The zCOSMOS survey. The dependence of clustering on luminosity and stellar mass at $z = 0.2$–1*

B. Meneux1,2, L. Guzzo3, S. de la Torre3,4,5, C. Porciani6,7, G. Zamorani8, U. Abbas9, M. Bolzonella5, B. Garilli5, Iovino10, L. Pozzetti9, E. Zucca7, S. J. Lilly1, O. Le Fèvre1, J.-P. Kneib5, C. M. Carollo7, T. Contini11, V. Mainieri12, A. Renzini13, M. Scudellaro5, S. Bardelli9, A. Bongiorno1, K. Caputi1, G. Coppa8,14, O. Cucciati4, L. de Ravel14, P. Franzetti5, P. Kampczyk7, C. Knobel1, K. Kovač5, F. Lamareille11, J.-F. Le Borgne11, V. Le Brun4, C. Maier7, R. Pellò11, Y. Peng3, E. Perez Montero11, E. Ricciardelli13, J. D. Silverman7, M. Tanaka12, L. Tasca4,5, L. Tresse4, D. Vergani8, D. Bottini5, A. Cappi8, A. Cimatti14, P. Cassata4, M. Fumana5, A. M. Keckmoer15, A. Leauthaud16, D. Maccagni5, C. Marinoni17, H. J. McCracken18, P. Memeo5, P. Oesch7, and R. Scaramella19

(Affiliations can be found after the references)

Received 9 April 2009 / Accepted 15 June 2009

ABSTRACT

Aims. We study the dependence of galaxy clustering on luminosity and stellar mass at redshifts $z \sim 0.2$–1, using the first 10K redshifts from the zCOSMOS spectroscopic survey of the COSMOS field.

Methods. We measured the redshift-space correlation functions $\xi_{p}(r,\pi)$ and $\xi(r)$ and the projected function, $\xi_{p}(r_{p})$ for subsamples covering different luminosity, mass, and redshift ranges. We explored and quantified in detail the observational selection biases from the flux-limited nature of the survey, using ensembles of realistic semi-analytic mock samples built from the Millennium simulation. We used the same mock data sets to carefully check our covariance and error estimate techniques, comparing the performances of methods based on the scatter in the mocks and on bootstrapping schemes. We finally compared our measurements to the cosmological model predictions from the mock surveys.

Results. At odds with other measurements at similar redshift and in the local Universe, we find a weak dependence of galaxy clustering on luminosity in all three redshift bins explored. A mild dependence on stellar mass is instead observed, in particular on small scales, which becomes particularly evident in the central redshift bin ($0.5 < z < 0.8$), where $\xi_{p}(r_{p})$ shows strong excess power on scales $>1 h^{-1}$ Mpc. This is reflected in the shape of the full $\xi(r_{p},\pi)$ that we interpret as produced by dominating structures almost perpendicular to the line of sight in the survey volume. Comparing to $z \sim 0$ measurements, we do not see any significant evolution with redshift of the amplitude of clustering for bright and/or massive galaxies.

Conclusions. This is consistent with previous results and the standard picture in the plastic band where the bias evolves more rapidly for the most massive haloes, and in which it is the baryons that form the visible fabric of the Universe, a major challenge in testing the galaxy formation paradigm is to build clear connections between these observed properties and those of the hosting dark-matter haloes. This is a difficult task, as any direct connection initially existing between the dark-matter mass and the baryonic component cooling within the halo is modified by all subsequent dynamical processes affecting the halo-galaxy system, such as merging or dynamical friction. This is confirmed by simulations, which also show however that galaxy luminosity and stellar mass do in fact retain memory of the “original” (not actual) halo mass, i.e. before it experiences a major merger or is accreted by a larger halo [Conroy et al. 2006; Wang et al. 2006, 2007]. This gives some hope that by measuring the dependence of the galaxy distribution on galaxy properties one is actually constraining the relationship between the dark and luminous components of galaxies.

Key words. cosmology: observations – large-scale structure of Universe – surveys – galaxy: evolution

1. Introduction

In the canonical scenario of galaxy formation, galaxies are thought to form through the cooling of baryonic gas within extended dark matter haloes [White & Rees 1978]. The mass of the hosting halo is expected to play a significant role in the definition of the visible properties of the galaxy, as the total mass in gas and stars, its luminosity, colour, star formation rate, and possibly, morphology.

Since it is the baryons that form the visible fabric of the Universe, a major challenge in testing the galaxy formation paradigm is to build clear connections between these observed properties and those of the hosting dark-matter haloes. This is a difficult task, as any direct connection initially existing between the dark-matter mass and the baryonic component cooling within the halo is modified by all subsequent dynamical processes affecting the halo-galaxy system, such as merging or dynamical friction. This is confirmed by simulations, which also show however that galaxy luminosity and stellar mass do in fact retain memory of the “original” (not actual) halo mass, i.e. before it experiences a major merger or is accreted by a larger halo [Conroy et al. 2006; Wang et al. 2006, 2007]. This gives some hope that by measuring the dependence of the galaxy distribution on galaxy properties one is actually constraining the relationship between the dark and luminous components of galaxies.
Measurements of first moments, such as the luminosity function (LF) or the stellar mass function, provide a way to understand how these are related to the total halo mass functions, which can be obtained from analytic predictions (e.g. Press & Schechter 1974) or n-body simulations (e.g. Warren et al. 2006). Similar investigations can be made on the second moment, i.e. the two-point correlation function (e.g. Springel et al. 2008). Studies of galaxy clustering in large local surveys have shown how clustering at z ∼ 0 does depend significantly on several specific properties. These include luminosity (Hamiton 1988; Maurogordato & Lachieze-Rey 1991; Iovino et al. 1993; Benoit et al. 1996; Guzzo et al. 2000; Norberg et al. 2001a; Norberg et al. 2002; Zehavi et al. 2005), colour or spectral type (Willmer et al. 1998; Norberg et al. 2002; Zehavi et al. 2002), morphology (Davis & Geller 1976; Giovaneli et al. 1986; Guzzo et al. 1997), stellar mass (Li et al. 2006), and environment (Abbas & Sheth 2006).

In recent years it has become possible to extend these investigations to high redshift, obtaining first indicative results on how these dependences evolve with time (Daddi et al. 2003; Coil et al. 2002; Zehavi et al. 2002), morphology (Davis & Geller 1976; Giovaneli et al. 1986; Guzzo et al. 1997), stellar mass (Li et al. 2006), and environment (Abbas & Sheth 2006).

The zCOSMOS survey (Lilly et al. 2007) is being performed to understand how these dependences evolve with time (Daddi et al. 2003; Coil et al. 2002; Zehavi et al. 2002), morphology (Davis & Geller 1976; Giovaneli et al. 1986; Guzzo et al. 1997), stellar mass (Li et al. 2006), and environment (Abbas & Sheth 2006).
Fig. 1. Distribution on the sky of the ∼10.000 galaxies with measured redshift (crosses) forming the zCOSMOS “10K” sample. The large blue dots mark the centres of independent VIMOS pointings, each including four quadrants on the sky (as described by the red solid lines).

criteria (Lilly et al. 2007) (zCOSMOS Faint). So far, the survey has observed about half of the total “Bright” sample. This is the so-called “10K” sample used for the analysis presented in this paper, and is based on the observations of 83 VIMOS pointings over 44 distinct telescope positions on the sky (Lilly et al. 2009). These are shown in Fig. 1 where the footprint of VIMOS (4 quadrants of ∼7 × 8 arcmin² separated by a cross about 2 arcmin² wide) is evident. About every 3rd galaxy was observed in the field. The final “20K” zCOSMOS sample will be twice larger, reaching a sampling around 60–70%. The correction of the complex angular selection function is discussed later in the context of our galaxy clustering measurement.

Observations were performed using the medium-resolution RED grism, corresponding to R ∼ 600 and covering the spectral range 5550–9650 Å. The average error on the redshift measurements was estimated from the repeated observations of 632 galaxies and found to be ∼100 km s⁻¹ (Lilly et al. 2009). This corresponds roughly to a radial distance error of 1 h⁻¹ Mpc. The reduction of the data to the redshift assignment was carried out independently at two institutes before a reconciliation process to solve discrepancies. The quality of each measured redshift was then quantified via a quality flag that provides us with a confidence level (see Lilly et al. 2007, 2009, for definition). For the present work, we only use redshifts with flags 1.5–4.5 and 9.3–9.5, corresponding to confidence levels greater than 98%.

