The ALHAMBRA survey: evolution of galaxy clustering since $z \sim 1$

P. Arnalte-Mur,[1][*] V. J. Martínez,[2,3,4] P. Norberg,[1] A. Fernández-Soto,[4,5] B. Ascaso,[6]
A. I. Merson,[7] J. A. L. Aguerri,[8] F. J. Castander,[9] L. Hurtado-Gil,[2,5]
C. López-Sanjuan,[10] A. Molino,[6] A. D. Montero-Dorta,[11] M. Stefanon,[12] E. Alfaro,[6]
T. Aparicio-Villegas,[13] N. Benítez,[6] T. Broadhurst,[14] J. Cabrera-Caño,[15] J. Cepa,[8,16]
M. Cerviño,[6,8,16] D. Cristóbal-Hornillos,[10] A. del Olmo,[6] R. M. González Delgado,[6]
C. Husillos,[6] L. Infante,[17] I. Márquez,[6] J. Masegosa,[6] M. Moles,[10] J. Perea,[6]
M. Pović,[6] F. Prada[6] and J. M. Quintana[6]

Affiliations are listed at the end of the paper

ABSTRACT

We study the clustering of galaxies as function of luminosity and redshift in the range $0.35 < z < 1.25$ using data from the Advanced Large Homogeneous Area Medium-Band Redshift Astronomical (ALHAMBRA) survey. The ALHAMBRA data used in this work cover 2.38 deg^2 in seven independent fields, after applying a detailed angular selection mask, with accurate photometric redshifts, $\sigma_z \lesssim 0.014(1 + z)$, down to $I_{AB} < 24$. Given the depth of the survey, we select samples in B-band luminosity down to $L^{th} \simeq 0.16L^*$ at $z = 0.9$. We measure the real-space clustering using the projected correlation function, accounting for photometric redshifts uncertainties. We infer the galaxy bias, and study its evolution with luminosity. We study the effect of sample variance, and confirm earlier results that the Cosmic Evolution Survey (COSMOS) and European Large Area ISO Survey North 1 (ELAIS-N1) fields are dominated by the presence of large structures. For the intermediate and bright samples, $L^{med} \gtrsim 0.6L^*$, we obtain a strong dependence of bias on luminosity, in agreement with previous results at similar redshift. We are able to extend this study to fainter luminosities, where we obtain an almost flat relation, similar to that observed at low redshift. Regarding the evolution of bias with redshift, our results suggest that the different galaxy populations studied reside in haloes covering a range in mass between $\log_{10}[M_h/(h^{-1}\,M_\odot)] \gtrsim 11.5$ for samples with $L^{med} \simeq 0.3L^*$ and $\log_{10}[M_h/(h^{-1}\,M_\odot)] \gtrsim 13.0$ for samples with $L^{med} \simeq 2L^*$, with typical occupation numbers in the range of $\sim$1–3 galaxies per halo.

Key words: methods: data analysis – methods: statistical – galaxies: distances and redshifts – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

The large-scale structure (LSS) of the Universe is one of the main observables that we can use to obtain information about the nature of dark matter and cosmic acceleration. The simplest way to study the LSS is to study the spatial distribution of galaxies in surveys covering cosmologically significant volumes. Although the galaxy distribution is closely related to the global matter distribution, they are not equal. The relation between both distributions is known as galaxy bias, and it depends on the processes of galaxy formation

and evolution. In the simplest case, one can consider the galaxy contrast to be proportional to the matter contrast. Then, the bias is simply the constant of proportionality, which is independent of scale. Being able to understand and model this bias is crucial for the correct interpretation of the cosmological information that can be obtained from the analysis of galaxy clustering.

As the bias encodes information about the galaxy formation and evolution process, it is logical to expect that it will be different for different galaxy populations. In other words, the clustering properties of galaxies should depend on some of their intrinsic properties, such as stellar mass, star formation rate or age, and should evolve with time. This phenomenon, known as galaxy segregation, is observed when studying the dependence of clustering on different

observables such as luminosity, colour or morphology. In general, it is observed that bright, red, elliptical galaxies are more strongly clustered (i.e. they have a larger bias) than faint, blue, spiral ones (see e.g. Davis & Geller 1976; Hamilton 1988; Madgwick et al. 2003; Skibba et al. 2009; Martínez, Arnalte-Mur & Stoyan 2010; Zehavi et al. 2011).

In this work, we focus on the dependence of the galaxy bias on luminosity, and the evolution of this relation with redshift. This dependence has been studied extensively in the local Universe using both the Two-degree Field Galaxy Redshift Survey (2dFGRS; Norberg et al. 2001, 2002) and Sloan Digital Sky Survey (SDSS; Tegmark et al. 2004; Zehavi et al. 2005, 2011). Guo et al. (2013) also studied this relationship at $z \sim 0.5$ using data from the Baryon Oscillation Spectroscopic Survey (BOSS). The bias shows a weak dependence on luminosity L for galaxies with $L < L^*$, where L^* is the characteristic luminosity parameter of the Schechter function. For $L \gtrsim L^*$, however, this relation steepens, and the bias clearly increases with luminosity.

These studies of galaxy clustering and luminosity segregation have been extended to redshifts in the range $z \sim 0.5$–1, using state-of-the-art spectroscopic surveys, such as the VIMOS-VLT Deep Survey (VVDS; Pollo et al. 2006; Abbas et al. 2010), the Deep Extragalactic Evolutionary Probe survey (DEEP2; Coil et al. 2006, 2008), the zCOSMOS survey (Meneux et al. 2009) or the VIMOS Public Extragalactic Redshift Survey (VIPERS; Marulli et al. 2013) or photometric surveys such as the Canada–France–Hawaii Legacy Survey (CFHTLS)-Wide survey (McCracken et al. 2008; Coupon et al. 2012). Recently, Skibba et al. (2014) used an intermediate method, somehow similar to the one presented in this work. They used low-resolution spectroscopy data (with a typical redshift precision of $\sigma_z/(1 + z) = 0.005$) from the Prism Multi-Object Survey (PRIMUS) to study galaxy clustering in the range $0.2 < z < 1$. Overall, these studies show strong evidence for luminosity segregation at these redshifts, with the relation between bias and luminosity being slightly steeper in this case than in local studies. However, the luminosity range covered by these surveys is more limited in these cases, and is restricted typically to $L^{th} \gtrsim 0.3L^*$.

The presence of this bias parameter can be understood in a natural way in the context of the halo model (e.g. Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002). In this model, the matter distribution is decomposed into a population of massive virialized dark matter haloes that form at the peaks of the density field, and galaxies form within these haloes. The bias parameter for dark matter haloes can be modelled, and depends on the properties of the halo such as its mass (e.g. Sheth, Mo & Tormen 2001; Mo & White 2002). Studying the clustering of a certain population of galaxies gives therefore information on the characteristics of the haloes that host them. In this context, luminosity segregation indicates that more luminous galaxies form preferentially in more massive haloes than fainter ones.

In this work, we use data from the Advanced Large Homogeneous Area Medium-Band Redshift Astronomical (ALHAMBRA) survey (Moles et al. 2008; Molino et al. 2014) to study galaxy clustering and luminosity segregation for redshifts in the range $0.35 < z < 1.25$, using the two-point correlation function. ALHAMBRA is a deep photometric survey which uses a total of 23 optical and near-infrared (NIR) bands in order to obtain accurate and reliable photometric redshifts (photo-z) for a large number of objects, in a nominal area of 4 deg^2 over eight independent fields. It is therefore well suited to study the large-scale distribution of galaxies in significant volumes over this redshift range, providing an opportunity to explore the clustering of fainter galaxies than it is possible using spectroscopic

surveys. Moreover, the use of several independent fields allows us to use ALHAMBRA to study the effect of sample variance in the clustering measurements, and in particular the effect of large structures present in the samples. López-Sanjuan et al. (2014) also exploited this independence of the ALHAMBRA fields to study the effect of sample variance on merger fraction studies.

In Section 2 we present the ALHAMBRA data used in this work (characterized in more detail in Molino et al. 2014), and our selection of samples. We also present here the mock catalogues created to test our clustering methods. In Section 3 we explain how we model the selection function of the survey, and in particular the masks created to reproduce the angular selection. Section 4 presents our method to estimate the projected correlation function (a real space quantity) taking into account the effect of the photometric redshifts, following Arnalte-Mur et al. (2009). We also present our error estimation method, and leave for Appendix A the detailed justification of our line-of-sight integration limit. Our results are presented in Section 5. We show the correlation functions obtained for our different samples, including the modelling in terms of a simple power-law model (Section 5.1), and of a Λ cold dark matter (ΛCDM) model in order to derive the bias parameter (Section 5.2). We also make use of the independence of the surveyed fields to study the effect of sample variance on our measurements (Section 5.3), and compare our results with those of previous surveys in a similar redshift range (Section 5.4). In Appendix B we present the tests done using the mock catalogues to test the reliability of the results, and Appendix C contains numerical tables of our results. Finally, in Section 6 we discuss our results and summarize our conclusions.

Unless noted otherwise, we use a fiducial flat ΛCDM cosmological model with parameters $\Omega_M = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.0458$ and $\sigma_8 = 0.816$ based on the 7-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*7) results (Komatsu et al. 2011). All the distances used are comoving, and are expressed in terms of the Hubble parameter $h \equiv H_0/100 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$. Absolute magnitudes are given as $M - 5 \log_{10}(h)$, even when not explicitly indicated.

2 DATA USED: THE ALHAMBRA SURVEY

The ALHAMBRA survey (Moles et al. 2008) is a photometric survey which covers a total of 4 deg^2 in the sky, using 20 medium-band filters in the optical range, and three standard broad-band filters (J, H and K_s) in the NIR. The survey was carried out using the 3.5-m telescope at the Centro Astronómico Hispano-Alemán (CAHA)[1] in Calar Alto (Almería, Spain). The camera used for the optical observations was the Large Area Imager for Calar Alto (LAICA),[2] and Omega-2000[3] was used for the NIR observations.

The optical filter system for the ALHAMBRA survey was specifically designed to optimize the output of the survey in terms of photo-z accuracy and number of objects with reliable z determination (Benítez et al. 2009b). It consists of a set of 20 contiguous, equal-width, medium-band filters of full width at half-maximum (FWHM) $\simeq 310$ Å covering the full optical range, between 3500 and 9700 Å (Aparicio Villegas et al. 2010). The survey is complemented by observations in the standard NIR filters J, H and K_s. The homogeneous spectral coverage of this system minimizes the variations in the selection functions of the different objects with redshift. The NIR observations help eliminate some degeneracies

---

[1] http://www.caha.es

[2] http://www.caha.es/CAHA/Instruments/LAICA

[3] http://www.caha.es/CAHA/Instruments/O2000

Table 1. Properties of the seven ALHAMBRA fields used in this work. We list the number of frames N_f included in the current catalogue in each case (where a completed field corresponds to eight frames), the area A_{eff} covered by the survey according to our angular selection mask, the number of galaxies N_g included in the catalogue used (at $I < 24$) and the resulting surface number density N_g/A_{eff}. We also list other surveys which have overlap with each of the fields, see Moles et al. (2008) for details.

Field	N_f	A_{eff} (deg^2)	N_g	N_g/A_{eff} (deg^{-2})
ALH-2/DEEP2	8	0.377	26 759	70 979
ALH-3/SDSS	8	0.404	28 331	70 126
ALH-4/COSMOS	4	0.203	16 877	83 138
ALH-5/HDF-N	4	0.216	16 629	76 986
ALH-6/GROTH	8	0.400	28 892	72 230
ALH-7/ELAIS-N1	8	0.406	29 530	72 734
ALH-8/SDSS	8	0.375	27 615	73 640
Total	48	2.381	174 633	73 344

in the photo-z determination while at the same time improving the determination of important galaxy properties such as stellar mass.

2.1 ALHAMBRA imaging data

The data used in this work correspond to the photometric catalogue described in Molino et al. (2014).[4] It contains data for a nominal area of 3 deg^2 distributed over seven fields in the sky, in order to minimize the effects of sample variance (see Table 1). The minimum distance between fields is 17°, so we can safely consider them as independent. The fields were primarily chosen because of their low extinction, and trying to have significant overlap with other surveys (Moles et al. 2008). Each field is typically composed of eight frames forming two strips of ∼15 arcmin × 1°, separated by a ∼15 arcmin gap. We discuss in detail the geometry of the different fields in Section 3.1. We developed our own pipelines for the reduction of the imaging data, including bias, flat-field and fringing corrections. The details of the data reduction can be found in Cristóbal-Hornillos et al. (2009) and Cristóbal-Hornillos et al. (in preparation) for both the optical and NIR data.

The detection of objects for inclusion in the catalogue is performed in synthetic images built using a combination of the ALHAMBRA filters in the range 7000 < λ < 9700 Å to match the *Hubble Space Telescope F814W* filter (hereafter denoted as our *I* band). Matched photometry is then obtained for these detected objects in the 23 ALHAMBRA filters. We restrict our analysis to the magnitude range $I < 24$, where the catalogue is photometrically complete and we do not expect any significant field-to-field variation in the depth (see section 3.8 in Molino et al. 2014).

We eliminate stars from the catalogue using the star–galaxy separation method described in Molino et al. (2014), which uses information on both the geometry and colours of the sources. In particular, we use the stellar flag given in the catalogue, and select only objects with STELLAR_FLAG <0.7. This method is only reliable for $I < 22.5$. However, at $I = 22.5$ the fraction of stars in the sample is ∼1 per cent, and we expect it to decrease at fainter magnitudes. Therefore the possible effect of stellar contamination

at $I > 22.5$ is negligible. The final catalogue used contains a total of $N_T = 174\,633$ galaxies.

2.2 Photometric redshifts

Photometric redshifts were estimated for this catalogue using an updated version of the Bayesian Photometric Redshift (BPZ) code (Benítez 2000), including a new prior and spectral template library (Benítez et al., in preparation), and a new technique for the re-calibration of the photometric zero-points. Molino et al. (2014) discussed in detail the methods used for the redshift estimation and the characteristics of the photo-z obtained. They performed a comparison for the ∼7000 galaxies with measured spectroscopic redshift (see their fig. 25) and showed that the global accuracy in the photo-z is $\sigma_z \lesssim 0.014(1 + z)$ for $I < 24.5$. We show the distribution of photo-z for this catalogue in Fig. 1. The median redshift of the catalogue is $z_{med} = 0.75$, with the bulk of the redshift distribution in the range $0.35 < z < 1.25$ that we study in this work.

As an additional test of the reliability of the photometric redshifts used, we made a comparison with the Cosmic Evolution Survey (COSMOS) photo-z catalogue described in Ilbert et al. (2009). This catalogue contains photometric redshift determinations with comparable accuracy and depth to those in the ALHAMBRA catalogue, and overlaps with the field ALH-4 (see Table 1). We matched both catalogues using a separation radius of 1 arcsec in the angular position, and obtained a sample of 12 832 objects common to both catalogues. We show the distribution of the relative redshift differences for this sample in Fig. 2, where we also quote the dispersion in the results measured using the normalized median absolute deviation (NMAD) method (see e.g. Brammer, van Dokkum & Coppi 2008).

