Publication Year	2015
Acceptance in OA	2020-03-17T14:04:35Z
Title	The galaxy-halo connection from a joint lensing, clustering and abundance analysis in the CFHTLenS/VIPERS field
Authors	Coupon, J., Arnouts, S., van Waerbeke, L., Moutard, T., Ilbert, O., van Uitert, E., Erben, T., GARILLI, BIANCA MARIA ROSA, Guzzo, L., Heymans, C., Hildebrandt, H., Hoekstra, H., Kilbinger, M., Kitching, T., Mellier, Y., Miller, L., SCODEGGIO, MARCO, Bonnett, C., Davidzon, I., DE LUCIA, GABRIELLA, Fritz, A., Fu, L., Hudelot, P., Hudson, M. J., Kuijken, K., Leauthaud, A., Le Fèvre, O., McCracken, H. J., Moscardini, L., Rowe, B. T. P., Schrabback, T., Semboloni, E., Velander, M., Branchini, Enzo Franco
Publisher's version (DOI)	10.1093/mnras/stv276
Handle	http://hdl.handle.net/20.500.12386/23321
Journal	MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY
Volume	449

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **449,** 1352–1379 (2015) doi:10.1093/mnras/stv276

The galaxy–halo connection from a joint lensing, clustering and abundance analysis in the CFHTLenS/VIPERS field

J. Coupon,[1★] S. Arnouts,[2] L. van Waerbeke,[3] T. Moutard,[2] O. Ilbert,[2] E. van Uitert,[4] T. Erben,[4] B. Garilli,[5] L. Guzzo,[6,7] C. Heymans,[8] H. Hildebrandt,[4] H. Hoekstra,[9] M. Kilbinger,[10] T. Kitching,[11] Y. Mellier,[10,12,13] L. Miller,[14] M. Scodeggio,[5] C. Bonnett,[15] E. Branchini,[16,17,18] I. Davidzon,[19,20] G. De Lucia,[21] A. Fritz,[5] L. Fu,[22] P. Hudelot,[12] M. J. Hudson,[23,24] K. Kuijken,[9] A. Leauthaud,[25] O. Le Fèvre,[2] H. J. McCracken,[12,13] L. Moscardini,[19,20,26] B. T. P. Rowe,[27] T. Schrabback,[4] E. Semboloni[4] and M. Velander[14]

Affiliations are listed at the end of the paper

Accepted 2015 February 8. Received 2015 January 15; in original form 2014 October 23

ABSTRACT

We present new constraints on the relationship between galaxies and their host dark matter haloes, measured from the location of the peak of the stellar-to-halo mass ratio (SHMR), up to the most massive galaxy clusters at redshift $z \sim 0.8$ and over a volume of nearly 0.1 Gpc^3. We use a unique combination of deep observations in the CFHTLenS/VIPERS field from the near-UV to the near-IR, supplemented by ∼60 000 secure spectroscopic redshifts, analysing galaxy clustering, galaxy–galaxy lensing and the stellar mass function. We interpret our measurements within the halo occupation distribution (HOD) framework, separating the contributions from central and satellite galaxies. We find that the SHMR for the central galaxies peaks at $M_{\rm h,peak} = 1.9^{+0.2}_{-0.1} \times 10^{12}\,{\rm M}_{\odot}$ with an amplitude of 0.025, which decreases to ∼0.001 for massive haloes ($M_{\rm h} > 10^{14}\,{\rm M}_{\odot}$). Compared to central galaxies only, the total SHMR (including satellites) is boosted by a factor of 10 in the high-mass regime (cluster-size haloes), a result consistent with cluster analyses from the literature based on fully independent methods. After properly accounting for differences in modelling, we have compared our results with a large number of results from the literature up to $z = 1$: we find good general agreement, independently of the method used, within the typical stellar-mass systematic errors at low to intermediate mass ($M_{\star} < 10^{11}\,{\rm M}_{\odot}$) and the statistical errors above. We have also compared our SHMR results to semi-analytic simulations and found that the SHMR is tilted compared to our measurements in such a way that they over- (under-) predict star formation efficiency in central (satellite) galaxies.

Key words: gravitational lensing: weak – galaxies: clusters: general – cosmology: observations – dark matter.

1 INTRODUCTION

The last few years have seen an increasing interest in statistical methods linking observed galaxy properties to their dark matter haloes, owing to the availability of numerous large scale multi-wavelength surveys. Those techniques are based on the assumption that the spatial distribution of dark matter is predictable and one is able to match its statistical properties with those of the galaxies. The halo model (see Cooray & Sheth 2002) is a quantitative representation of the distribution of dark matter, characterized by three main ingredients: the halo mass function describing the number density of haloes per mass, the halo bias tracing the clustering amplitude and the halo density profile.

Galaxies are born and evolve in individual haloes where the baryonic gas condensates, cools and forms stars. Galaxies are gravitationally bound to dark matter and share a common fate with their host, e.g. during mergers. Although we understand qualitatively individual physical processes likely to be involved in galaxy evolution, a number of key answers are missing.

★E-mail: jean.coupon@unige.ch

Published by Oxford University Press on behalf of the Royal Astronomical Society

Observations show that a fraction of galaxies experienced star formation quenching and have become passive, shaping the galaxy population into a bimodal blue/red distribution (Faber et al. 2007; Ilbert et al. 2013). The number of these passive galaxies is higher today than in the past and increases with increasing halo mass. Might feedback processes in massive haloes be responsible for this, or is there a universal critical stellar mass above which star formation ceases, independently of the halo mass? Studying the connection between galaxies and their host haloes is crucial to answer these questions.

Another enigmatic question is the low stellar mass fraction in low-mass haloes, seen in early studies connecting galaxies to their host haloes (Yang, Mo & van den Bosch 2003; Vale & Ostriker 2006; Zheng, Coil & Zehavi 2007). In fact, when measuring the stellar-to-halo mass ratio (SHMR) as a function of time, we observe that stellar mass is building up asymmetrically, first in massive haloes, later on in low-mass haloes (Conroy, Wechsler & Kravtsov 2006; Behroozi et al. 2013b). This asymmetry in the SHMR is one corollary of the so-called galaxy *downsizing* effect (Cowie et al. 1996). In low-mass haloes, stellar winds and supernovae may slow down star formation until the potential well grows deep enough to retain the gas and increase the star formation rate (SFR). Again, it becomes necessary to relate galaxy properties to their host halo mass.

A number of studies have related galaxy properties to dark matter haloes using the *abundance matching* technique (Marinoni & Hudson 2002; Conroy et al. 2006; Behroozi, Conroy & Wechsler 2010; Guo et al. 2010; Moster et al. 2010), which employs the stellar mass (or luminosity) function and the halo mass function to match halo–galaxy properties based on their cumulative abundances. The *conditional luminosity function* technique proposed by Yang et al. (2003) includes a parametrized $M_\star - M_{\rm h}$ relationship whose parameters are fitted to the luminosity function. Both this formalism and recent abundance matching studies feature a scatter in $M_\star$ at fixed $M_{\rm h}$, which is an important ingredient to account for, given the steep relation between the two quantities at high mass.

More recently, models adopting a similar approach to abundance matching consist of directly populating dark matter haloes from N-body simulations, to reproduce the observed stellar mass functions as a function of redshift, using a parametrized SFR model to account for redshift evolution (Moster, Naab & White 2012; Behroozi et al. 2013b).

Except in some rare cases where central or satellite galaxies can be individually identified (e.g. George et al. 2011; More et al. 2011), in studies based on luminosity or stellar mass distributions, the satellite galaxies' properties cannot be disentangled from those of the central galaxies. To remedy the problem, abundance matching techniques either assume an ad hoc fraction of satellites or use a subhalo mass function estimated from numerical simulations. Unfortunately, as subhaloes may be stripped and disappear after being accreted on to larger haloes, the subhalo mass function at the time considered might not correspond to the distribution of satellites, and one must consider the mass of subhaloes at the time of accretion, further extrapolated to the time considered. Obviously these complications limit the amount of information one can extract about galaxy satellites.

Galaxy clustering, on the other hand, allows separation of the contributions from central and satellite galaxies due to the different typical clustering scales. To model the clustering signal of a given galaxy population, the halo occupation distribution (HOD) formalism assumes that the galaxy number per halo is solely a function of halo mass and that the galaxy satellite distribution is correlated to that of the dark matter (Berlind & Weinberg 2002; Kravtsov et al. 2004).

One achievement of HOD modelling was to demonstrate from simulations (Berlind et al. 2003; Moster et al. 2010) that only a handful of parameters was necessary to fully describe galaxy–halo occupation. This parametric HOD was fitted to a number of observations over a large range of redshifts and galaxy properties. Among the more remarkable results are the local Universe galaxy clustering and abundance matching studies performed on the Sloan Digital Sky Survey (SDSS; see e.g. Zehavi et al. 2011) and at higher redshifts (Foucaud et al. 2010; Wake et al. 2011; Coupon et al. 2012; de la Torre et al. 2013; Martinez-Manso et al. 2015).

However, some underlying assumptions on the distribution of dark matter haloes implied in the HOD formalism are observationally challenging to confirm and one has to rely on N-body simulations. Fortunately, additional techniques may be used to relate galaxy properties to halo masses, among which gravitational lensing is one of the most powerful probes: by evaluating the distortion and magnification of background sources, one is able to perform a direct estimation of the dark matter halo profile (for a review, see Bartelmann & Schneider 2001). The low signal-to-noise ratio associated with individual galaxies, however, forces us to 'stack' them (e.g. binned together within narrow stellar mass ranges), using a technique known as *galaxy–galaxy lensing* (Brainerd, Blandford & Smail 1996; Hudson et al. 1998; Hoekstra, Yee & Gladders 2004; Mandelbaum et al. 2005a; Yoo et al. 2006; van Uitert et al. 2011; Cacciato, van Uitert & Hoekstra 2014; Velander et al. 2014; Hudson et al. 2015).

Clearly, each of the above methods brings a different piece of information and combining all observables together is particularly interesting, although doing so properly is challenging. In a recent study using COSMOS data, Leauthaud et al. (2012) have successfully combined galaxy clustering, galaxy–galaxy lensing and the stellar mass function (see also Cacciato et al. 2009; Mandelbaum et al. 2013; Miyatake et al. 2013; More et al. 2014), fitted jointly and interpreted within the HOD framework: the authors have used a global central galaxy $M_\star - M_{\rm h}$ relationship (as opposed to measuring the mean $M_{\rm h}$ per bin of stellar mass) and extended it in a consistent way to satellite galaxies.

In this paper, we apply this advanced formalism using a new data set covering a uniquely large area of $\sim$25 deg^2 with accurate photometric redshifts in the redshift range $0.5 < z < 1$ and stellar masses $> 10^{10}$ M$_\odot$. Our galaxy properties' measurements are calibrated and tested with 70 000 spectroscopic redshifts from the VIPERS survey and a number of publicly available data sets. Our data span a wide wavelength range of ultraviolet (UV) deep data from *GALEX*, optical data from the Canada–France–Hawaii Telescope (CFHT) Legacy Survey and K_s-band observations with the CFHT WIRCam instrument. This large statistical sample allows us to measure with high precision the stellar mass function, the galaxy clustering, and we use the CFHTLenS shear catalogue to measure galaxy–galaxy lensing signals. The galaxy clustering is measured on the projected sky for the photometric sample and in real space for the spectroscopic sample.