The zCOSMOS survey benefits the large multi-wavelength coverage of the COSMOS field (Capak et al. 2007), which with the latest additions now comprises 30 photometric bands (Ilbert et al. 2009) extending well into the infrared. These include, in particular, accurate K-band and Spitzer-IRAC photometry over the whole area, which have allowed us to derive relevant physical properties as rest-frame luminosity and stellar mass with unprecedented accuracy (Bolzonella et al. 2009; Pozzetti et al. 2009; Zucca et al. 2009).

3. Mock survey catalogues

In this paper we make intense use of mock surveys constructed from the Millennium simulation (Springel et al. 2005). This was done a) to understand the effect of our selection criteria on the measured quantities (Sect. 6.3) and b) to estimate the measurement errors and covariance of the data (Sect. 7).

We used two sets of light cones, constructed as explained in Kitzbichler & White (2007) and Blaizot et al. (2005) by combining dark-matter halo trees from the Millennium run to the Munich semi-analytic model of galaxy formation (De Lucia & Blaizot 2007). The two sets contain 24 1.4 × 1.4 deg² mocks built by Kitzbichler & White (2007) and 40 1 × 1 deg² mocks built by De Lucia & Blaizot (2007), which we name KW24 and DLB40, respectively. The main difference between the two sets, in addition to the different survey area, is that the DLB40 set contains all galaxies irrespective of any criteria down to the simulation limit that corresponds roughly to 10⁸ M⊙, up to redshift z = 1.7, whereas the KW24 set only contains galaxies brighter than I ≤ 22.5. This implies that the DLB40 set allowed us to select in stellar mass down to very low masses and to test selection effects. The observing strategy of the zCOSMOS 10K sample was only applied to the KW24 set, allowing us to do careful error analysis of our measurements.

The Millennium run contains N = 2160³ particles of mass 8.6 × 10⁵ h⁻¹ M⊙ in a cubic box of size 500 h⁻¹ Mpc. The simulation was built with a ΛCDM cosmological model with Ω_m = 0.25, Ω_Λ = 0.75, σ_8 = 0.9 and H₀ = 73 km s⁻¹ Mpc⁻¹.
4. Luminosity- and mass-selected subsamples

4.1. Luminosity selection

Absolute magnitudes were derived for the 10K galaxies using the code ALF ([Ilbert et al. 2005; Zucca et al. 2009]), which is based on fitting a spectral energy distribution (SED) to the observed multi-band photometry. There are various sources of uncertainties to take into account (errors on apparent magnitudes, number of available photometric bands, method used, etc.). A direct comparison with absolute magnitudes derived with the independent code ZEBRA ([Feldmann et al. 2006]) shows consistent estimates with a small dispersion of $\sigma \approx 0.05$ magnitudes, in particular in the B-band. This can reasonably be considered as the typical error on our absolute magnitudes.

For our analysis, the goal is to define luminosity-limited samples that are as close as possible to truly volume-limited samples, i.e. with a constant number density. This might be done within a few independent redshift ranges. The size of the redshift slices in which to split the sample has to be chosen as a compromise between two aspects: a) large enough to have sufficient statistics and provide a good measurement of clustering and b) not too large to avoid significant evolution within each redshift bin.

We know, however, that the luminosity of galaxies evolves through the redshift range covered by the zCOSMOS survey ($0.2 < z < 1.1$), with a clear change in the characteristic parameters of the LF ([Ilbert et al. 2005]). This evolution does depend on the morphological/spectral type of the galaxy considered. To be able to select a nearly volume-limited sample within a given redshift interval, we need to take the corresponding evolution into account. This can only be done realistically in a statistical way by looking at the population-averaged evolution of the global LF.

We therefore considered the observed LF measured from the same data ([Zucca et al. 2009]) and modelled its change with redshift as a pure luminosity evolution (i.e. keeping a constant slope) of the LF ([Ilbert et al. 2005]). This evolution does depend on the redshift range considered, i.e. $z = [0.2–0.5]$, $z = [0.5–0.8]$, and $z = [0.8–1.0]$. The details of the resulting samples are described in Table 1.

We therefore defined our luminosity-limited samples by an absolute magnitude cut at $z = 0$, $M_{B,\text{cut}}$, including all galaxies with $M_B(z) = 5 \log(h) \leq M_{B,\text{cut}} - z$. The resulting selection loci for different values of $M_{B,\text{cut}}$ are plotted over the data in the luminosity-redshift plane in the right panel of Fig. 2. As is evident in the figure, the faintest allowed threshold $M_{B,\text{cut}}$ depends on the redshift range considered, i.e. $z = [0.2–0.5]$, $z = [0.5–0.8]$, and $z = [0.8–1.0]$. The details of the resulting samples are described in Table 1.

Table 1. Properties of the luminosity-selected samples.

| Sample | Redshift range | Mean $M_{B,\text{cut}}$ ($z = 0$) | Number of galaxies |
|--------|----------------|-------------------------------|-------------------|
| L1.1   | 0.2–0.5        | 0.37                          | 1892              |
| L1.2   | 0.2–0.5        | 0.37                          | 1311              |
| L1.3   | 0.2–0.5        | 0.37                          | 811               |
| L1.4   | 0.2–0.5        | 0.37                          | 469               |
| L2.1   | 0.5–0.8        | 0.67                          | 1848              |
| L2.2   | 0.5–0.8        | 0.67                          | 1025              |
| L2.3   | 0.5–0.8        | 0.67                          | 441               |
| L3.1   | 0.8–1.0        | 0.91                          | 500               |
| L3.2   | 0.8–1.0        | 0.91                          | 447               |

Table 2. Properties of the mass-selected samples.

| Sample | Redshift range | Mean $M/M_{\odot}$ | Number of galaxies |
|--------|----------------|--------------------|-------------------|
| M1.1   | 0.2–0.5        | 0.66               | 2159              |
| M1.2   | 0.2–0.5        | 0.66               | 1445              |
| M1.3   | 0.2–0.5        | 0.66               | 827               |
| M1.4   | 0.2–0.5        | 0.66               | 275               |
| M2.1   | 0.5–0.8        | 0.66               | 2351              |
| M2.2   | 0.5–0.8        | 0.66               | 2276              |
| M2.3   | 0.5–0.8        | 0.67               | 1366              |
| M2.4   | 0.5–0.8        | 0.67               | 477               |
| M3.1   | 0.8–1.0        | 0.90               | 755               |
| M3.2   | 0.8–1.0        | 0.90               | 344               |

We therefore defined our luminosity-limited samples by an absolute magnitude cut at $z = 0$, $M_{B,\text{cut}}$, including all galaxies with $M_B(z) = 5 \log(h) \leq M_{B,\text{cut}} - z$. The resulting selection loci for different values of $M_{B,\text{cut}}$ are plotted over the data in the luminosity-redshift plane in the right panel of Fig. 2. As is evident in the figure, the faintest allowed threshold $M_{B,\text{cut}}$ depends on the redshift range considered, i.e. $z = [0.2–0.5]$, $z = [0.5–0.8]$, and $z = [0.8–1.0]$. The details of the resulting samples are described in Table 1.

4.2. Mass selection

Stellar mass has become a quantity routinely measured in recent years, thanks to surveys with multi-wavelength photometry, extending to the near-infrared ([Rettura et al. 2006]), although some uncertainties related to the detailed modelling of stellar
Fig. 3. The observed relationship between stellar mass and luminosity for galaxies in the 10K sample, within the three redshift ranges studied in this paper. The left panel shows an aspect of the galaxy bimodality, with red galaxies more massive and brighter than blue ones.

Fig. 4. Estimate of how the completeness in stellar mass changes as a function of redshift, due to the survey flux limit ($I_{AB} < 22.5$). The shaded grey area and green contours describe the loci of constant completeness. They are derived from the DLB40 mock samples of $1 \times 1 \text{deg}^2$ and defined as the fraction of observed ($I_{AB} \leq 22.5$) galaxies over the total number in a given cell with size $\Delta z = 0.01$ and $\log(M) \geq \log(M_{\text{cut}})$. The superimposed red points correspond to the actual data of the 10K sample. The yellow points and dashed line show the 95% M/L ratio completeness level derived independently by Pozzetti et al. (2009) (see text), directly from the observed data.

The observed relationship between stellar mass and luminosity for galaxies in the 10K sample, within the three redshift ranges studied in this paper. The left panel shows an aspect of the galaxy bimodality, with red galaxies more massive and brighter than blue ones.