We compare the dispersion obtained in this way to a simple estimate based on the redshift errors quoted in both catalogues. In each case, we estimate the typical redshift uncertainty (in both ALHAMBRA and COSMOS) as the mean of the 1σ errors quoted for each object in the sample. Our estimate for the dispersion in the difference shown in Fig. 2 is then $\sigma_{diff} = \sqrt{\sigma_{ALH}^2 + \sigma_{COS}^2}$. We obtain that this value of the dispersion obtained from the quoted errors is a good estimate of that observed. However, for our faintest samples ($I > 23$), we need to increase this estimate by a factor of ∼1.3, suggesting that the photo-z uncertainty could be slightly underestimated
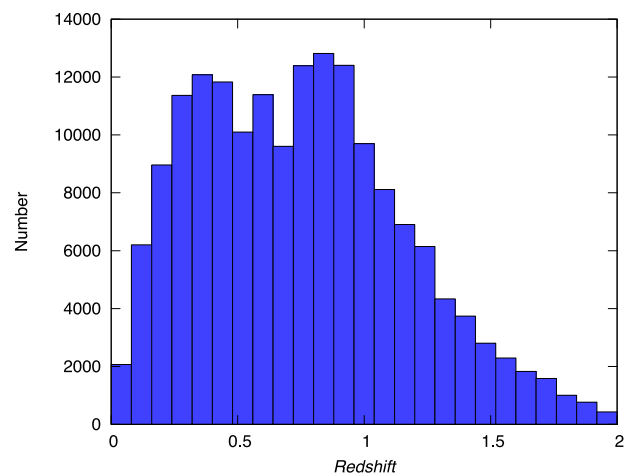

Figure 1. Redshift distribution of the 174 633 galaxies in the ALHAMBRA catalogue used in this work. The distribution shown corresponds directly to a histogram of the 'best' photo-z for each galaxy, in bins of width 0.08.

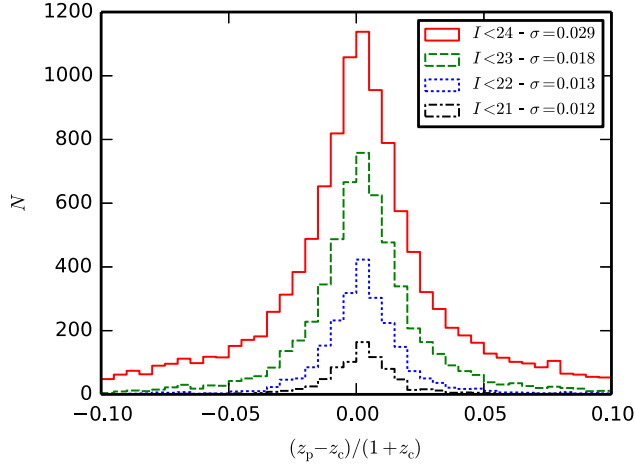[4] This ALHAMBRA catalogue is publicly available at http://www.alhambra survey.com/

Figure 2. Distribution of the relative differences between the ALHAMBRA photometric redshift (z_p) and the COSMOS photometric redshift (z_c) for the objects matched between the two catalogues. We show this distribution for different *I*-band magnitude selections, as shown in the label. We quote in each case the dispersion σ estimated using the NMAD method.

for these galaxies in both ALHAMBRA and COSMOS. Hereafter, we quote the error estimates for our samples (e.g. in Table 2) as the mean of the quoted BPZ error for the objects in the sample. For consistency, we correct this value by the factor of 1.3 for all samples,

although we only see an indication for the underestimation of the errors at the faintest ones.

2.3 Selection of samples in redshift and luminosity

To study the dependence of clustering properties on both luminosity and cosmic time, we build a series of subsamples splitting the catalogue in redshift and absolute magnitude.

The size of the redshift bins has to be larger than the distance we will integrate over the radial direction, π_{max}. As shown in Arnalte-Mur et al. (2009), using smaller bins may introduce systematic effects in the correlation functions we want to measure. Taking this fact into account, and the limitations in volume covered and galaxy density, we decided to use the four redshift bins $0.35 < z_p < 0.65$, $0.55 < z_p < 0.85$, $0.75 < z_p < 1.05$, $0.95 < z_p < 1.25$. We allow for overlap between consecutive bins in order to better trace the redshift evolution in our analysis, but one should bear in mind that results for different bins will be therefore correlated. Our low-redshift limit $z_p = 0.35$ was set in order for the scales of interest to be well sampled given the angular size of the fields. At this redshift, the typical size of a field, $1°$, corresponds to a projected comoving separation of $17\,h^{-1}$ Mpc. We fixed our high-redshift limit at $z_p = 1.25$ as, for higher redshifts, the quality of the photo-z and the number density of objects are significantly reduced.

In addition to the redshift selection, we also apply a set of cuts in the rest-frame *B*-band absolute magnitude M_B. We use this band for the selection as the region of the spectrum corresponding to it is well

Table 2. Characteristics of the different samples selected in redshift and luminosity. For each sample we quote the redshift range, *B*-band absolute magnitude threshold at $z = 0$ $M_B^{th}(0)$ (see equation 1), number of galaxies N, mean number density $\bar{n}$, median redshift z_{med}, median absolute magnitude M_B^{med}, median luminosity L^{med} as function of $L^*(z_{med})$, typical redshift error $\sigma_z/(1 + z)$ and typical line-of-sight distance error $r(\sigma_z)$ (see the text for details).

| Sample	Redshift range	$M_B^{th}(0)$	N	$\bar{n}$ ($10^{-3}\,h^3$ Mpc^{-3})	z_{med}	M_B^{med}	$L^{med}/L^*(z_{med})$	$\sigma_z/(1 + z)$	$r(\sigma_z)$ (h^{-1} Mpc)
Z05M0	0.35–0.65	−16.8	29 496	32.4 ± 0.5	0.521	−18.32	0.16	0.0197	69.2
Z05M1	0.35–0.65	−17.6	19 096	21.0 ± 0.4	0.523	−18.91	0.27	0.0143	50.4
Z05M2	0.35–0.65	−18.1	13 837	15.2 ± 0.3	0.524	−19.27	0.37	0.0133	46.5
Z05M3	0.35–0.65	−18.6	9530	10.46 ± 0.22	0.522	−19.60	0.50	0.0135	47.4
Z05M4	0.35–0.65	−19.1	6012	6.60 ± 0.17	0.522	−19.99	0.73	0.0131	46.1
Z05M5	0.35–0.65	−19.6	3295	3.62 ± 0.12	0.528	−20.41	1.06	0.0086	30.2
Z05M6	0.35–0.65	−20.1	1627	1.79 ± 0.08	0.523	−20.80	1.52	0.0075	26.3
Z05M7	0.35–0.65	−20.6	657	0.72 ± 0.05	0.519	−21.29	2.40	0.0068	23.8
Z07M1	0.55–0.85	−17.6	33 146	23.1 ± 0.6	0.739	−19.06	0.26	0.0172	61.1
Z07M2	0.55–0.85	−18.1	24 664	17.2 ± 0.5	0.740	−19.39	0.35	0.0139	49.3
Z07M3	0.55–0.85	−18.6	16 979	11.8 ± 0.4	0.740	−19.75	0.49	0.0115	40.9
Z07M4	0.55–0.85	−19.1	10 713	7.5 ± 0.3	0.741	−20.13	0.69	0.0101	35.8
Z07M5	0.55–0.85	−19.6	6031	4.20 ± 0.20	0.740	−20.49	0.97	0.0090	32.0
Z07M6	0.55–0.85	−20.1	2811	1.96 ± 0.13	0.740	−20.92	1.44	0.0079	27.9
Z07M7	0.55–0.85	−20.6	1130	0.79 ± 0.06	0.738	−21.35	2.13	0.0069	24.4
Z09M2	0.75–1.05	−18.1	34 712	18.2 ± 0.5	0.910	−19.51	0.34	0.0170	60.0
Z09M3	0.75–1.05	−18.6	24 248	12.7 ± 0.4	0.916	−19.85	0.46	0.0137	48.5
Z09M4	0.75–1.05	−19.1	15 178	7.94 ± 0.23	0.916	−20.22	0.66	0.0115	40.6
Z09M5	0.75–1.05	−19.6	8413	4.40 ± 0.15	0.917	−20.59	0.93	0.0103	36.3
Z09M6	0.75–1.05	−20.1	3830	2.00 ± 0.09	0.916	−21.00	1.35	0.0089	31.5
Z09M7	0.75–1.05	−20.6	1387	0.73 ± 0.04	0.901	−21.39	1.95	0.0083	29.5
Z11M3	0.95–1.25	−18.6	23 773	10.29 ± 0.19	1.100	−20.02	0.47	0.0186	65.3
Z11M4	0.95–1.25	−19.1	15 745	6.82 ± 0.12	1.108	−20.34	0.63	0.0152	53.1
Z11M5	0.95–1.25	−19.6	8677	3.76 ± 0.07	1.110	−20.70	0.88	0.0130	45.5
Z11M6	0.95–1.25	−20.1	3868	1.67 ± 0.05	1.111	−21.10	1.26	0.0114	40.0
Z11M7	0.95–1.25	−20.6	1285	0.56 ± 0.02	1.114	−21.52	1.85	0.0103	36.0
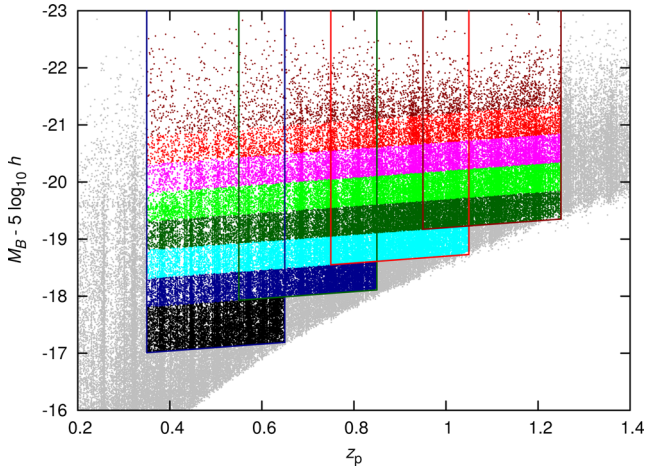
Figure 3. Selection of samples in the absolute *B*-band magnitude M_B versus photometric redshift diagram. The different coloured dots show the eight magnitude cuts, while the lines mark the boundaries of our redshift bins. See the main text and Table 2 for the details of the sample selection.

sampled by the ALHAMBRA filters (including the NIR filters) for the whole redshift range studied. Moreover, as this same band (or similar ones as *g*) is used for luminosity selection by other surveys at these redshifts, this will allow for more direct comparisons. The M_B for each object is obtained as a by-product of the photo-*z* estimation, and includes the appropriate *K*-correction at the best value of z_p. We use 'threshold samples', meaning that we will impose a faint luminosity threshold, but not a bright limit. In this way, we obtain approximately volume-limited samples, but also we can study the luminosity dependence of clustering, and its evolution. Following Meneux et al. (2009) and Abbas et al. (2010), we apply an absolute magnitude threshold depending linearly on redshift as

$$M_B^{\rm th}(z) = M_B^{\rm th}(0) + A z_p, \qquad (1)$$

in order to follow the evolution of samples corresponding approximately to the same galaxy population. The value of the constant *A* characterizes the typical luminosity evolution of the galaxies in the catalogue. We use here a value of $A = -0.6$, which we selected to produce samples with similar number density across the whole redshift range. This value is also similar to the observed evolution of the typical luminosity parameter M^* derived from luminosity function studies at similar redshifts (Ilbert et al. 2005; Zucca et al. 2009).

We show in Fig. 3 the actual cuts made in the redshift–absolute magnitude plane to define our samples, and list the properties of all the samples used in Table 2. We estimate the error in the mean number density $\bar{n}$ of each sample using a block bootstrap method based on the seven independent fields. For each sample, we compute the typical z_p error $\sigma_z/(1 + z)$ as described in Section 2.2, and the line-of-sight distance that corresponds to this uncertainty, $r(\sigma_z)$, measured at the median redshift $z_{\rm med}$ of the sample. We also measure the median absolute luminosity, $M_B^{\rm med}$, and express it in terms of the typical luminosity parameter L^* at $z_{\rm med}$. We compute $L^*(z)$ from a linear fit to the results of Ilbert et al. (2005).

2.4 Mock catalogues

To test our methods for clustering and error estimation, and to provide a test bench for future ALHAMBRA studies, we use a set of mock catalogues, based on the Millennium dark matter simulation

(Springel et al. 2005). We populate the dark matter haloes with galaxies using the Lagos et al. (2011) version of the semi-analytic galaxy formation model GALFORM (Cole et al. 2000). In addition to other physical parameters, we compute the photometry for each of the galaxies in the model using the 24 ALHAMBRA filters, including the synthetic *I* band and, for completeness, also using the five SDSS broad-band filters *ugriz*. A light-cone is built from the simulation's snapshots up to $z = 2$, reproducing the photometric depth of the survey. In order to properly model the evolution of structures along the line of sight, the galaxy positions are interpolated between snapshots. The procedure used to generate the light-cone mocks is presented in detail in Merson et al. (2013). The cosmological model used for the mocks is set by that of the Millennium simulation, which uses the parameters $\Omega_M = 0.25$, $\Omega_\Lambda = 0.75$, $\sigma_8 = 0.9$. We will use these parameters when doing tests with the mocks in Appendices A1 and B.

We generate a 200 deg² light-cone, which is divided in 50 non-overlapping mock ALHAMBRA realizations. Each of these realizations reproduces the ideal geometry of the full survey, containing eight fields covering 0.5 deg² each, for a total of 4 deg² per realization. The fields in each realization are as separated as possible within our light-cone geometry. Each field is formed by two strips of 15 arcmin × 1°, separated by a 15 arcmin gap, approximately reproducing the geometry of the ALHAMBRA fields, as described in Section 3.1.

To simulate the photometric redshifts for the galaxies in the mock we proceeded as follows. We first use the original rest-frame photometry and spectroscopic redshifts in the mock to assign to each galaxy a spectral type from the same BPZ template library used to estimate photo-*z* in the real data.[5] Then, we measure consistent ALHAMBRA photometry for these spectral types by using the ALHAMBRA filter curve response. Finally, we estimate the photometric redshifts together with the spectral types and absolute magnitudes associated with the previous photometry by running BPZ in normal mode. These photometric redshifts are found to be very realistic as their performance is very similar to those obtained for real data, although with a somewhat larger uncertainty (∼30 per cent). All the details can be found in Ascaso et. al (in preparation).

3 MODELLING THE SELECTION FUNCTION

To study the clustering of the galaxies in a survey, it is crucial to understand and to model its selection function. In this work, we separate the angular and radial parts of the selection function, with our angular selection function (or 'mask') defining the geometry of the survey on the sky. We assume a uniform depth inside the mask, as the catalogue considered does not reach the photometric limit of the survey.

