHAM approach returns a model which is built directly on the considered simulation box.

Next generation high-redshift surveys as SDSS-IV/eBOSS (Dawson et al. 2015, in prep.), Subaru Prime Focus Spectrograph (PFS; Sugai et al. 2012; Smee et al. 2014), DESI (Schlegel et al. 2015), 4MOST<sup>2</sup> and Euclid<sup>3</sup>

<sup>1</sup> <http://www.cfht.hawaii.edu/Science/CFHTLS/>

<sup>2</sup> <https://www.4most.eu/cms/>

<sup>3</sup> <http://sci.esa.int/euclid/>

(Laureijs et al. 2011; Sartoris et al. 2015) will use emission-line galaxies (ELGs) as BAO tracers to explore the Universe large-scale structure out to  $z \sim 2$ . Observing ELGs, learning how to model their clustering properties and understanding how they populate their host halos are therefore crucial points that we need to understand in order to select the targets for future experiments. From the observational point of view, the recent increment of available ELG spectroscopic data (Guzzo et al. 2014; Comparat et al. 2015) allows one to measure their clustering properties over about  $12 \text{ deg}^2$  at  $z = 0.8$  (corresponding to a comoving volume of  $V \sim 10.6 \times 10^6 h^{-3} \text{ Mpc}^3$  in the Planck cosmology; see Section 3 for details), which represents a dramatic improvement.

Comparat et al. (2013) demonstrated that neither a standard HOD nor a traditional SHAM technique are able to reproduce the angular clustering of ELGs on small scales. In fact, both techniques are based on the assumption that the galaxy sample to model is complete, but this is not the case of the ELGs, which are highly incomplete in stellar mass. One could instead use semi-analytic models of galaxy formation and hydrodynamic simulations, but they lack of mass resolutions to model emission line galaxies.

The aim of this work is to provide a modified version of the standard SHAM prescription, directly based on the latest MultiDark N-body simulation with Planck cosmology, that accounts for the ELG incompleteness and returns suitable mock galaxy catalogs able to accurately predict the ELG angular and redshift-space clustering, respectively, on small and larger scales. These mock catalogs are released to the public.

The paper is organized as follows. Section 2 describes the data sets and the MultiDark simulation box used in our analysis. In Section 3 we present our ELG clustering and weak lensing measurements. In Section 4 we explain how we model the ELG clustering and we present our main results. Section 5 discusses the implications of our ELG clustering analysis in a galaxy evolution perspective, and Section 6 summarizes our main results.

Throughout the paper, we assume the Planck cosmology (Planck Collaboration et al. 2014) and magnitudes in the AB system (Oke & Gunn 1983).

## 2 DATA AND SIMULATION

### 2.1 Data sets

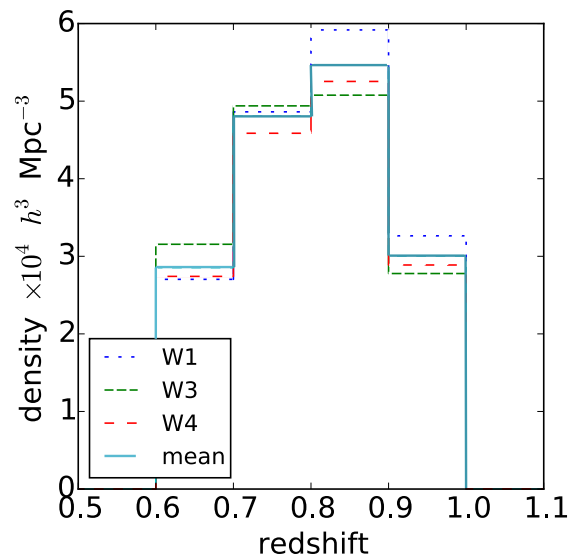
We build our ELG galaxy sample using the Canada-France-Hawaii Telescope Legacy Survey (CFHT-LS) Wide T0007<sup>4</sup> photometric redshift catalog (Ilbert et al. 2006; Coupon et al. 2009). We apply a  $g$ -band magnitude cut,  $20 < g < 22.8$  (Fukugita et al. 1996), to select galaxies with bright emission lines and low dust at  $z < 1$ . We also apply a color selection,  $-0.5 < (u - r) < 0.7(g - i) + 0.1$ , to remove the low-redshift galaxies. For details on the selection function, see Comparat et al. (2015). Then, to obtain the largest possible area, we convert the  $i$ -selection into the new Megacam

$i$ -band filter<sup>5</sup>. For the W1, W3 and W4<sup>6</sup> fields, we derive an average density of about 500 ELGs per  $\text{deg}^2$ , 70% of which have a photometric redshift in the range  $0.6 < z < 1$ . The densities of each field are reported in Table 1, and the errors on the photometric redshift are  $\sigma_z < 0.05(1+z)$  for  $i < 22.5$  and  $z < 1$ . The  $ugri$  ELG selection is brighter than  $i < 22.5$ .

We match the photometric targets to the available spectroscopic surveys – BOSS DR12, DEEP2, VIPERS (Bolton et al. 2012; Alam et al. 2015; Newman et al. 2013; Guzzo et al. 2014) – within  $1''$  radius; see Table 2. Based on KS-tests, the VIPERS, BOSS and DEEP2 spectroscopic selections constitute fair sub-samples of the complete selection: the hypothesis that they are drawn from the same distribution cannot be rejected at the 90% confidence level. For these samples, we create random catalogs with the same redshift distribution of the data and 30 times denser. Figure 2 displays the ELG spectroscopic redshift distribution per unit volume for the three Wide fields (dashed, dotted, dot-dashed histograms), and their mean (solid line). Two thirds of the galaxy density is located in the redshift range  $0.7 < z < 0.9$ , while both the intervals  $0.6 < z < 0.7$  and  $0.9 < z < 1$  contain one sixth of the sample. According to the ELG selection function in Comparat et al. (2015), we select only galaxies at  $z > 0.6$  since we are not interested in low-redshift objects. We have investigated further the impact of the higher redshift cut,  $z < 1$ , on the angular clustering by imposing to the ELG sample different redshift thresholds:  $z < 1, 1.2, 1.4, 1.6$ . In all these samples the lower redshift cut is fixed at  $z > 0.6$  and we have imposed the  $i < 22.5$  magnitude cut to eliminate bad photometric redshifts. We find that including also ELGs at  $z \geq 1$ , we are slightly enhancing the galaxy num-

<sup>5</sup> <http://www4.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/en/megapipe/docs/filt.html>

<sup>6</sup> <http://www.cfht.hawaii.edu/Science/CFHTLS/T0007/T0007-docsu10.html>

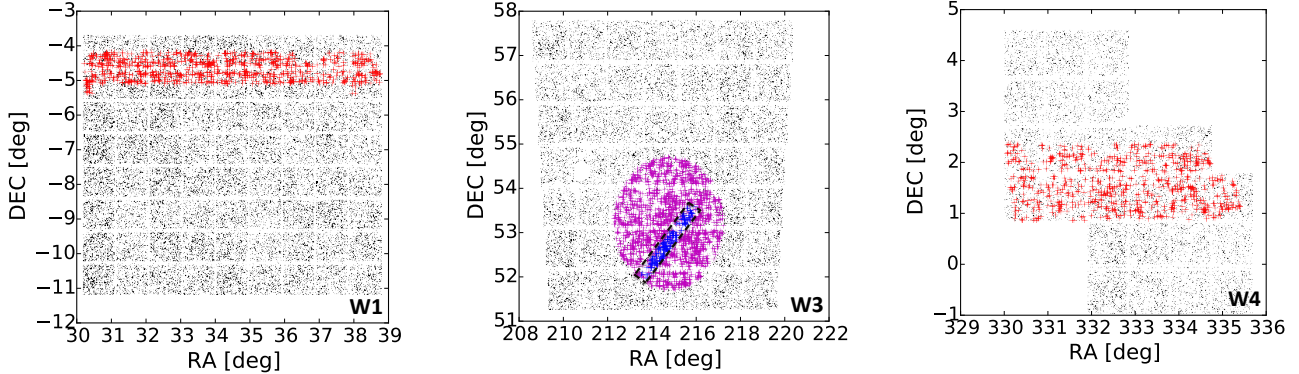


**Figure 2.** ELG weighted spectroscopic redshift distribution per unit volume for the W1, W3, and W4 Wide fields (dashed and dotted histograms), and their mean value (solid line).