The flux-limited nature of surveys like zCOSMOS ($I_{AB} < 22.5$) mean that the lowest-mass samples are affected to varying degrees by incompleteness related to the scatter in the mass-luminosity relation (Fig. 3). This introduces a bias against objects that would be massive enough to enter the mass-selected samples, but too faint to fulfill the apparent-magnitude limit of the survey. These missed high mass-to-light ratio galaxies will be those dominated by low-luminosity stars, i.e. the red and faint objects. Clearly, if this is not accounted for in some way, it would inevitably affect the estimated clustering properties, with respect to a truly complete, mass-selected sample (Meneux et al. 2008).

It is therefore necessary to understand the effective completeness level in detail in the stellar mass of the samples that we defined for our analysis. Meneux et al. (2008) have used two different methods to explore and quantify the completeness limit in stellar mass as a function of redshift. The first is based on the observed scatter in the mass-luminosity relation, obtained from the data themselves and extrapolated to fainter fluxes. The second instead makes use of mock survey samples, under the hypothesis that they provide a realistic description of the mass-luminosity relation and its scatter: the DLB40 set of mock survey catalogues that are complete in stellar mass are “observed” under the same conditions as the real data, i.e. selected at $I \leq 22.5$. The completeness is then simply defined, for a given redshift range and mass threshold, as the ratio of the number of galaxies brighter than the zCOSMOS flux...
Table 3. The completeness in stellar mass of mass-selected mock sub-samples reproducing the properties and selection criteria of our 10K data samples.

| Sample | Redshift range | Stellar mass (log(M/M_☉)) | Completeness range |
|--------|----------------|----------------------------|--------------------|
| M1.1   | 0.2–0.5        | ≥ 9.0                      | 0.783              |
| M1.2   | 0.2–0.5        | ≥ 9.5                      | 0.972              |
| M1.3   | 0.2–0.5        | ≥ 10.0                     | 1.000              |
| M1.4   | 0.2–0.5        | ≥ 10.5                     | 1.000              |
| M2.1   | 0.5–0.8        | ≥ 9.0                      | 0.349              |
| M2.2   | 0.5–0.8        | ≥ 9.5                      | 0.652              |
| M2.3   | 0.5–0.8        | ≥ 10.0                     | 0.919              |
| M2.4   | 0.5–0.8        | ≥ 10.5                     | 0.996              |
| M3.1   | 0.8–1.0        | ≥ 10.0                     | 0.571              |
| M3.2   | 0.8–1.0        | ≥ 10.5                     | 0.882              |

limit over those at any flux. Interestingly, even if this method is model-dependent (in particular, on the prescription of galaxy formation used in the semi-analytic models), this approach leads to similar completeness limits to the first one. The results of this second exercise are shown as a function of redshift and mass threshold and for a flux limit I ≤ 22.5, in Fig. 4. Completeness is estimated in narrow redshift ranges (Δz = 0.01) for different mass thresholds M_{min} increasing from 10^9 to 10^{11.7} h^{-2} M_☉ with a step of 10^{10} h^{-2} M_☉. A large fraction of low-mass objects are clearly missed at high redshift.

We also add in Fig. 4 the completeness limit estimated from the observed scatter in the M/L relation of the data, and defined at each redshift as the lower boundary, M_{min}(z), including above it 95% of the mass distribution (Pozzetti et al. 2009). It is very encouraging to notice the very good agreement between this independent estimate from the data and that based on the DLB40 set of mock catalogues. This adds confidence in the use of the simulated samples. Table 3 summarises the completeness estimates derived from these mock catalogues for each of the 10 zCOSMOS galaxy samples defined in Table 2. The sample M2.1 shows the strongest incompleteness: 65.1% of the galaxies more massive than 10^9 h^{-2} M_☉ are fainter than I = 22.5 at z = [0.5–0.8] and then, not included in our sample. In Sect. 6.3 we discuss the effects of this incompleteness on the galaxy clustering measurement.

5. Estimating the two-point correlation function

The two-point correlation function is the simplest estimator for quantifying galaxy clustering, because it is related to the second moment of the galaxy distribution, i.e. its variance. In practice, it describes the excess probability ξ(r) of observing a pair of galaxies at a given separation r, with respect to that of a random distribution (Peebles 1980). Here we estimate the redshift-space correlation function ξ_s(r_p, π), which allows one to incorporates the effect of peculiar motions on the pure Hubble recession velocity. In this case, galaxy separations are split into the tangential and radial components, r_p and π, respectively (Davis & Peebles 1983; Fisher et al. 1994).

The real-space correlation function ξ_r(r) can be recovered by projecting ξ_s(r_p, π) along the line of sight, as

\[ w_p(r_p) = 2 \int_0^\infty \xi_s(r_p, π) dπ = 2 \int_0^\infty \xi_r \left( r_p^2 + y^2 \right)^{1/2} dy. \]  

For a power-law correlation function, \( \xi_r(r) = \left( r/r_0 \right)^{-\gamma} \), this integral can be solved analytically and fitted to the observed \( w_p(r_p) \) to find the best-fitting values of the correlation length \( r_0 \) and slope \( \gamma \) (e.g. Davis & Peebles 1983). In computing \( w_p(r_p) \), a finite upper integration limit has to be chosen in practice. Its value has to be high enough as to include most of the clustering signal dispersed along the line of sight by peculiar motion. However, it must not be too high to avoid adding only noise, which is dominant above a certain π. Previous works (Pollo et al. 2005) have shown that, for similar data, the best results are obtained with an integration limit \( \pi_{max} \) between 20 and 40 h^{-1} Mpc. Our tests show that the scatter in the recovered \( w_p(r_p) \) is obtained using the lowest value in this range. This can introduce a 5–10% underestimate in the recovered large-scale amplitude, which can be accounted for when fitting a model to \( w_p(r_p) \). In the following, we in general use \( \pi_{max} = 20 h^{-1} \) Mpc and show examples of how the amplitude is biased by this choice for the real data.

To estimate ξ(r_p, π) from each galaxy sample, we used the standard estimator of Landy & Szalay (1993):

\[ ξ(r_p, π) = \frac{N_G(N_G - 1) GG(r_p, π) - N_G - 1 GR(r_p, π) + 1}{N_G N_G GR(π, π) - N_G RR(π, π)}, \]  

where \( N_G \) is the mean galaxy density (or, equivalently, the total number of objects) in the sample, \( N_R \) the mean density of a catalogue of random points distributed within the same survey volume and with the same selection function as the data, \( GG(π, r_p) \) the number of independent galaxy-galaxy pairs with separation between \( r_p \) and \( r_p + dr_p \) and between π and π + dπ, \( RR(π, r_p) \) the number of independent random-random pairs within the same interval of separations, and \( GR(r_p, π) \) represents the number of galaxy-random cross pairs.

6. Observational biases and selection effects

6.1. Correction of VIMOS angular footprint and varying sampling

To properly estimate the correlation function from the 10K zCOSMOS data, we need to correct for its spatial sampling rate, which is on average ~30%, but which varies with the position on the sky due to the VIMOS footprint and the superposition of multiple passes (see Fig. 1). The correction scheme used here is an evolution of the one discussed in Pollo et al. (2005), but with a simplified weighting scheme. The main differences between this sample and the VVDS-Deep data of Pollo et al. (2005) are that: a) this sample is 1.5 magnitudes brighter; b) the spectra are longer, and up to four of them can be packed along the same column on the CCD. The net result of these differences can be appreciated in Fig. 1 where the sample is characterised by a well-sampled central region, but also by rather sparsely sampled VIMOS pointings in the outskirts of the field. In particular, these external pointings clearly show target galaxies concentrated along rows. This effect is produced by the significant length of the spectra on the CCD in medium-resolution mode: not more than two spectra and thus fewer objects along the dispersion direction; and c) there are as many as eight repeated observations (“passes”) covering each point on the sky in the central area of the COSMOS field.

We tested three different algorithms to correct for the angular selection function of the survey and obtained comparable results. Other weighting schemes use in particular the angular correlation...
functions of the 10K sample and the photometric catalogue to correct for the nonuniform spatial sampling rate. These methods are discussed in the parallel clustering analyses by de la Torre et al. (2009) and Porciani et al. (in preparation). In the latter paper in particular, comparative tests of the three algorithms are presented.

Since the subsamples analysed in this work are essentially volume-limited (above the luminosity/mass completeness limits), we did not need to apply any further minimum-variance weighting scheme (as e.g. the $J_3$ weighting, Fisher et al. 1994). This is normally necessary for purely flux-limited surveys in which the selection function varies significantly as a function of redshift, such that different parts of the volume are sampled by galaxies with different luminosities and number densities (e.g. Li et al. 2006).