3.1 Angular selection mask

The angular selection mask is defined in the first instance by the coverage of the ALHAMBRA survey. It consists of independent fields of ∼0.5 deg² each, with a specific geometry set by the configuration of the detectors in the optical camera used, LAICA. The camera has four 15.5 × 15.5 arcmin² detectors, distributed in a square leaving a space of 13.6 arcmin between them. Each of the ALHAMBRA fields consists of two pointings made with this configuration, resulting in two strips of 15.5 × 58.5 arcmin² with a gap of 13.6 arcmin

---

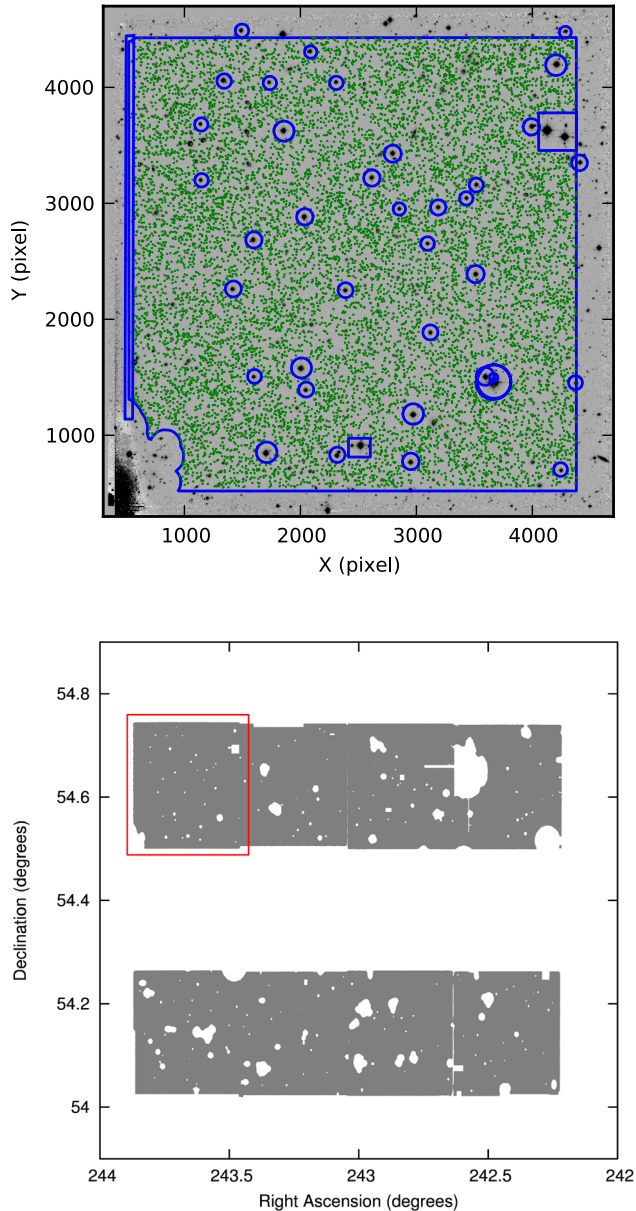[5] We do this assignment running BPZ with the ONLY_TYPE option.

Figure 4. Illustration of the ALHAMBRA angular mask for field ALH-7. Top: synthetic *I*-band image for one of the eight frames in the field, showing an area of $\sim 16 \times 16\,\text{arcmin}^2$. Green dots mark the position of the objects included in the catalogue, and the blue lines show the limits of the angular selection mask. Bottom: angular mask for the ALH-7 field. The shaded area corresponds to the regions of the survey that are included in the calculations. The red rectangle marks the area shown in the top image.

between them (see bottom panel of Fig. 4). For this work, fields ALH-4 and ALH-5 correspond to only one pointing each, and thus are formed by four disjoint $15.5 \times 15.5\,\text{arcmin}^2$ frames.

Based on that geometry, we define a set of masks describing the sky area which has been reliably observed. We start with the *flag images* described in Molino et al. (2014) that give information on the areas in which the detection of the objects in the synthetic *I*-band images was performed. They exclude areas with low exposure time (less than 60 per cent of the maximum in each frame), which mainly correspond to regions next to the borders of each frame, or corresponding to large saturated stars.

To avoid possible variations in depth, which could potentially introduce a spurious clustering pattern, we remove some additional regions from the survey area, taking a conservative approach. We mask out regions around bright stars, using the Tycho-2 catalogue (Høg et al. 2000). The masked regions are circles of radius 33 arcsec centred on each star. For the brightest stars ($V < 11$), we extend this radius to 111 arcsec. We define these radii by observing the typical maximum extension of the stellar haloes in the *I*-band detection images. Furthermore, we select objects showing saturated detections in the ALHAMBRA catalogues (using the SATUR_FLAG parameter; see Molino et al. 2014 for details), and mask a region around each of them with a radius twice that of the object itself.

Finally, we mask by hand some obvious defects in the image (typically extended stellar spikes), and some small overlap between contiguous frames. The latter is needed to avoid double-counting objects from the overlap regions when computing the clustering for the combined field. To avoid position-dependent differences in the photo-z quality we mask by hand regions which present bad photometric quality in at least three of the ALHAMBRA bands (but not necessarily in the *I* band used for detection). This uses the IRMS_OPT_FLAG and IRMS_NIR_FLAG parameters in the catalogue (see Molino et al. 2014, for details).

We defined and combined the different masks using the MANGLE[6] software (Hamilton & Tegmark 2004; Swanson et al. 2008), which allows for an easy manipulation of angular masks, and for some additional routines like generating random catalogues. These angular masks will be publicly available from http://www.alhambrasurvey.com/. Fig. 4 illustrates the resulting mask for ALH-7.

The total effective area after applying this mask is $A_{\text{eff}} = 2.381\,\text{deg}^2$, distributed over the different fields as shown in Table 1. Overall, this procedure masks an additional ~ 15 per cent of the area not yet masked by the original *flag images*. This explains the difference in area between this work and Molino et al. (2014).

3.2 Radial selection function

We model the radial selection function for our different samples directly using the observed number density of galaxies as function of comoving distance (or, equivalently, redshift), $n(d)$. We show in Fig. 5 the number density of our different samples selected in luminosity (solid lines), measured using a smoothing length of $200\,h^{-1}\,\text{Mpc}$. Given our redshift-dependent luminosity cut, the number density for each of the samples is approximately constant over the redshift range considered, as expected for nearly volume-limited samples.

However, apart from the small-scale variations due to the presence of structures, we observe some long-range variations in $n(d)$. We assume the latter are part of our selection function, and model them by fitting a third-order polynomial to $n(d)$ over the full range spanned by each of the samples. This model is smooth enough not to include possible variations in $n(d)$ due to LSSs, to prevent a systematic underestimation of the clustering signal.

We use this smooth model for our clustering measurements as described below. However, we performed some tests assuming either a model with constant $n(d)$, or using directly the measured $n(d)$ as our radial selection. Our results do not change significantly in either case.

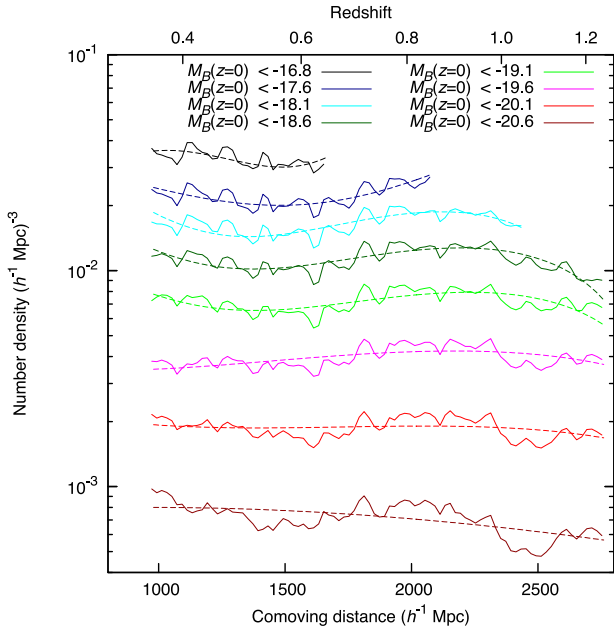[6] http://space.mit.edu/~molly/mangle/

Figure 5. Number density as function of comoving distance (or, equivalently, redshift) for our different cuts in absolute magnitude. We show the function directly measured from the data with a smoothing length of $200\,h^{-1}$ Mpc (continuous lines) and our third-order polynomial fit (dashed lines) in each case. Lines from top to bottom correspond to samples with fainter to brighter luminosity cuts.

One particularity of the radial density of ALHAMBRA as measured here is the presence of a series of regularly spaced 'peaks'. They can be seen more clearly in Fig. 5 for the fainter samples (higher n), or as a series of vertical 'strips' in the distribution of galaxies in Fig. 3. The presence of these peaks is the consequence of using only the best value z_p of the photometric redshift estimate for each galaxy, instead of the full probability density function $p(z)$ (Benítez 2000). We tested whether this issue could introduce any systematic bias in our measurements by creating a new 'realization' of the photometric redshifts: we assigned to each galaxy a new value of z_p drawn from a Gaussian distribution centred at the original value, and with a width given by the quoted error. Additionally, we randomly selected 5 per cent of galaxies to be 'outliers', and assigned them a random value of z_p within the studied range. We computed the projected correlation function for our samples using this new 'realization', and obtained only small changes contained within the quoted errors. We therefore conclude that the presence of these peaks in $n(d)$ does not significantly bias our results.

4 THE PROJECTED CORRELATION FUNCTION CALCULATION IN PHOTOMETRIC REDSHIFT CATALOGUES

The two-point correlation function $\xi(r)$ measures the excess probability of finding two points separated by a vector r compared to that probability in a homogeneous Poisson sample (Peebles 1980; Martínez & Saar 2002). If the point process considered is homogeneous and isotropic, the correlation function can be expressed simply in terms of the distance between the points, i.e. $r \equiv |r|$. However, this is not the case when studying a sample from a redshift galaxy survey. Although the galaxy distribution is intrinsically isotropic, the way in which it is measured is not, as the line-of-sight component of each position is derived from the observed redshift.

A way around this issue is the use of the projected correlation function, $w_p(r_p)$, first introduced by Davis & Peebles (1983) to deal with the redshift-space effects present in spectroscopic samples (Kaiser 1987; Hamilton 1998). As shown in Arnalte-Mur et al. (2009), this same approach can be used to deal with samples of photometric redshifts, and we use it in this paper. In this approach, we first separate the redshift-space distance between any pair of galaxies in two components: parallel (π) and perpendicular (r_p) to the line of sight.[7] We compute the correlation function as function of these components, $\xi(r_p, \pi)$, and define the projected correlation function $w_p(r_p)$ as

$$w_p(r_p) \equiv 2 \int_0^{+\infty} \xi_s(r_p, \pi)\,\mathrm{d}\pi. \tag{2}$$

We estimate $\xi(r_p, \pi)$ following Landy & Szalay (1993). We first generate an auxiliary random Poisson process following the same selection function as our sample, as defined in Section 3. We compute, for a given bin in the distance components (r_p, π), the number of pairs in our galaxy catalogue (DD), in our random catalogue (RR) and the number of crossed pairs between both catalogues (DR). The correlation function is estimated as

$$\xi(r_p, \pi) = 1 + \left(\frac{N_R}{N_D}\right)^2 \frac{DD(r_p, \pi)}{RR(r_p, \pi)} - 2\frac{N_R}{N_D}\frac{DR(r_p, \pi)}{RR(r_p, \pi)}, \tag{3}$$

where N_D is the number of galaxies in our sample, and N_R is the number of points in the auxiliary random catalogue. In this work, we always fix $N_R = 20N_D$. We tested that our results do not change if we increase the number of random points used to $N_R = 50N_D$.

The projected correlation function defined in equation (2) does not depend on the line-of-sight component of the separation π and thus, to first order, is not affected by the uncertainty on the photometric redshift determination. However, in a real survey, we cannot use this definition, as we cannot calculate the integral in equation (2) up to infinity. We calculate instead

$$w_p(r_p, \pi_{max}) \equiv 2 \int_0^{\pi_{max}} \xi_s(r_p, \pi)\,\mathrm{d}\pi, \tag{4}$$

which introduces a bias in the result, which is now dependent on the redshift-space effects. The upper limit π_{max} has to be chosen in each case with the aim of minimizing this bias, but also of avoiding the introduction of too much additional noise in the calculation.

In Appendix A we explore this issue in detail for the case of photometric redshift surveys like ALHAMBRA, using both an analytical model including Gaussian photo-z errors and the full mock catalogues described in Section 2.4. We study the bias introduced by the finite integration limit, and calculate the minimum value of π_{max} needed given the statistical uncertainty in our measurements. Accounting for this study, we use throughout $\pi_{max} = 200\,h^{-1}$ Mpc, which is appropriate for the ALHAMBRA samples considered here. As a further test, we study the change of our results with π_{max} in Appendix A2. Hereafter, we omit the explicit dependence of w_p on the value of π_{max}, and just write $w_p(r_p) \equiv w_p(r_p, \pi_{max} = 200\,h^{-1}$ Mpc).

4.1 Integral constraint

The integral constraint (Peebles 1980) is a bias in the estimation of the correlation function due to the use of a finite volume. It is related

---

[7] Taking s_1 and s_2 to be the position vectors of the two galaxies, these components are defined as $\pi \equiv |s \cdot l|/|l|$ and $r_p \equiv \sqrt{s \cdot s - \pi^2}$, where $s \equiv s_2 - s_1$ and $l \equiv s_2 + s_1$.

to the fact that the correlations are measured with respect to the mean density of the sample considered (the particular survey) instead of with respect to the global mean (that of the parent population). We can derive the effect of this constraint on w_p based on that of the three-dimensional correlation function ξ. When ξ is measured using an estimator such as that of equation (3), it can be shown that the bias introduced by the integral constraint is given, at first order, by (Bernardeau et al. 2002; Labatie et al. 2012)

$$\xi(\mathbf{r}) = \xi^{\text{true}}(\mathbf{r}) - K, \tag{5}$$

where

$$K \equiv \frac{1}{V^2} \int_V \int_V \mathrm{d}^3\mathbf{r}_1 \mathrm{d}^3\mathbf{r}_2 \xi^{\text{true}}(\mathbf{r}_2 - \mathbf{r}_1), \tag{6}$$

and V is the volume of the survey. Using equation (4) this translates into a bias on the estimated projected correlation function $w_p(r_p, \pi_{\max})$ which depends also on $\pi_{\max}$:

$$w_p(r_p, \pi_{\max}) = w_p^{\text{true}}(r_p, \pi_{\max}) - 2K\pi_{\max}. \tag{7}$$

To correct the measured values of w_p for the integral constraint using equation (7), one needs to know the true underlying correlation function. Here we choose an alternative approach, by including the integral constraint correction in the models we fit to the data. In practice, we follow Roche et al. (1999) and make use of the auxiliary Poisson catalogue to compute numerically the double integral in equation (6) as

$$K \simeq \frac{\sum_i RR(r_i)\xi^{\text{model}}(r_i)}{\sum_i RR(r_i)} = \frac{\sum_i RR(r_i)\xi^{\text{model}}(r_i)}{N_R(N_R - 1)}, \tag{8}$$

where we use the same notation as in equation (3), and where the sum is over bins in distance extending up to the largest separations in the survey. In all cases, however, we check that the value of the integral constraint correction is small compared with the errors on w_p (as can be seen in Fig. 6), so our results are not sensitive to the details of the estimation of K.