<sup>4</sup> <http://www.cfht.hawaii.edu/Science/CFHTLS/>

**Table 1.** ELG photometric data per CFHT-LS Wide field after applying the bright star and bad field mask.

Field	W1	W3	W4	all
Center $\alpha, \delta$	35°, -7°	215°, 54°	333°, 2°	-
area [deg <sup>2</sup> ]	63.75	44.22	23.3	131.27
N	32,808	22,195	11,025	66,028
N [deg <sup>-2</sup> ]	514.64	501.92	473.18	502.99
$z_{phot}$ quartiles	0.78 / 0.88 / 1.03	0.77 / 0.88 / 1.05	0.78 / 0.88 / 1.03	0.78 / 0.88 / 1.03
$z_{phot}$ Deciles D10, D90	0.7 / 1.24	0.68 / 1.29	0.69 / 1.25	0.69 / 1.26
N ( $0.6 < z_{phot} < 1$ )	23,433	15,242	7,861	46,536
N ( $0.6 < z_{phot} < 1$ ) [deg <sup>-2</sup> ]	367.58	344.69	337.38	354.51
$z_{phot}$ quartiles	0.75 / 0.83 / 0.9	0.74 / 0.81 / 0.89	0.75 / 0.83 / 0.89	0.75 / 0.82 / 0.89

**Figure 1.** Photometric (black points) and spectroscopic (VIPERS: red crosses in the right and left panels; BOSS: magenta crosses forming the oval in the central panel; DEEP2: blue crosses in the dashed box in the central panel) coordinates of our ELG sample in the three CFHT-LS Wide fields.

ber density of our sample and consequently suppressing the amplitude of  $w(\theta)$ , but we do not see any substantial change in the angular clustering trend with respect to the  $z < 1$  case. We therefore restrict the analysis to the redshift range  $0.6 < z < 1$ .

## 2.2 MultiDark simulations

The MultiDark Planck simulation (MDPL, Klypin et al. 2014; [www.MultiDark.org](http://www.MultiDark.org)) contains  $3840^3$  particles in a  $1 h^{-1} \text{Gpc}$  box, and was created adopting Planck  $\Lambda \text{CDM}$  cosmology (Planck Collaboration et al. 2014). Halos are identified based on density peaks including substructures using the Bound Density Maximum (BDM) halo finder (Klypin & Holtzman 1997; Riebe et al. 2013).

We use the MDPL halo catalogs to build a mock light-cone that matches the mean ELG redshift distribution shown in Figure 2. Given the high density of the ELG tracers and their expected low-mass host halos, the MDPL box is an excellent compromise between numerical resolution and volume. We apply the SURvey GenerAtor (SUGAR; Rodriguez-Torres et al. (2015), in prep.) to the 11 snapshots available from MDPL to construct a light-cone with a volume ten times the observations that covers the redshift range  $0.6 < z < 1$  ( $\sim 1 h^{-1} \text{Gpc}$  depth). The procedure used is analogous to the method presented by Blaizot et al. (2005)

and Kitzbichler & White (2007), and can be summarized as follows:

(i) Set the properties of the light-cone: angular mask, radial selection function (number density) and number of snapshots within the redshift range considered. Each slice of the light-cone is constructed by selecting all halos from every MDPL snapshot. The thickness of a slice at redshift  $z_i$  is given by  $[(z_i + z_{i-1})/2, (z_i + z_{i+1})/2]$

(ii) Place an observer (i.e.,  $z = 0$ ) inside the box and shift the cartesian coordinates of the box in such a way that the observer occupies the central point of the box at  $z = 0.8$

(iii) Convert from cartesian  $(x, y, z)$  to spherical  $(\alpha, \delta, r_c)$  coordinates, where  $r_c$  is the comoving distance in real space. The redshift of each point will be:

$$r_c(z) = \int_a^b \frac{cdz'}{H_0 \sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda}} \quad (1)$$

(iv) From each snapshot, select the (sub)halos so that  $(z_i + z_{i-1})/2 < z < (z_i + z_{i+1})/2$  and  $\alpha/\delta$  lie inside the sky window. Since the ELG observational data represent halos with typical masses  $\sim 10^{12} h^{-1} M_\odot$ , in the light-cone we include all halos for which the simulation is complete

**Table 2.** ELG spectroscopic data.

survey	match	good $z$	$0.6 < z < 1$	area [deg <sup>2</sup> ]	$\bar{z}$
VIPERS W1	1,223	942	760	5.478	0.803
BOSS W3	2,145	1,876	1,357	6.67	0.803
DEEP2 W3	225	222	156	0.5	0.803
VIPERS W4	1,148	846	680	5.120	0.795
All	4,741	3,886	2,953	17.668	0.803

i.e.,  $\log(M_h/h^{-1}M_\odot) > 11.2$

(v) Using the halo velocities,  $v_p$ , we compute the peculiar velocity contribution for each object along the line-of-sight and derive its distance in redshift-space as

$$s = r_c + (v_p \cdot r_c)/(aH(z)), \quad (2)$$

where  $a = (1+z)^{-1}$  is the scale factor and  $H(z)$  is the Hubble parameter at redshift  $z$

(vi) Finally, select objects from the light-cone using our selection function.

Throughout the paper we will designate our lightcone as ‘‘MDPL-LC’’. Section 4 describes in detail the halo selection and the (Sub)Halo-Abundance Matching modeling adopted to determine the halo occupation distribution of our ELG sample.

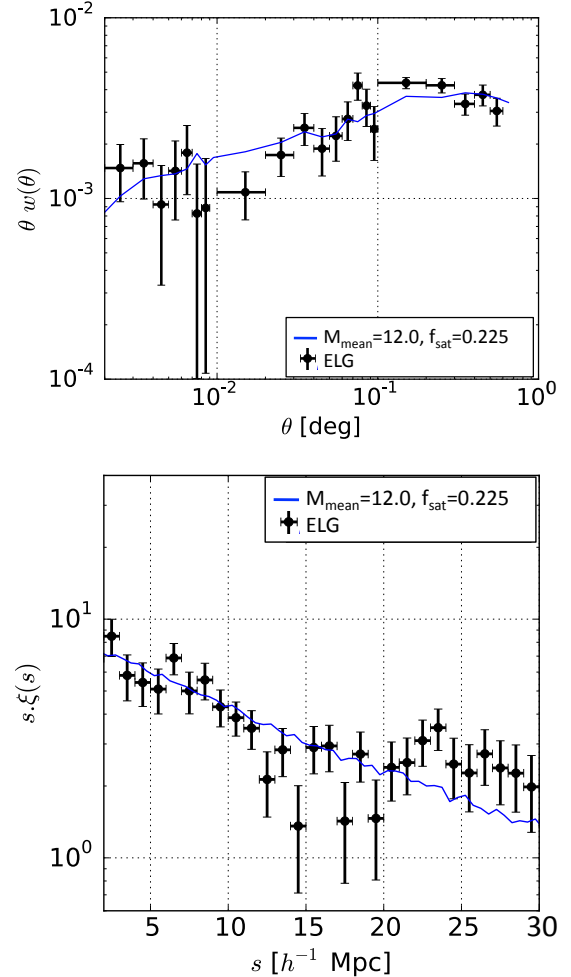
### 3 MEASUREMENTS

Using the ELG sample described in Section 2.1, we measure both galaxy clustering and galaxy-galaxy lensing. The following provides a detailed description of our measurements.