6.2. Construction of reference random samples

A significant source of uncertainty that we encountered in estimating two-point functions from our 10K subsamples is related to the construction of the random sample and in particular to its redshift distribution. We soon realised that the strongly clustered nature of the COSMOS field along the line of sight, with several dominating structures at different redshifts, required some particular care so as not to generate systematic biases in the random sample. These superclusters are already evident as vertical stripes in Fig. 2 and even more clearly so in the redshift histogram of Fig. 5. We point out the big “walls” at $z = 0.35, 0.75$, and 0.9, which are also clearly identified by the density field reconstruction of Kovač et al. (2009).

A standard way to generate a random redshift coordinate accounting for the radial selection function of the data uses a Gaussian-filtered version of the data themselves. This is normally obtained using smoothing kernels with a dispersion $\sigma$ (in co-moving coordinates) in the range $150$ to $250$ $h^{-1}$ Mpc. The results of applying this technique to the current 10K data are shown in Fig. 5. One notes how for smoothing scales of 150 and $250$ $h^{-1}$ Mpc, the curves still retain memory of the two largest galaxy fluctuations. These are only erased when a very strong smoothing filter ($450$ $h^{-1}$ Mpc) is adopted. However, in this case the smoothed curve is unable to correctly follow the global shape of the distribution, overestimating the number density in the lowest and highest redshift ranges. The situation for our specific analysis, however, is somewhat simpler than this general case. Our luminosity-limited or mass-limited samples are in principle “volume-limited”, i.e. samples that – if properly selected – should have a constant density within the specific redshift bin. One such case is shown in the zoom of Fig. 6 where the redshift distribution in the range $z = [0.5, 0.8]$ is plotted.

An alternative way to generate the radial distribution of the random sample is to integrate the galaxy LF in steps along the redshift direction, computing at each step a value for the density of galaxies expected at that redshifts. Ideally, the LF can be measured from the sample itself and would include any detected evolution of its parameters. This is what we did here, using the evolving LF parameters presented in the companion dedicated paper (Zucca et al. 2009). The dashed red line in Fig. 6 shows the result one obtains if smoothing with a kernel of $450$ $h^{-1}$ Mpc, compared to the one obtained from the integration of the LF. The latter is fully consistent with what is expected from a truly volume-limited sample with the given selection criteria, with the number of objects increasing as the square of the radial co-moving distance.

**Fig. 5.** Overall radial distribution of the zCOSMOS 10K sample, compared to three different smoothed distributions. These are obtained by filtering the observed data with a Gaussian kernel of increasing $\sigma = 150, 250$, and $450$ $h^{-1}$ Mpc. The first two smoothed curves retain information on the two large structures located at $\sim 1000$ and $\sim 1800$ $h^{-1}$ Mpc along the line of sight, while the third one overestimates the number density of galaxies at low and high redshifts.
6.3. Effect of mass incompleteness on \(w_p(r_p)\)

As discussed in Sect. 4.3 when we constructed our mass-limited samples, a fraction of the galaxies more massive than the formal mass threshold are in fact lost because of the limiting \(I_{AB} < 22.5\) flux cut of the survey. This becomes more and more important with increasing redshift. As said, this population of missing galaxies is inevitably dominated by red objects with high mass-to-light ratio (Meneux et al. 2008), which are known to cluster more strongly than the average population (Meneux et al. 2006; Coil et al. 2008; McCracken et al. 2008). Having defined our clustering tools, we can now further extend the analysis of Sect. 4.3 and use the DLB40 mock samples to directly quantify the effect this has on the measured \(w_p(r_p)\). We thus computed
The derivation of realistic errors on the galaxy correlation function to very small masses ($\sim 10^8 M_\odot$), with and without applying the apparent-magnitude cut. Clearly, we are making a very strong hypothesis here, i.e. that the simulated samples have intrinsic clustering properties (and their relation to the galaxy’s M/L ratio), which are similar to those of real data.

The ratio of these two estimates (“true” over “observed”) averaged over the 40 mock catalogues is shown in Fig. 7. For a mass selection that is 100% complete within the given redshift bin, we would measure $\langle R \rangle = 1$ at all separations. We can see that the only mass range for which this is strictly happening at any redshift is the one with $\log(M/M_\odot) \geq 10.5$. For lower mass samples, we see a clear reduction of the clustering amplitude. However, we can also see that, for most samples, the shape of $w_p(r_p)$ is distorted mainly only below $1 h^{-1}$ Mpc. Above this scale, the mass incompleteness introduces an amplitude reduction up to $\sim 20\%$ in the worst cases. This will have to be considered when comparing our measurements with models (although keeping in mind that these estimates come from simulated data, not from real observations). For general comparisons, however, the amount of amplitude reduction of $w_p(r_p)$ is typically negligible on scales larger than $\sim 1 h^{-1}$ Mpc, given the statistical errors of the data measurements.

7. Systematic and statistical errors on correlation estimates

The derivation of realistic errors on the galaxy correlation function has been the subject of debate since its early measurements (see e.g. Bernstein 1994). In particular, it is well known that the measured values of the two-point correlation function are not independent on different scales. This means that, the bins of $w_p(r_p)$ have a degree of correlation among them, which needs to be taken into account when fitting a model to the observed values. This can be done if we are able to reconstruct the $N \times N$ covariance (or correlation) matrix of the $N$ bins (Fisher et al. 1994).

In a recent paper, Norberg et al. (2009) compare in detail three different methods for estimating the covariance matrix of a given set of measurements. These use a) the ensemble variance from a set of mock catalogues, reproducing as accurately as possible the clustering properties and selection function of the real data; b) a set of bootstrap resamplings of the volume containing the data; and c) a so-called jack-knife subset of volumes of the survey. In this latest case, the survey volume is divided into $N_C$ subvolumes and the statistics under study recomputed each time excluding one of the subparts. In the “block-wise” incarnation of the bootstrap technique (Porciani & Giavalisco 2002), instead, $N$ subvolumes are selected each time with repetition, i.e. excluding some of them, but counting two or more times some others as to always get a global sample with the same total volume. We note, however, that there are historically two possible ways of resampling internally the data set. The classical “old” bootstrap (Ling et al. 1986) entailed boot-strapping the sample “galaxy-by-galaxy”. This means each time randomly picking a sample of $N_C$ galaxies among our data set of $N_C$ galaxies, allowing repetitions. In this way, within one bootstrap realisation a galaxy can be selected more than once, while some others are never selected. This technique has been shown to generally lead to some underestimation of the diagonal errors (Fisher et al. 1994). Here we also directly test this aspect.

The advantage of using mock samples is that, under the assumption that these are a realistic realisation of the real data, they allow us to obtain a true ensemble average and standard deviation from samples with the same size as the data sample, including both Poissonian noise and cosmic variance. Unfortunately, the covariance properties derived from mock samples are not necessarily a good description of those of the real data, thus making the use of the derived covariance matrix (e.g. in model fitting) doubtful. Conversely, depending on the sample size, jack-knife or volume-bootstrap covariance matrices can exacerbate the peculiarities of some subregions, again not adequately representing the true covariance properties of the data.

For the present investigation, we put considerable effort in understanding how to best estimate a sensible covariance matrix for our $w_p(r_p)$ measurements. The available mock samples were crucial for allowing us to perform direct comparisons of the performances of the different techniques. After some initial attempts, we excluded the jack-knife method because of the limited size of the survey volume. We then directly compared the covariance matrices derived through the bootstrap technique and from the KW24 mock catalogues. For the bootstrap method, we decided
to directly test how galaxy- and volume-bootstrap were performing. We concentrated on the redshift range \( z = [0.5–0.8] \) by selecting simulated galaxies brighter than \( M_B - 5 \log(h) \leq -19.5 - z \). After computing the correlation function \( w(r_p) \) for all 24 mock samples, we constructed for each of them a) 100 galaxy-galaxy-bootstrap samples and b) 100 volume-volume-bootstrap samples. In the latter case, we considered 8 equal subvolumes, defined as redshift slices within the redshift range considered. The number of subvolumes was chosen as the best compromise between having enough of them and not having volumes that were too small. With this choice, their volume is \( \sim 1.4 \times 10^5 h^{-3} \) Mpc\(^3\) for the samples with \( z = [0.5–0.8] \) and \( z = [0.8–1.0] \) and \( \sim 0.6 \times 10^5 h^{-3} \) Mpc\(^3\) for \( z = [0.2–0.5] \). The two bootstrap techniques led to a total of 4800 samples and corresponding estimates of \( w(r_p) \). We then calculated the covariance (and correlation) matrices for each of these two cases, along with the one derived from the correlation function of the 24 mocks themselves.