4.2 Error estimation

To estimate the statistical error on our $w_p(r_p)$ measurements, we use the standard block bootstrap method (see e.g. Norberg et al. 2009), making use of the fact that the survey consists of seven totally independent fields. We generate $N_b = 1000$ bootstrap realizations for each calculation, using the fields as bootstrap regions. Each of these realizations is created by selecting seven fields at random, allowing for repetition. We then compute the projected correlation function for each bootstrap realization using equations (3) and (4). We obtain the error of w_p at each bin in r_p as the standard deviation of the measurements from the N_b bootstrap realizations. To account for the covariance between bins in r_p when fitting a model to our data, we repeat the χ^2 fitting for the N_b realizations, using only the derived diagonal errors. Our estimate of the error on each model parameter is then the standard deviation of the best values obtained for the N_b realizations.

We test in Appendix B this error estimation and model fitting procedure for the case of ALHAMBRA using the mock galaxy catalogues described in Section 2.4. We show that it produces an unbiased estimate of the galaxy bias and of its uncertainty.

We also compared our bootstrap error estimate with the standard jackknife method (see e.g. Norberg et al. 2009). We obtained that the error on $w_p(r_p)$ estimated using both methods is consistent for $r_p \gtrsim 1\,h^{-1}$ Mpc. For $r_p \lesssim 1\,h^{-1}$ Mpc the jackknife method slightly underestimates the error with respect to the bootstrap estimate.

5 CORRELATION FUNCTIONS FOR ALHAMBRA SAMPLES

We show the resulting projected correlation functions $w_p(r_p)$ for the different samples selected in redshift and luminosity in Fig. 6. When comparing the results for samples at a given redshift bin we see clearly the effect of segregation by luminosity: bright galaxies are systematically more clustered than faint ones. This effect can be readily seen in all four redshift bins. Moreover, we see that all results show approximately a power-law behaviour for scales $r_p \gtrsim 0.2\,h^{-1}$ Mpc. We focus here on these scales, and leave the study of smaller scales for a later work.

5.1 Power-law modelling of the correlation functions

In order to study the change of the clustering properties with luminosity and redshift, we fit the obtained projected correlation function $w_p(r_p)$ of each sample using a power-law model. Following the standard practice, we assume the real-space correlation function, $\xi(r)$, is given by

$$\xi^{\text{pl}}(r) = \left(\frac{r}{r_0}\right)^{-\gamma}. \tag{9}$$

When transforming this model, using equation (2), to a model for $w_p(r_p)$, we also obtain a power law which expressed in terms of the parameters r_0 and γ above is given by

$$w_p^{\text{pl}}(r_p) = r_p \left(\frac{r_0}{r_p}\right)^{\gamma} \frac{\Gamma(1/2)\Gamma[(\gamma - 1)/2]}{\Gamma(\gamma/2)}, \tag{10}$$

where $\Gamma(\cdot)$ is Euler's Gamma function. Fitting the power-law model of equation (10) to our observed data, we can study the change of both the slope γ and the correlation length r_0 with the properties of each sample.

In practice, we modify this power-law model by adding the effect of the integral constraint described in Section 4.1. Following equation (7) and leaving explicit the dependence on the model parameters (r_0, γ), the model projected correlation function is

$$w_p^{\text{model}}(r_p|r_0, \gamma) = w_p^{\text{pl}}(r_p|r_0, \gamma) - 2\,K(r_0, \gamma)\pi_{\max}, \tag{11}$$

where $w_p^{\text{pl}}(r_p|r_0, \gamma)$ is given by equation (10), and the integral constraint term $K(r_0, \gamma)$ is obtained from equation (8) using the power-law model for the three-dimensional correlation function of equation (9). We fit the model of equation (11) to the projected correlation function measured for our different samples in the range $0.2 < r_p < 17\,h^{-1}$ Mpc. We obtain the best-fitting parameters γ, r_0 in each case using a standard χ^2 minimization method, and their error using the method described in Section 4.2.

The best-fitting models obtained are shown as solid lines in Fig. 6. The effect of the integral constrain produces a slight deviation from a straight line (in the log–log plot) at larger scales, very small compared with the errors. We plot in Fig. 7 the resulting parameters γ, r_0 for each of our redshift bins, as function of the median B-band luminosity expressed as function of $L^*(z)$.

From the bottom panel of Fig. 7 we conclude that the slope γ is approximately constant, with a value $\gamma \sim 1.75$. This is in agreement with previous studies at similar redshifts (Coil et al. 2006; Marulli et al. 2013), although Pollo et al. (2006) found significantly steeper slopes for the brightest samples. The results for r_0 shown in the top panel of Fig. 7, however, show clear evidence of luminosity segregation, as already observed qualitatively in Fig. 6. In all cases, luminous galaxies are more clustered than faint ones. However, the change of r_0 with redshift is not monotonic. While the results at
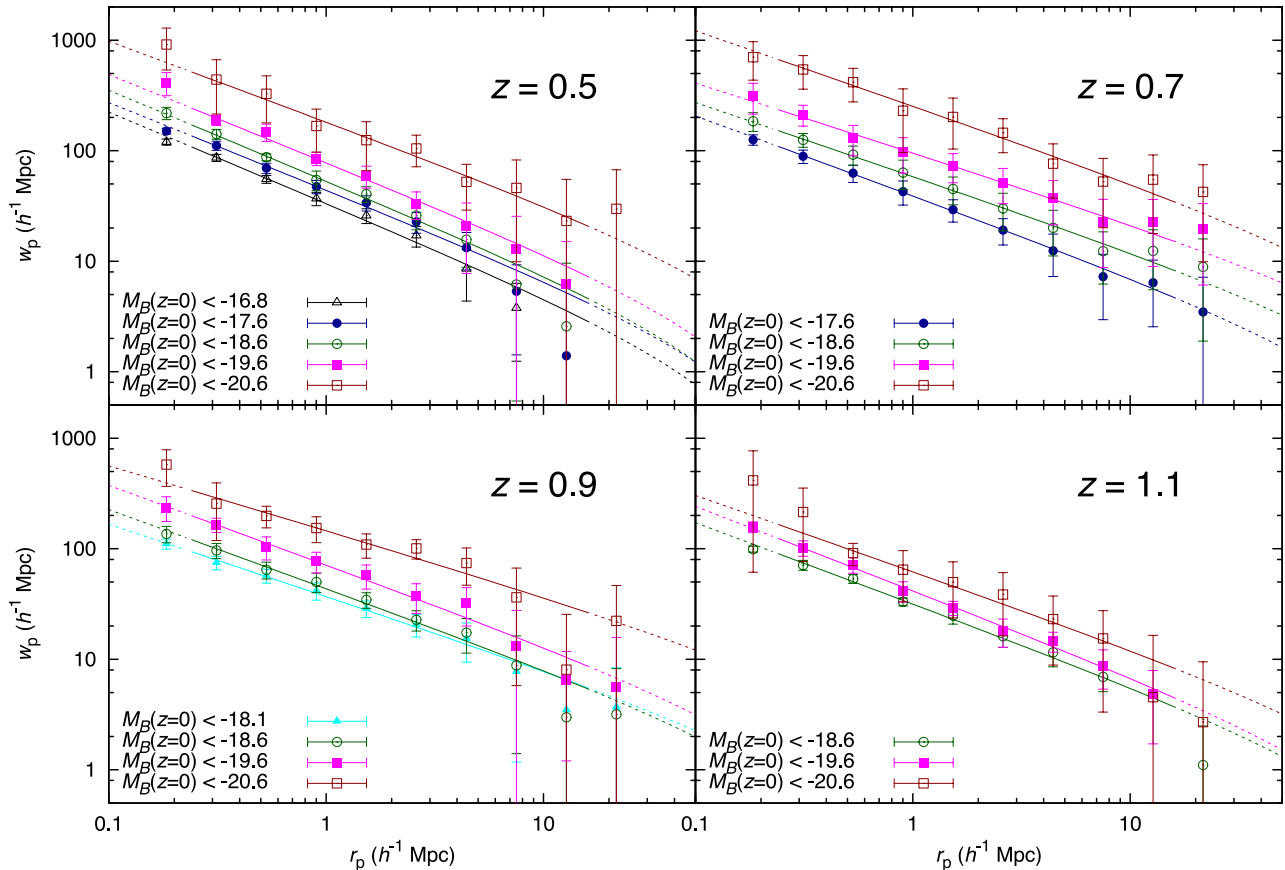
Figure 6. Projected correlation functions for the samples selected in absolute magnitude M_B and redshift (see Table 2). We omit some of the samples for clarity. The solid lines show the corresponding best-fitting power laws, according to equation (11), in the range in which the fit was done. Dashed lines show the extrapolation of these models to larger or smaller scales.

$z = 0.5$ and 0.9 are very similar, the bin at $z = 0.7$ shows a stronger clustering.

The bin at $z = 1.1$ shows a behaviour clearly different to the other three redshift bins. On one side, the r_0 values for this bin are consistently smaller than those of the lower redshift bins. On the other side, its dependence on luminosity is much weaker. However, it is difficult to interpret the results for this last bin, as there is a possible selection bias affecting it. The reason for this bias is that, for this redshift range, the rest-frame 4000 Å break is crossing the observer-frame I band used for the selection of our catalogue. This means that the selection function is changing inside the redshift bin, and in particular this will affect the selection of red passive galaxies (which we expect to show a stronger clustering). We do not study further this redshift bin in this work, but will study it in more detail in Hurtado-Gil et al. (in preparation), where we focus on the clustering as function of spectral type.

For the three bins at $z \leq 1$, we analyse the clustering properties in detail in the next sections. First, we separate the evolution of the clustering of the underlying matter density field from that of the bias of our different samples in Section 5.2. Then, we study the effect sample variance has on our results, and develop a more robust clustering measurement in Section 5.3.

5.2 Dependence of bias on luminosity and redshift

We study the bias b of our samples by comparing the observed projected correlation function w_p for each sample to that of the matter distribution at the corresponding median redshift w_p^m. We assume a simple linear model, in which bias is constant and independent of scale,

$$w_p(r_p) = b^2 w_p^m(r_p). \tag{12}$$

We restrict our study to the bias in the range $1 < r_p < 10 \, h^{-1}$ Mpc, corresponding mainly to the two-halo term of the correlation function. We leave a more detailed study using the full halo occupation distribution (HOD) formalism (Scoccimarro et al. 2001; Berlind & Weinberg 2002) for a future work. We note, however, that previous works have shown that the value of the bias obtained using our method is consistent to that using the HOD modelling (Zehavi et al. 2011).

We use a model for w_p^m based on ΛCDM and using values of the cosmological parameters consistent with the *WMAP*7 results (Komatsu et al. 2011). In particular, we use a normalization of the power spectrum $\sigma_8 = 0.816$. We obtain the matter power spectrum at each redshift using the CAMB software (Lewis, Challinor & Lasenby 2000), including the non-linear corrections of HALOFIT (Smith et al. 2003). We then Fourier transform the power spectrum to obtain the real-space correlation function $\xi(r)$ of matter, and finally obtain the projected correlation function using equation (2). We perform a χ^2 fit to the model in equation (12) as described in Section 4.2, keeping our model $w_p^m(r_p)$ fixed and with b as the only free parameter. In this case, we use for each sample the value of the integral constraint K obtained from the best-fitting power-law model, and correct the observed $w_p(r_p)$ according to equation (7).
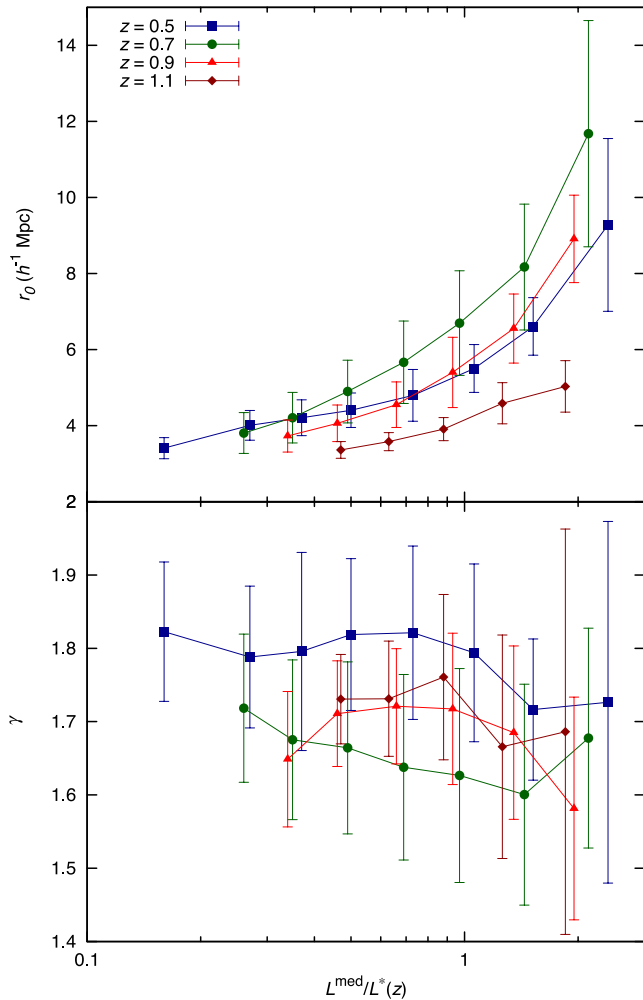
Figure 7. Parameters r_0 and γ obtained from the power-law fits for the different samples, as a function of the rest-frame *B*-band median luminosity, for each of the redshift bins.

Figure 8. Top: galaxy bias for the different samples from the fit to equation (12), as a function of the median luminosity. Bottom: galaxy bias as function of median redshift for the different luminosity cuts. We omit some of the samples for clarity. The horizontal error bars indicate the full extent of each redshift bin. The solid lines correspond to the bias of haloes above a given mass according to the model of Mo & White (2002). The label for each of these lines indicates the minimum halo mass in terms of $\log_{10}[M_h/(h^{-1}\,M_\odot)]$.

The top panel of Fig. 8 shows the value of the bias obtained as function of the median luminosity of the sample for each of the three redshift bins considered. Not surprisingly we see again the effect of luminosity segregation for all redshift bins, like for r_0 (Fig. 7). In the bottom panel of Fig. 8, we show the bias as function of redshift for a few of our luminosity-selected samples. For comparison, we show the bias of haloes of different masses according to the model of Mo & White (2002). For the samples with faintest luminosities, the evolution of bias with redshift is not significant. For the brightest samples, however, the bias does change with redshift. This evolution is not monotonic, as it seems to have at maximum at $z \sim 0.7$. Given our uncertainties, this result is not very significant. However, we study in the next section whether this behaviour is due to the effects of sample variance, and in particular to the contribution of any particular ALHAMBRA field.