This paper is organized as follows: in Section 2, we describe the observations, the photometric redshift and stellar mass estimates. In Section 3, we present the measurements of the stellar mass function, the galaxy clustering (both from the photometric and spectroscopic samples) and galaxy–galaxy lensing signals. In Section 4, we describe the HOD model, and the Markov chain Monte Carlo (MCMC) model fitting results are given in Section 5. In Section 6, we discuss our results and conclude. Throughout the paper, we adopt the following cosmology: $H_0 = 72$ km s^{-1} Mpc^{-1} and $\Omega_\mathrm{m} = 0.258$, $\Omega_\Lambda = 0.742$ (Hinshaw et al. 2009) unless otherwise stated. To compute stellar masses, we adopt the initial mass function (IMF) of Chabrier (2003) truncated at 0.1 and 100 M$_\odot$, and the stellar population synthesis (SPS) models of Bruzual & Charlot (2003). All magnitudes are given in the AB system. The dark matter halo masses are denoted as M_h and defined within the virial radius enclosing a mean overdensity Δ_vir compared to the mean density background, taking the formula from Weinberg & Kamionkowski (2002). At $z = 0.8$, $\Delta_\mathrm{vir} = 215$. All masses are expressed in unit of M$_\odot$. Measured quantities are denoted as $\widetilde{X}$ and theoretical quantities as X. We call *cosmic variance* the statistical uncertainties caused by the density fluctuations of dark matter and we define the *sample variance* as the sum of the cosmic variance and Poisson noise variance.

2 DATA

In this work, we combine several data sets to build a volume-limited sample of galaxies more massive than $M_\star = 10^{10}$ M$_\odot$ in the redshift range $0.5 < z < 1$. Our galaxy selection is based on NIR ($K_s < 22$) observations, collected in the two fields of the VIMOS Public Extragalactic Redshift Survey (‘VIPERS-*W*1’ and ‘VIPERS-*W*4’), overlapping the (Canada–France–Hawaii Telescope Legacy Survey) CFHTLS-Wide imaging survey, and covering a total unmasked area of 23.1 deg^2. We refer to Arnouts et al. (in preparation) for a complete description of the multiwavelength UV and NIR observations, reduction and photometry.

Our background galaxy selection used for the measurement of the lensing signal is based on the CFHTLS-Wide i-band selection in the area that overlaps with the NIR observations.

2.1 The CFHTLS-Wide survey

The CFHTLS[1] is a photometric survey performed with MegaCam (Boulade et al. 2003) on the CFHT telescope in five optical bands $u^\star$, g, r, i, z ($i < 24.5$–25, 5σ detection in 2 arcsec apertures) and covering four independent patches in the sky over a total area of 154 deg^2. In this analysis, we use the photometric and shear catalogues produced by the CFHTLenS[2] team (Heymans et al. 2012). The CFHTLenS photometry is performed with SExtractor (Bertin & Arnouts 1996) on the PSF-homogenized images (Hildebrandt et al. 2012; Erben et al. 2013). Magnitudes are based on the `MAG_ISO` estimator where the isophotal apertures are derived from the i-band detection image. This approach optimizes the colour measurements and leads to an improvement in the photometric redshift accuracy (Hildebrandt et al. 2012). To estimate the total magnitude of each source, a global shift is applied to the `MAG_ISO` magnitude in all the bands based on the difference between `MAG_ISO` and `MAG_AUTO` magnitudes, as measured in the i-band detection image (Hildebrandt et al. 2012).

As the magnitude errors are measured with SExtractor directly from the local background in the PSF-homogenized image, we need to correct for the noise correlation introduced by the convolution process. To do so, we multiply the CFHTLenS magnitude errors in all bands by the ratio of the i-band detection image errors to the i-band PSF-homogenized image errors. The correction factor ranges from 3 to 5, where the strongest correction occurs when the seeing difference between the i band and the worse-seeing image is the largest. As the i-band image is usually the best-seeing image, this procedure may slightly overestimate the correction in the other bands, however we neglect it here.

In addition, magnitude errors must be rescaled to account for image resampling. Two independent tests have been performed to accurately estimate the correction factor: we measured the dispersion of magnitudes between the i-band detection (un-convolved) magnitudes and the CFHTLS-Deep magnitudes, and between duplicated observations of the same object in the overlapping regions of adjacent tiles. We find that the errors must be rescaled by a factor of 2.5.

The footprints of the CFHTLS MegaCam tiles overlapping the VIPERS survey are shown as grey squares in Fig. 1.

2.2 The Near-IR observations

We have conducted a K_s-band follow-up of the VIPERS fields with the WIRCam instrument at CFHT (Puget et al. 2004) for a total allocation time of ∼120 h. The integration time per pixel was 1050 s and the average seeing of all the individual exposures was 0.6 ±0.09 arcsec. The data have been reduced by the Terapix team:[3] the images were stacked and resampled on the pixel grid of the CFHTLS-T0007 release (Hudelot et al. 2012). The images reach a depth of $K_s = 22$ at $\sim 3\sigma$ (Arnouts et al., in preparation). The photometry was performed with SExtractor in dual image mode with a $gri - \chi^2$ image (Szalay, Connolly & Szokoly 1999) as the detection image. To correct for the noise correlation introduced by image resampling, we multiply the errors by a factor 1.5, obtained from the dispersion between the WIRCam K_s-band magnitudes and the magnitudes measured in the deeper ($K < 24.5$) UKIDSS Ultra Deep Survey (UDS; Lawrence et al. 2007). We also used the UDS survey to confirm that our sample completeness based on $gri - \chi^2$ detections reaches 80 per cent at $K_s = 22$. Using the WIRCAM/CFHTLS-Deep data with an i-band cut simulating the CFHTLS-Wide data depth, we have checked that this incompleteness is caused by red galaxies above $z = 1$ and does not affect our sample selected in the range $0.5 < z < 1$. The K_s `MAG_AUTO` estimates are then simply matched to their optical counterparts based on position.

In addition to this data set, we also use the CFHTLS-D1 WIRDS data (Bielby et al. 2012), a deep patch of 0.49 deg^2 observed with WIRCam J, H and K_s bands and centred on $02^\mathrm{h}26^\mathrm{m}59^\mathrm{s}$, $-04°30'00''$. All three bands reach 50 per cent completeness at AB magnitude 24.5.

The WIRCam observations are shown in Fig. 1 as the red regions. After rejecting areas with poor WIRCam photometry and those with CFHTLenS mask flag larger than 2, the corresponding effective area used in this work spans over 23.1 deg^2, divided into 15 and 8.1 deg^2 in the VIPERS-*W*1 and VIPERS-*W*4 fields, respectively.

2.3 The UV-*GALEX* observations

When available, we make use of the UV deep imaging photometry from the *GALEX* satellite (Martin et al. 2005; Morrissey et al. 2005). We only consider the observations from the Deep Imaging Survey (DIS), which are shown in Fig. 1 as blue circles (Ø∼1.°1). All the *GALEX* pointings were observed with the NUV channel with exposure times of $T_\mathrm{exp} \geq 30$ ksec. FUV observations are available for 10 pointings in the central part of *W*1.

---

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

[2] http://cfhtlens.org/

[3] http://terapix.iap.fr/

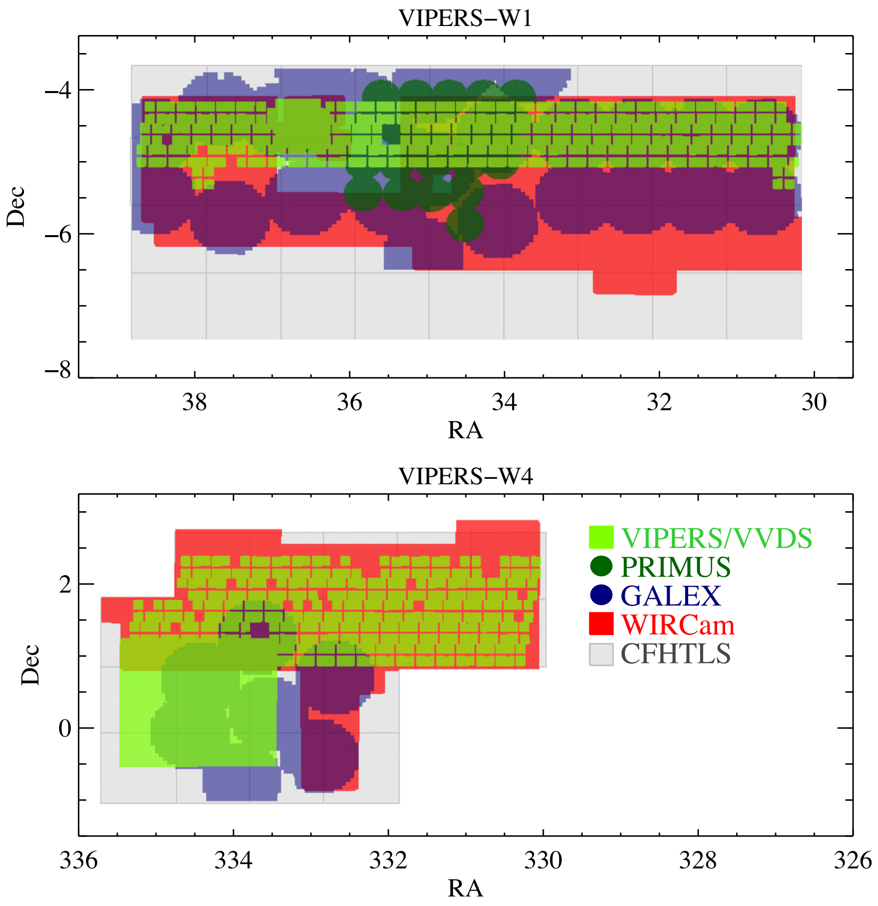

Figure 1. Footprints of the different data sets used in this work. Our selection is based on WIRCam data shown in red and covering approximately 25 deg^2 (23.1 deg^2 after masking). The CFHTLS MegaCam pointings are shown in grey, the *GALEX* DIS observations as large blue circles (in purple if overlapped with WIRCam), the spectroscopic surveys VIPERS/VVDS in light green and PRIMUS in dark green. The SDSS/BOSS coverage is almost complete. The data outside the WIRCam footprint are not used, and shown here only for reference.

Due to the large PSF (FWHM $\sim$ 5 arcsec), source confusion becomes a major issue in the deep survey. To extract the UV photometry we use a dedicated photometric code, EMPHOT (Conseil et al. 2011) which will be described in a separate paper (Vibert et al., in preparation). In brief, EMPHOT uses U band (here the CFHTLS u band) detected objects as a prior on position and flux. The uncertainties on the flux account for the residual in the [simulated−observed] image. The images reach a depth of $m_{\rm NUV} \sim 24.5$ at $\sim 5\sigma$. As for the WIRCAM data, the *GALEX* sources are matched to the optical counterparts based on position.

The NUV observations cover only part of the WIRCam area with $\sim$10.8 and 1.9 deg^2 in VIPERS-*W*1 and VIPERS-*W*4, respectively. The UV photometry slightly improves the precision of photometric redshifts and the stellar mass estimates in the *GALEX* area. However, by comparing our measurements inside and outside the *GALEX* area, we have checked that the addition of UV photometry does not make a significant change for the galaxies of interest in this study. Therefore, in the final sample, we mix galaxies inside the *GALEX* area with those outside.

2.4 Spectroscopic data

To optimize the calibration and the validation of our photometric redshifts, we make use of all the spectroscopic redshifts available in the WIRCam area.

The largest sample is based on the VIPERS spectroscopic survey (Garilli et al. 2014; Guzzo et al. 2014) and its first public data release PDR1.[4] VIPERS aims to measure redshift space distortions and explore massive galaxy properties in the range $0.5 < z < 1.2$. The survey is located in the *W*1 and *W*4 fields of the CFHTLS-Wide survey and will cover a total area of 24 deg^2 when completed, with a sampling rate of $\sim$40 per cent down to $i < 22.5$. In Fig. 1, we show the layout of the VIMOS pointings as the light-green squares. The PDR1 release includes redshifts for $\sim$54 204 objects. After keeping galaxy spectra within the WIRCam area (44 474) and with the highest confidence flags between 2.0 and 9.5 (95 per cent confidence, see Guzzo et al. 2014), we are left with 35 211 galaxies, which corresponds to a spectroscopic success rate of 80 per cent.