#### 3.1 Galaxy Clustering

We estimate both the angular,  $w(\theta)$ , and the redshift-space,  $\xi(s)$  (hereafter  $\xi^s$ ), two-point correlation functions following the procedures described by Landy & Szalay (1993), Coupon et al. (2012) and de la Torre et al. (2013).

To compute  $\xi^s$  on the VIPERS and the BOSS ELG samples (see Table 2), independently, we create linear bins in separations of  $1 h^{-1}\text{Mpc}$  at  $s < 10 h^{-1}\text{Mpc}$ , and  $4 h^{-1}\text{Mpc}$  for  $10 < s < 40 h^{-1}\text{Mpc}$ . We then correct the impact of redshift errors and catastrophic redshifts to recover the correlation function down to  $1 h^{-1}\text{Mpc}$ . The ELG we are targeting are observed using three plates overlapping the same area of the sky. This configuration guarantees that all the targets are observed at the end of the process and there is no fiber collision (Blanton et al. 2003; Reid et al. 2014). We correct for the finite size of VIPERS following de la Torre et al. (2013), and define the survey completeness in terms of target sampling rate (TSR) and the spectroscopic success rates (SSR). The first quantity weights our ability of obtaining spectra from the potential targets meeting the survey selection in the parent photometric sample. For each galaxy in the spectroscopic catalog, we count the number of objects that lie within a given radius in the spectroscopic ( $N_{targeted}$ ) and in the photometric ( $N_{parent}$ ) sample. The



**Figure 3.** Two-point angular (top panel) and redshift-space (bottom panel) ELG correlation functions (points), together with our best-fit model (blue line), which corresponds to the point highlighted by a star in Figure 6.

TSR is then given by the ratio  $TSR = N_{targeted}/N_{parent}$ . The SSR represents our ability of determining galaxy redshifts from observed spectra. To compute it, we replace the bad redshifts in the spectroscopic catalog with good photometric redshifts. Galaxies with good spectra are assigned  $flag=0$  ( $N_{good}$ ); galaxies with replaced redshifts are assigned  $flag=1$  ( $N_{targeted}$ ). The SSR is then computed as the ratio  $SSR = N_{good}/N_{targeted}$ . The contribution of each galaxy in the clustering pair counts is then weighted by  $w = 1/(TSR * SSR)$ . Finally, we combine the VIPERS and

BOSS measurements weighted by the projected density of each field. The resulting redshift-space correlation function is displayed in the bottom panel of Figure 3; by fitting a power-law model,  $\xi(s) = (s/s_0)^\alpha$ , in the separation range  $2 < s < 30 h^{-1}\text{Mpc}$ , we find  $s_0 = (5.3 \pm 0.2) h^{-1}\text{Mpc}$  and  $\alpha = -1.6 \pm 0.1$ .

Analogously, we calculate the angular 2PCF,  $w(\theta)$ , using photometric redshifts from the W1, W3 and W4 CFHT-LS fields. The points are corrected from the integral constraint following Tinker et al. (2010) and Coupon et al. (2012) to account for the restricted area of observation. On scales  $\theta < 0.05^\circ$ , all three fields provide consistent measurements. At larger scales, the clustering signals in the W1 and W4 fields do not decrease as rapidly as expected, probably pointing to possible systematics that need to be investigated further. We therefore use only the measurement on the W3 field, which appears the most robust (see Figure 3, top panel). The  $w(\theta)$  of the W3 field is in perfect agreement with Figure 9 (panel 4) in Comparat et al. (2013). This result was computed on the Stripe 82 region (Stoughton et al. 2002), with three times larger area. At the mean redshift of the sample,  $z = 0.8$ , one degree corresponds to  $18.847 h^{-1}\text{Mpc}$ ; thus,  $w(\theta)$  spans the range from  $\sim 40 h^{-1}\text{kpc}$  up to  $\sim 20 h^{-1}\text{Mpc}$ . We investigate further the impact on the clustering amplitude of including the tails of the photometric distribution, i.e.  $z_{\text{phot}} < 0.6$  and  $z_{\text{phot}} > 1$ . This inclusion does not produce any substantial change in  $w(\theta)$ , except for some additional noise. We also test how photometric uncertainties affect the clustering errors computed via mock resampling. To this purpose, following Coupon et al. (2012), we perturb our original redshift distribution by applying a photometric scatter with mean  $\sigma_z = 0.035(1+z)$ . We then quantify the number of photometric objects that, due to this scatter, enter the ELG selection in the range  $0.6 < z < 1$  from the lower and higher tails of the distribution, and the objects that exit. We find that only 2.5% photometric redshifts enter the ELG selection in the range  $0.6 < z < 1$  from the upper and lower tails, and their effect on the clustering is negligible.

To estimate the errors on our galaxy clustering measurements, since the simulated light-cone area is larger than the data ( $\sim 560 \text{ deg}^2$ ), we divide the best MDPL-LC model into independent (i.e., non-overlapping) realizations of our ELG data (8 for the photometric and 24 for the spectroscopic samples), and obtain sample variance diagonal errors that we use rather than Poisson errors. Including the photometric uncertainties in our jackknife resamplings does not provoke any significant change in the error estimates. We neglect a full-covariance analysis because the number of sub-samples we have is too small to produce reliable covariance estimates. Including also the off-diagonal elements of the covariance matrices would result in large fluctuations of the error bars. Of course, excluding covariances we are adopting a simplified approach, but it provides a good sense of how the SDSS BOSS ELG clustering behaves. On the other hand, the ELG sample considered here is too sparse to derive tight constraints from our clustering analysis. New-generation large-volume spectroscopic surveys as eBOSS, DESI and 4MOST, will provide new data with unprecedented statistics, sky coverage and imaging quality. Using those data, a fully covariant approach will return reliable and accurate error estimates.

We compare the combined  $\xi^s$  measurement from BOSS and VIPERS to previous measurements by Marulli et al. (2013) to provide a first interpretation. Our result matches both the clustering signal of galaxies selected in the stellar mass range  $9.5 < \log(M_*/h^{-1}M_\odot) < 11$ , and the clustering of galaxies selected by absolute magnitude in the interval  $-22 < M_B - 5 \log(h) < -20.5$ . Using the stellar-to-halo mass relation from Leauthaud et al. (2012), Shan et al. (2015) and Coupon et al. (2015), we can deduce a rough estimate of the halo masses populated by our ELG sample i.e.,  $11.6 < \log(M_h/h^{-1}M_\odot) < 12.7$ . These halo masses are typical of Milky-Way size halos, being much less massive than those hosting the LRG sample, see Nuza et al. (2013).

In the angular clustering measurement, the change of slope occurs at  $\theta \sim 0.01^\circ$ , corresponding to  $\sim 200 h^{-1}\text{kpc}$ . Using MDPL, we derive the relation between halo mass and virial radius at  $z \sim 0.8$ ; halos with virial radius  $\sim 200 h^{-1}\text{kpc}$  occupy the mass range  $M_h = (0.5 - 1) \times 10^{12} h^{-1}M_\odot$ . Since a single galaxy per halo would not induce such a change in the  $w(\theta)$  slope, this result implies a satellite fraction of approximately 22.5% (see Section 4). Figure 3 displays a good agreement between our clustering measurements and predictions for ELG halos of mass  $10^{12} h^{-1}M_\odot$  with this satellite fraction.