In Fig. 8 we compare the standard deviations derived from the two bootstrap techniques, to those derived from the 24 mocks. In each case, these values correspond by definition to the square root of the diagonal elements of the covariance matrix. In the plot we show the mean (over the 24 mocks) of the ratio of \( \sigma_{ij} \) from the bootstrap to the “true” one from the ensemble of mock surveys. This shows clearly how the rms values obtained with the single-galaxy-bootstrap grossly underestimate the true variance, up to one order of magnitude on large scales. Bootstrapping by volumes produces a better result, providing a realistic estimate of \( \sigma_{ij} \) between 0.1 and 1 \( h^{-1} \) Mpc, and a 20–25% underestimate on larger scales.

Each element of the correlation matrix \( r_{ij} \) is obtained from the corresponding element of the covariance matrix \( \sigma_{ij} \) as

\[
r_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}.
\]

By definition, the off-diagonal terms of the correlation matrix will then range between \(-1\) and \(1\), indicating the degree of correlation between different scales of the function \( w_p(r_p) \). Considering the redshift range \( z = [0.5–0.8] \), we show in Fig. 9 the mean of the 24 correlation matrices derived by resampling the galaxies (left panel) or by resampling 8 equal subvolumes (centre), 100 times each. These are compared to the correlation matrix directly derived from the 24 mocks (right panel). The first case shows a mainly diagonal correlation matrix where off-diagonal terms are mostly noise. In the second case they instead decrease smoothly from 1 to 0 as a function of bin separation. The matrix derived from the 24 mocks shows high correlation on all scales.

In order to directly compare the properties of the correlation matrices derived with the 3 methods, we compute the principal components and the amplitudes of the corresponding eigenvalues \( \lambda_i \) (\(i = 1–12\)) for each of the 24 + 24 + 1 matrices. The sum of the eigenvalues of a correlation matrix is always equal to its dimension, i.e. 12 in our case. We report in Table 4 the values of the five main eigenvalues obtained with the 24 mocks (first column) compared to the averages over the 24 mocks of those obtained with the two resampling methods. The numbers show that the correlation matrix derived from the 24 mocks essentially contains four principal components and is mostly dominated by one of them. This indicates a strong correlation in the data. The bootstrap matrices, instead, show more than five non-negligible components, with the fifth one the same order of magnitude as the second in the mock matrix. This implies a lower correlation. We note, however, that volume resampling tends to produce a matrix whose structure is closer to that of the mocks, with 1–2 dominant components. This is another indication of how volume-bootstrapping, although not perfectly reproducing the intrinsic covariance properties of the sample, better estimates the variance and correlation in the data than does a galaxy-galaxy-bootstrap.

These experiments are extended and further discussed in our parallel accompanying papers, in particular Porciani et al. (in preparation). The bottom-line result of our extensive investigations is that a volume-bootstrap provides a good enough reconstruction of the intrinsic covariance matrix of the data set, if enough resamplings are used. This is obtained at the expense of a slightly less accurate account of cosmic variance on large scale than what can be obtained from the scatter among mock samples, where wavelengths longer than the survey size can be sampled. However, we have shown (Fig. 3) that this effect on scales \( \sim 10 \) \( h^{-1} \) Mpc is limited to \( \sim 20\% \).

![Fig. 9](image.png)

Table 4. The five main eigenvalues of the correlation matrix derived with the bootstrap resampling of galaxies (first column) and subvolumes (Col. 2), and from the ensemble variance of the 24 mocks (Col. 3).

| Eigenvalue | Bootstrap | Mocks |
|------------|-----------|-------|
|            | Galaxies  | Volumes |
| \(\lambda_1\) | 3.87784   | 8.28193 |
| \(\lambda_2\) | 1.95834   | 2.17866 |
| \(\lambda_3\) | 1.36841   | 0.71635 |
| \(\lambda_4\) | 1.11316   | 0.38252 |
| \(\lambda_5\) | 0.91683   | 0.20340 |

For the two latest cases, each mock sample is used in turn as “data” and the reported eigenvalues are the obtained as the average over the 24 mocks.
8. Dependence of galaxy clustering on luminosity

8.1. Luminosity dependence at fixed redshift

Figure 10 shows the projected correlation function $w_p(r_p)$ estimated for our nine luminosity-selected subsamples at different redshifts. Error bars correspond to the 1σ dispersion provided by 200 volume-bootstrap resamplings, as extensively discussed in Sect. 7.

No clear dependence on luminosity is observed within any of the three redshift ranges. Also, in the shape of $w_p(r_p)$, there is some hint of the usual “shoulder”, i.e. a change in slope around $1 \, h^{-1} \text{Mpc}$, but no clear separation between the expected 1-halo term on small scales and the 2-halo component above this scale (see the Introduction for definitions). In particular, in the intermediate-redshift bin, all subsamples show a rather flat large-scale slope, with no evidence of the usual breakdown above $\sim 2 \, h^{-1} \text{Mpc}$.

To try to understand the origin of the observed flat shape, it is interesting to look directly at the contour plots of the bi-dimensional redshift-space correlation function $\xi(r_p, \pi)$. These are shown in Fig. 11 for the three luminosity-selected subsamples L1.4, L2.2 and L3.1 (see Table 1 for definitions), which include galaxies brighter than $M_p - 5 \log(h) \leq -19.5 - z$. The three contour plots show some interesting features. First, one clearly notices the much stronger distortion along the line of sight $\pi$, in the central panel. At the same time, a much more extended signal is also observed along the perpendicular direction $r_p$ in the same redshift range. It is tempting to interpret both these effects as produced in some way by the two dominating structures that we showed in Fig. 6. The excess signal along the line of sight is very plausibly due to the distortions by “Fingers of God”, due to an anomalous number of virialised systems (groups and clusters) within these structures. At the same time, the extension along $r_p$ indicates that there is also an excess of pairs perpendicular to the line of sight, with respect to an isotropic distribution. In fact, we know (Guzzo et al. 2007; Scoville et al. 2007b) that the large-scale structure at $z \approx 0.73$ extends over a large part of the COSMOS area. This evidently biases the observed number of pairs along $r_p$, for simple geometrical reasons. We cannot exclude that part of the large-scale compression observed in $\xi(r_p, \pi)$ is also generated by an excess of galaxy infall onto this structure, thus producing what is known as the Kaiser effect (Kaiser 1987). This effect is proportional to the growth of structure (see e.g. Guzzo et al. 2008, for a recent direct estimate at similar redshift) and can be extracted when the underlying clustering can be assumed to be isotropic. In this case it is in practice impossible to disentangle this dynamical distortion from the geometrical anisotropy generated by having one dominating structure elongated perpendicular to the line of sight.

The flatter shape in $w_p(r_p)$ in Fig. 10 is also consistent with the overdense samples of Abbas & Shell (2007), who notice not only a higher amplitude for the most overdense (10% and 30%) samples of mock and SDSS galaxies, but also a flattening in the correlation function compared to the full sample. This is another line of evidence favouring the hypothesis that the zCOSMOS field is centred on an overdensity.

The plots of Fig. 11 also explicitly show the reasons for our choice of $\pi_{\max} = 20 \, h^{-1} \text{Mpc}$ as the upper integration limit in computing $w_p(r_p)$, a value that provides a reasonable compromise between including most of the signal and excluding the noisiest regions in the upper part the diagrams. In the central redshift bin, however, some real clustering power may still be present above this scale, for small $r_p$'s. In Fig. 12 we show directly how $w_p(r_p)$ changes, when $\pi_{\max}$ is extended from 20 to $30 \, h^{-1} \text{Mpc}$. We see that, somewhat counter intuitively, below $1 \, h^{-1} \text{Mpc}$, no extra amplitude is gained, while – as indicated by the mock experiments (see Sect. 5) – the scatter is increased. Conversely, one can see the slight scale-dependent bias on the amplitude at larger separations, which gets up to $\sim 10\%$ at $15 \, h^{-1} \text{Mpc}$ when increasing $\pi_{\max}$.