In addition to VIPERS, we also consider the following spectroscopic surveys:

(i) the VIMOS-VLT Deep Survey (VVDS) F02 and Ultra-Deep Survey (Le Fèvre et al. 2005, 2014) which consist of 11 353 galaxies down to $i < 24$ (Deep) and 1125 galaxies down to $i < 24.5$ (Ultra-Deep) over a total area of 0.75 deg^2 in the VIPERS-*W*1 field. We also use part of the VIMOS-VLT F22 Wide Survey with 12 995 galaxies over 4 deg^2 down to $i < 22.5$ (Garilli et al. 2008, shown as the large green square in the southern part of the VIPERS-*W*4 field

---

[4] http://vipers.inaf.it/rel-pdr1.html

Table 1. Magnitude zero-point offsets measured per CFHTLS MegaCam pointing in VIPERS-*W*1 and VIPERS-*W*4 (mean and standard deviation). *J*- and *H*-band zero-points were computed for the pointings overlapping WIRDS data.

Filter	VIPERS-*W*1	VIPERS-*W*4
FUV	0.18 ± 0.11	0.02 ± 0.16
NUV	0.11 ± 0.09	0.15 ± 0.10
u	0.10 ± 0.03	0.13 ± 0.03
g	− 0.02 ± 0.01	− 0.01 ± 0.01
r	0.02 ± 0.01	0.01 ± 0.01
i	− 0.01 ± 0.01	− 0.00 ± 0.01
z	− 0.02 ± 0.01	− 0.01 ± 0.01
J	0.08 ± 0.05	–
H	0.02 ± 0.05	–
K	0.02 ± 0.03	0.01 ± 0.05

in Fig. 1). In total, we use 5122 galaxies with secure flags 3 or 4 from the VVDS surveys within the WIRCam area;

(ii) the PRIMUS survey (Coil et al. 2011) which consists of low resolution spectra ($\lambda/\Delta\lambda \sim 40$) for galaxies down to $i \sim 23$ and overlapping our VIPERS-*W*1 field. PRIMUS pointings are shown as the dark green circles in Fig. 1. We keep 21 365 galaxies with secure flags 3 or 4;

(iii) the SDSS-BOSS spectroscopic survey based on data release DR10 (Ahn et al. 2014) down to $i < 19.9$, overlapping both VIPERS-*W*1 and VIPERS-*W*4 fields, totalling 4675 galaxies with `zWarning=0` (99 per cent confidence redshift) within our WIRCam area.

In total, the spectroscopic sample built for this study comprises 62 220 unique galaxy spectroscopic redshifts with the highest confidence flag. We use the spectroscopic redshift value, when available, instead of the photometric redshift value. The galaxies with a spectroscopic redshift represent 6.5 per cent of the full sample, and 12 per cent after selection in the range $0.5 < z < 1$, where most of the galaxies are from the VIPERS sample.

2.5 Photometric redshifts

To compute the photometric redshifts, we use the template fitting code LEPHARE[5] (Arnouts et al. 1999; Ilbert et al. 2006). We adopt similar extinction laws and parameters as Ilbert et al. (2009) used in the COSMOS field (Scoville et al. 2007), and identical priors as in Coupon et al. (2009) based on the VVDS redshift distribution and maximum allowed g-band absolute magnitude. We note that the use of priors is essential for the $z > 1$, low signal-to-noise ratio (or no NIR flux), galaxies used as lensed (background) sources (see Section 3). A probability distribution function (PDF) in steps of 0.04 in redshift is computed for every galaxy.

We use the full spectroscopic sample to adjust the magnitude relative zero-points in all the passbands on a MegaCam pointing-to-pointing basis. For the pointings with no spectroscopic information, we apply a mean correction obtained from all the pointings with spectra. The mean zero-point offsets and standard deviations in all passbands are given in Table 1 for the two fields separately. We further add the zero-point scatter in quadrature to the magnitude errors in each band. We recall that these zero-point corrections may not represent absolute calibration offsets but rather *relative* (i.e. depending on colours) ones and tied to the adopted spectral energy distribution (SED) template set. We come back to the impact of this issue on stellar mass measurements in Section 3.5.

Our SED templates are based on the library used in Ilbert et al. (2009), however the fewer bands used in this study compared to COSMOS necessitate adapting the templates to reduce redshift-dependent biases. The initial templates are based on the SEDs from Polletta et al. (2007), complemented by a number of starburst SEDs from the Bruzual & Charlot (2003) SPS library. Using 35 211 spectroscopic redshifts from VIPERS, we adapt the templates with LEPHARE using the following procedure. First, a best-fitting template from the original set is found for each galaxy and normalized to unity, and the photometry is then corrected into the rest frame given the spectroscopic redshift value. The rest-frame photometry for all galaxies with identical best-fitting templates is combined and the adapted template is constructed from the sliding-window median values as a function of wavelength. The process is repeated iteratively. Given the high number of galaxies with spectroscopic redshifts, we found that only two iterations were necessary to reach convergence. Interestingly, although the improvement in the photometric redshift bias is significant, the new templates appear very similar 'by eye' compared to the original ones, which implies that small features in the SED templates may lead to large photometric errors, as also noted by Ilbert et al. (2006).

In Fig. 2, we show the accuracy of the photometric redshifts by comparing with the spectroscopic redshift sample from VIPERS ($i < 22.5$, left-hand panel) and VVDS Deep/Ultra-Deep ($22.5 < i < 24.5$, right-hand panel). We observe a dispersion[6] of $\sigma/(1 + z) \sim 0.03$–0.04 and a fraction of catastrophic redshifts ($|\Delta z| \geq 0.15(1 + z)$) of $\eta \sim 1$–4 per cent. The dispersion in both magnitude ranges is significantly better than previous results in the CFHTLS-Wide (Coupon et al. 2009), due to the choice of isophotal magnitudes and PSF homogenization (Hildebrandt et al. 2012) at faint magnitude, and the contribution of NIR data above $z \sim 1$. We note that the faint sample is compared to the VVDS redshifts where deep NIR data from WIRDS are available over a small part ($<$1 deg^2) of the field, and with a magnitude distribution biased towards bright galaxies compared to the photometric sample. Therefore, we foresee degraded photometric redshift performance elsewhere, mainly relevant for $z > 1$ galaxies. However, as shown in Appendix C, no systematic bias affecting our lensing measurements is introduced by the use of sources beyond $z = 1$.

2.6 Stellar mass estimates

To compute stellar masses, we adopt the same procedure as Arnouts et al. (2013) and described in detail in their Appendix . In brief, we use the photometric or spectroscopic (when available) redshift and perform a χ^2 minimization on a SED library based on the SPS code from Bruzual & Charlot (2003). The star formation history is either constant or described with an exponentially declining function, with e-folding time $0.01 \leq \tau \leq 15$. We use two metallicities (Z$_\odot$, 0.2 Z$_\odot$) and adopt the Chabrier (2003) IMF. As discussed in Arnouts et al. (2013), the use of various dust extinction laws is critical to derive robust SFR and stellar mass; and in this work, we adopt their choices for differing attenuation curves: a starburst (Calzetti

---

[5] http://www.cfht.hawaii.edu/arnouts/lephare.html

[6] Defined as the normalized median absolute deviation (Hoaglin, Mosteller & Tukey 1983): 1.48×Median($|z_{\rm s} - z_{\rm p}|/(1 + z_{\rm s})$), and robust to outliers.

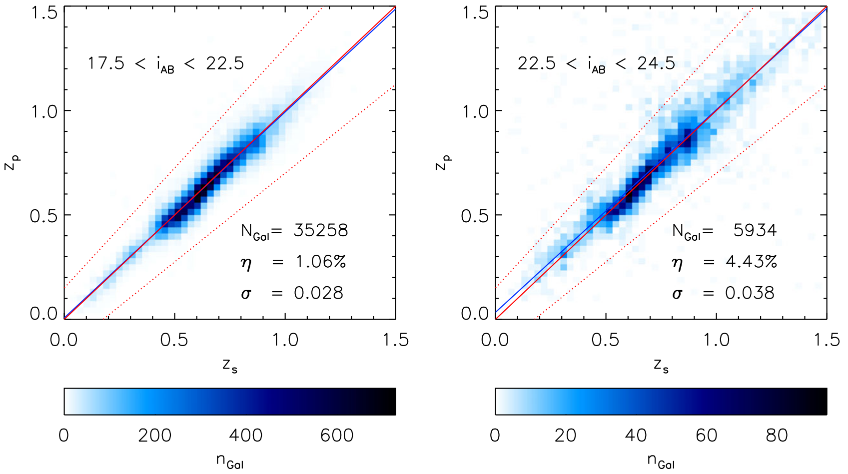

Figure 2. Photometric redshifts measured with *ugrizK* (left) or *ugrizJHK* (right) photometry versus VIPERS and VVDS spectroscopic redshifts. Left: $17.5 < i < 22.5$, where the sample is dominated by galaxies between $0.5 < z < 1.2$ due to the VIPERS selection. Right: $22.5 < i < 24.5$, from the VVDS Deep and Ultra-Deep surveys. The limits for the outliers are shown as red dotted lines.

Table 2. Sample mass definitions in $\log(M_\star/\,\mathrm{M}_\odot)$ and number of galaxies in each sample. The parent sample comprises a total of 352 585 galaxies.

	Clustering–$w(\theta)$		Clustering–$w_\mathrm{p}(r_\mathrm{p})$		Lensing	
Sample	Mass cut	Number	Mass cut	Number	Mass cut	Number
1	10.00–10.40[a]	23 886	10.60–10.90[a]	2154	10.00–10.40[a]	23 886
2	10.40–10.60	36 560	10.90–11.20[b]	1964	10.40–10.65	45 032
3	10.60–10.80	31 900	11.20–12.00	816	10.65–10.80	23 427
4	10.80–11.00	24 451	–	–	10.80–10.95	19 293
5	11.00–11.20	13 538	–	–	10.95–11.15	16 317
6	11.20–12.00	6326	–	–	11.15–12.00	8654

Notes. [a]$0.5 < z < 0.7$.
[b]$0.5 < z < 0.8$.

et al. 2000), an SMC-like (Prevot et al. 1984) and an intermediate slope ($\lambda^{-0.9}$) law. We consider reddening excess in the range $0 \le E(B - V) \le 0.5$. When fixing the redshift, the typical 68 per cent stellar mass statistical uncertainty, as derived by marginalising the likelihood distribution, ranges from $\sigma(M_\star) \sim 0.05$ to 0.15 for galaxies with $K_s \le 22$ and $z < 1$. This stellar mass uncertainty is an underestimate, since we neglect photometric redshift uncertainties.[7]

In addition to statistical errors, in Section 3.5, we investigate the different sources of systematic effects in the stellar mass estimates, arising from our lack of knowledge of galaxy formation and evolution. The choice of differing dust treatments (and resulting dust attenuation laws) is one of them: Ilbert et al. (2010) have measured a shift of 0.14 dex, with a large scatter, between stellar masses estimated with the Charlot & Fall (2000) dust prescription and the Calzetti et al. (2000) attenuation law. The dust parametrization leads to systematics larger than the statistical errors in the stellar mass function. Even more critical is the choice of the SPS model and the IMF (see more detailed systematic errors analysis in Behroozi et al. 2010; Marchesini et al. 2009; Fritz et al. 2014), leading to systematic differences in stellar mass estimates up to 0.2 dex. One must keep these limitations in mind when comparing results from various authors using different methods, and we come back to these issues when presenting our results.

3 MEASUREMENTS

We aim to compute high signal-to-noise measurements of four distinct observables: the stellar mass function $\phi(M_\star)$, the projected galaxy clustering $w(\theta)$, the real-space galaxy clustering $w_\mathrm{p}(r_\mathrm{p})$ and the galaxy–galaxy lensing $\Delta\Sigma(r)$.

To do so, we select volume-limited samples in the redshift range $0.5 < z < 1$, where the high sampling rate of VIPERS and our NIR data guarantee both robust photometric redshift and stellar mass estimates. As for the stellar mass function, we adopt a lower mass limit of $M_\star = 10^{10}\,\mathrm{M}_\odot$ and employ the V_max estimator to correct for galaxy incompleteness near $z = 1$. The total volume probed in this study is 0.06 Gpc^3.