### 3.2 Weak lensing

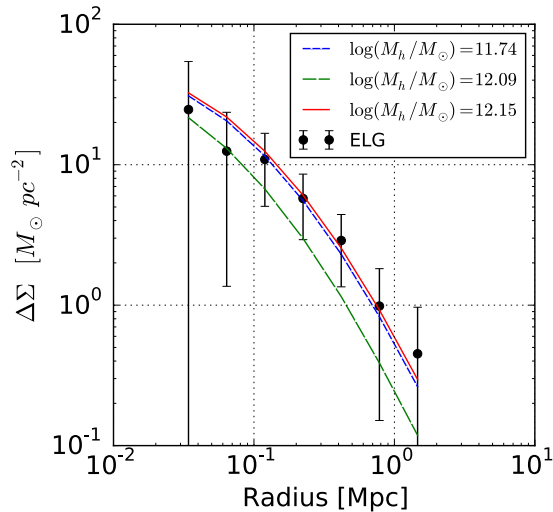
We use the latest weak lensing catalogs produced by the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012b; Erben et al. 2013b) on the W1 and W3 fields to measure the galaxy-galaxy lensing around 47,485 ELG lenses. This measurement allows one to constrain the halo masses. We follow Gillis et al. (2013) and apply only the multiplicative correction,  $m_s$ , to the shear measurement and avoid the c2 correction. We measure the tangential shear,  $\gamma^t$ , around the photometric ELG sample as a function of the radial distance from the lenses using the (Velandar et al. 2011) estimator:

$$\Delta\Sigma = \left[ \frac{\sum_{ls} w_{ls} \gamma_{ls}^t \Sigma_c}{\sum_{ls} w_{ls}} \right] / \left[ \frac{\sum_{ls} w_{ls} (1 + m_s)}{\sum_{ls} w_{ls}} \right], \quad (3)$$

where the sum runs over the lens - source pairs ( $ls$ ) and the  $w_{ls}$  values are the weight obtained by lensfit.

Since the lenses are at the higher tail of the redshift distribution and the ELGs are expected to live in low-mass halos, we recover a low signal-to-noise ratio around 2 for  $R < 1 \text{ Mpc}$ .

We model the measurement using a truncated Navarro, Frank & White (NFW) halo profile (Baltz et al. 2009) and the mass-concentration relation from Neto et al. (2007) to truncate halos at half their concentration (Wright & Brainerd 2000). The best-fit model suggests typical halo masses of  $M_{200} = 1.25 \pm 0.45 \times 10^{12} h^{-1}M_\odot$ . The lower and upper mass limits are, respectively,  $M_{200} = 5.61 \pm 7.20 \times 10^{11} h^{-1}M_\odot$  and  $M_{200} = 1.41 \pm 0.51 \times 10^{12} h^{-1}M_\odot$ ; see Fig. 4. This measurement is in good agreement with the first interpretations based on the clustering (see Section 3).



**Figure 4.** ELG surface density ( $\Delta\Sigma$ ) as a function of the physical scale for different lens models.

#### 4 HALO OCCUPATION FOR EMISSION LINE GALAXIES

The (Sub)Halo-Abundance Matching (SHAM; e.g., Conroy et al. 2006; Trujillo-Gomez et al. 2011) technique is a straightforward method to link observed galaxies with dark-matter-only simulated halos. It relies in a monotonic correspondence between halo and galaxy number densities, which is based on the assumption that more luminous galaxies reside in more massive halos. Such association is performed by choosing suitable proxies for both halos and galaxies (e.g., the halo maximum circular velocity and the galaxy luminosity or stellar mass) and includes some scatter (see Trujillo-Gomez et al. (2011) for details). The advantage of using  $N$ -body simulations, compared to analytical models, is given by the accuracy achieved in the predictions of the clustering for a given halo population. Many state-of-the-art clustering measurements have been modeled using a SHAM technique that maps the observations onto suitable high-resolution  $N$ -body simulations, allowing the interpretation of the halo occupation distribution and bias (de la Torre et al. 2013; Nuza et al. 2013; Carretero et al. 2015). Watson et al. (2015) recently presented a method to upgrade SHAM models to account for differences between quenched and star-forming galaxies.

In the specific case of the emission-line galaxies, the traditional SHAM approach cannot be applied since it requires a complete galaxy sample, and ELGs are far from being complete in any parameter space, even in terms of their emission line luminosity, see Comparat et al. (2013). We therefore must modify the standard SHAM procedure to take into account the ELG incompleteness and match their clustering amplitude. To this purpose, we selected halos and subhalos by mass (for the subhalos we considered only the mass of the bound particles, to avoid ambiguities) to be able to compare directly with the weak lensing measurements. In the future, provided a high signal-to-noise ratio in the clustering measurement, we will properly select

(sub)halos by their maximum circular velocity at accretion, (e.g., Behroozi et al. 2013b).

In order to model both the 1-halo and the 2-halo terms in the ELG two-point correlation functions and the weak lensing measurement, we use the MultiDark Planck  $1h^{-3}\text{Gpc}^3$  box (see Section 2.2), which represents the best compromise between high resolution and volume, as previously described in Section 3.

We parametrize the probability of selecting a halo hosting an ELG as follows:

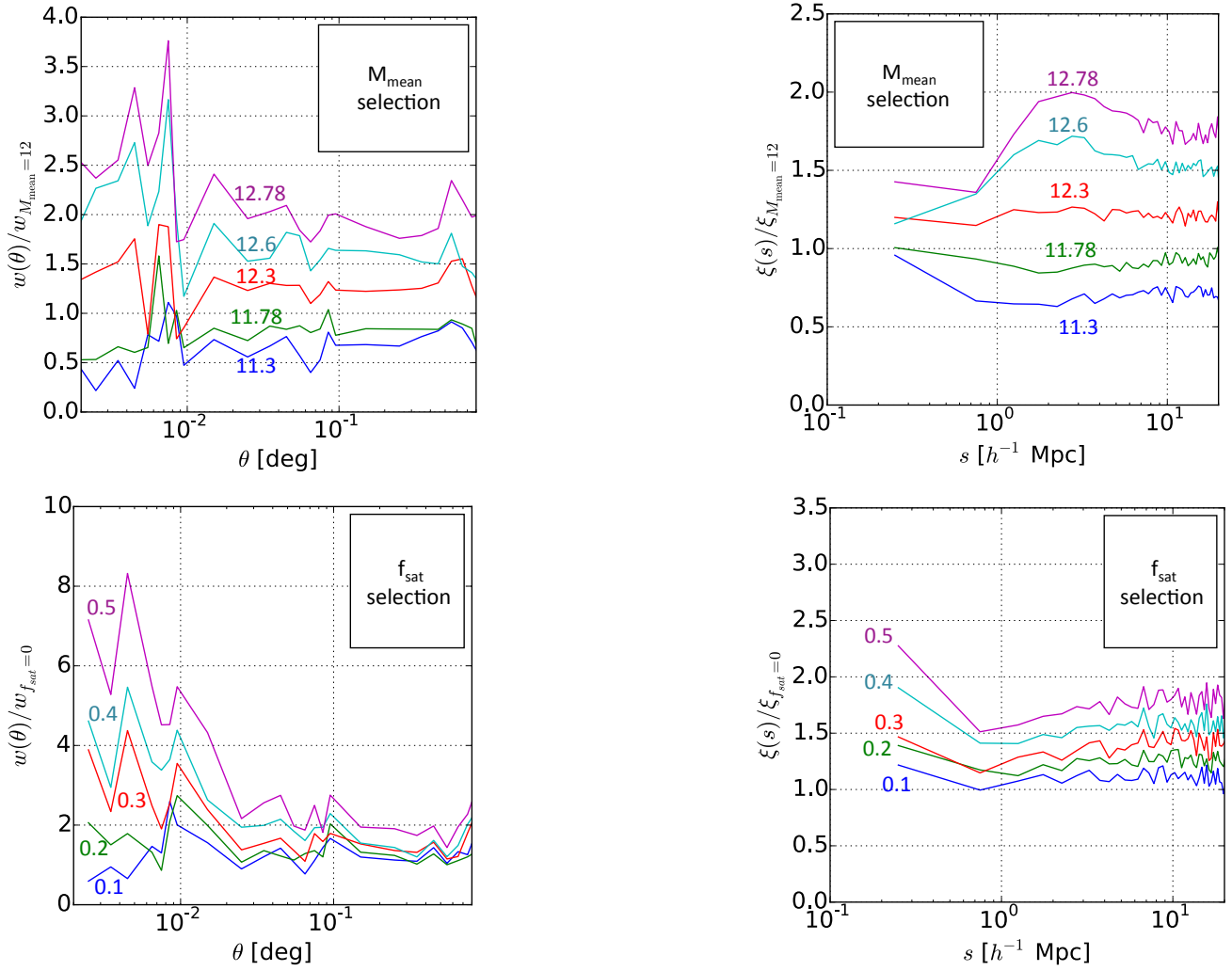
$$P(M_h, M_{mean}, \sigma_M, f_{sat}) = f_{sat} \mathcal{N}(M_h, M_{mean}, \sigma_M, \text{flag} = \text{sat}) + (1 - f_{sat}) \mathcal{N}(M_h, M_{mean}, \sigma_M, \text{flag} = \text{cen}) \quad (4)$$