8.2. Redshift evolution at fixed (evolving) luminosity

In Sect. 4.1 we discussed how our luminosity selection was devised such as to account for the average evolution in the luminosity of galaxies, assuming this to be the dominant effect in modifying the mean density of objects at any given luminosity. Under this assumption, it is then interesting to test how $w_p(r_p)$ changes with redshift for galaxies within the same (de-evolved) luminosity interval. This is shown in Fig. 13 for $M_B - 5 \log(h) \leq -19.5 - z$, i.e. for the same three samples for which $\xi(r_p, \pi)$ is plotted in Fig. 11. No coherent evolution of the amplitude and shape of the projected correlation function is observed among the three samples, characterised by mean redshifts 0.37, 0.61, and 0.91. The three curves are consistent within the error bars, with – as expected – the intermediate-redshift bin (triangles) showing a systematically higher amplitude than the other two. Again, this is easily interpreted as a local effect, resulting from the extreme large-scale clustering observed in this redshift range. On the other hand, the overall lack of apparent amplitude evolution of luminous galaxies with redshift is consistent with previous results from the VVDS-Deep (Pollo et al. 2006), DEEP2 (Coil et al. 2006), and SDSS (Zehavi et al. 2005) surveys for galaxies brighter than $\sim M^*$. The only evolutionary effect shown in
Fig. 11. Iso-correlation contours of $\xi(r_p, \pi)$ (here smoothed with a Gaussian kernel) for galaxies brighter than $M_B - 5 \log(h) \leq -19.5 - z$, computed in 3 redshift ranges. The amplitude is colour-coded according to the side bar, while the black contour corresponds to $\xi(r_p, \pi) = 1$. White values correspond to $\xi(r_p, \pi) < 0.4$.

Fig. 12. Sensitivity of the projected function $w_p(r_p)$ to the upper integration limit $\pi_{\text{max}}$ for one luminosity-selected sample in the central redshift bin.

9. Dependence of galaxy clustering on stellar mass

The relation of clustering properties to galaxy stellar masses is in principle more informative and straightforward for interpreting because stellar mass is a more fundamental physical parameter than luminosity.

9.1. Mass dependence at fixed redshift

Also in the case of stellar mass dependence, it is interesting to look at the shape of the full correlation function $\xi(r_p, \pi)$ in the three redshift ranges. We show in Fig. 14 the result obtained for the 3 samples M1.3, M2.3, and M3.2 (see Table 2), which include galaxies more massive than $10^{10} h^{-2} M_\odot$. Also in this case, the central panel is significantly different from the other two, with $\xi(r_p, \pi)$ remaining positive out to much larger scales in both $r_p$ and $\pi$ directions. We note that the small-scale stretching along $\pi$ seems to be less extended than the one obtained for the corresponding luminosity-selected sample, although it is hard to say whether this difference is significant.

Figure 15 shows the projected correlation function $w_p(r_p)$ of the 10 mass-selected samples. The plotted points are not corrected for the residual stellar mass incompleteness (see Sect. 6.3). Errors are estimated as in the luminosity case using 200 bootstrap resamplings of 8 equal subvolumes of each data.

Fig. 13. Sensitivity of the projected function $w_p(r_p)$ to the upper integration limit $\pi_{\text{max}}$ for one luminosity-selected sample in the central redshift bin.
Fig. 13. Evolution of the projected function $w_p(r_p)$ of galaxies with $M_B - 5 \log(h) \leq -19.5 - z$ between redshift $z = 0.2$ and $z = 1$.

Fig. 14. Example of full redshift-space correlation function $\xi(r_p, \pi)$ (here smoothed with a Gaussian kernel) for the subsample containing galaxies more massive than $10^{10} h^{-2} M_\odot$, computed in the same 3 redshift ranges as in Fig. 11. $\xi(r_p, \pi)$ is computed in cells of 1 $h^{-1}$ Mpc side in both $r_p$ and $\pi$ and the iso-correlation levels are colour-coded according to the side bar. The thick black contour corresponds to $\xi(r_p, \pi) = 1$ and the white values to $\xi(r_p, \pi) < 0.4$.

Fig. 15. The projected correlation function $w_p(r_p)$ as a function of galaxy stellar masses in the zCOSMOS 10K sample, in three redshift ranges.
set. The figure shows a weak mass dependence of clustering in the low- and high-redshift bins, in particular at small separations. At the same time, a strong dependence at all separations is evident in the intermediate-redshift slice. There, the slope of $w_p(r_p)$ remains extremely flat out to the largest explored scales, even more strongly than in the luminosity-selected cases. Finally, in the low- and high-redshift bins there is evidence of a steeper “1-halo term” contribution at $r_p < 1 h^{-1}$ Mpc, where a sudden steepening is then observed. The slope below $0.2 h^{-1}$ Mpc seems to depend directly on the limiting mass, with more massive galaxies showing a steeper correlation function. In summary, no clear overall trend can be shown among the three redshift ranges, with the central volume again displaying peculiar clustering properties that apparently dominate any possible cosmological effect.

9.2. Clustering evolution at fixed stellar mass

It is also interesting to directly compare the evolution of $w_p(r_p)$ with redshift, when a specific class of stellar mass is selected. As mentioned earlier in this section, this is particularly interesting because it should not require accounting for strong galaxy evolutionary trends as in the case of luminosity. We are assuming here that stellar mass does not grow significantly between $z \sim 1$ and $z \sim 0.2$, which we know is only partially true. A factor of ~2 growth in stellar mass is in fact expected on average between $z = 1$ and $z = 0$, which would have little effect on the estimated correlation function, however, if taken into account.

In Fig. 16 we show $w_p(r_p)$ computed for the same three samples with $\log(M/M_\odot) \geq 10$ of Fig. 14. Similarly to the luminosity samples, we do not see any clear evolution with redshift of the amplitude and shape of the projected correlation function. The three curves are consistent with each other within $1\sigma$.

10. Comparison with independent measurements and models

10.1. Redshift evolution of $w_p(r_p)$

An accurate $z \sim 0$ reference measurement of $w_p(r_p)$ as a function of stellar mass has been obtained by the SDSS (Li et al. 2006). Meneux et al. (2008) do find evidence for evolution of the amplitude of $w_p(r_p)$ for galaxies less massive than $10^{10.5} h^{-2} M_\odot$, when comparing this to the measurements from VVDS-Deep at $z \sim 0.9$. The SDSS and zCOSMOS stellar masses were derived with the same initial mass function (Chabrier 2003) and normalised to $h = 1$. They are directly comparable. The SDSS clustering measurements were obtained within differential stellar mass ranges (Li et al. 2006), while ours correspond to galaxies that are more massive than a given threshold. However, from Fig. 2 we see that the zCOSMOS sample includes a very small number of galaxies that are more massive than $10^{11} h^{-2} M_\odot$, due to the much smaller volume than for the SDSS. Any of our mass-selected samples has therefore, in practice, an upper bound at this value of mass. This implies that we can coherently compare two of the SDSS measurements of $w_p(r_p)$ (for their galaxy samples with stellar masses in the ranges [10.0–10.5] and [10.5–11.0]) to those from our samples M3.1 and M3.2, which include galaxies more massive than $10^{10}$ and $10^{10.5} h^{-2} M_\odot$. 
respectively, within the redshift range \( z = [0.8–1.0] \). This comparison (Fig. 17) does not show a clear evolution with redshift. For both samples, the large-scale amplitude of \( w_p(r_p) \) is virtually the same as in the local SDSS samples. Considering a simple evolution of structures, this implies that the bias for galaxies more massive than \( 10^{10} \, h^{-2} \, M_\odot \) has evolved significantly between \( z \sim 1 \) and today, as to keep their apparent clustering amplitude substantially unchanged. This implies in practice that the bias \( b(z) \) must evolve in such a way that \( b(0)D(0) = b(z)D(z) \), where \( D(z) \) is the linear growth rate of density fluctuations. In the standard model, this implies that, at the approximate mean redshifts of our redshift bins, \( z \simeq 0.35, 0.75, 0.9 \), the bias of massive galaxies must have been, respectively, 1.2, 1.44, and 1.53 times its value at the current epoch. Meneux et al. (2008) observed the same effect at \( z \sim 0.8 \) in the VVDS data but only for galaxies more massive than \( \sim 10^{10.5} \, h^{-2} \, M_\odot \), with lower-mass objects showing a weaker bias evolution. A lack of any evolution of the clustering of the most massive galaxies was also noticed in the NDWFS (Brown et al. 2008) and 2SLAQ surveys (Wake et al. 2008).

### 10.2. Observed and predicted shape of \( w_p(r_p) \) at \( 0.5 < z < 1 \)

The only available measurement of clustering as a function of stellar mass at redshifts comparable to those explored by our sample comes from the VVDS-Deep survey (Meneux et al. 2008) at \( 0.5 < z < 1.2 \). VVDS-Deep goes 1.5 magnitude deeper (although over a smaller area of \( \sim 0.5 \, \text{deg}^2 \)) which allows the analysis to be extended beyond \( z = 1 \). To provide a qualitative, yet meaningful comparison of these two data sets, we can recompute the correlation function for the 10K data within the widest usable redshift range overlapping with the VVDS interval, i.e. \( 0.5–1.0 \). We applied the same stellar mass selection limits, keeping in mind the residual incompleteness that will affect the highest redshift part of the sample. The result is shown in Fig. 18, where the VVDS and zCOSMOS mass-selected samples are directly compared. The result is shown in Fig. 18, where the VVDS and zCOSMOS mass-selected samples are directly compared. The difference in shape and amplitude in the \( w_p(r_p) \) derived from the two data sets is rather striking. The zCOSMOS points show in general a much flatter relation than those from the VVDS. The amplitudes for a given mass selection also seem to be incompatible at several standard deviations between the two samples, especially above 1 \( h^{-1} \, \text{Mpc} \).