The stellar mass bins for the clustering and lensing measurements are defined to keep approximately a constant signal-to-noise ratio across the full mass range (which may lead to differing mass cuts depending on the observable), and guarantee complete galaxy samples (see Appendix A). We summarize our samples' properties in Table 2.

To measure each of the observables described below, we use the parallelized code SWOT, a fast tree-code for computing two-point correlations, histograms and galaxy–galaxy lensing signals from large data sets (Coupon et al. 2012). The stellar mass function is expressed in comoving units, whereas the clustering and

---

[7] We will see in Section 4.2 that our model accounts for such an extra source of uncertainty in stellar mass through a stellar-mass-dependent parametrization of the stellar-to-halo mass scatter.

galaxy–galaxy lensing signal are measured in physical units. We estimate statistical covariance matrices from a jackknife resampling of 64 subregions with equal area (0.35 $\deg^2$ each), by omitting a subsample at a time and computing the properly normalized standard deviation (see more details in Coupon et al. 2012). This number was chosen to meet both requirements of using large enough subregions to capture the statistical variations at large scale, while keeping a sufficient number of subsamples to compute a robust covariance matrix. Nevertheless, we expect the projected galaxy clustering errors to be slightly underestimated on scales larger than the size of our subregions, ~0.5 deg, and the noise in the covariance matrix to potentially bias the best-fitting χ^2 value.

A random sample with 1 million objects is constructed using our WIRCam observations layout and the union of the WIRCam and CFHTLenS photometric masks. For real-space clustering, measured from VIPERS spectroscopic redshifts, the random sample is constructed using the layout of the VIPERS PDR1 geometry (and photometric masks) plus a random redshift drawn in the range $0.5 < z < 1$ from a distribution following $\mathrm{d}V/\mathrm{d}z$, to match our volume-limited samples. The subregions for the measurements of statistical errors are constructed by SWOT based on the random catalogue: the field is divided into 64 areas with an equal number of random objects.

3.1 The stellar mass function

The stellar mass function $\widetilde{\phi}(M_\star) = \mathrm{d}n/\mathrm{d}\log M_\star$ is measured per unit of comoving volume in 10 equally spaced logarithmic mass bins of width 0.2 dex, centred on the mass mean weighted by the number of galaxies. To correct for the incompleteness in the low-mass galaxy sample ($10^{10} < M_\star/\mathrm{M}_\odot < 10^{10.4}$) occurring near $z = 1$ (see Appendix A), we up-weight low-redshift galaxies by a factor $1/V_{\max}$ defined as

$$V_{\max} = \Omega \int_{0.5}^{z_{\max}} \frac{\mathrm{d}V}{\mathrm{d}z} \mathrm{d}z\,, \tag{1}$$

where Ω is the solid angle of the survey, 23.1 $\deg^2$, V the comoving volume per unit area and $z_{\max}$ the maximum redshift for a galaxy to be observed given a $K_s < 22$ magnitude cut, calculated with LEPHARE.

We have performed a number of tests to check our internal error estimates. In the top panel of Fig. 3, we show our stellar mass function error estimates (square root of the covariance matrix diagonal) as a function of stellar mass compared to the GETCV code estimate of Moster et al. (2011) at $z = 0.8$. The latter code computes the theoretical expectations of cosmic variance[8] assuming a prediction for dark matter clustering and galaxy biasing (Bardeen et al. 1986). We add to the GETCV cosmic variance the theoretical Poisson error and show the resulting (total) sample variance as the thick line in the bottom panel. Our jackknife estimate is represented as the black points, for which we find that the cosmic variance part (after subtracting Poisson noise) needs to be multiplied by a factor of 2 to agree with theoretical expectations (we then multiply the covariance matrix by a factor of 4). We have not found a definitive explanation for the underestimation of the errors from the jackknife resampling, however it is likely caused by the strong correlation between bins (a combined effect of stellar mass scatter and large-scale structure correlations).

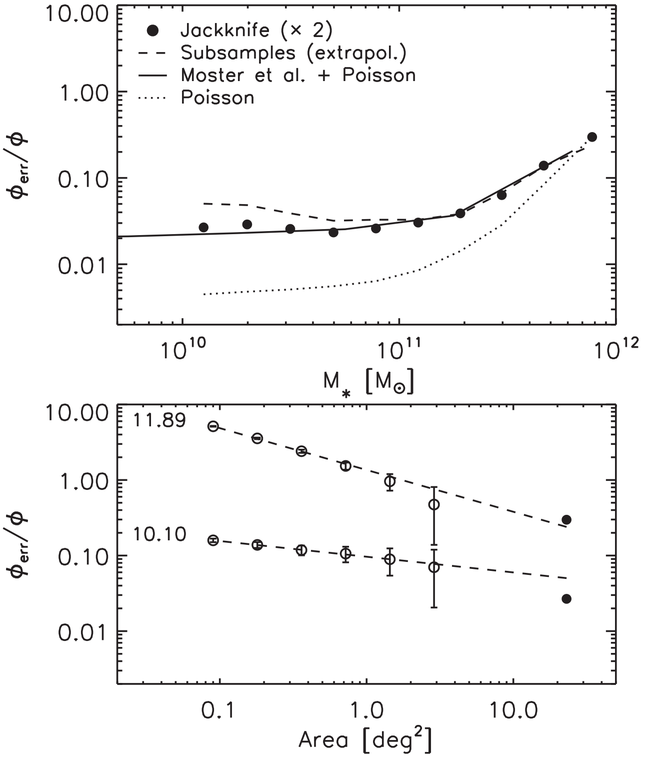

Figure 3. Stellar mass function statistical errors as function of stellar mass (top) and area (bottom). In the top panel, we show the jackknife estimator based on 64 subregions and multiplied by a factor of 2, compared to the theoretical cosmic variance plus Poisson error derived from the Moster et al. (2011) GETCV code (the Poisson error only is shown as the dotted line). The bottom panel shows an alternative internal estimate based on the standard deviation of subregions as a function of their size, in two mass bins ($\log M_\star/\mathrm{M}_\odot = 10.10$ and 11.89), extrapolated to the size of the full survey (dashed lines in both panels). As in the top panel, the black dots are the jackknife estimates, for which the cosmic variance part has been multiplied by 2.

In the bottom panel of Fig. 3, we show an alternative internal estimator as function of area, based on the standard deviation of subsamples with sizes varying from 0.1 to 2.9 $\deg^2$ (the black dots represent our Jackknife estimates in the two mass bins $\langle \log M_\star/\mathrm{M}_\odot \rangle = 10.10$ and 11.89). We use a power-law fit (the amplitudes of the error bars are arbitrarily scaled to the square root of the number of subsamples, ranging from $\sqrt{256}$ to $\sqrt{8}$) to extrapolate to the full size of the survey. The extrapolated values are shown as the dashed line in the top panel of Fig. 3. The bin correlations between small subsamples may tilt the slope of the fit and lead to an overestimate of the extrapolated error estimate, as observed in the low-mass bin. In the high-mass bins, characterized by an uncorrelated sampling variance dominated by Poisson noise, the extrapolated estimate is consistent with both the jackknife estimate and the theoretical Poisson noise.

3.2 Projected galaxy clustering

We measure the two-point correlation function $\widetilde{w}(\theta)$ in 10 logarithmically spaced bins centred on the pair-number weighted averaged separation over the range $0\overset{\circ}{.}002 < \theta < 2^\circ$. The modelled $w(\theta)$ is compared to the measured $\widetilde{w}(\theta)$ by projecting the theoretical spatial

---

[8] We note that the highest mass bin galaxy bias was estimated a posteriori from our HOD results, since it was not provided by the authors of GETCV, although the contribution of cosmic variance is negligible compared to the Poisson error in this bin, populated by rare massive galaxies.

clustering $\xi(r)$ on to the sample redshift distribution computed as the sum of photometric redshift PDFs (see Section 2.5).

We use the Landy & Szalay (1993) estimator following a similar procedure to that described in Section 3.3 of Coupon et al. (2012). Owing to the limited size of the survey, our measurements are affected by the *integral constraint*, an effect that biases the clustering signal low. Here, we adopt a refined way to correct for it: the correction is calculated directly for every parameter set from the modelled $w(\theta)$ (instead of a pre-determined power law) and integrated over the survey area using random pairs as in Roche et al. (2002), leading to better agreement between the data and the model at large scales. Here, the typical values of the integral constraint range from 10^{-3} to 3×10^{-3}.

We have checked, using the galaxy mocks prepared for the VIPERS sample (de la Torre et al. 2013), that our jackknife error estimates could reproduce within 20 per cent the correct sample variance amplitude of $\widetilde{w}(\theta)$ (this result is in agreement with a number of tests from the literature, e.g. Zehavi et al. 2005; Norberg et al. 2009), and we do not apply any correction.

3.3 Real-space galaxy clustering

We measure the real-space galaxy clustering for the VIPERS spectroscopic sample by integrating the weighted redshift-space correlation function along the line of sight to alleviate redshift-space distortion effects:

$$\widetilde{w_{\rm p}}(r_{\rm p,phys}) = 2\int_0^{\pi_{\rm max}} \widetilde{\xi}(r_{\rm p,phys}, \pi_{\rm phys}){\rm d}\pi_{\rm phys}, \tag{2}$$

where $r_{\rm p,phys}$ and $\pi_{\rm phys}$ are the coordinates perpendicular and parallel to the line of sight, respectively. $r_{\rm p,\,phys}$ is expressed in physical coordinates and divided into 10 logarithmically spaced bins centred on the pair-weighted averaged separation over the range $0.2 < r_{\rm p,phys}/{\rm Mpc} < 10$, and $\pi_{\rm phys}$ is divided into linear bins up to $\pi_{\rm max} = 40$ Mpc. The value of $\pi_{\rm max}$ is consistently used in the derivation of the modelled $w_{\rm p}$. As for $\widetilde{w}(\theta)$, $\widetilde{\xi}(r_{\rm p,phys}, \pi_{\rm phys})$ is computed using the Landy & Szalay estimator and the covariance matrix estimated from the jackknife resampling of 64 subregions.

Each galaxy is weighted to account for the undersampling of the spectroscopic sample: we use the global colour sampling rate (CSR), target sampling rate (TSR) and success sampling rate (SSR), as described in Davidzon et al. (2013), to account for the VIPERS colour selection, the sparse target selection and measurement success as function of signal-to-noise ratio, respectively. In addition, we also use number-count normalized (to prevent global CSR, TSR and SSR double weighting) spatial weights computed for each VIPERS panel by de la Torre et al. (2013) to correct for the position-dependent sampling. Here, the SSR is the most affected quantity, as a function of position in the sky, due to the differing observing conditions at the times of observation.

Small pair incompleteness due to ‘slit collision’ is corrected by a factor $1 + \widetilde{w_{\rm A}}$, such that:

$$1 + \widetilde{w}_{\rm p,corr} = \frac{1 + \widetilde{w}_{\rm p}}{1 + \widetilde{w}_{\rm A}}, \tag{3}$$

where

$$1 + \widetilde{w}_{\rm A} = 1 - \frac{0.03}{r_{\rm p,phys}} \tag{4}$$

is derived from the projected correlation as function of angular scale by de la Torre et al. (2013) and translated into physical scales at $z = 0.8$. We note that given our conservative small-scale cut of $r_{\rm p,phys} > 0.2$, the correction remains below 15 per cent.