5.3 Analysis of the impact of sample variance on the clustering results

The use of seven independent fields in the ALHAMBRA survey is an opportunity to study the effect of sample variance. Regarding our clustering measurements, we have already used the fact that we have data in several independent fields to estimate the errors in our
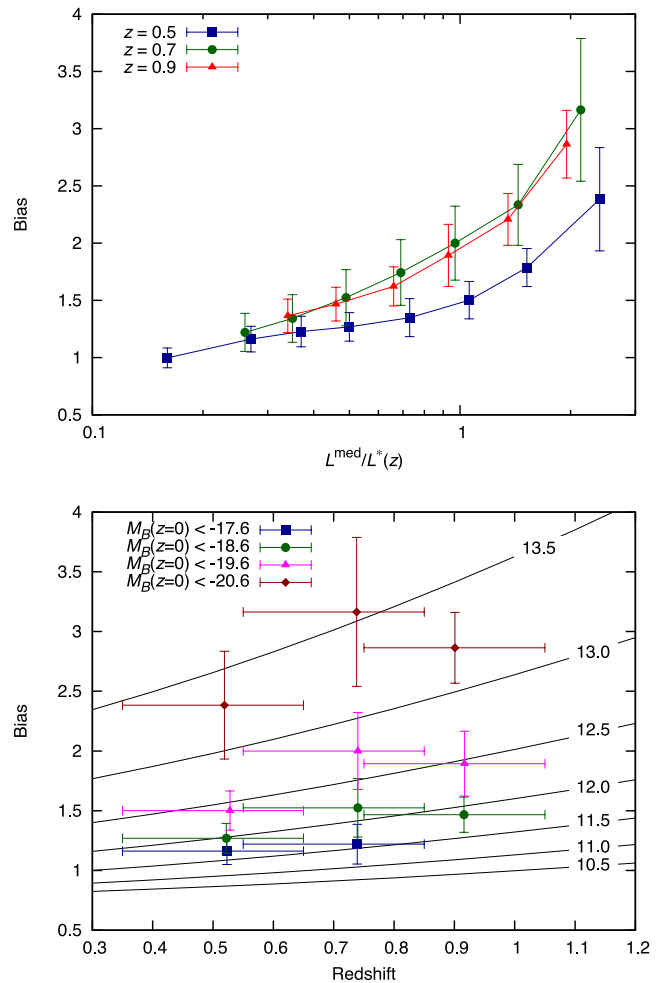
results, as explained in Section 4.2. However, this is based only on a global measure of the variance of the measurements (through the use of the bootstrap technique).

We can go one step further and study the impact of individual fields on our final measurements. Given the relatively small volume of the survey and, especially, the typical size of the fields, the presence of a large structure in one of the fields could significantly affect our clustering measurement in a given redshift bin. Similar studies have been performed with other surveys. For example, when using data from the SDSS, Zehavi et al. (2011) studied the effect on their results of including or avoiding the SDSS 'Great Wall' (Gott et al. 2005). Wolk et al. (2013) performed a similar study for the case of higher order statistics.

To study the impact of these large structures in our measurements we use the jackknife ensemble fluctuation statistic introduced by Norberg et al. (2011). This statistic is designed as an objective way of identifying 'outlier regions': those that, due to the presence of a superstructure, dominate the clustering signal of the whole survey. In the case of ALHAMBRA, it seems natural to take as

jackknife regions our seven independent fields. We present here a basic description of this statistic as used in our case for the projected correlation function, but a more detailed description can be found in Norberg et al. (2011).

For a given sample, we start by computing the projected correlation function removing from the survey a given field i, $w_p^i(r_p)$, and the corresponding rescaled quantity

$$\Delta_i(r_p) = \frac{w_p^i(r_p) - w_p^{full}(r_p)}{w_p^{full}(r_p)}, \qquad (13)$$

where $w_p^{full}(r_p)$ refers to the projected correlation function measured from the full sample. This *jackknife resampling fluctuation* $\Delta_i(r_p)$ therefore quantifies the relative change in w_p due to the exclusion of a given field. To assess the significance of this change, we define the quantity $\sigma_{tot-i}^2(r_p)$ as the rms error of this resampling fluctuations, omitting field i,

$$\sigma_{tot-i}^2(r_p) = \frac{1}{N_{fields} - 1} \sum_{j \neq i}^{N_{fields}-1} \Delta_j^2(r_p). \qquad (14)$$

In the case of the ALHAMBRA catalogue used here, $N_{fields} = 7$. We finally define the *jackknife ensemble fluctuation* δ_i as the resampling fluctuation normalized to its error

$$\delta_i(r_p) = \frac{\Delta_i(r_p)}{\sigma_{tot-i}(r_p)}. \qquad (15)$$

This is a direct measure of how significant the change in the clustering result for a given sample is when a given field i is either included or excluded. Norberg et al. (2011) define an 'outlier region' as that for which $|\delta_i| > 2.5$, where δ_i is averaged over the range of scales of interest. We adopt this same limit to define 'outlier fields' in the case of ALHAMBRA. This choice is somehow arbitrary, as full N-body simulations would be needed to test the needed value in this case, as done in Norberg et al. (2011).

We computed the jackknife ensemble fluctuation δ_i, averaged over the range $1 < r_p < 10 \, h^{-1}$ Mpc (the same range used to estimate the bias) for the samples selected by $M_B(z = 0) < -19.6$ in our three redshift bins, corresponding to $L^{med} \simeq L^*(z)$. However, as the effects we measure here are due to sample variance, we obtain consistent results when using a different luminosity cut. We show the results, for the different ALHAMBRA fields, in Fig. 9. As expected, in most cases we obtain values $|\delta_i| \lesssim 1$ corresponding to the expected variance. However, we can use the criterion explained above to identify outliers in an iterative way.

The first outlier we identify is the ALH-4 field, for which we obtain the largest value of $|\delta_i|$, $\delta_i = -5.01$ for the redshift bin centred at $z = 0.9$. Once this outlier field is identified, we exclude it from the calculation, and repeat the measurement of δ_i. Using these new values, we identify an additional outlier: the ALH-7 field, for which we now obtain $\delta_i = -3.45$ for the redshift bin centred at $z = 0.7$. The original value for this field and redshift bin, when we included also ALH-4 in the calculation, was $\delta_i = -2.74$. We repeat the process again, excluding both the ALH-4 and ALH-7 fields from the calculation, and find now in all cases values of $|\delta_i| \leq 1.73$, which we interpret as all fields being equally consistent with each other.

The most obvious outlier is the ALH-4/COSMOS field. The large negative value of δ_i obtained means that the inclusion of this field in the survey produces a very significant increase in the measured clustering for this bin. This is consistent with the fact that previous studies of clustering in the COSMOS survey at similar redshifts have obtained values significantly larger than other similar surveys
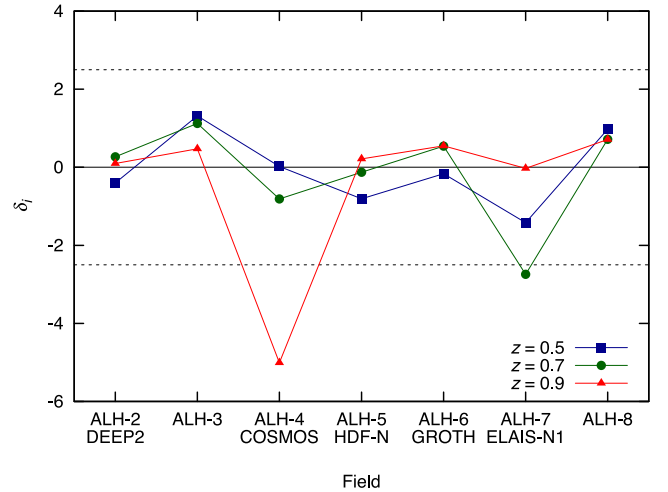

Figure 9. Ensemble fluctuation δ_i averaged over the range $r_p \in [1, 10] \, h^{-1}$ Mpc for the different redshift bins, as function of the excluded field. These results correspond to the samples selected with $M_B(z = 0) < -19.6$, for which $L^{med} \sim L^*$. The dashed lines denote our limits $|\delta_i| = 2.5$ to identify a field as an 'outlier'.

(McCracken et al. 2007; Meneux et al. 2009; de la Torre et al. 2010; Skibba et al. 2014). The excess clustering can be explained by the presence of large overdense structures in this field (Guzzo et al. 2007; Scoville et al. 2007; Kovač et al. 2010). In fact, taking into account the particular area covered by the ALH-4 field, we obtain that the four largest structures found by Scoville et al. (2007, see their table 3) are partially included in our sample. The central redshifts estimated for these structures are $z = 0.73, 0.88, 0.93, 0.71$, so all of them have substantial overlap with the redshift bin $0.75 < z < 1.05$ where we identify this field as an outlier. The particularly large overdensity of this field is also observed in ALHAMBRA. The surface density of galaxies is significantly larger in this field than in the rest, as shown in Table 1. Moreover, the redshift distribution $N(z)$ of this field shows a broad peak centred at $z \sim 0.8$ when compared to the global ALHAMBRA $N(z)$ (see fig. 32 in Molino et al. 2014).

The second 'outlier' is the ALH-7/European Large Area ISO Survey North 1 (ELAIS-N1) field. Unfortunately, this field is not as well studied as COSMOS and, to the best of our knowledge, there are no previous studies of clustering or identification of large structures at these redshifts. However, we also find a peak in the density of clusters and groups in this field at $z \sim 0.7$ using the same ALHAMBRA data set (see Ascaso et al., in preparation, for details), indicating the presence of a large structure at this particular redshift, which could explain the particularly large clustering observed here.

Fig. 10 shows the bias of our samples (measured as described in Section 5.2) as function of their median luminosity and redshift, when we completely omit from the calculation the 'outlier fields' ALH-4 and ALH-7. We can compare this figure directly to Fig. 8, where we considered the whole survey. We obtain results very similar to the whole survey for the bin centred at $z = 0.5$. This was expected from the results in Fig. 9: the low values of $|\delta_i|$ for the fields ALH-4 and ALH-7 in this case indicated that removing them would not significantly change the result. However, we see significant differences for the bins where the removed fields were 'outliers', at $z = 0.7$ and 0.9. In this case, the bias obtained is smaller now. The dependence of the bias on luminosity, however, does not change significantly except for the overall normalization. This is due to
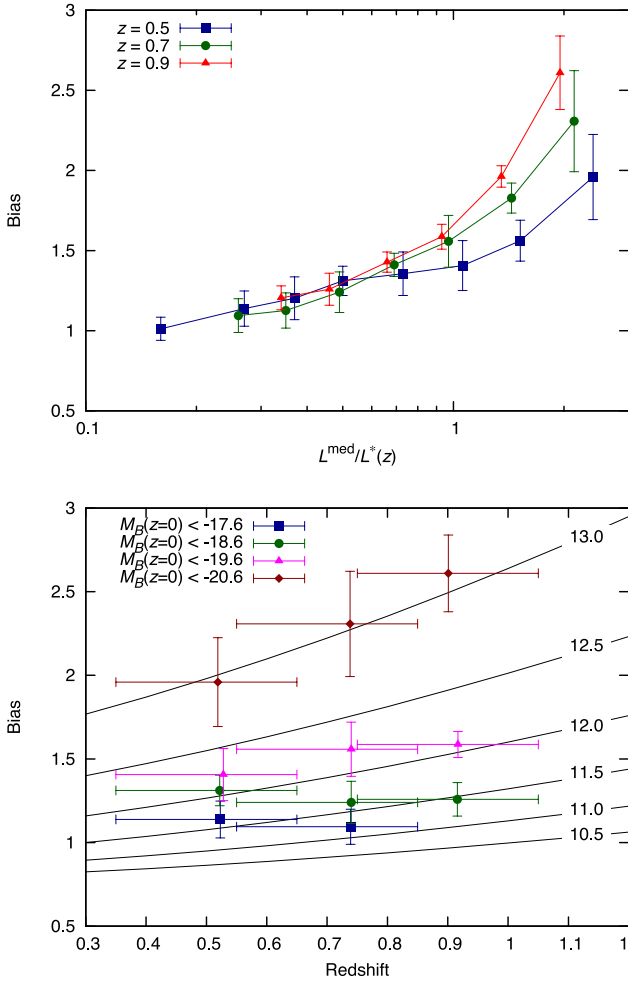
Figure 10. Both panels are identical to these in Fig. 8, for the case in which we totally omit from the calculation the 'outlier' fields ALH-4/COSMOS and ALH-7/ELAIS-N1.

the fact that, for a given redshift bin, we expect sample variance to affect in the same way all the samples regardless of the luminosity selection.

The error on the bias computed using the bootstrap method has also been greatly reduced. This was also expected: as we eliminated the greatest outliers, the variance of the remaining measurements is reduced. However, we note that the original error estimate for the full survey was also affected by the presence of the 'outlier' fields, as these imply a very non-Gaussian error distribution.

From the bottom panel of Fig. 10 we can analyse the evolution of the bias in this case. For the faintest samples we obtain now an even weaker evolution of the bias. For the brightest ones we see again a clear variation of bias with redshift, but the observed trend is somewhat different to that seen in Fig. 8. Now, for our three bins at $z < 1$, we see a roughly monotonic trend, with bias increasing with increasing redshift.

Overall, the evolution observed in Fig. 10 is similar to the bias evolution for haloes above a given mass, according to the model of Mo & White (2002). According to that model, the bias we obtain for our different samples correspond to populations of haloes with minimum masses in the range $11.5 \lesssim \log_{10}[M_{\rm h}/(h^{-1}\,{\rm M}_\odot)] \lesssim 13.0$. The bias of galaxies with $L^{\rm med} \simeq L^*$ roughly corresponds to that of a halo population with $\log_{10}[M_{\rm h}/(h^{-1}\,{\rm M}_\odot)] \gtrsim 12.2$.
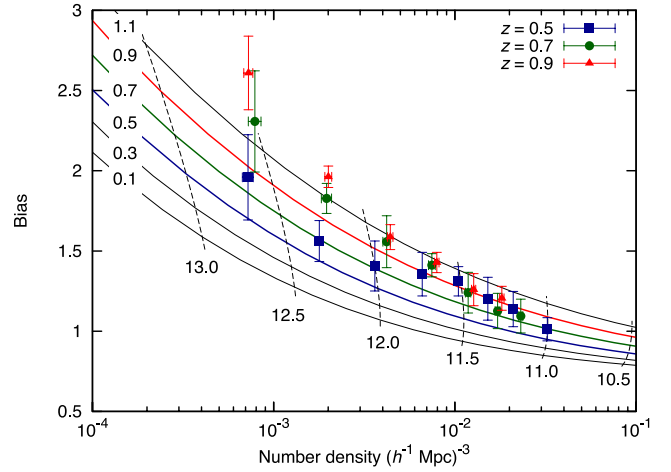
Figure 11. Galaxy bias as a function of the number density of galaxies for our different samples (points). Galaxy bias is obtained from the fit to equation (12), for the case in which we omit the 'outlier' fields ALH-4/COSMOS and ALH-7/ELAIS-N1. The lines show the prediction of the model of Mo & White (2002) for haloes above a given mass. Continuous lines show the prediction for fixed values of the redshift (indicated by the labels in the left). Dashed lines correspond to the prediction for fixed values of the minimum halo mass (indicated by the labels in the bottom, in terms of $\log_{10}[M_{\rm h}/(h^{-1}\,{\rm M}_\odot)]$). Comparing these predictions for haloes to the observed values, we obtain that the typical mean occupation numbers for the ALHAMBRA galaxies are in the range $\sim$1–3.