### 10.3. Comparison to analytic and semi-analytic models

At this point, it is relevant to compare the available observations from zCOSMOS and VVDS with model predictions in a standard \( \Lambda \)CDM scenario. We can do this in two ways. We first used the HALOFIT public code (Smith et al. 2003), which uses the halo model to compute the expected nonlinearly evolved power spectrum at \( z = 0.8 \), which we take as a reasonable mean redshift for the two samples. Our conclusions would not differ at all if predictions for \( z = 0.7 \) or 0.9 were used. The corresponding projected function \( w_p(r_p) \) is then computed by Fourier-transforming the power spectrum and projecting the resulting real-space correlation function. The result gives the expected \( w_p(r_p) \) of the mass density field at \( z = 0.8 \). Second, we can compute the expected value and the scatter expected in the same redshift range \( 0.5 < z < 1 \) for \( w_p(r_p) \) using the available semi-analytic mock surveys built from the Millennium run. To this end, we used the DLB40 mocks for which we have full control over stellar masses, selecting simulated galaxies with \( I_{AB} \leq 22.5 \) and \( \log(M/M_\odot) \geq 10 \) and reproducing the sampling rate of the 10K data. In Fig. 19 we plot the HALOFIT prediction both for the dark matter and for an arbitrarily biased population of halos with \( b = \sqrt{5/6} \sim 1.61 \), together with the DLB40 mean \( w_p(r_p) \) and the corresponding 1\,\sigma and 3\,\sigma scatter corridor from the 40 mock surveys. As a consistency check, we note the rather good agreement between the analytic HALOFIT result and the expected value from the full N-body plus semi-analytic simulation. On these model predictions, we overplot the corresponding zCOSMOS and VVDS estimates. The zCOSMOS points agree with the models at better than 68% confidence in both shape and amplitude on scales smaller than 1 \( h^{-1} \, \text{Mpc} \). On larger scales, however, the observed \( w_p(r_p) \) would require a strongly scale-dependent bias to be compatible with the model predictions. This scale-dependence would also behave contrary to what models and very general considerations suggest, implying a bias that grows with scale, rather than declines. From the plot we see in fact that the 10K data are compatible with \( b \simeq 1.6 \) on small scales, but would require \( b \approx 2.45 \) on 10\,\text{\,h}^{-1} \, \text{Mpc} \) scales. The shaded area shows that this large-scale excess is marginally compatible with the model predictions, representing a very strong positive fluctuation. A few percent of volumes this size would show this high clustering amplitude (on these scales and for this kind of galaxies), in a \( \Lambda \)CDM Universe.
Fig. 18. Direct comparison of the dependence of clustering on stellar mass in the VVDS-Deep and zCOSMOS samples, over a similar redshift range.

Fig. 19. The zCOSMOS (solid circles) and VVDS-Deep (open circles) $w_p(r_p)$ of galaxies with mass larger than $10^{10} h^{-2} M_\odot$ and $\langle z \rangle \sim 0.8$, compared to model predictions. These include the nonlinear mass projected correlation function computed using HALOFIT (lower solid line, Smith et al. 2003), and, for reference, the corresponding $w_p(r_p)$ for a population of haloes with bias $b^2 = 2.6$ (dashed curve). The latter curve is a very good description of the full nonlinear $w_p(r_p)$ (light solid blue line), obtained averaging the 40 DLB40 mocks from the Millennium run after applying the same sampling, magnitude and mass selections of the 10K sample. The shaded areas (thick and thin shaded) give, respectively, the 1$\sigma$ and 3$\sigma$ confidence corridors around the mean.
Fig. 20. The redshift-space, angle-averaged correlation function \(\xi(s)\) of the mass-selected samples in the redshift bin \([0.5, 1]\). A mild systematic mass dependence is visible.

It is interesting to note, at the same time, how the VVDS measurements lie on the opposite side of the distribution, at about 1.5–2\(\sigma\) from the mean, but with a shape that is compatible with the model prediction over the whole range (corresponding to a linear bias \(b \sim 1.2\)). Based on the results of Abbas & Sheth (2007), one would also conclude that the central volume of the COSMOS survey is dominated by overdense regions, while these should be slightly under-represented in the corresponding volume of the VVDS survey.

These results show how a full HOD model fitting to the \(w_p(r_p)\) measured from the 10K data – originally planned for this paper – would add no meaningful information to the current analysis. Our first experiments with HOD models based on the universal halo mass function indicate that rather unrealistic sets of parameters are required to reproduce the observed function. An interesting possibility would be to use in such modelling a halo mass function that depends on local environment (e.g. Abbas & Sheth 2005, 2006), to consider the evidence that a large part of this sample is dominated by an overdensity. We plan to explore this possibility using the larger 20K zCOSMOS sample that is now nearly complete.

11. Discussion

Together with previous analyses (McCracken et al. 2007; Kováč et al. 2009), these results suggest that a significant fraction of the volume of Universe bounded by the COSMOS field is indeed characterised by particularly extreme density fluctuations. We have seen how, in statistical terms, these seem to lie at the 3\(\sigma\) limit of the distribution of amplitudes expected in volumes of a few \(10^6 h^{-3} \text{Mpc}^3\). We should consider, however, that these conclusions are drawn from measurements that are strongly affected by the angular distribution of structure. The McCracken et al. (2007) result is based on the angular correlation function, while here we studied the projected function \(w_p(r_p)\). Although making use of the redshift information, the latter is in practice a clustering measure dominated by galaxy pairs lying almost perpendicular to the line of sight. The underlying assumption when measuring \(w_p(r_p)\) is that the geometrical distribution of structures within the sample being analysed is completely isotropic; in other words, that there are superclusters aligned along several directions, such that the only remaining radial signal is produced by galaxy-peculiar velocities. The very reason for using \(w_p(r_p)\) is indeed to get rid of the distortions introduced in the shape of \(\xi(s)\) (the redshift-space, angle-averaged correlation function) by galaxy motions. If this is true, and only in this case, then \(w_p(r_p)\) is fully equivalent to an integral over the real-space correlation function \(\xi(r)\), so it carries the same cosmological information. However, if, as in the case we have encountered here, there is one or more dominating structures extending preferentially along one direction, then the use of \(w_p(r_p)\) to infer cosmological information is inappropriate.

One may thus wonder whether more robust cosmological information could instead be inferred by looking directly at the simplest, angle-averaged redshift-space correlation function \(\xi(s)\). The expectation is that the average over all directions reduces the weight of the excess pair counts produced by just a few structures oriented along one preferred direction. In such case any analytic modelling (e.g. with HOD models) should also include an appropriate model for the linear and nonlinear redshift distortions (Scoccimarro 2004; Tinker et al. 2007). More simply, we can use the available mock samples in redshift-space to compute the nonlinear redshift-space \(\xi(s)\) and its variance and compare it to the data, as we did for \(w_p(r_p)\). In
Fig. 20: we first plot $\xi(s)$ for the 10K sample, computed for the usual four mass ranges in the broad redshift range $0.5 < z < 1$. The four data sets show a smooth power-law behaviour, with some evidence of a mass dependence of the clustering amplitude, in particular at the upper mass limit. The overall shape is well described by a rather flat power-law $\xi(s) \propto (s_0/s)^{\gamma}$, with slope $\gamma \sim 1$ and a correlation length $s_0$ between 6 and $\sim 10$ $h^{-1}$ Mpc. These values for the shape and amplitude of $\xi(s)$ are similar to those measured for luminous red galaxies at $z = 0.55$ in the 2SLAQ survey (see Fig. 7 in Ross et al. 2007) and for luminous early-type galaxies in the 2dFGRS (Norberg et al. 2002). This is consistent with the most massive objects in the 10K sample being predominantly red, early-type galaxies that show moderate or no evolution in the overall clustering amplitude with redshift.