3.4 Galaxy–galaxy lensing

The gravitational lensing signal produced by the foreground matter overdensity is quantified by the tangential distortion of background sources behind a sample of stacked ‘lens’ galaxies, also known as the weighted galaxy–galaxy lensing estimator (e.g. Mandelbaum et al. 2006; Yoo et al. 2006). The excess surface density of the projected dark matter halo relates to the measured tangential shear through:

$$\widetilde{\Delta\Sigma}(r_{\rm p,phys}) = \Sigma_{\rm crit} \times \widetilde{\gamma}_t(r_{\rm p,phys}), \tag{5}$$

(see also Appendix B). We measure the signal in 10 logarithmically spaced bins centred on the number-weighted averaged separation, in the range $0.02 < r_{\rm p,\,phys}/{\rm Mpc} < 1$. $r_{\rm p,\,phys}$ is expressed in physical coordinates.[9]

The critical surface density $\Sigma_{\rm crit}$ is given by

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G_{\rm N}} \frac{D_{\rm OS}}{D_{\rm OL}\, D_{\rm LS}}, \tag{6}$$

with $D_{\rm OS}$ the observer-source angular diameter distance, $D_{\rm OL}$ the observer-lens (foreground galaxy) distance and $D_{\rm LS}$ the lens-source distance. $G_{\rm N}$ is the gravitational constant and c the speed of light. All distances are computed in physical coordinates using the photometric (spectroscopic when available) redshift. For photometric redshift values, a cut $z_{\rm source} - z_{\rm lens} > 0.1 \times (1 + z_{\rm lens})$ is adopted. The background source galaxy sample includes all galaxies detected in the i band with a non-zero lensing weight (Miller et al. 2013). Here, we do not restrict our redshift sample to $z_{\rm p} < 1.2$, but consider galaxies at all redshifts, taking advantage of the improved photometric redshift estimates in our sample, increasing the background source sample by 30 per cent compared to other CFHTLenS lensing studies, without introducing any systematic bias (see Appendix C).

The galaxy shape measurement was performed on individual exposures using the LENSFIT analysis pipeline (Miller et al. 2007; Kitching et al. 2008; Miller et al. 2013) and systematics checks were conducted by Heymans et al. (2012) for cosmic shear (the projected large-scale structure lensing power spectrum). The lensing (inverse-variance) weights account for shape measurement uncertainties (Miller et al. 2013). Following Velander et al. (2014), who performed extensive systematics checks of the CFHTLenS shear catalogue specifically for galaxy–galaxy lensing (see their Appendix C), we do not reject those CFHTLS-Wide pointings that did not pass the requirements for cosmic shear, and we applied appropriate shape measurement corrections as described in their Section 3.1.

We compute the boost factor (to account for dilution due to sources physically associated with the lens, see Sheldon et al. 2004; Mandelbaum et al. 2006) by randomizing the source positions, and correct the final signal for it. On small scales (<0.1 Mpc), the boost factor reaches up to 20 per cent for the most massive galaxies.

Here, the relatively low source density implies that our errors are dominated by the source galaxy shape noise, originating from ellipticity measurement uncertainties and intrinsic shape dispersion, rather than sample variance. Indeed, when compared to the sum of inverse-variance lensing weights, we have checked that our jackknife estimate was similar at all scales (with small off-diagonal

---

[9] Note that the galaxy–galaxy lensing signal is measured in physical units, whereas a number of authors assume comoving units, which would require multiplying the excess surface density by a factor of $(1 + z)^{-2}$ compared to our definition.

correlation), confirming the negligible impact of cosmic variance (see Appendix C).

Nevertheless, a correlation exists *between* the mass bins due to the re-use of background source galaxies. We neglect this contribution in the computation of the combined χ^2, but we note that this correlation is likely to lead to underestimation of our parameter confidence limits.

3.5 Systematic errors in stellar mass measurements

In this section, we are concerned with systematic errors affecting the stellar mass measurements caused by the uncertainties in the assumed cosmology (i.e. volume and distance estimates), the dust modelling, and potential biases in the photometry.

To assess the impact of systematics on the measurements of the observables, we propagate the errors affecting the stellar masses by changing one parameter configuration at a time, then re-computing all stellar masses and the observables, and finally measuring the difference with the reference measurements. We repeat the process for the three different kinds of systematics listed below:

(i) assumed cosmology. We explore three Λ cold dark mater (ΛCDM) parameter sets: in addition to the *Wilkinson Microwave Anisotropy Probe* (*WMAP)* cosmology used in this study with $H_0 = 72$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.258$, $\Omega_\Lambda = 0.742$ (Hinshaw et al. 2009), a 'concordance' cosmology model with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$ and the *Planck* cosmology with $H_0 = 67$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.320$, $\Omega_\Lambda = 0.680$ (Planck Collaboration XVI 2014) are tested. In each case, the stellar masses and the observables are consistently re-computed with the same cosmology. We note that the term 'systematics' here refers to the choice for one or another set of parameters that produces a systematic shift in stellar mass and not to systematic errors associated with the measurement of cosmological parameters;

(ii) lens galaxy dust extinction modelling. We compute five different stellar masses for each galaxy by varying one aspect at a time: two different extinction law configurations (among our choice of three laws, see Section 2.6) and three different $E(B - V)$ maximum allowed values (ranging from 0.2 to 0.7);

(iii) photometric calibration. As zero-point offsets do not correct for absolute calibration uncertainties (but do for colours), nor correct for photometric measurement biases (e.g. missing flux of bright objects), a change in the photometric calibration may cause a shift in the best-fitting template and further bias the stellar mass measurements. We re-compute stellar masses applying ad hoc global shifts (in all bands) of -0.05 and $+0.05$ magnitude, which correspond to typical offsets caused by various calibration strategies or photometry measurements (Moutard et al., in preparation).

Results are shown in Fig. 4. For each observable (top left: stellar mass function, top right: projected clustering, bottom left: real-space clustering and bottom right: galaxy–galaxy lensing), we display the re-computed measurements divided by the reference quantities, in each of the 'Cosmology', 'Extinction' and 'Calibration' panels as well as the sum in quadrature of all these effects ('Total'). The grey area corresponds to the maximum value among the differing re-computations, not the standard deviation, as each of the solutions is equally likely to be opted for. Except for the stellar mass function, we only display the results in the most massive bins (where we observe the most significant changes), although the calculations were repeated in all mass bins.

To allow comparison with the statistical errors, we overplot the error bars from our jackknife error estimates. For the stellar mass function (whose jackknife error estimate is multiplied by a factor of 2, see Section 3.1), the systematic errors compared to the statistical errors are striking, with the former being larger by one order of magnitude compared to the latter. The increase of the systematic errors towards the high-mass regime is a direct consequence of the shift in stellar mass and the steep slope of the SMF at the massive end.

It is interesting to note that the different cosmologies lead also to large systematic errors compared to statistical errors. Although many authors in galaxy evolution studies claim to account for cosmological parameter uncertainties by presenting h-free results, we recall that, in a flat Universe, both $\Omega_{\rm m}$ and H_0 enter in the computation of the comoving volume and luminosity distances and, in the precision era of *WMAP* and *Planck*, happen to contribute equally to the distance uncertainties. Comparing our results to the recent literature is therefore not as simple as scaling the different quantities with respect to h, and we must properly account for the more complex dependence of distances on $\Omega_{\rm m}$ and H_0.

In comparison, the projected and real-space galaxy clustering as well as galaxy–galaxy lensing are relatively less prone to systematic errors. For the effect of cosmology, the measurement of projected clustering has no dependence on galaxy distances, and the only difference originates from the modified galaxy selection caused by the stellar mass shift. Interestingly, although the real-space clustering and the galaxy–galaxy lensing do depend on galaxy distance measurements, the change in cosmology also has little impact at the level of our statistical errors. We can draw similar conclusions on the effects of dust extinction modelling and photometric calibration.

Obviously, the stellar mass function is the measured quantity suffering from the largest systematic error contribution, compared to the statistical errors. In particular, we will see in Section 5 that most of the constraints on the central galaxy $M_\star - M_{\rm h}$ relationship emanate from the stellar mass function and taking into account these systematic uncertainties when comparing our results with the literature is necessary.

Ideally, one would like to estimate a best-fitting model for each of the re-computed quantities. Unfortunately, this would be computationally very expensive. Instead, we create two sets of measurements: a 'statistical error' set based on our jackknife error estimate and a 'total error' set for which we add in quadrature the systematic errors (assuming they are Gaussian distributed) and the statistical errors. We present in Section 5 separate results for both.

3.6 Impact of photometric redshift uncertainties

The dispersion of photometric redshifts may also cause systematic effects of several kinds, first on the stellar mass function, as a contribution to the stellar mass scatter, which shifts towards higher masses the high-mass end where the slope is steep, an effect known as *Eddington bias*. Secondly, the projected clustering amplitude is biased low due to the scattering of galaxies falling outside the mass bins.

We will see in Section 4 that our model properly accounts for these systematic effects caused by photometric redshift dispersion, through the parametrization of the stellar mass scatter. However, catastrophic failures and photometric redshift biases may be more problematic. We have demonstrated in Section 2.5 that our catastrophic error rate was not higher than 4 per cent, and based on results from Section 3.2 of Coupon et al. (2012), such a low contamination rate should have no impact on clustering results at our statistical

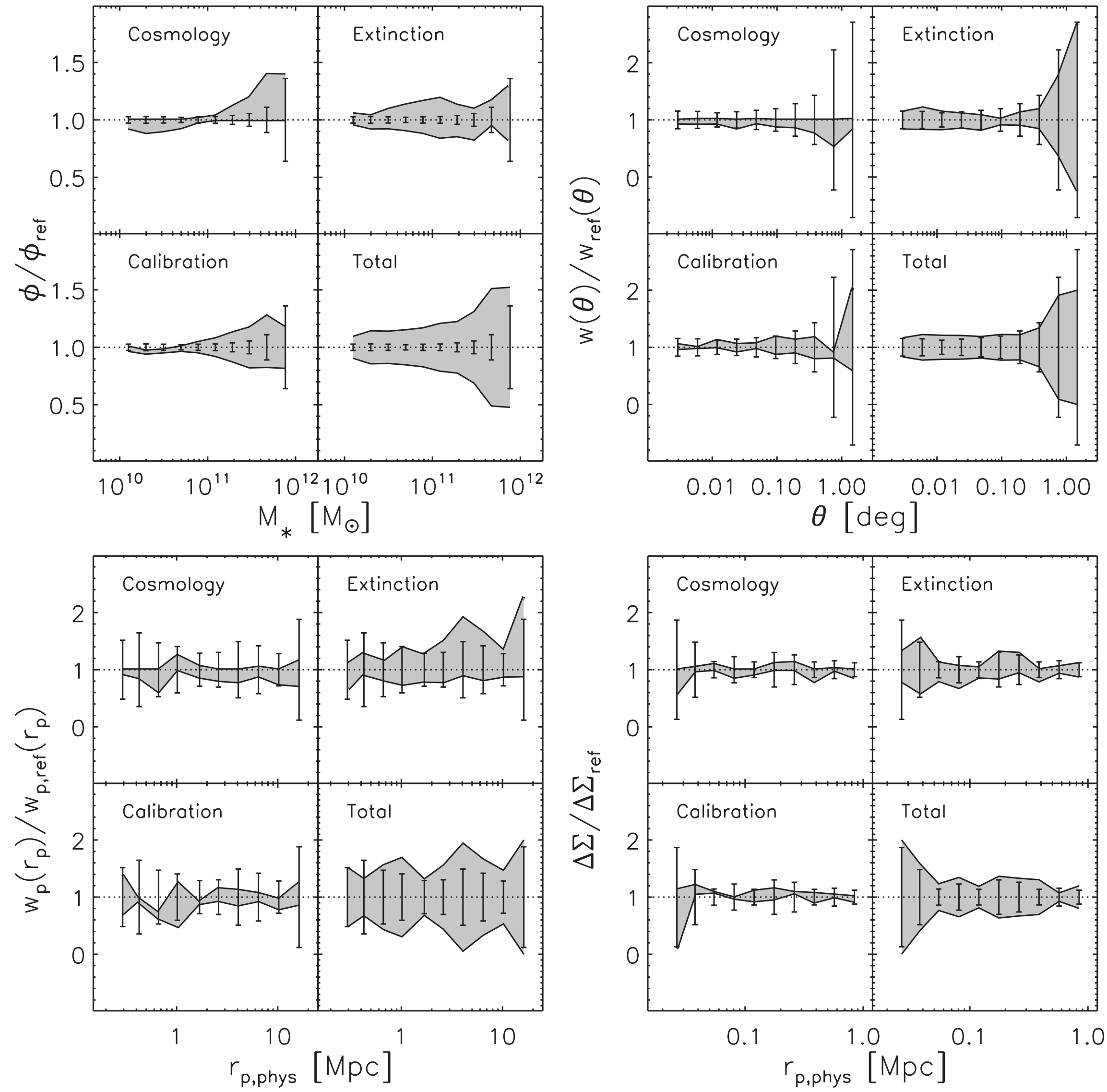

Figure 4. Systematic errors affecting the galaxy stellar mass function (top left), the projected correlation function (top right), the real-space correlation function (bottom left) and the galaxy–galaxy lensing signal (bottom right). In each panel, the grey area symbolizes the envelope (maximum value) of the re-computed measurement compared to the reference. The error bars are statistical errors from the internal jackknife estimator. The 'Total' panel represents the symmetric sum in quadrature of all three contributions. Here, we only show the most massive bins for the clustering and lensing measurements, however we repeated the tests in all mass bins.