where  $\mathcal{N}$  is a Gaussian distribution with the variable being  $M_h$ , the halo mass. The parameters are:  $M_{mean}$ , the mean halo mass of the sample including both host and satellite halos;  $\sigma_M$ , the dispersion around the mean halo mass;  $f_{sat}$ , the satellite fraction. The additional parameter “flag” enables to identify among the halos the ones that are centrals (flag=cen) or the ones that are satellites (flag=sat).

To qualitatively understand the dependence of clustering on  $M_{mean}$  and  $f_{sat}$ , we impose (i) a maximum halo mass threshold to the MDPL-LC by removing all halos with  $M_h > M_{max}$  and we apply the standard SHAM procedure. The higher-mass ( $M_{max} > 10^{13} h^{-1} M_\odot$ ) models reproduce well the observed  $w(\theta)$ , and that the lower-mass models ( $M_{max} < 10^{13} h^{-1} M_\odot$ ) match the large-scale clustering, but not the small-scale amplitude witnessed below  $\theta \sim 0.01^\circ$ . The top row in Figure 5 displays the ratio between the angular (left panel) and the monopole (right) correlation functions of the lower-mass models and the model with  $M_{mean} = 10^{12} h^{-1} M_\odot$ . We see a mild variation in  $w(\theta)$  as a function of the physical scale, and a flatter trend in the monopole.

We next (ii) fix the halo mass by selecting all the halos in the mass bin  $M_h = (1 \pm 0.5 \times 10^{12} h^{-1} M_\odot)$ , and vary the satellite fraction. We split this halo catalog into two catalogs, one containing only central halos ( $f_{sat} = 0$ ) and one with satellites; then downsample both mocks to match the ELG  $n(z)$ . The bottom panels in Figure 5 present the variation of the angular and monopole clustering as a function of the scale. At small scales the amplitude of  $w(\theta)$  with more than 30% satellite fraction is strongly enhanced compared to the 10 – 20% cases. In the monopole there is almost no variation with the scale. We then combine these two products to build galaxy mock catalogs that contain a  $f_{sat}$  fraction of satellites (taken from the satellite-only mock) and  $(1 - f_{sat})$  centrals (from the central-only mock). Satellite fractions between 20% and 30% account for the clustering signal on both small and large scales; see Figure 6. All the selections above are done on the halo mass defined as  $M_{200}$ , which correspond to an overdensity threshold of  $\Delta_{200} = 200\rho_c$  (Prada et al. 2012), where  $\rho_c$  is the critical density of the Universe.

To produce a mock catalog, we randomly select halos from the light-cone according to the probability distribution  $P$ , defined in Eq. 4, until the ELG redshift distribution  $n(z)$  in Figure 2 is achieved. We then construct a grid of mocks by selecting  $M_{mean}$  in the range  $10^{11.2} - 10^{12.7} h^{-1} M_\odot$ ,  $\sigma_M$  between the values  $M_{mean}/[1., 2., 4.] h^{-1} M_\odot$  (the sampling space is three times larger), and the satellite fraction in the interval  $0 < f_{sat} < 0.5$ , to obtain predictions for both  $\xi^s$  and



**Figure 5.** *Left column, top panel:* ratio of the angular correlation functions of the MDPL-LC halos selected by mass, to  $w(\theta)$  computed at  $M_{mean} = 10^{12} h^{-1} M_{\odot}$ . The curves in the plot go from lower mass (bottom line) to higher mass (top line). *Left column, bottom panel:* ratio of the angular correlation functions of the MDPL-LC halos with varying satellite fraction, to  $w(\theta)$  computed at  $f_{sat} = 0$ . The lines in the plot go from lower  $f_{sat}$  (bottom line) to higher  $f_{sat}$  (top line). *Right column:* same results for the monopole. The top row presents our first experiment (see the text for details) on the lightcone: we impose different halo mass thresholds to the MDPL-LC and apply a standard SHAM. The bottom row displays SHAM in the mass bin  $M_h = (1 \pm 0.5 \times 10^{12} h^{-1} M_{\odot})$  with varying satellite fractions.

$w(\theta)$ . Finally, we compare these model predictions with our measurements by computing a combined  $\chi^2$  on scales  $2 < s < 22 h^{-1} \text{Mpc}$  for the monopole, and  $0.002^\circ < \theta < 0.55^\circ$  for the angular clustering, as follows:

$$\chi^2 = \frac{N_{\xi} \chi_{\xi}^2 + N_w \chi_w^2(\theta)}{N_{\xi} + N_w}, \quad (5)$$

where

$$\chi_w^2(\theta) = \frac{1}{N_w} \sum_i^{N_w} \frac{|w_{observed}(\theta_i) - w_{halos}(\theta_i)|^2}{\sigma^2(w_{observed}(\theta_i))}, \quad (6)$$

and

$$\chi_{\xi}^2 = \frac{1}{N_{\xi}} \sum_i^{N_{\xi}} \frac{|\xi_{observed}(s_i) - \xi_{halos}(s_i)|^2}{\sigma^2(\xi_{observed}(s_i))}. \quad (7)$$

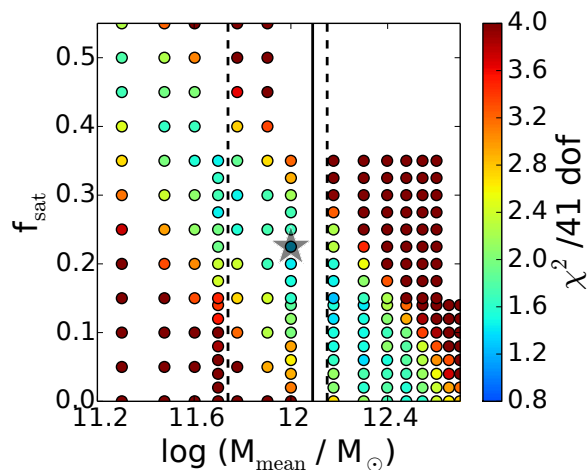
The possible models accounting for the ELG clustering

are degenerate with respect to the mean halo mass and the satellite fraction. In fact, Figure 6 shows that a plethora of  $(\log M_{mean}, f_{sat})$  models fit the data: from  $(11.3, 0.45)$  by  $(12, 0.2)$  to  $(12.5, 0)$ . Given the 41 degrees of freedom we have, we consider acceptable those models with  $\chi^2 < 1.25$ . Models with a higher  $\chi^2$  value are rejected at the 90% level.