In Fig. 21 instead, we compare $\xi(s)$ of our “reference” sample with $\log(M/M_\odot) \geq 10$ to the mean and scatter (at 1$\sigma$ and 3$\sigma$ confidence, respectively) of the similarly-selected set of DLB40 mocks. Despite the angular average, we note a behaviour that is similar to what is observed in $w_p(r_p)$, although now the agreement extends to slightly larger scales. The observed clustering is compatible with the predictions of the standard model (to better than the 68% level) on scales smaller than $\sim 2 h^{-1}$ Mpc. On larger scales, $\xi(s)$ also shows excess power with respect to the models, which places the zCOSMOS volume at the upper 3$\sigma$ limit of the statistical distribution obtained from the mocks. This exercise shows that, even after angle-averaging our clustering estimator, the amount of structure present in this specific volume of the Universe remains outstanding in comparison to the model expectations. The conclusion can only be that we have either been very unlucky in the selection of the COSMOS field and picked up a fluctuation with a probability of $\sim 1\%$ to be found in such a volume or fluctuations with this amplitude are in reality more common than what the standard cosmology predicts.

12. Summary

We used the 10K zCOSMOS spectroscopic sample to study galaxy clustering as a function of galaxy luminosity and stellar mass, in the range of redshift $[0.2, 1]$. To this end, we built luminosity and mass-selected samples from the 10K catalogue sampling three separate redshift ranges. We used mock catalogues to quantify the effect of stellar mass incompleteness on the measured clustering, as a function of redshift. We carefully checked our covariance and error estimate techniques, comparing the performances of methods based on the scatter in the mocks and on bootstrapping schemes. We adopted the latter, based on 200 re-samplings of 8 subvolumes of the survey, as the most appropriate description of the covariance properties of the data.

By measuring the redshift-space correlation functions $\xi(s)$ and $\xi(r_p, \pi)$ and the projected function $w_p(r_p)$ for these subsamples, we found the following results.

- Surprisingly, we do not see any clear dependence on luminosity of the correlation function at all redshifts. This is at odds with results in the local Universe by the 2dFGRS and with measurements at similar redshift by the VVDS and DEEP2 surveys, which found a significant steepening of $w_p(r_p)$ with luminosity.
- We find a mildly more evident (although not striking) dependence of $w_p(r_p)$ on stellar mass, especially on small scales. The central redshift bin ($0.5 < z < 0.8$) displays in general a more evident effect, with a very flat shape of $w_p(r_p)$...
on scales $r_p = [1-10] \ h^{-1} \ Mpc$. The overall shape of the corresponding map of $\xi(r_p, \pi)$ shows strong distortions that we interpret as the effect of dominant structure extending preferentially perpendicular to the line of sight.

- From comparison to the SDSS measurements at $z \sim 0$, we do not see any significant evolution with redshift of the amplitude of clustering for bright and/or massive galaxies. Together with previous results from VVDS, this is consistent with a more rapid evolution of the linear bias for the most massive objects with respect to the general population. In the $\z$COSMOS sample, this invariance in the clustering amplitude between $z \sim 1$ and $z \sim 0$ seems to remain valid down to lower masses than in the VVDS, an effect easily explained by the overall larger clustering amplitude observed in general for $z > 0.5$ in this sample. This is shown by a much flatter shape (higher amplitude) of $w(r_p)$ of $\z$COSMOS galaxies with respect to VVDS galaxies, when selected with the same criteria.

- This particularly high level of structure is confirmed by comparing the measured $w(r_p)$ and $\xi(s)$ at $0.5 < z < 1$ with model predictions, concentrating on the sample with $\log(M/M_\odot) \geq 10$. On scales smaller than $\sim 1-2 \ h^{-1} \ Mpc$, the observations agree very well with the model expectation values in the standard CDM scenario for a linear bias $b \sim 1.6$. On these scales, the measured values are compatible to better than 68% with the DLB40 mocks. On larger scales, however, the observed clustering amplitude is reproduced in only a few percent of the mocks. In other words, if the shape of the power spectrum is that of ACDM and the bias has no “innatural” scale-dependence, COSMOS has picked up a volume of the Universe that is rare, $2-3 \sigma$ of the power spectrum.

Acknowledgements. We thank the anonymous referee for a detailed review of the manuscript that helped to improve the paper. We thank G. De Lucia, J. Blaizot, S. Phelegs, and A. Sánchez for their thorough comments on an early version of the manuscript. This work was supported by Grant ASI/COFIN/WP3110 I/026/07/0.

References

Abbàs, U., & Sheth, R. K. 2005, MNRAS, 364, 1327
Abbàs, U., & Sheth, R. K. 2006, MNRAS, 372, 1749
Abbàs, U., & Sheth, R. K. 2007, MNRAS, 378, 641
Abbàs, U., de la Torre, S., Le Fèvre, O., et al. 2009, A&A, submitted
Arnouts, S., Walcher, C. J., Le Fèvre, O., et al. 2007a, A&A, 476, 137
Bell, E. F., Naab, T., McIntosh, D. H., et al. 2006, ApJ, 640, 241
Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A., & Schaeffer, R. 1996, ApJ, 472, 452
Bernstein, G. M. 1994, ApJ, 424, 569
Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476
Bolzonella, M., Kovač, K., Pozzetti, L., et al. 2009, A&A, submitted
Blanton, M. R., Sesar, B., Kauffmann, G., Jing, Y. P., et al. 2006, MNRAS, 368, 21
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, submitted
Ling, E. N., Barrow, J. D., & Frenk, C. S. 1986, MNRAS, 223, 21
Maurogordato, S., & Lachièze-Rey, M. 1991, ApJ, 369, 30
McCracken, H. J., Peacock, J. A., Guzzo, L., et al. 2007, ApJS, 172, 314
McCracken, H. J., Ilbert, O., Mellier, Y., et al. 2008, A&A, 479, 321
Meneux, B., Le Fèvre, O., Guzzo, L., et al. 2006, A&A, 452, 387
Meneux, B., Guzzo, L., Garilli, B., et al. 2008, A&A, 478, 299
Norberg, P., Baugh, C. M., Hawkins, E., et al. 2001, MNRAS, 328, 64
Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002, MNRAS, 332, 827
Norberg, P., Baugh, C. M., Gaztanaga, E., & Croton, D. J. 2009, MNRAS, 396, 19
Peebles, P. J. E. 1980, The Large Scale Structure of the Universe (Princeton: Princeton University Press)
Phèps, S., Peacock, J. A., Meisenheimer, K., & Wolf, C., 2006, A&A, 457, 145
Pollo, A., Meneux, B., Guzzo, L., et al. 2005, A&A, 439, 887
Pollo, A., Guzzo, L., Le Fèvre, O., et al. 2006, A&A, 451, 409
Porciani, C., & Giavalisco, M. 2002, ApJ, 565, 24
Pozzetti, L., Bolzonella, M., Lamareille, F., et al. 2007, A&A, 474, 443
Pozzetti, L., Bolzonella, M., Zacki, E., et al. 2009, A&A, submitted
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Rettura, A., Rosati, P., Strazzullo, V., et al. 2006, A&A, 458, 717
Ross, N. P., da Angela, J., Shanks, T., et al. 2007, MNRAS, 381, 573
Scoccimarro, R. 2004, Phys. Rev. D, 70, 083007
Scoville, N., Abraham, R. G., Aussel, H., et al. 2007a, ApJS, 172, 38
Scoville, N., Aussel, H., Benson, A., et al. 2007b, ApJS, 172, 38
Skibba, R. A., Sheth, R. K., & Martino, M. C. 2007, MNRAS, 382, 1940
Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137
Stringer, M. J., Benson, A. J., Bundy, K., Ellis, R. S., & Quetin, E. L. 2009, MNRAS, 143
Tinker, J. L., Norberg, P., Weinberg, D. H., & Warren, M. S. 2007, ApJ, 659, 877
Wake, D. A., Sheth, R. K., Nichol, R. C., et al. 2008, MNRAS, 387, 1045
Wang, L., Li, C., Kauffmann, G., & De Lucia, G. 2006, MNRAS, 371, 537
Wang, L., Li, C., Kauffmann, G., & De Lucia, G. 2007, MNRAS, 377, 1419
Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, ApJ, 646, 881
Warren, M. S., & Rees, M. J. 1978, MNRAS, 183, 341
Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, ApJ, 630, 1
Zheng, Z., Coil, A., & Zehavi, I. 2007, ApJ, 667, 760
Zucca, E., Ilbert, O., Baldieri, S., et al. 2006, A&A, 455, 879
Zucca, E., Bardelli, S., Bolzonella, M., et al. 2009, A&A, submitted

1 Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse, 85748 Garching-bei-München, Germany
e-mail: bmeneux@mpe.mpg.de

2 Universitäts-Sternwarte, Scheinerstrasse 1, Munich 81679, Germany

3 INAF – Osservatorio Astronomico di Brera, Via Bianchi 46, 23807 Merate (LC), Italy