error level. To check this statement on the calibration sample (which means the conclusions are limited to the photometric sample with similar properties to the spectroscopic sample), we use the VIPERS galaxies with spectroscopic redshift and re-compute all stellar masses, as well as each observable, using the corresponding photometric redshift. We show the measurements in Fig. 5 (solid lines) divided by the reference measurement made with spectroscopic redshifts and where the error bars are from the statistical jackknife estimator. From left to right, we display the results for the stellar mass function, the projected clustering and the galaxy–galaxy lensing signal, all in the mass range $10^{10} < M_\star/\,\mathrm{M}_\odot < 10^{12}$ and redshift range $0.5 < z < 1$.

We conclude that for galaxies with similar properties to VIPERS galaxies, none of the observables measured with photometric redshifts display a large bias with respect to the spectroscopic redshift ones. This represents a reassuring sanity check for the calibration procedure. Only the projected clustering presents a slightly low systematic value, expected from the dispersion of redshifts and accounted for in the model, through the projection of the modelled 3D clustering on the redshift distribution constructed from the sum of photometric redshift PDFs (assuming that estimated PDFs are representative of the true PDFs).

4 MODEL AND FITTING PROCEDURE

We use the HOD formalism to connect galaxy properties to dark matter halo masses. Here, we assume that the number of galaxies per halo is solely a function of halo mass, split into central and satellite contributions. The fitting procedure then consists of finding a set of parameters to describe the HOD that best reproduces the observables.

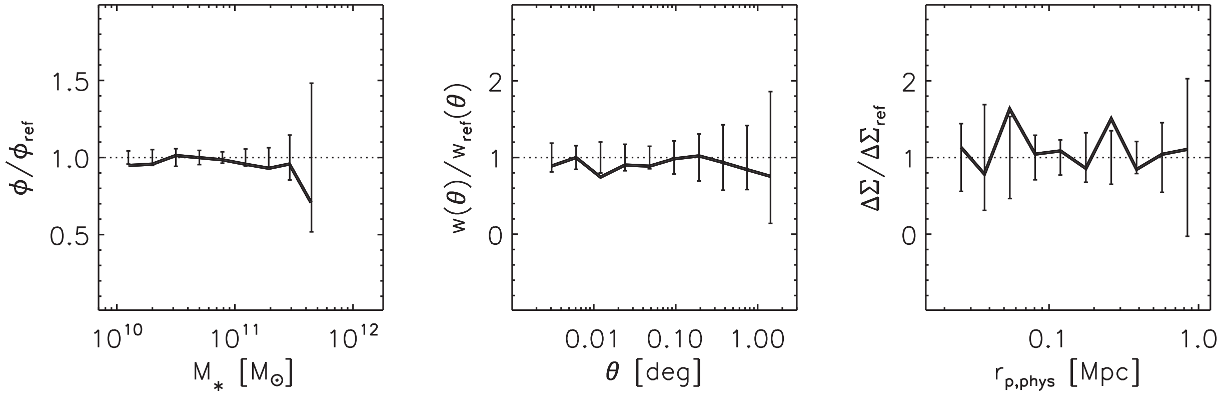

Figure 5. Measurements made with photometric redshifts divided by those made with spectroscopic redshifts. From left to right: the stellar mass function, the projected clustering and the galaxy–galaxy lensing signal, all made with VIPERS galaxies in the mass range $10^{10} < M_\star/\mathrm{M}_\odot < 10^{12}$ and redshift range $0.5 < z < 1$. Error bars represent the statistical error estimates from jackknife resampling.

A key ingredient of the HOD model is the statistical description of the spatial distribution of dark matter. We assume that the matter power spectrum, the halo mass function and the dark matter halo profile are all known quantities over the scales and redshift range ($0.5 < z < 1$) explored in this study. All the technical details about the halo model are given in appendix A.1 of Coupon et al. (2012), with the exception of the large-scale halo bias, for which we use in this study the fitting formula proposed by Tinker et al. (2010).

The exact way to parametrize the HOD is often at the origin of the differences between HOD studies in the literature. In this paper, we follow Leauthaud et al. (2011) who adopted two advanced features:

(i) the HOD is a conditional function of the stellar mass given the halo mass (this formalism is an extension of the conditional luminosity function technique developed by Yang et al. 2003). In this formalism, the central galaxy $M_\star - M_{\rm h}$ relationship is a parametrized function representing the mean stellar mass given its host halo mass, $\langle M_\star | M_{\rm h}\rangle$;

(ii) all observables, namely the stellar mass function, the projected clustering, the real-space clustering and the galaxy–galaxy lensing signal are fitted jointly.

4.1 The stellar-to-halo mass relationship

To describe the central galaxy $M_\star - M_{\rm h}$ relationship, we adopt the parametrized function $f_{\rm SM\text{-}HM}$ proposed by Behroozi et al. (2010), and defined via its inverse:

$$\log_{10}\left(f^{-1}_{\rm SM-HM}\right) = \log_{10}(M_{\rm h}(M_\star))$$

$$= \log_{10}(M_1) + \beta\,\log_{10}\left(\frac{M_\star}{M_{\star,0}}\right) + \frac{\left(\frac{M_\star}{M_{\star,0}}\right)^{\delta}}{1+\left(\frac{M_\star}{M_{\star,0}}\right)^{-\gamma}} - \frac{1}{2}. \qquad (7)$$

M_1 controls the scaling of the relation along the halo mass coordinate, whereas $M_{\star,0}$ controls the stellar mass scaling. β, δ and γ control the low-mass, high-mass and curvature of the relation, respectively.

4.2 The central occupation function

For central galaxies contained in a threshold sample ($M_\star > M^{\rm t}_\star$), the HOD is defined as a monotonic function increasing from 0 to 1, with a smooth transition centred on the halo mass value corresponding to $M^{\rm t}_\star$:

$$\langle N_{\rm cen}(M_{\rm h}|M^{\rm t}_\star)\rangle$$

$$= \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\log_{10}(M^{\rm t}_\star)) - \log_{10}(f_{\rm SM-HM}(M_{\rm h}))}{\sqrt{2}\sigma_{\log M_\star}(M^{\rm t}_\star)}\right)\right]. \qquad (8)$$

The parameter $\sigma_{\log M_\star}$, expresses the scatter in stellar mass at fixed halo mass, which we parametrize as:

$$\sigma_{\log M_\star}(M^{\rm t}_\star) = \sigma_{\log M_\star,0}\left(\frac{M^{\rm t}_\star}{10^{10}\,\mathrm{M}_\odot}\right)^{-\lambda}, \qquad (9)$$

to account for the change in intrinsic stellar mass dispersion as a function of stellar mass.

4.3 The satellite occupation function

We describe the satellite HOD for a threshold sample $M^{\rm t}_\star$ with a simple power law as a function of halo mass $M_{\rm h}$:

$$\langle N_{\rm sat}(M_{\rm h}|M^{\rm t}_\star)\rangle = \left(\frac{M_{\rm h} - M_{\rm cut}}{M_{\rm sat}}\right)^{\alpha}, \qquad (10)$$

for which we fix the cut-off mass scale $M_{\rm cut}$ such that

$$M_{\rm cut} = f^{-1}_{\rm SM-HM}(M^{\rm t}_\star)^{-0.5}. \qquad (11)$$

This assumption is based upon the values reported by Coupon et al. (2012) for their equivalent parameter 'M_0'. We have checked that the exact parametrization of $M_{\rm cut}$ had very little importance compared to the other parameters and did not change any of our conclusions, in agreement with the loose constraints observed by Coupon et al. (2012).

As in Leauthaud et al. (2011), the normalization $M_{\rm sat}$ of the satellite HOD follows the halo mass scaling driven by the central $M_\star - M_{\rm h}$ relationship, with some degree of freedom controlled by a power law:

$$\frac{M_{\rm sat}}{10^{12}\,\mathrm{M}_\odot} = B_{\rm sat}\left(\frac{f^{-1}_{\rm SM-HM}(M^{\rm t}_\star)}{10^{12}\,\mathrm{M}_\odot}\right)^{\beta_{\rm sat}}. \qquad (12)$$

4.4 Total occupation functions and observables

Finally, the total HOD is

$$\langle N_{\rm tot}(M_{\rm h}|M^{\rm t}_\star)\rangle$$

$$= \langle N_{\rm cen}(M_{\rm h}|M^{\rm t}_\star)\rangle + \langle N_{\rm sat}(M_{\rm h}|M^{\rm t}_\star)\rangle, \qquad (13)$$

and since our measurements are made in bins of stellar mass, we transform the threshold HOD functions into binned functions by writing:

$$\begin{aligned}&\langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_1}, M_\star^{\rm t_2})\rangle\\ &= \langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_1})\rangle - \langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_2}\rangle . \qquad (14)\end{aligned}$$

Equivalent relations hold for central and satellite binned HODs.

The stellar mass function, the projected two-point correlation function, the real-space correlation function and the galaxy–galaxy lensing signals are computed from the halo model and the HOD as detailed in Appendix B.

4.5 Systematic errors in the model

As detailed in the previous sections, the HOD formalism relies on an accurate description of the dark matter spatial distribution. Here, we evaluate the impact of our model uncertainties and assumptions on the best-fitting HOD parameters. Ideally, one would like to repeat the fitting procedure to test each of the different assumptions of the model, but to avoid such a time-consuming exercise, we take the simple approach of modifying one feature at a time, and tuning the HOD parameters by hand to reproduce the modelled quantities derived from the best-fitting parameters reported in Section 5. We explore two stellar mass bins ($M_\star = 10^{10}, 10^{11.5}\,{\rm M}_\odot$) and we focus on the two parameters M_1 and $B_{\rm sat}$, controlling the halo-mass scaling of the $M_\star - M_{\rm h}$ relationship, and the normalization of the satellite HOD, respectively. The results are shown in Table 3, and we detail below our calculations for each assumption listed.

The power spectrum normalization parameter, σ_8, is currently known to a precision of a few per cent. This parameter has a strong impact on the large-scale galaxy clustering, and a larger value would lead to an increased number of massive structures, hence shifting the massive end of the halo mass function towards more massive haloes. Choosing *Planck* over *WMAP*7 cosmology (as for the tests in Section 3.5), would result in a 5 per cent increase in σ_8, leading to relatively small changes in best-fitting HOD parameters, of the order of a few per cent.

Halo bias uncertainties originate from the measurement of the bias-to-halo mass relation $b(M_{\rm h})$ using *N*-body simulations, affected by low-mass resolution, small volume, or the limitations of halo identification techniques. In the low-clustering regime, the typical errors on the bias are as small as a few per cent (Tinker et al. 2010), however the rather shallow slope of bias versus halo mass (see e.g. fig. 18 of Coupon et al. 2012) translates into a larger uncertainty in the deduced halo mass. In the high-mass regime, errors are mainly dominated by the sample variance of simulations, up to ~10 per cent, but have fewer impact on the deduced halo mass owing to the steeper slope in this regime.