To further investigate the relationship between our galaxy samples and the halo populations, we show in Fig. 11 the bias of our samples as a function of their number density. We compare our results to the prediction for populations of haloes above a given minimum mass from the model of Mo & White (2002), shown as the continuous (for fixed redshift) and dashed (for fixed minimum mass) lines in the plot. We can estimate roughly the halo occupation number (i.e. the mean number of galaxies per halo) for a given sample by comparing its number density to that of the halo population at the same redshift and with similar bias. For the different ALHAMBRA samples, we obtain that the occupation numbers are typically in the range $\sim$1–3, although with a large uncertainty due to our uncertainty in the bias measurement.

5.4 Comparison to previous results from other surveys

We compared our results with previous studies using the largest galaxy surveys to date covering similar redshifts. Coupon et al. (2012) studied galaxy clustering in the range $0.2 < z < 1.2$ using data from CFHTLS-Wide,[8] a broad-band photometric survey covering $\sim$155 deg^2. The bias was derived in each case by fitting a HOD model to the angular correlation function of each sample. Marulli et al. (2013) measured the clustering using spectroscopic data from VIPERS[9] covering $\sim$15 deg^2, in the range $0.5 < z < 1.1$. They measured the bias from the measured projected correlation function in the same way as we do here (equation 12), and showed that their results were in rough agreement with other (smaller) spectroscopic surveys at similar redshifts such as DEEP2 (Coil et al. 2006) and VVDS (Pollo et al. 2006). In both cases, the depth of the

[8] http://www.cfht.hawaii.edu/Science/CFHLS/
[9] http://vipers.inaf.it/

data used was $i < 22.5$. We note that the area covered by VIPERS is a subset of that covered by CFHTLS-Wide. The ALH-6 field also overlaps with CFHTLS-Wide.

We also included in this comparison the results in the range $0.5 < z < 1.0$ of Skibba et al. (2014) using data from PRIMUS. PRIMUS (Coil et al. 2011) is a survey which uses a low-resolution spectrograph resulting in a typical redshift precision of $\sigma_z/(1 + z) = 0.005$ to a depth of $i < 23$. The data cover five independent fields (including the COSMOS field) covering a total[10] of $7.80 \, \mathrm{deg}^2$. Skibba et al. measured the bias of different samples selected in redshift, luminosity and colour using the projected correlation function in the same way as we describe above (equation 12).

In Fig. 12 we plot the bias obtained in our different redshift bins as a function of the threshold luminosity L^{th} used to select the different samples in ALHAMBRA, CFHTLS-Wide, VIPERS and PRIMUS. L^{th}/L^* is measured at the median redshift of the sample, taking into account the use of different selection parameters A in equation (1). We note that L^{th} refers to the B band in the case of ALHAMBRA and VIPERS, and to the g band in the case of CFHTLS-Wide and PRIMUS. In each case, we compare the ALHAMBRA results with the CFHTLS-Wide results for the bin centred at the same redshift. As Marulli et al. (2013) used bins centred at different redshifts, we plot in each case the one or two closest bins to the ALHAMBRA one. In the case of PRIMUS, the actual redshift range of each sample is slightly different with mean redshifts in the range 0.60–0.74, so we plot their results in the central panel. In each case, we renormalize the bias by the value of σ_8 considered. Changes in bias due to other differences in the cosmology used are much smaller than our errors. For reference, we also plot as a continuous line the relation derived for low redshifts by Zehavi et al. (2011) from the SDSS data, which is very similar to that obtained by Norberg et al. (2001) from the 2dFGRS. We plot the ALHAMBRA results both for the full survey (dashed lines) and for the case in which we have removed the two 'outlier fields' ALH-4 and ALH-7 (solid lines).

We obtain a good agreement between our results and both the CFHTLS-Wide and VIPERS ones, especially considering the significantly smaller area surveyed by ALHAMBRA. When looking at the $z = 0.7$ and 0.9 bins, we see how the result obtained after omitting the outlier fields is in better agreement with the other data than the original results. This confirms the idea that using the jackknife ensemble fluctuation to identify outlier regions results in a good measurement of the typical clustering properties (bias in this case) of the samples. We point out that the comparison presented here was performed only after the full analysis of ALHAMBRA data was finished, so it did not influence the design of the method described in Section 5.3.

Our results are also in very good agreement with the PRIMUS results. We note that PRIMUS obtained slightly larger values of the bias than CFHTLS-Wide or VIPERS, and they attributed this fact to the presence of the COSMOS field in their sample. This is compatible with their results lying between our results with and without the outliers fields included.

We note, however, that the dependence of bias on luminosity appears to be slightly steeper in ALHAMBRA than in previous works. This is noticeable at the bright end of the $z = 0.9$ bin. It is difficult to assess the significance of this discrepancy, as our bias error estimate is affected by the removal of the 'outlier'

---

[10] This corresponds to the study at $z > 0.5$, where they excluded two additional fields from their analysis. The total area covered by the survey is $9.05 \, \mathrm{deg}^2$.
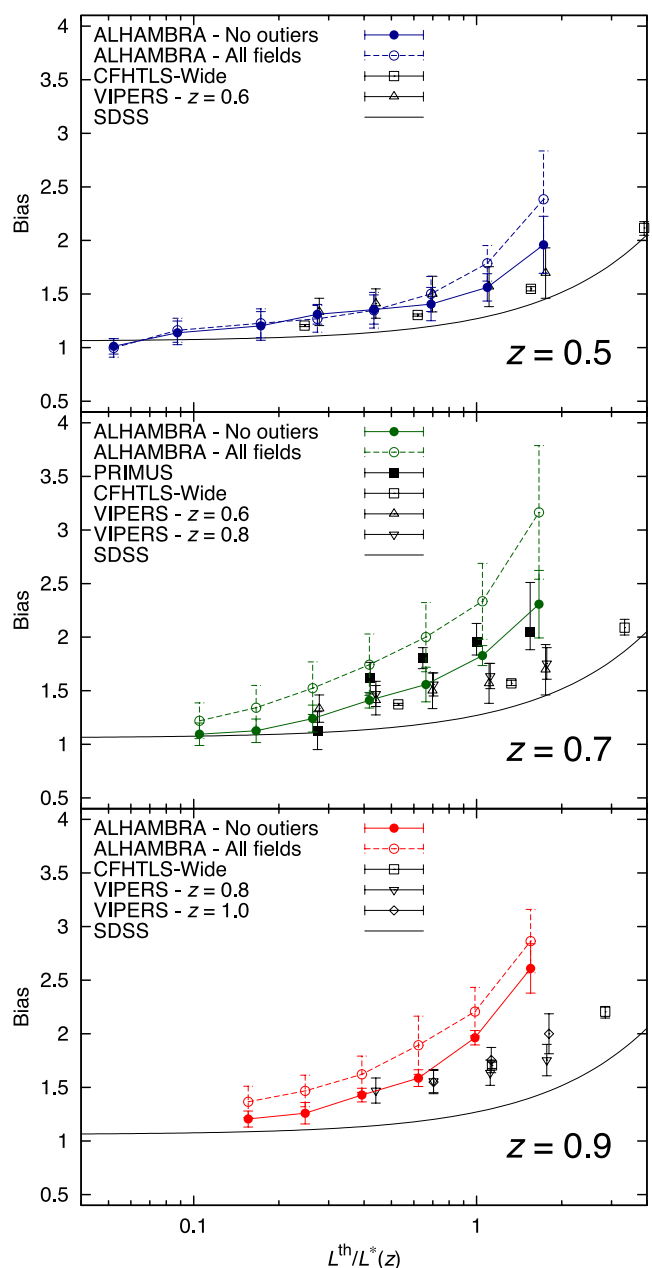

Figure 12. Galaxy bias comparison between ALHAMBRA (this work), VIPERS (Marulli et al. 2013), CFHTLS-Wide (Coupon et al. 2012) and PRIMUS (Skibba et al. 2014). The solid line in each panel corresponds to the low-redshift SDSS results of Zehavi et al. (2011). The bias measurements have been renormalized to the fiducial value $\sigma_8^{fid} = 0.816$ used in this work.

fields, and the different measurements are highly correlated. With these caveats in mind, we estimate that the discrepancy for the most extreme case is at the $\lesssim 2\sigma$ level. Given its small area, the ALHAMBRA survey is not designed to provide an accurate measurement of low number density samples, nor is the error analysis necessarily adequate for them either. The lowest number density samples (i.e. bright galaxies) require large survey areas to be properly estimated.

Fig. 12 shows the complementarity between the different surveys covering this redshift range to study the dependence of galaxy bias on luminosity and redshift. Large area surveys such as

CFHTLS-Wide and VIPERS can measure very accurately the bias of relatively bright samples, $L \gtrsim 0.3L^*(z)$, thus setting the overall normalization of the $b(L)$ relation at each redshift. Despite its smaller volume, ALHAMBRA can extend this relation to luminosities $\simeq 1.5$ mag fainter, with our study of the outliers showing that the result is robust to sample variance, except for the overall normalization. This larger luminosity range in ALHAMBRA allows us to see clearly the transition from a nearly flat relation at the faint end to a steep one at the bright end.

6 DISCUSSION AND CONCLUSIONS

In this work, we have studied the clustering of galaxies in the ALHAMBRA survey and its dependence on luminosity and redshift, in the range $0.35 < z < 1.25$. To this end, we have used the projected correlation function $w_p(r_p)$, taking into account the uncertainties associated with the use of photometric redshifts, following the method described in Arnalte-Mur et al. (2009). We have compared the measured $w_p(r_p)$ to the prediction from our fiducial ΛCDM model to estimate the bias for the different samples selected in redshift and luminosity. We also used the method introduced in Norberg et al. (2011) to study the effect on the clustering measurements of superstructures located in particular ALHAMBRA fields.

The use of the projected correlation function for the case of high-quality photometric redshifts was tested in Arnalte-Mur et al. (2009) using a simulated halo catalogue. Here, we have tested the method using more realistic galaxy mock catalogues (Appendix B), and have applied it to real data from the ALHAMBRA survey. We obtain results that are consistent with larger-area surveys (Section 5.4), and in particular the VIPERS spectroscopic survey, while reaching 1.5 mag deeper. This confirms the reliability of the method, and shows that surveys using a large number of medium-band filters can provide very useful data sets for the study of galaxy clustering. In addition to further results from ALHAMBRA, this indicates good prospects for the planned Javalambre-Physics of the Accelerating Universe Astrophysical Survey[11] (J-PAS; Benítez et al. 2009a) and Physics of the Accelerating Universe[12] (PAU; Castander et al. 2012) surveys, which will use a similar technique covering larger cosmological volumes.

One of the main characteristics of the ALHAMBRA survey is the mapping of eight independent fields in the sky (although only seven are available in the current data set), which provide a useful tool to study the effect of sample variance. We have studied this issue in two complementary ways. On one side, we have used the independence of the fields to obtain a global measure of the clustering uncertainty using the block bootstrap technique described in Section 4.2. On the other side, we used the jackknife ensemble fluctuation statistic δ_i (Norberg et al. 2011) to assess the impact of particular superstructures in the clustering measurements. This method is based on measuring the clustering omitting one region (field in our case) at a time and comparing it to the global result. In this way, we have identified the fields ALH-4/COSMOS (at $z \sim 0.9$) and ALH-7/ELAIS-N1 (at $z \sim 0.7$) as 'outliers', as the inclusion or omission of each of them changes our results significantly. We therefore provide also the results for the bias of our samples when we omit these two fields from the calculation, which give a better description of the 'typical' clustering properties of the samples, as

evidenced by the comparison with the VIPERS and CFHTLS-Wide surveys.

One may want to discuss which is the 'correct' result for the bias from this work: that obtained using the full sample (Fig. 8) or that obtained omitting the outlier fields (Fig. 10). However, it is the combination of both approaches what gives a more complete view of the information about clustering contained in the survey. On one side, the results obtained after removing the outliers provide information about the typical dependence of galaxy bias on redshift and luminosity. This is confirmed by the comparison to surveys covering larger volumes, discussed in Section 5.4. On the other side, the results for the global sample show how this typical behaviour can be affected by the inclusion or omission of particular fields containing extreme superstructures. However, the relatively small number of fields covered by ALHAMBRA, and the fact that we only identify either none or one field as an outlier in each of the redshift bins, does not allow us to assess how rare these superstructures are.

Our clustering results give a detailed picture of the dependence of galaxy bias on both luminosity and redshift, summarized in Figs 10 and 12. The depth and photometric redshift reliability of the ALHAMBRA survey allow us to extend the study of the bias to fainter luminosities than previous surveys at similar redshifts. In this way, the full dependence of bias with luminosity is more clearly seen. Moreover, our results in Section 5.3 show that this dependence is reliable, and not significantly affected by sample variance. At the faint end this relation is nearly flat, up to $L^{\mathrm{med}} \simeq L^*$ for $z = 0.5$, and up to $L^{\mathrm{med}} \simeq 0.5L^*$ for higher redshifts. At brighter luminosities, the bias increases, following a dependence on L which, for $z = 0.7$ and 0.9, is significantly steeper than the relation found at low redshift by the SDSS and 2dFGRS surveys.

Regarding the evolution of bias, we see very little dependence of bias with redshift for the faint samples ($L^{\mathrm{med}} \lesssim 0.8L^*$), while the evolution is strong for the brighter samples. In the latter case, for samples with a approximately fixed number density, bias decreases with cosmic time. This behaviour is consistent with that expected from the halo model, where the bias of the more massive haloes shows much stronger evolution than that of the less massive ones, as illustrated in Figs 8 and 10.

The comparison of our results with the predicted bias of haloes according to the model of Mo & White (2002) suggests that the galaxies studied reside in haloes covering a range in mass between $\log_{10}[M_{\mathrm{h}}/(h^{-1}\,\mathrm{M}_{\odot})] \gtrsim 11.5$ (for the samples selected with $M_B(z=0) < -17.6$) and $\log_{10}[M_{\mathrm{h}}/(h^{-1}\,\mathrm{M}_{\odot})] \gtrsim 13.0$ (for the samples selected with $M_B(z=0) < -20.6$). The samples with $L^{\mathrm{med}} \simeq L^*$ ($M_B(z=0) < -19.6$) are found to correspond to haloes with mass $\log_{10}[M_{\mathrm{h}}/(h^{-1}\,\mathrm{M}_{\odot})] \gtrsim 12.2$. From the joint comparison of the bias and number density of our samples to the theoretical prediction for haloes, we obtain that the mean number of galaxies per halo is in the range ~ 1–3.

We excluded from this detailed study of the luminosity dependence of the galaxy bias the redshift bin centred at $z = 1.1$. As explained in Section 5.1, this is due to the fact that our I-band selection could be biasing the sample in that redshift range, affecting in a different way active and passive galaxies.

In this paper, we have focused the study of galaxy clustering in ALHAMBRA on the effect of luminosity segregation and evolution up to $z \sim 1$. In a companion paper (Hurtado-Gil et al., in preparation) we use this same data set to study the segregation by spectral type in a similar redshift range. We also plan to extend this work to further redshifts by the use of a NIR-selected catalogue, which will allow us to study the clustering of extremely red objects (EROs; Nieves-Seoane et al., in preparation).