The combination with the weak lensing results breaks this degeneracy and rules out the higher- and lower-mass models. However, among these latter, there is one with  $\chi^2 = 1$  and parameters:  $\log M_{mean} = 12$ ,  $\sigma_M = M_{mean}/2$ ,  $f_{sat} = 22.5\%$  (star symbol in Figure 6). The angular and redshift-space correlation functions of this best-fit mock are displayed in Figure 3 (blue line), together with the ELG measurements. The weak lensing measurement are perfectly compatible with this best-fit model.

We provide our best-fit MDPL mock catalog to the





**Figure 6.** The two parameters driving the model: fraction of satellite ( $f_{sat}$ ) and mean halo mass ( $M_{mean}$ ). The spread around the mean halo mass is fixed at the value  $\sigma_M = M_{mean}/2$ . The vertical black lines represent the constraints by weak lensing (dashed: lower and upper limits; solid: mean), which rule out the majority of the low-mass and high-mass models. Our best-fit model is highlighted by the star symbol.

ELG clustering measurements at <http://projects.ift.uam-csic.es/skies-universes/>.

## 5 RESULTS AND DISCUSSION

### 5.1 ELG clustering trends as a function of magnitude, flux, luminosity and stellar mass

We employ the complete VIPERS data sample at  $z \sim 0.8$ , which has about 30,000 reliable redshifts in the range  $0.6 < z < 1$ , to investigate trends of the clustering amplitude (bias) with observed or rest frame broad band magnitude or emission line flux. To this purpose, we measure the emission line properties in the VIPERS spectra and find a significant [OII] flux in about two thirds of them; the rest does not show emission lines (Comparat et al., in prep.). We bin the data according to apparent and absolute magnitude, [OII] flux and luminosity, and measure the clustering in each sample (the binning scheme was set to contain between 9000 and 10,000 data points). Figure 7 shows our ELG results in the observed (bottom row) and rest frame (top row). Consistently with previous analyses (e.g., Marulli et al. 2013; Mostek et al. 2013), we find that the brighter the selection in the  $i$ -band, either observed or rest-frame, the higher the bias. Analogously, the fainter the  $g$ -band limit, either observed or rest-frame, the higher the bias. The anti-correlation between [OII] flux and bias is only seen in the observed frame (the difference is  $\sim 1.4$ ); in the rest frame it is not significant. It would be interesting to further investigate the correlation between [OII] luminosity and  $g$ -band magnitude in the small-scale clustering, but with the resolution of current data we are not able to push the analysis to scales  $\sim 200h^{-1}\text{kpc}$ , which is the typical virial radius of a halo of mass  $10^{12}h^{-1}M_\odot$ . New data from eBOSS will be able to address this issue. The results above indicate that

if we have a  $g$ -selected ELG sample and [OII] fluxes for a certain number of its galaxies, in order to maximize its clustering signal, we should select the ELGs with brighter  $i$ -band magnitudes. To investigate the clustering dependence on stellar mass, we map the host halo masses for ELGs at  $z \sim 0.8$ ,  $M_h \sim 10^{12}h^{-1}M_\odot$ , onto stellar mass values using the stellar-to-halo-mass relation by Leauthaud et al. (2012), see their Figure 11. Our data are right before the “knee” at  $M_* \sim 3.5 \times 10^{10}h^{-1}M_\odot$ .

### 5.2 Star formation efficiency

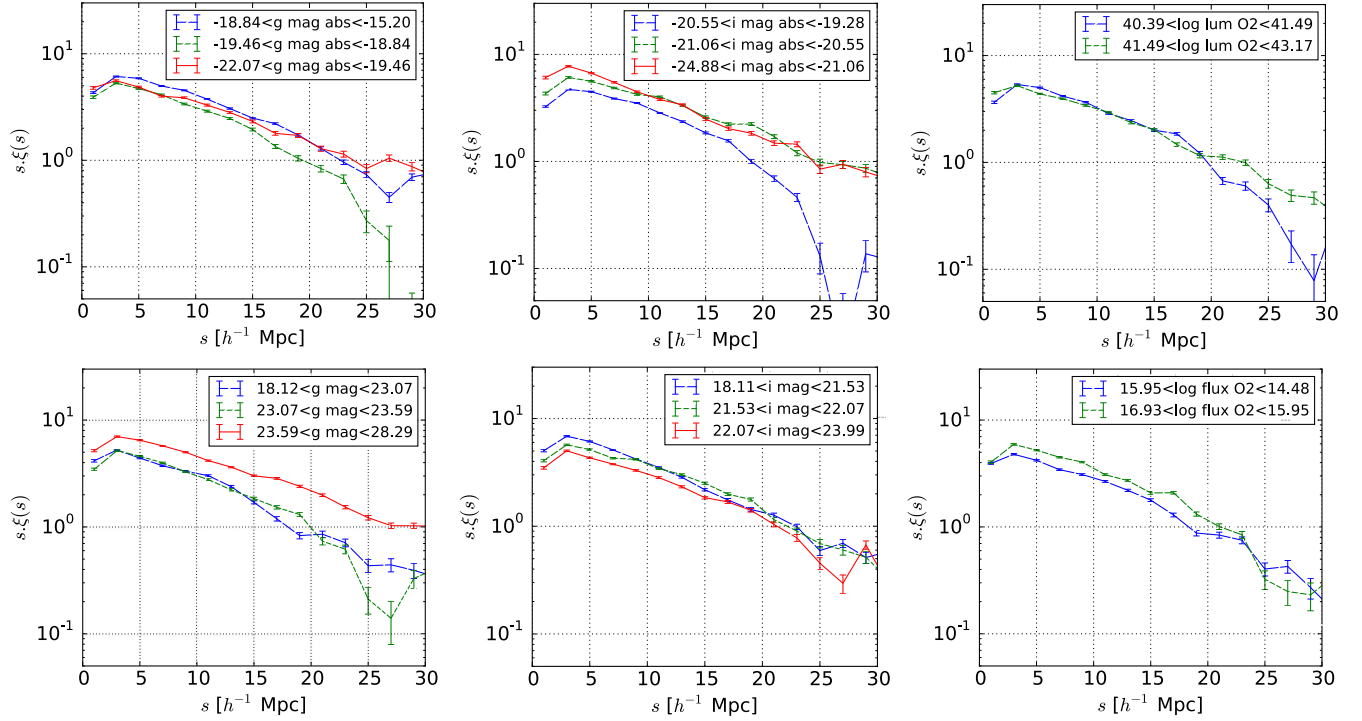
From our analysis, the typical halo masses hosting ELGs at  $z \sim 0.8$  are  $M_h \sim (1 \pm 0.5) \times 10^{12}h^{-1}M_\odot$ , and  $22.5\% \pm 2.5\%$  of them are satellites belonging to a larger halo, whose central is a quiescent galaxy. Figure 8 provides a schematic representation of the possible ELG configurations. A total of 22.1% ELGs are single satellites belonging to a parent halo with mass  $M_{hQ} \sim 2.5 \times 10^{13}h^{-1}M_\odot$ ; only in 1.3% of the cases the parent halo hosts more than one satellite ELG. The maximum number of satellites,  $n = 1.8$ , is achieved in the highest-mass case, where  $M_{hQ} \sim 6.8 \times 10^{13}h^{-1}M_\odot$ . These results imply that the mean number of ELG satellites is only slightly larger than unity ( $\sim 1.01$ ). The quiescent galaxies at the center of the parent halos are not included in the sample, since the stellar masses for ELGs from the SHMR discussed above are too low for halos of  $10^{13}h^{-1}M_\odot$ .