The assembly bias (Zentner et al. 2014, and references therein) refers to the correlation between clustering amplitude and time of halo formation, whereas in our model the bias is assumed to vary only with halo mass. The effect is stronger when selecting a population of galaxies based on a parameter correlated with halo formation history, such as the SFR, but moderate when considering the full galaxy population selected by stellar mass only. In this case, and in the mass regime explored in this study, Zentner et al. (2014) found that the systematics caused by assembly bias on HOD parameters do not exceed 10–15 per cent.

In our model, the dark matter halo profile is assumed to follow a Navarro, Frenk & White (1997, NFW) profile. While lensing observations tend to favour NFW profiles (Umetsu et al. 2011; Coupon, Broadhurst & Umetsu 2013; Okabe et al. 2013), the mass–concentration relation – driving the slope of the profile – remains uncertain. We have used a simple mass–concentration relation based on theoretical predictions (updated from Takada & Jain 2003) and empirical redshift evolution (Bullock et al. 2001), but more recent relations such as the work from Muñoz-Cuartas et al. (2010) have been measured. Compared to our concentration values, the difference with Muñoz-Cuartas et al. rises from 11 per cent at $M_{\rm h} \sim 10^{12}\,{\rm M}_\odot$ to 30 per cent at $\sim 10^{15}\,{\rm M}_\odot$ (with a minimum of 2 per cent at $\sim 10^{13}\,{\rm M}_\odot$). These systematics affect the slope of the small-scale clustering and galaxy–galaxy lensing. We estimate that if all of our constraints came from lensing, this may result in a 28 per cent systematic error in M_1.

We assume that the satellite distribution in the halo follows the dark matter density profile. However, this assumption may not be always true and Budzynski et al. (2012) tested this hypothesis from a stacked analysis of massive clusters from the SDSS. They found a typical factor of 2 (with ~50 per cent scatter) lower concentration of the satellite distribution compared to dark matter, whereas Muzzin et al. (2007) measured a value closer to dark matter around $z \sim 0.3$, and van der Burg et al. (2014) a relatively high concentration at $z = 1$. These trends may show a redshift evolution of the concentration or can simply be inherent to the difficulty of observationally measuring the satellite distribution. In Table 3, we report the impact on $B_{\rm sat}$ after setting the satellite concentration a factor of 2 higher than that of dark matter. The effect on $B_{\rm sat}$ does not exceed 11 per cent.

Finally, in our model we neglect the lensing contribution of the subhaloes hosting the satellite galaxies. This effect, first introduced by Mandelbaum et al. (2005b) under the term 'stripped satellite central profile', assumes that a fraction of the satellite haloes survive inside the host halo and further contribute to the lensing signal at

Table 3. Estimated systematic errors from the model on the central halo mass, $\log_{10}M_1$, and the satellite normalization, $B_{\rm sat}$. The total is the sum in quadrature of the errors.

	Error on $\log_{10}M_1(\sim 12.7)$		Error on $B_{\rm sat}(\sim 10)$		
Assumption	$M_\star({\rm M}_\odot) = 10^{10}$	$10^{11.5}$	$M_\star({\rm M}_\odot) = 10^{10}$	$10^{11.5}$	Affected quantities
σ_8	0.05	0.05	1	0.5	SMF, clustering (small and large scales)
$b(M_{\rm h})$ relation	0.08	0.1	–	–	Clustering (large scale)
Assembly bias[a]	<0.04	<0.04	~1.5	~1.5	SMF, clustering (small and large scales)
$c(M_{\rm h})$ relation	0.11	0.03	0.1	0.4	Clustering (small scale), lensing
Satellite concentration	–	–	1.1	0.9	Clustering (small scale), lensing
Stripped subhaloes	0.09	0.07	–	–	Lensing
Total	0.17	0.14	2.1	1.9	All

Note. [a]From Zentner, Hearin & van den Bosch (2014).

small scales. As a result, the lensing contribution of those subhaloes adds up to the central-galaxy halo term in such a way that the best-fitting *host* halo mass gets reduced compared to a model in which the contribution of subhaloes is neglected. Hudson et al. (2015) quantify the systematic change in best-fitting halo mass as a systematic decrease by a factor of $\sim(1 + f_{\rm sat})$, where $f_{\rm sat}$ is the fraction of satellites in the sample. Assuming a satellite fraction between 20 and 30 per cent, this leads to a systematic error of up to 0.09 in $\log_{10}M_1$. This number must be read as if all the constraints would come from lensing only. In our study where the stellar mass function and the clustering signal-to-noise ratio is higher than that of the lensing, this effect plays relatively little role, and our results would not significantly change if we accounted for it.

The sum in quadrature of these model systematics is shown as 'Total' in Table 3. Intermediate stellar mass bins ($\sim 10^{10}$ M$_\odot$) seem to be most affected, with an error of 0.17 for $\log_{10}M_1$ (~50 per cent in M_1) and 2.1 (~20 per cent) for $B_{\rm sat}$. We will see below that these values dominate over the typical statistical and systematic errors from the measurements in this mass regime. However, as each of these systematic errors affects the observables in a different way and we fit all of them jointly, one must see these numbers as pessimistic estimates. The high-stellar mass bin ($\sim 10^{11.5}$ M$_\odot$) is equally affected but in a regime where statistical errors are large, hence leading to a smaller impact.

4.6 MCMC sampling

We write the combined log-likelihood as the sum of each observable χ^2:

$$-2\ln\mathcal{L} = \chi^2_{\phi} + \sum_{\rm spl}\chi^2_{w(\theta)} + \sum_{\rm spl}\chi^2_{w_{\rm p}(r_{\rm p})} + \sum_{\rm spl}\chi^2_{\Delta\Sigma}\,, \qquad (15)$$

where individual χ^2's are computed as

$$\chi^2 = \sum_{i,j}[\widetilde{X}_i - X_i](C^{-1})_{ij}[\widetilde{X}_j - X_j]\,, \qquad (16)$$

using the covariance matrices evaluated for each measurement as described in Section 3 ($\widetilde{X}$ and X represent the measured and modelled observables, respectively). Each observable χ^2 is summed over the samples ('spl') as described in Table 2. The 'i' and 'j' subscripts refer to the stellar mass (stellar mass function) or transverse separation (clustering and lensing) binning of each measurement.

We find the best-fitting parameters and posterior distribution (assuming flat priors for all parameters) employing the MCMC sampling technique with the Metropolis–Hastings sampler from the software suite COSMOPMC (Wraith et al. 2009). We check for individual chain convergence and chain-to-chain mixing using the Gelman & Rubin (1992) rule from the R-language CODA package.[10] We find a typical chain-to-chain mixing coefficient (potential scale reduction factor) to be equal to 1.01, and the acceptance rate around 30 per cent.

In practice, we first evaluate a diagonal Fisher matrix at the maximum likelihood point found using the Amoeba algorithm (Press et al. 2002) and run 10 chains in parallel with the inverse Fisher matrix as the MCMC sampler covariance matrix. The acceptance rate is usually very low due to the noisy diagonal Fisher matrix affected by some strong correlations between parameters. Once the chains have converged (after typically 5000–10 000 steps) we compute the final likelihood covariance matrix after rejecting the burn-in phase of the chains (a few thousand steps). This covariance matrix is used as the input sampler covariance matrix of a second and final MCMC run, in which 10 chains of 30 000 steps each are computed in parallel and combined together assuming a burn-in phase of 2000 steps and checking for proper mixing.

Table 4. HOD best-fitting parameters and 68 per cent confidence limits (CL) for the statistical errors (top) and total errors (bottom).

Parameter	Mean	Upper CL	Lower CL
Jackknife resampling errors			
$\log_{10}M_1$	12.84	0.020	− 0.026
$\log_{10}M_{\star,0}$	10.98	0.015	− 0.019
β	0.48	0.017	− 0.021
δ	0.63	0.094	− 0.073
γ	1.60	0.166	− 0.202
$\sigma_{\log M_{\star},0}$	0.337	0.045	− 0.035
λ	0.21	0.047	− 0.044
$B_{\rm sat}$	10.87	0.443	− 0.416
$\beta_{\rm sat}$	0.83	0.038	− 0.035
α	1.17	0.020	− 0.021
Total errors			
$\log_{10}M_1$	12.67	0.124	− 0.083
$\log_{10}M_{\star,0}$	10.90	0.082	− 0.067
β	0.36	0.077	− 0.051
δ	0.75	0.193	− 0.151
γ	0.81	0.477	− 0.386
$\sigma_{\log M_{\star},0}$	0.394	0.100	− 0.074
λ	0.25	0.082	− 0.083
$B_{\rm sat}$	9.96	0.938	− 0.845
$\beta_{\rm sat}$	0.87	0.078	− 0.065
α	1.14	0.040	− 0.038

We run the full MCMC procedure twice. The first run is performed using the statistical covariance matrices from the jackknife estimator and the second MCMC run uses the total error covariance matrices, which are constructed from the statistical covariance matrices plus the systematic error estimates added in quadrature to the diagonal, as described in Section 3.5.

5 RESULTS

Best-fitting parameters with 68 per cent confidence intervals are given in the top panel of Table 4 for the statistical- and total-error MCMC runs. The 1D and 2D likelihood distributions are shown in Fig. D1. The reduced χ^2_ν for the statistical-error fit is $\chi^2/(N_{\rm points} - N_{\rm parameters}) = 260/(160 - 10) = 1.7$, which is an overestimate given the correlations neglected in the computation of the log-likelihood. Firstly, we recall that the lensing and clustering measurements are affected by a sample-to-sample correlation due to the scatter in stellar mass. The re-use of background galaxies in the lensing measurements causes an additional sample-to-sample correlation. Secondly, the projected and real-space clustering are correlated, as both observables bring similar information. This mostly affects the satellite distribution parameter errors, which could be slightly underestimated. Finally, the few number of subsamples (64) used in the computation of a noisy covariance matrix may have biased the inverse estimate and contributed to an increase in χ^2_ν.

5.1 Measurements and best-fitting models

The measured stellar mass function and best-fitting model are displayed in Fig. 6. Statistical error bars and corresponding best-fitting model are shown as thick black lines, whereas total (statistical plus

---

[10] http://cran.r-project.org/web/packages/coda/citation.html

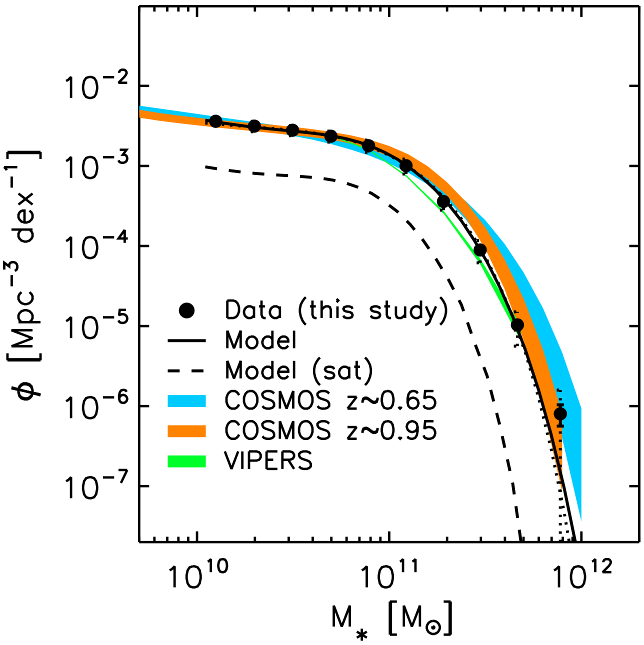

Figure 6. Measured stellar mass function and best-fitting model in the range $0.5 < z < 1$. The statistical errors from the jackknife estimate are shown as black thick lines, whereas the total (statistical plus systematic) error bars as dotted lines. The COSMOS (Ilbert et al. 2013) and VIPERS (Davidzon et al. 2013) mass functions are displayed with their respective statistical errors as shaded areas.

systematic) errors and corresponding best-fitting model are represented in dotted lines. We compare our measurements with the COSMOS mass function evaluated in the ranges $0.5 < z < 0.8$ and $0.8 < z < 1.1$ by Ilbert et al. (2013), and the VIPERS stellar mass function (Davidzon et al. 2013), measured in the range $0.5 < z < 1$ (Davidzon, private communication).