REFERENCES

Abbas U. et al., 2010, MNRAS, 406, 1306
Aparicio Villegas T. et al., 2010, AJ, 139, 1242
Arnalte-Mur P., Fernández-Soto A., Martínez V. J., Saar E., Heinämäki P., Suhhonenko I., 2009, MNRAS, 394, 1631
Benítez N., 2000, ApJ, 536, 571
Benítez N. et al., 2009a, ApJ, 691, 241
Benítez N. et al., 2009b, ApJ, 692, L5
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Castander F. J. et al., 2012, Proc. SPIE, 8446, 84466D
Coil A. L., Newman J. A., Cooper M. C., Davis M., Faber S. M., Koo D. C., Willmer C. N. A., 2006, ApJ, 644, 671
Coil A. L. et al., 2008, ApJ, 672, 153
Coil A. L. et al., 2011, ApJ, 741, 8
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coupon J. et al., 2012, A&A, 542, A5
Cristóbal-Hornillos D. et al., 2009, ApJ, 696, 1554
Davis M., Geller M. J., 1976, ApJ, 208, 13
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
de la Torre S. et al., 2010, MNRAS, 409, 867
Gott J. R., III, Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463
Guo H. et al., 2013, ApJ, 767, 122
Guzzo L. et al., 2007, ApJS, 172, 254
Hamilton A. J. S., 1988, ApJ, 331, L59
Hamilton A. J. S., 1998, in Hamilton D., ed., Astrophysics and Space Science Library, Vol. 231, The Evolving Universe. Kluwer, Dordrecht, p. 185
Hamilton A. J. S., Tegmark M., 2004, MNRAS, 349, 115
Høg E. et al., 2000, A&A, 355, L27
Ilbert O. et al., 2005, A&A, 439, 863
Ilbert O. et al., 2009, ApJ, 690, 1236
Kaiser N., 1987, MNRAS, 227, 1
Komatsu E. et al., 2011, ApJS, 192, 18
Kovač K. et al., 2010, ApJ, 708, 505
Labatie A., Starck J., Lachièze-Rey M., Arnalte-Mur P., 2012, Stat. Methodology, 9, 85

Lagos C. D. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011, MNRAS, 416, 1566
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
López-Sanjuan C. et al., 2014, A&A, 564, A127
McCracken H. J. et al., 2007, ApJS, 172, 314
McCracken H. J., Ilbert O., Mellier Y., Bertin E., Guzzo L., Arnouts S., Le Fèvre O., Zamorani G., 2008, A&A, 479, 321
Madgwick D. S. et al., 2003, MNRAS, 344, 847
Martínez V. J., Saar E., 2002, Statistics of the Galaxy Distribution. Chapman & Hall/CRC, Boca Raton, FL
Martínez V. J., Arnalte-Mur P., Stoyan D., 2010, A&A, 513, A22
Marulli F. et al., 2013, A&A, 557, A17
Meneux B. et al., 2009, A&A, 505, 463
Merson A. I. et al., 2013, MNRAS, 429, 556
Mo H. J., White S. D. M., 2002, MNRAS, 336, 112
Moles M. et al., 2008, AJ, 136, 1325
Molino A. et al., 2014, MNRAS, preprint (arXiv:1306.4968)
Norberg P. et al., 2001, MNRAS, 328, 64
Norberg P. et al., 2002, MNRAS, 332, 827
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, 396, 19
Norberg P., Gaztañaga E., Baugh C. M., Croton D. J., 2011, MNRAS, 418, 2435
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Pollo A. et al., 2006, A&A, 451, 409
Roche N., Eales S. A., Hippelein H., Willott C. J., 1999, MNRAS, 306, 538
Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20
Scoville N. et al., 2007, ApJS, 172, 150
Seljak U., 2000, MNRAS, 318, 203
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Skibba R. A. et al., 2009, MNRAS, 399, 966
Skibba R. A. et al., 2014, ApJ, 784, 128
Smith R. E. et al., 2003, MNRAS, 341, 1311
Springel V. et al., 2005, Nature, 435, 629
Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391
Tegmark M. et al., 2004, ApJ, 606, 702
Wolk M., McCracken H. J., Colombi S., Fry J. N., Kilbinger M., Hudelot P., Mellier Y., Ilbert O., 2013, MNRAS, 435, 2
Zehavi I. et al., 2005, ApJ, 621, 22
Zehavi I. et al., 2011, ApJ, 736, 59
Zucca E. et al., 2009, A&A, 508, 1217

APPENDIX A: OPTIMAL VALUE OF π_{max} FOR THE ESTIMATION OF THE PROJECTED CORRELATION FUNCTION

A1 Theoretical determination of the minimum π_{max} needed

As explained in Section 4, it is not possible to estimate the integral of equation (2) without choosing a finite upper limit π_{max}, and computing instead $w_p(r_p, \pi_{max})$, as defined in equation (4). This introduces a bias that has to be accounted for in the modelling. At the same time, if we extend the measurement to large values of π where the signal-to-noise ratio of $\xi(r_p, \pi)$ is small, we would be introducing additional noise in the measurement. In this appendix, we study the relation of this bias with the photo-z errors of the catalogue used, and what is the minimum value of π_{max} needed in the case of ALHAMBRA. To this end, we use the mock catalogues described in Section 2.4, which include photo-z for the galaxies with similar properties to the real data, and a simple analytic model. In this appendix, we use the cosmological parameters

used in the creation of the mocks, $\Omega_M = 0.25$, $\Omega_\Lambda = 0.75$, $\sigma_8 = 0.9$.

From equations (2) and (4), the bias introduced by the finite integration is given by

$$\Delta w(r_p, \pi_{max}) \equiv w_p(r_p) - w_p(r_p, \pi_{max}) = 2 \int_{\pi_{max}}^{+\infty} \xi(r_p, \pi)\, d\pi.$$
(A1)

In principle, given a model for $\xi(r_p, \pi)$, one could do the same finite integration in the model, and obtain a prediction directly for $w_p(r_p, \pi_{max})$. However, in this case $w_p(r_p, \pi_{max})$ is not a real space quantity any longer, and it depends on the way in which redshift-space distortions (due to peculiar velocities and photo-z) are included in the model. If we want to avoid this and keep the statistic used as a real space quantity, we should choose a value of π_{max} such that the bias $\Delta w(r_p, \pi_{max})$ is negligible. Given the difficulties to model in detail the effect of the photo-z distribution in $\xi(r_p, \pi)$, we follow here the latter approach.

We use a galaxy sample selected in redshift and absolute magnitude from the mock catalogues with the limits $0.5 < z_p < 0.8$, $M_B - 5 \log_{10} h < -17.95$, which is similar to our sample 'Z07M1'. Following the same method described in Section 4, we obtain the correlation function $\xi(r_p, \pi)$ for the 50 realizations, and for the combined 200 deg^2 mock.

We compare the mock results with a simple analytic model obtained using the following steps. First, we obtain the matter power spectrum $P_m(k)$ at the median redshift of the sample using CAMB (Lewis et al. 2000). We then obtain the real-space galaxy power spectrum $P_g(k)$ using a simple HOD model, as described in Abbas et al. (2010). We include the large-scale redshift-space effects following Kaiser (1987) to obtain the redshift-space correlation function $\xi_s(r_p, \pi)$. Finally, we include the effect of the photometric redshifts assuming a simple model in which the redshift errors follow a Gaussian distribution. In this model, the observed correlation function is given by the convolution

$$\xi_{phot}(r_p, \pi) = \int_{-\infty}^{+\infty} \xi_s(r_p, \pi') f_{\sigma_{pw}}(\pi - \pi')\, d\pi',$$
(A2)

where $f_\sigma(x)$ is the Gaussian distribution of width σ. In this case, the width of the distribution is given by the pairwise photometric redshift uncertainty $\sigma_{pw} = \sqrt{2} r(\sigma_z)$. As explained in Section 2.3, $r(\sigma_z)$ is the comoving separation corresponding to a given photometric redshift uncertainty σ_z at the median redshift of the sample. We choose the HOD parameters of this model (including the bias) to reproduce the observed $w_p(r_p)$ of the mock.

We use both the results from the mock and the analytical model to compute the finite integration bias $\Delta w(r_p, \pi_{max})$ defined in equation (A1), and find the minimum value of π_{max} for which this bias is sufficiently small. We express this requirement in terms of the statistical error on $w_p(r_p)$, by requiring the bias to be smaller than 20 per cent of the estimated statistical uncertainty,

$$\frac{\Delta w(r_p, \pi_{max})}{w_p(r_p)} \le 0.2 \sigma_{w_p}^r(r_p),$$
(A3)

where $\sigma_{w_p}^r(r_p)$ is the relative uncertainty in the measurement of $w_p(r_p)$ in a single mock ALHAMBRA realization obtained from the dispersion of the measurements in the 50 mocks.

In Fig. A1 we plot the minimum value of π_{max} needed to fulfil condition (A3), as function of r_p, for our different models: the computation from the combined 200 deg^2 mock, and the analytic Gaussian model with different values of the photometric redshift error between $\sigma_z/(1 + z) = 0.010$ and 0.025. In the case of the
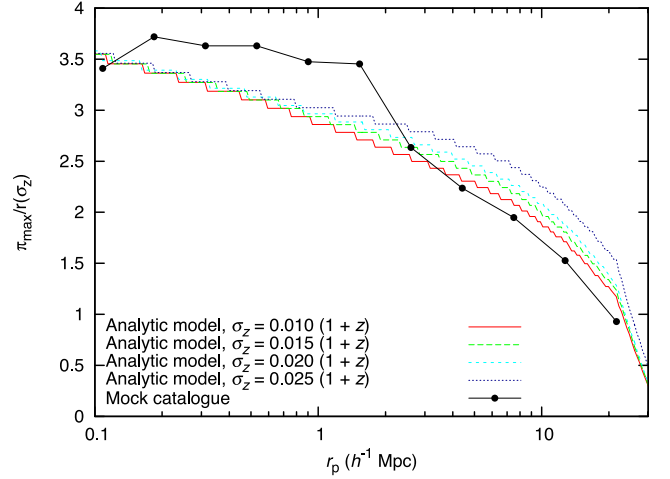

Figure A1. Minimum value of π_{max} needed to fulfil our condition (A3), as function of transverse separation r_p, for our different models. The π_{max} in each case is expressed in terms of $r(\sigma_z)$, the comoving separation corresponding to the photometric redshift error σ_z of the model, at the median redshift of the bin considered (in this case, $z_{med} \simeq 0.65$). The dashed lines correspond to the theoretical model described in the text, which include the effect of the photometric redshifts using a Gaussian distribution as shown in equation (A2). The solid line and points correspond to the measurement in the mock catalogue (combining the 50 ALHAMBRA realizations). For the selected mock sample, we estimate $\sigma_z = 0.025(1 + z)$, using the method described in Section 2.2.

mock catalogue we estimate $\Delta w(r_p, \pi_{max})$ using as the reference $w_p(r_p)$ in equation (A1) the projected correlation function obtained in real space. We estimate the typical redshift uncertainty in the mock sample using the BPZ confidence limits in the same way as explained in Section 2.2, in particular including the correction factor of 1.3, and obtain $\sigma_z/(1 + z) = 0.025$. In all cases, we plot the value of π_{max} in terms of $r(\sigma_z)$ for each particular model.

From Fig. A1 we see that overall the required value of π_{max} decreases with r_p. This is a consequence of our condition (equation A3), given by the fact that the relative error of w_p increases with r_p. Comparing the different analytical models we see that the different lines are almost coincident for $r_p \lesssim 1\,h^{-1}$ Mpc, meaning that the required value of π_{max} scales linearly with $r(\sigma_z)$. At larger scales, there is a slight deviation from this proportionality. Regarding the result obtained from the mock, we see how the non-Gaussianity of the photo-z error distribution has an impact on the observed correlation function $\xi(r_p, \pi)$. This is clearly seen at the smaller scales, $r_p \lesssim 2\,h^{-1}$ Mpc, where the needed value of π_{max} is significantly larger than that predicted by the analytical Gaussian models. Here, the value of the relative error $\sigma_{w_p}^r$ is small, so our condition (equation A3) is more stringent and the extended wings of the photo-z distribution have the effect of slowing down the convergence of the integral in equation (A1). At larger scales, $r_p \gtrsim 2\,h^{-1}$ Mpc, our condition (equation A3) is much weaker (because $\sigma_{w_p}^r$ is large), so the details of the wings of the photo-z distribution are less relevant. Actually, as the mock photo-z distribution is slightly more peaked at the centre than the equivalent Gaussian distribution, we obtain values of π_{max} slightly lower than in the analytic case.

Overall, we see that, to fulfil our condition (equation A3) over the full range of scales of interest $0.2 < r_p < 20\,h^{-1}$ Mpc, the minimum value of π_{max} needed is $\pi_{max} \simeq 3.5$–$4 r(\sigma_z)$. This result is in agreement with previous, less detailed estimates (Arnalte-Mur

et al. 2009). The particular value of $\pi_{\max}$ needed for each ALHAMBRA sample will depend on the details of the correlation function and its error, with the most significant effect being the change in the correlation function error $\sigma_{w_p^r}^r(r_p)$ appearing in equation (A3). As this error depends, at first order, on the sample volume, we repeated the calculation rescaling it according to the volumes of the actual ALHAMBRA samples used. We obtained only minor changes in the required value of $\pi_{\max}$ in all cases. We can therefore estimate the minimum value of $\pi_{\max}$ needed for each ALHAMBRA sample from the value of $r(\sigma_z)$ in each case (see Table 2). Taking $\pi_{\max} = 4r(\sigma_z)$, we obtain values in the range $\pi_{\max} \sim 100$–$280\,h^{-1}$ Mpc. As increasing the value of $\pi_{\max}$ also introduces additional noise in the measurement, a compromise should be made in deciding the actual value of $\pi_{\max}$ to use. For simplicity, we decided to use a constant value of $\pi_{\max}$ for all our samples, fixing it at $\pi_{\max} = 200\,h^{-1}$ Mpc. According to the criterion described here, this value of $\pi_{\max}$ is adequate for all our samples, except for the faintest samples of each redshift bin. However, as shown below in Appendix A2, we find the bias introduced in these cases to be still acceptable.