The typical masses for halos hosting ELGs suggest that we are sampling halos ( $\sim 10^{12}h^{-1}M_\odot$ ) that form stars in the most efficient way, according to the stellar-to-halo mass ratio discussed by Behroozi et al. (2013a) (see their Figure 1, bottom panel). This result opens a new science field and, hopefully, in the near future, integrated models combining N-body simulations with semi-analytic models (SAMs) will be able to probe star formation and shed some light on the correlations between [OII] flux and magnitude in the clustering of galaxies.

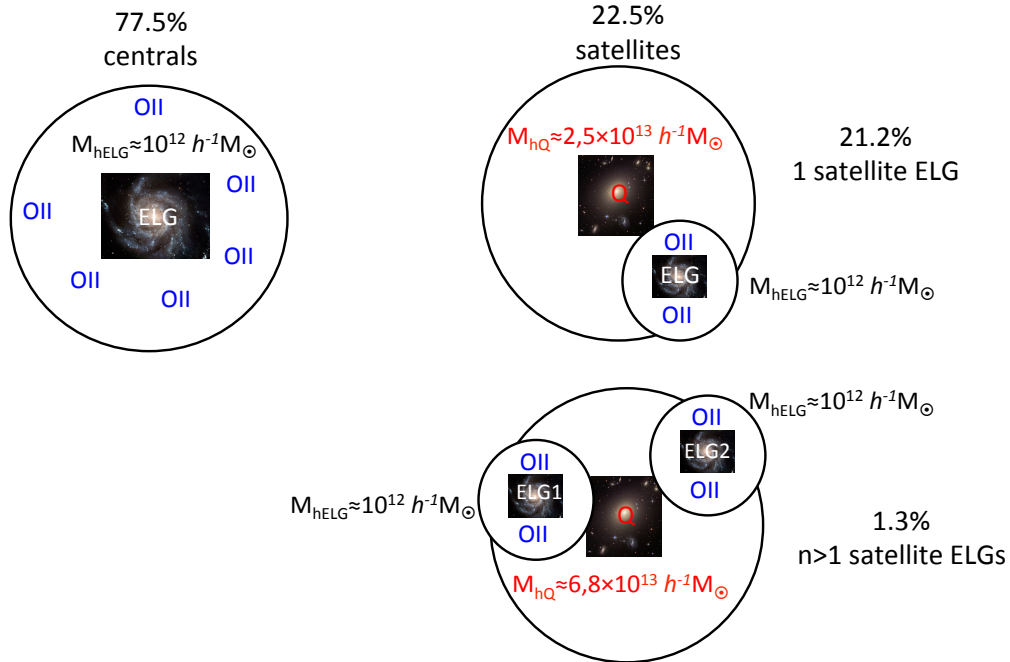
## 6 SUMMARY

We have presented an analysis of the halo occupation distribution for emission line galaxies, which jointly accounts for three measurements: the angular correlation function, the monopole, and the weak lensing signal around ELGs (see Section 3). Our procedure can be summarized in the following points:

- Apply the SUGAR (Rodríguez-Torres et al. (2015), in prep.) algorithm to the 11 snapshots available from the MDPL simulation to construct a light-cone (Section 2.2), with the same geometry and angular footprint of the ELG data.
- Modify the traditional SHAM technique (Section 4), to account for the ELG incompleteness, by selecting model galaxies by mass, until we match the observed ELG  $n(z)$ . In this way, our mock is constrained by the observed ELG redshift distribution, and represents a reliable model.
- Parametrize the probability of selecting a halo hosting an ELG with Eq. 4, in terms of the mean halo mass of the



**Figure 7.** VIPERS clustering trends as a function of the  $g$ -band and  $i$ -band magnitudes (top row: rest frame; bottom row: observed frame), [OII] luminosity (top row) and [OII] flux (bottom row).



**Figure 8.** Schematic diagram of possible ELG configurations. ELGs at  $z \sim 0.8$  typically live in halos of mass  $M_h \sim (1 \pm 0.5) \times 10^{12} h^{-1} M_\odot$  and 22.5% are satellites belonging to larger halos, whose central galaxy is quiescent. Among these satellite configurations, 21.2% of parent halos with  $M_{hQ} \sim 2.5 \times 10^{13} h^{-1} M_\odot$  host one satellite ELG, and only 1.3% of parents host more than one satellite ELG. The maximum number of satellites,  $n = 1.8$ , is achieved in the highest-mass case,  $M_{hQ} \sim 6.8 \times 10^{13} h^{-1} M_\odot$ . See the text for details.

sample ( $M_{mean}$ ), the dispersion around the mean ( $\sigma_M$ ), and the satellite fraction ( $f_{sat}$ ). The additional parameter “flag” enables to distinguish central and satellite halos.

- We perform two experiments (see Section 4) on the MDPL light-cone to derive information on which are the halo mass and satellite fraction ranges of values we need to input in our modified SHAM model to correctly fit the ELG clustering signal.

- Construct a grid of models based on these values, and jointly fit both angular and redshift-space clustering (see Section 4). Our best-fit models (see Figure 6) are degenerate with respect to  $M_{mean}$  and  $f_{sat}$ . The combination with the weak lensing analysis (see Section 3.1) breaks this degeneracy and rules out the highest and lowest mass models. Our best-fit ( $\chi^2 = 1$ ) model is shown in Figure 3 together with the ELG measurements, and is given by  $\log M_{mean} = 12$ ,  $f_{sat} = 22.5\%$ ,  $\sigma_M = M_{mean}/2$ .

To conclude, we have built and released to the community a reliable galaxy mock catalog that correctly fits the clustering amplitude of the *ugri* ELG sample constructed by matching spectroscopic redshifts from BOSS DR12, VIPERS and DEEP2 (for details see Section 2). With these tools, we can begin building many realizations of the density field to predict errors on the BAO measurement.

The measured halo masses for halos hosting emission-line galaxies indicate that we are sampling the halos that form stars in the most efficient way, according the stellar-to-halo mass ratio discussed by Behroozi et al. (2013a) (see their Figure 1, bottom panel). This is an important point for the future, and opens the path to further studies to understand the correlation between clustering and the strength of emission lines. With the resolution available from current data, we are not able to push the analysis to the typical scales ( $\sim 200h^{-1}\text{kpc}$ ) of halos of  $10^{12}h^{-1}M_{\odot}$ ; however, next-generation surveys, as eBOSS and DESI, will provide better resolution, and in the near future we should be able to build robust combinations of N-body simulations and SAMs that will address those questions.