A2 Test of the robustness of our results with respect to changes in $\pi_{\max}$

We performed an additional test of the effect of our choice of $\pi_{\max}$ in our results. We repeated the calculation of $w_p(r_p)$ for all our samples using a substantially larger value, $\pi_{\max} = 350\,h^{-1}$ Mpc, in the integration of equation (4). We did not observe any significant difference in our results but obtained, as expected, larger uncertainties due to the additional noise included in the integration. As an example of the results obtained in this case, we plot in Fig. A2 the bias as function of luminosity for our samples obtained using different values 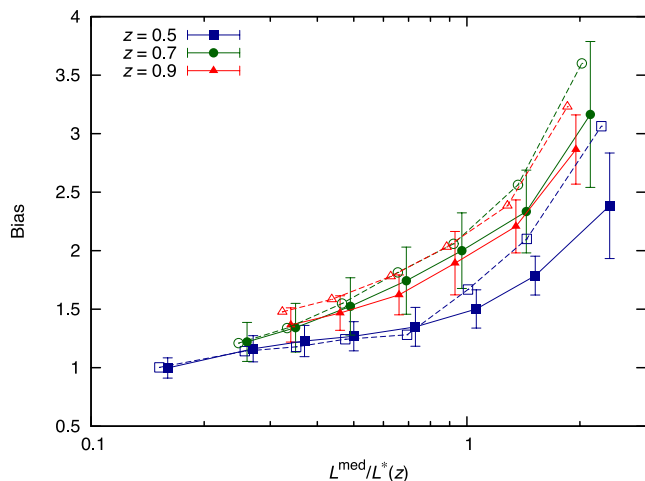of $\pi_{\max}$. The solid lines and filled symbols correspond to the results when using $\pi_{\max} = 200\,h^{-1}$ Mpc, and match the results presented in the top panel of Fig. 8. The dashed lines and open symbols are our results when we use $\pi_{\max} = 350\,h^{-1}$ Mpc. The results in both cases are consistent, especially noting that the errors in the case of using $\pi_{\max} = 350\,h^{-1}$ Mpc (not shown in the figure) are consistently larger than those shown for our main results.

APPENDIX B: ANALYSIS OF THE RELIABILITY OF THE RECOVERED BIAS

We used the mock catalogues to test the full process used to estimate the bias of a given galaxy sample and its uncertainty. We perform this test using the same mock galaxy sample used in Appendix A1, selected in redshift and absolute magnitude as $0.5 < z_p < 0.8$, $M_B - 5\log_{10}h < -17.95$. We estimated the projected correlation function $w_p(r_p)$ and its error for this sample in each of the 50 mock ALHAMBRA realizations available, following the method described in Section 4. In the calculation, we used a value of $\pi_{\max} = 3r(\sigma_z)$, as discussed in Appendix A. Given that the redshift uncertainty in this mock sample, $\sigma_z/(1 + z) = 0.025$, is somewhat larger than that in the data (see Table 2), this results in a value of $\pi_{\max} = 270\,h^{-1}$ Mpc, larger than the value used for the data. However, as we have shown in Appendix A1, the optimal value of $\pi_{\max}$ scales with σ_z, so this value provides an adequate comparison with the calculation done with the data.

We then estimated the bias b and its uncertainty for each realization using the method described in Section 5.2, using a model for the matter correlation function matching the Millennium cosmology used in the mocks. The obtained values are shown in Fig. B1. The mean value obtained for the bias is $\langle b \rangle = 1.230$, and the standard deviation of the values from the 50 realizations is $\sigma_{\rm real}(b) = 0.162$. These values are shown as the dashed line and shaded area in

Figure A2. Bias parameter obtained for the different samples from the fit to equation (12), using different values of $\pi_{\max}$ in the estimation of the projected correlation function $w_p(r_p)$ (equation 4). The solid lines and filled symbols with error bars correspond to the calculation using $\pi_{\max} = 200\,h^{-1}$ Mpc and are the same as shown in the top panel of Fig. 8. The dashed lines and empty symbols (slightly displaced in the horizontal direction for clarity) correspond to the calculation using $\pi_{\max} = 350\,h^{-1}$ Mpc. We do not show the error bars in this case, but they are consistently larger than those shown for $\pi_{\max} = 200\,h^{-1}$ Mpc.

Figure B1. Bias parameter obtained for the galaxy sample defined by $0.5 < z_p < 0.8$, $M_B - 5\log h < -17.95$ in each of the 50 mock ALHAMBRA realizations. The bias and its uncertainty in each case are estimated using the projected correlation function in the same way as is done for the real data in Sections 4 and 5.2. The dashed line shows the mean value obtained from the 50 realizations, $\langle b \rangle = 1.230$, and the shaded area corresponds to a region of $\pm 1\sigma_{\rm real}$ around it, where $\sigma_{\rm real} = 0.162$ is the standard deviation of the 50 values. The blue solid line corresponds to the mean bias parameter obtained from the spherically averaged correlation function $\xi(r)$ of the 50 realizations.

Fig. B1. The values of the uncertainty estimated show a broad distribution, with a mean value of $\langle\sigma(b)\rangle = 0.159 \pm 0.007$. This shows that the block bootstrap method used provides an unbiased estimation of the galaxy bias uncertainty.

Finally, we compare these values obtained for the mock catalogues to the real-space result. To this end, we compute the spherically averaged real-space correlation function $\xi(r)$ for each of the 50 realizations, using the real position of each galaxy, instead of that estimated from its photometric redshift. We then obtain the real-space bias using a method analogous to that described in Section 5.2, using $\xi(r)$ instead of $w_p(r_p)$. The mean value of the bias obtained in this way is $\langle b_r \rangle = 1.244 \pm 0.012$ (shown as the continuous blue line in Fig. B1). Therefore, we can estimate the bias in our measurement as $\Delta b = \langle b \rangle - \langle b_r \rangle = -0.014$. This corresponds to $0.09\sigma_{\text{real}}(b)$, and is therefore consistent with the condition imposed in Appendix A1 to determine the value of $\pi_{\max}$ used (equation A3). We therefore conclude that the method used to recover the real-space clustering (and in particular the galaxy bias) from the ALHAMBRA photometric redshift catalogues using the projected correlation function is reliable, as we recover the direct real-space result within the expected accuracy.

APPENDIX C: TABLES OF NUMERICAL RESULTS

We present in Table C1 the parameters obtained from the fits of different models to the projected correlation function of our different samples. The parameters listed are the correlation length r_0 and exponent γ obtained from the fit to the power-law model in equation (11), and the bias b obtained from the fit to the model in equation (12). We list both the results obtained using the full survey (with seven fields) and those obtained when we exclude the 'outlier fields' ALH-4/COSMOS and ALH-7/ELAIS-N1 (see Section 5.3 for details).

For completeness, we show in Fig. C1 the parameters r_0 and γ obtained from the power-law fits when we exclude the 'outlier fields' from the calculation. This figure can be directly compared to Fig. 7.

Table C1. Parameters obtained from the fits of different models to the projected correlation function of our different samples. The correlation length r_0 and exponent γ are obtained from the fit to the power-law model in equation (11) (see Section 5.1 for details). The bias is obtained from the fit to the model in equation (12) (see Section 5.2 for details). Results are listed both for the full survey and when excluding the 'outlier fields' ALH-4/COSMOS and ALH-7/ELAIS-N1. The sample name listed matches that in Table 2. The results shown in this table correspond to those shown in Figs 7, 8, 10 and C1.

| Sample	Fit to the full survey			Fit omitting the 'outlier' fields ALH-4 and ALH-7		
	$r_0(h^{-1}\,\text{Mpc})$	γ	Bias	$r_0(h^{-1}\,\text{Mpc})$	γ	Bias
Z05M0	3.4 ± 0.3	1.82 ± 0.09	1.00 ± 0.09	3.45 ± 0.25	1.82 ± 0.09	1.01 ± 0.07
Z05M1	4.0 ± 0.4	1.79 ± 0.10	1.16 ± 0.11	3.9 ± 0.4	1.80 ± 0.10	1.14 ± 0.11
Z05M2	4.2 ± 0.5	1.80 ± 0.14	1.23 ± 0.13	4.1 ± 0.5	1.82 ± 0.14	1.20 ± 0.13
Z05M3	4.4 ± 0.5	1.82 ± 0.10	1.27 ± 0.12	4.5 ± 0.4	1.79 ± 0.09	1.31 ± 0.09
Z05M4	4.8 ± 0.7	1.82 ± 0.12	1.35 ± 0.17	4.7 ± 0.6	1.85 ± 0.11	1.36 ± 0.14
Z05M5	5.5 ± 0.6	1.79 ± 0.12	1.50 ± 0.16	4.9 ± 0.4	1.89 ± 0.11	1.41 ± 0.16
Z05M6	6.6 ± 0.8	1.72 ± 0.10	1.79 ± 0.17	5.37 ± 0.24	1.89 ± 0.05	1.56 ± 0.13
Z05M7	9.3 ± 2.3	1.73 ± 0.25	2.4 ± 0.5	7.1 ± 1.1	1.89 ± 0.12	2.0 ± 0.3
Z07M1	3.8 ± 0.5	1.72 ± 0.10	1.22 ± 0.17	3.3 ± 0.3	1.82 ± 0.08	1.09 ± 0.11
Z07M2	4.2 ± 0.7	1.68 ± 0.11	1.34 ± 0.21	3.5 ± 0.3	1.81 ± 0.08	1.13 ± 0.11
Z07M3	4.9 ± 0.8	1.66 ± 0.12	1.52 ± 0.24	3.9 ± 0.4	1.85 ± 0.10	1.24 ± 0.13
Z07M4	5.7 ± 1.1	1.64 ± 0.13	1.74 ± 0.29	4.4 ± 0.3	1.80 ± 0.08	1.41 ± 0.07
Z07M5	6.7 ± 1.4	1.63 ± 0.15	2.0 ± 0.3	4.9 ± 0.6	1.81 ± 0.13	1.56 ± 0.16
Z07M6	8.2 ± 1.7	1.60 ± 0.15	2.3 ± 0.4	6.0 ± 0.4	1.71 ± 0.11	1.83 ± 0.09
Z07M7	11.7 ± 3.0	1.68 ± 0.15	3.2 ± 0.6	6.8 ± 1.3	1.78 ± 0.22	2.3 ± 0.3
Z09M2	3.7 ± 0.4	1.65 ± 0.09	1.37 ± 0.15	3.4 ± 0.3	1.79 ± 0.04	1.20 ± 0.07
Z09M3	4.1 ± 0.5	1.71 ± 0.07	1.47 ± 0.15	3.7 ± 0.4	1.83 ± 0.05	1.26 ± 0.10
Z09M4	4.6 ± 0.6	1.72 ± 0.08	1.62 ± 0.17	4.2 ± 0.3	1.80 ± 0.05	1.43 ± 0.06
Z09M5	5.4 ± 0.9	1.72 ± 0.10	1.9 ± 0.3	4.8 ± 0.3	1.80 ± 0.04	1.59 ± 0.08
Z09M6	6.6 ± 0.9	1.69 ± 0.12	2.21 ± 0.23	5.91 ± 0.22	1.77 ± 0.10	1.96 ± 0.07
Z09M7	8.9 ± 1.2	1.58 ± 0.15	2.9 ± 0.3	7.8 ± 1.0	1.74 ± 0.09	2.61 ± 0.23
Z11M3	3.36 ± 0.22	1.73 ± 0.06	–	3.29 ± 0.23	1.79 ± 0.05	–
Z11M4	3.58 ± 0.24	1.73 ± 0.08	–	3.44 ± 0.24	1.80 ± 0.07	–
Z11M5	3.9 ± 0.3	1.76 ± 0.11	–	3.52 ± 0.24	1.81 ± 0.11	–
Z11M6	4.6 ± 0.5	1.67 ± 0.15	–	3.9 ± 0.4	1.55 ± 0.08	–
Z11M7	5.0 ± 0.7	1.7 ± 0.3	–	4.3 ± 0.5	1.9 ± 0.3	–
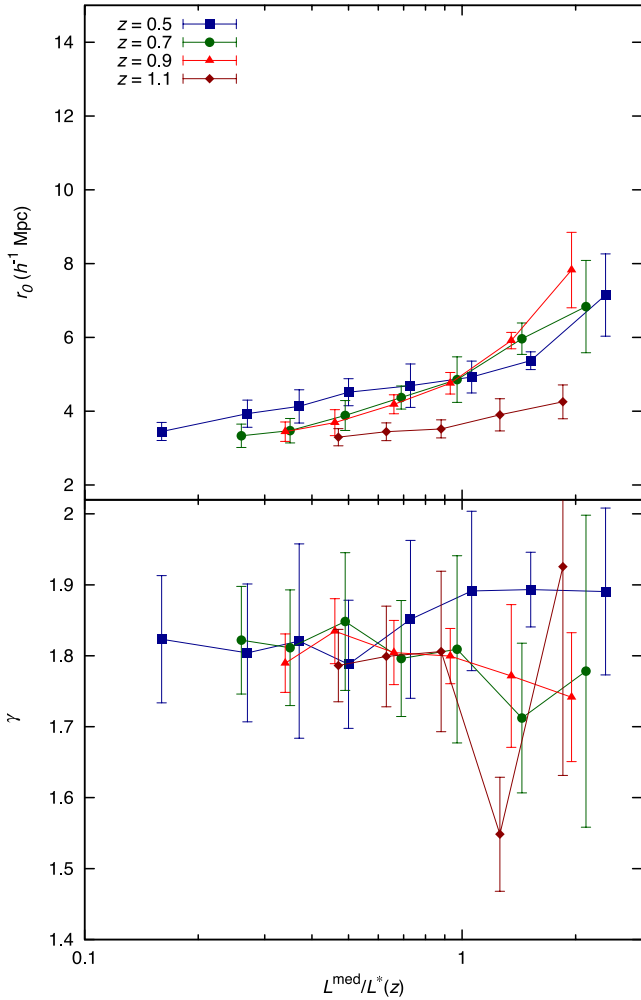
Figure C1. Parameters r_0 and γ obtained from the power-law fits (Section 5.1) for the different samples, as a function of the rest-frame *B*-band median luminosity, for the case in which we omit from the calculation the 'outlier' fields ALH-4/COSMOS and ALH-7/ELAIS-N1 (see Section 5.3).

[1]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Observatori Astronòmic, Universitat de València, C/Catedràtic José Beltrán 2, E-46980 Paterna, Spain*
[3]*Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot, Spain*
[4]*Unidad Asociada Observatorio Astronómico (IFCA-UV), E-46980 Paterna, Spain*
[5]*Instituto de Física de Cantabria (CSIC-UC), E-39005 Santander, Spain*
[6]*IAA-CSIC, Glorieta de la Astronomía s/n, E-18008 Granada, Spain*
[7]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[8]*Instituto de Astrofísica de Canarias, Vía Láctea s/n, E-38200 La Laguna, Tenerife, Spain*
[9]*Institut de Ciències de l'Espai (IEEC-CSIC), Facultat de Ciències, Campus UAB, E-08193 Bellaterra, Spain*
[10]*Centro de Estudios de Física del Cosmos de Aragón, Plaza San Juan 1, E-44001 Teruel, Spain*
[11]*Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA*
[12]*Physics and Astronomy Department, University of Missouri, Columbia, MO 65211, USA*
[13]*Observatório Nacional-MCT, Rua José Cristino 77, CEP 20921-400 Rio de Janeiro-RJ, Brazil*
[14]*Department of Theoretical Physics, University of the Basque Country UPV/EHU, E-4808014 Bilbao, Spain*
[15]*Departamento de Física Atómica, Molecular y Nuclear, Facultad de Física, Universidad de Sevilla, E-41012 Sevilla, Spain*
[16]*Departamento de Astrofísica, Facultad de Física, Universidad de la Laguna, E-38206 La Laguna, Spain*
[17]*Departamento de Astronomía, Pontificia Universidad Católica, 782-0436 Santiago, Chile*

This paper has been typeset from a TeX/LaTeX file prepared by the author.