## ACKNOWLEDGMENTS

GF is supported by the Ministerio de Educación y Ciencia of the Spanish Government through FPI grant AYA2010-2131-C02-01. JC acknowledges financial support from MINECO (Spain) under project number AYA2012 - 31101. GY acknowledges financial support from MINECO (Spain) under project number AYA2012-31101 and grant FPA2012-34694. EJ acknowledges the support of CNRS, and the Labex OCEVU. SEN acknowledges support by the Deutsche Forschungsgemeinschaft under the grant NU 332/2-1. GF, FP, SART, AK, SN and CC acknowledge financial support from the Spanish MICINN Consolider-Ingenio 2010 Programme under grant MultiDark CSD2009 - 00064, MINECO Centro de Excelencia Severo Ochoa Programme under grant SEV-2012-0249, and MINECO grant AYA2014-60641-C2-1-P. GF, JC and FP wish to thank the Lawrence Berkeley National Laboratory for the hospitality during the creation of this work. FP acknowledges the spanish MEC “Salvador de Madariaga” program, Ref. PRX14/00444.

The MultiDark Planck simulation has been performed in the Supermuc supercomputer at the Libniz Supercomputing Center (LRZ, Munich) thanks to the cpu time awarded by PRACE (proposal number 2012060963).

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## References

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543  
 Alam S. et al., 2015, *ApJS*, 219, 12  
 Anderson L. et al., 2014, *MNRAS*, 441, 24  
 Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, *Astrophys. J.*, 600, 681  
 Baltz E. A., Marshall P., Oguri M., 2009, *JCAP*, 1, 15  
 Bartelmann M., 1999, in *Evolution of Large Scale Structure : From Recombination to Garching*, Banday A. J., Sheth R. K., da Costa L. N., eds., p. 213  
 Behroozi P. S., Wechsler R. H., Conroy C., 2013a, *ApJ*, 762, L31  
 Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013b, *Astrophys. J.*, 763, 18  
 Berlind A. A., Weinberg D. H., 2002, *Astrophys. J.*, 575, 587  
 Blaizot J., Wadadekar Y., Guiderdoni B., Colombi S. T., Bertin E., Bouchet F. R., Devriendt J. E. G., Hatton S., 2005, *MNRAS*, 360, 159  
 Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, *AJ*, 125, 2276  
 Bolton A. S. et al., 2012, *AJ*, 144, 144  
 Bray A. D. et al., 2015, *ArXiv e-prints*: 1502.01348  
 Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, *MNRAS*, 447, 646  
 Coil A. L. et al., 2008, *Astrophys. J.*, 672, 153  
 Comparat J. et al., 2013, *MNRAS*, 433, 1146  
 Comparat J. et al., 2015, *A&A*, 575, A40  
 Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *Astrophys. J.*, 647, 201  
 Cooper M. C. et al., 2008, *MNRAS*, 383, 1058

- Cooray A., Sheth R., 2002, *Physics Letters B*, 372, 1
- Coupon J. et al., 2015, *MNRAS*, 449, 1352
- Coupon J. et al., 2009, *A&A*, 500, 981
- Coupon J. et al., 2012, *A&A*, 542, A5
- Dawson K. S. et al., 2013, *AJ*, 145, 10
- de la Torre S. et al., 2013, *A&A*, 557, A54
- Eisenstein D. J. et al., 2011, *AJ*, 142, 72
- Erben T. et al., 2013a, *MNRAS*, 433, 2545
- Erben T. et al., 2013b, *MNRAS*, 433, 2545
- Favole G., McBride C. K., Eisenstein D. J., Prada F., Swanson M. E., Chuang C.-H., Schneider D. P., 2015, *ArXiv e-prints*: 1506.02044
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
- Garilli B. et al., 2014, *A&A*, 562, A23
- Gillis B. R. et al., 2013, *MNRAS*, 431, 1439
- Gunn J. E. et al., 2006, *AJ*, 131, 2332
- Guo H. et al., 2015, *ArXiv e-prints*: 1505.07861
- Guo H. et al., 2014, *MNRAS*, 441, 2398
- Guzzo L. et al., 2014, *A&A*, 566, A108
- Hearin A. P., Watson D. F., 2013, *MNRAS*, 435, 1313
- Hearin A. P., Watson D. F., Becker M. R., Reyes R., Berlind A. A., Zentner A. R., 2014, *MNRAS*, 444, 729
- Heymans C. et al., 2012a, *MNRAS*, 427, 146
- Heymans C. et al., 2012b, *MNRAS*, 427, 146
- Ilbert O. et al., 2006, *A&A*, 457, 841
- Kitzbichler M. G., White S. D. M., 2007, *MNRAS*, 376, 2
- Klypin A., Holtzman J., 1997, *ArXiv Astrophysics e-prints*: astro-ph/9712217
- Klypin A., Prada F., Yepes G., Hess S., Gottlöber S., 2013, *ArXiv e-prints*: 1310.3740
- Klypin A., Yepes G., Gottlöber S., Prada F., Hess S., 2014, *ArXiv e-prints*: 1411.4001
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, *Astrophys. J.*, 609, 35
- Landy S. D., Szalay A. S., 1993, *Astrophys. J.*, 412, 64
- Laureijs R. et al., 2011, *ArXiv e-prints*: 1110.3193
- Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, *Astrophys. J.*, 738, 45
- Leauthaud A. et al., 2012, *Astrophys. J.*, 744, 159
- Marulli F. et al., 2013, *A&A*, 557, A17
- Mostek N., Coil A. L., Cooper M., Davis M., Newman J. A., Weiner B. J., 2013, *Astrophys. J.*, 767, 89
- Neto A. F. et al., 2007, *MNRAS*, 381, 1450
- Newman J. A. et al., 2013, *ApJS*, 208, 5
- Nuza S. E. et al., 2013, *MNRAS*, 432, 743
- Oke J. B., Gunn J. E., 1983, *Astrophys. J.*, 266, 713
- Planck Collaboration et al., 2014, *A&A*, 571, A16
- Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, *MNRAS*, 423, 3018
- Reid B. A., Seo H.-J., Leauthaud A., Tinker J. L., White M., 2014, *MNRAS*, 444, 476
- Riebe K. et al., 2013, *Astronomische Nachrichten*, 334, 691
- Sartoris B. et al., 2015, *ArXiv e-prints*: 1505.02165
- Schlegel D. J. et al., 2015, in *American Astronomical Society Meeting Abstracts*, Vol. 225, American Astronomical Society Meeting Abstracts
- Scoville N. et al., 2007, *ApJS*, 172, 150
- Shan H. et al., 2015, *ArXiv e-prints*: 1502.00313
- Skibba R. A. et al., 2015, *Astrophys. J.*, 807, 152
- Smee S. A. et al., 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 2
- Smee S. A. et al., 2013, *AJ*, 146, 32
- Stoughton C. et al., 2002, *AJ*, 123, 485
- Sugai H. et al., 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 0
- Swanson M. E. C., Tegmark M., Blanton M., Zehavi I., 2008, *MNRAS*, 385, 1635
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *Astrophys. J.*, 724, 878
- Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, *Astrophys. J.*, 742, 16
- Velander M., Kuijken K., Schrabback T., 2011, *MNRAS*, 412, 2665
- Wang Y., Yang X., Mo H. J., van den Bosch F. C., 2007, *Astrophys. J.*, 664, 608
- Watson D. F. et al., 2015, *MNRAS*, 446, 651
- Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, *Physics Letters B*, 530, 87
- Wright C. O., Brainerd T. G., 2000, *Astrophys. J.*, 534, 34
- York D. G. et al., 2000, *AJ*, 120, 1579
- Zehavi I. et al., 2011, *Astrophys. J.*, 736, 59
- Zehavi I. et al., 2005, *Astrophys. J.*, 630, 1
- Zheng Z. et al., 2005, *Astrophys. J.*, 633, 791
- Zheng Z., Coil A. L., Zehavi I., 2007, *Astrophys. J.*, 667, 760