The clustering measurements and best-fitting models are shown in Fig. 7. The projected two-point correlation functions $w(\theta)$ are displayed in the top panels. The mass ranges are given in each top-right corner in units of $\log(M_\star/\mathrm{M}_\odot)$. Similarly, the real-space two-point correlation functions $w_{(r_\mathrm{p})}$ are displayed in the bottom panels.

The galaxy–galaxy lensing measurements and best-fitting models are shown in Fig. 8. The most massive lensing bin features a few data points lower than the model around the transition between the central and the satellites term.

For all observables, we report good agreement between the data and the model. The constraints on the shape of the central $M_\star - M_\mathrm{h}$ relationship (parametrized by $\log_{10}M_1$, $\log_{10}M_\mathrm{star0}$, β, γ and δ), are mostly driven by the high signal-to-noise stellar mass function measurements. Satellite HOD parameters (B_sat, β_sat and α) are mainly constrained by the clustering and lensing measurements. The amplitude of clustering at small scale is directly proportional to the relative number of satellites, hence giving strong leverage on the satellite galaxy HOD. Additional information is given on scales $r \sim 0.1$ Mpc from lensing, through the satellite lensing signal. The dispersion in $M_\star$ at fixed M_h, parametrized in amplitude by $\sigma_{\log M_\star,0}$ and in power-law slope by λ, is mainly constrained by the high-mass end of the stellar mass function and the amplitude of the galaxy–galaxy lensing signal in the most massive bins, resulting in a high-mass ($M_\star \sim 10^{11}\,\mathrm{M}_\odot$) scatter of approximately $\sigma_{\log M_\star} \simeq 0.2$ in both the jackknife and total error cases, and a medium mass ($M_\star \sim 10^{10}\,\mathrm{M}_\odot$) scatter of $\sigma_{\log M_\star} \simeq 0.35$.

Because the stellar mass function is most affected by the inclusion of systematics in the error budget, we note a significant increase in uncertainties associated with the parameters driving the central $M_\star - M_\mathrm{h}$ relationship. From Table 4, we report an increase from a factor ∼3 in the error in γ, up to a factor ∼6 in the error in $\log_{10}M_1$. HOD parameters describing the satellite occupation function such as B_sat, β or α show substantially less sensitivity to the addition of systematic errors in the error budget (a maximum of factor ∼2 increase is found). This is explained by the relatively smaller contribution of systematic versus statistical errors affecting the clustering and lensing measurements, compared to the stellar mass function.

The occasional large differences between best-fitting parameters from statistical alone and total errors, seen in Table 4, do not lead to significantly different derived quantities, owing to the strong correlations between parameters. This is confirmed by the almost indistinguishable dotted lines and thick lines in Figs 6–8, and is most probably a consequence of having symmetrically added the systematic errors to the statistical errors.

5.2 Central $M_\star - M_\mathrm{h}$ relationship and the SHMR

In Fig. 9, we show the best-fitting central galaxy $M_\star - M_\mathrm{h}$ relationship (left-hand panel) as parametrized by equation (7), and the SHMR (right-hand panel). The SHMR is shown as function of host halo mass and is derived for the central galaxy in dark grey (from the $M_\star - M_\mathrm{h}$ relationship), the satellites in light grey (integrated over the galaxies above a mass threshold of $M_\star > 10^{10}\,\mathrm{M}_\odot$), and the total in black.

The shaded areas represent the 68 per cent confidence limits, and in the bottom-left panel, we have shown the results obtained with statistical errors in light blue and with total errors in black. As for the stellar mass function, the statistical uncertainties grow by a factor of ∼2–4 in the lower mass regime, when incorporating systematics.

The central SHMR peak position is indicated by a black arrow located at $M_\mathrm{h,peak} = 1.92^{+0.17}_{-0.14} \times 10^{12}\,\mathrm{M}_\odot$. The SHMR peak value is $\mathrm{SHMR}_\mathrm{peak} = 2.2^{+0.2}_{-0.2} \times 10^{-2}$. When accounting for satellites, the peak position and value do not significantly differ from the estimates for centrals only. However, a remarkable result highlighted in this figure is the increasing contribution of stellar mass enclosed in satellites as function of halo mass. When reaching cluster-size haloes, this contribution reaches over 90 per cent (and presumably higher when accounting for satellite galaxies with masses lower than $10^{10}\,\mathrm{M}_\odot$). However, we stress that we do not take into account the intracluster light, which is challenging to quantify using ground-based photometric data.

5.3 Comparison with the literature

In Figs 10 and 11, we compare our best-fitting $M_\star - M_\mathrm{h}$ relationship for central galaxies with a number of results from the literature. As described in Section 4, our relation describes the mean stellar mass at fixed halo mass which is, due to the scatter in stellar mass, not equivalent to the mean halo mass at fixed stellar mass. This issue becomes particularly important when the slope of the stellar or halo mass distribution is steep (i.e. at high mass). Therefore, we have recomputed our results using the latter definition and we consistently compare our results with the literature in each case.

When required, we convert halo masses to our virial definition using the recipe given by Hu & Kravtsov (2003) in their Appendix C and, following Ilbert et al. (2010), we divide stellar masses by a factor of 1.74 and 1.23 to convert from Salpeter (1955) and 'Diet'

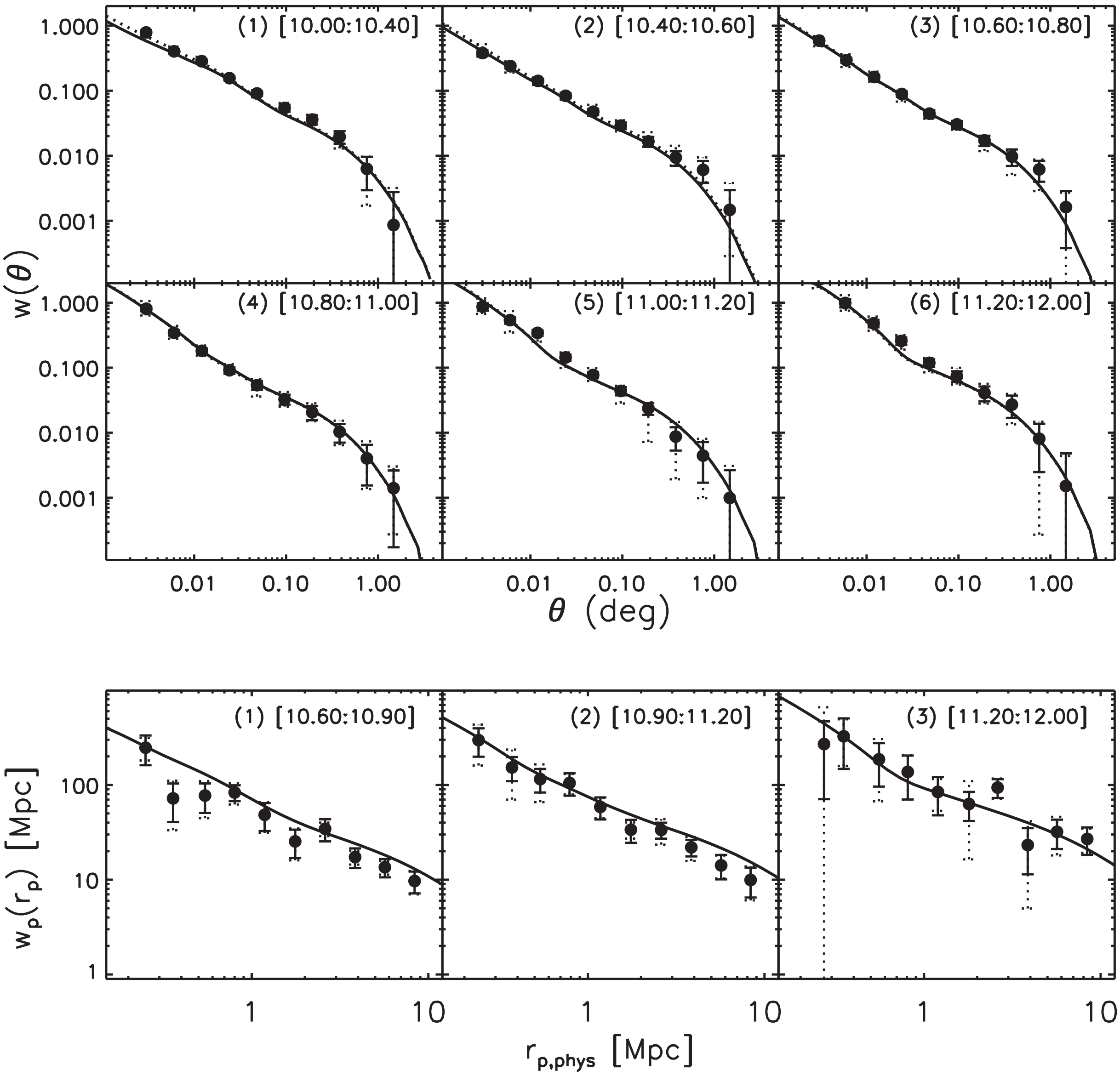

Figure 7. Galaxy clustering measurements (data points with error bars) and best-fitting models (thick lines). The top panels show the projected $w(\theta)$ from the photometric sample (the measurements are corrected for the integral constraint), and the bottom panels show the spectroscopic real-space $w_{\mathrm{p}}(r_{\mathrm{p}})$. The thick error bars associated with thick lines represent the statistical errors and subsequent best-fitting models, whereas dotted lines are for total errors. The mass ranges in the top-right corner of each panel are given in $\log(M_\star/\mathrm{M}_\odot)$.

Salpeter IMFs, respectively, to our Chabrier IMF stellar masses. We apply no correction to Kroupa (2001) IMF stellar masses.

The mean redshift, measured from the sum of the photometric redshift PDFs, is found to be $\langle z\rangle = 0.82$ for our measurements in the range $0.5 < z < 1.0$ ($M_\star > 10^{10.40}\,\mathrm{M}_\odot$) and $\langle z\rangle = 0.65$ in the range $0.5 < z < 0.7$ ($10^{10} < M_\star/\,\mathrm{M}_\odot < 10^{10.40}$). We point out that the lensing signal is more sensitive to lower redshift lens galaxies characterized by a higher signal-to-noise ratio (due to the more numerous background sources), and is likely to be more representative of a lower redshift population, but this effect is assumed to be small compared to the lensing statistical errors.

5.3.1 $\langle M_\star|M_{\mathrm{h}}\rangle$ results

We first compare the results for $\langle M_\star|M_{\mathrm{h}}\rangle$ in Fig. 10. The black shaded area shows our results for the central galaxy relationship with 68 per cent confidence limits from the total errors. The total errors consist of the statistical uncertainties from jackknife resampling in addition to three sources of systematic effects from the measurements: the cosmology chosen among widely-used ΛCDM models, the fine-tuning of our dust extinction law modelling, and potential biases in the photometry/calibration. We recall that this list of systematic uncertainties is not exhaustive and, for example, ignores the choice of SPS models, which may be responsible for even larger systematic effects. An estimate of the systematic errors from the model, as detailed in Section 4.5, is also shown in the bottom-right corner.

Behroozi et al. (2013b), shown as the light-blue shaded area, put constraints on the $M_\star$–M_{h} relationship by populating dark matter haloes in simulations and comparing abundances using observed stellar mass functions from a number of surveys. They

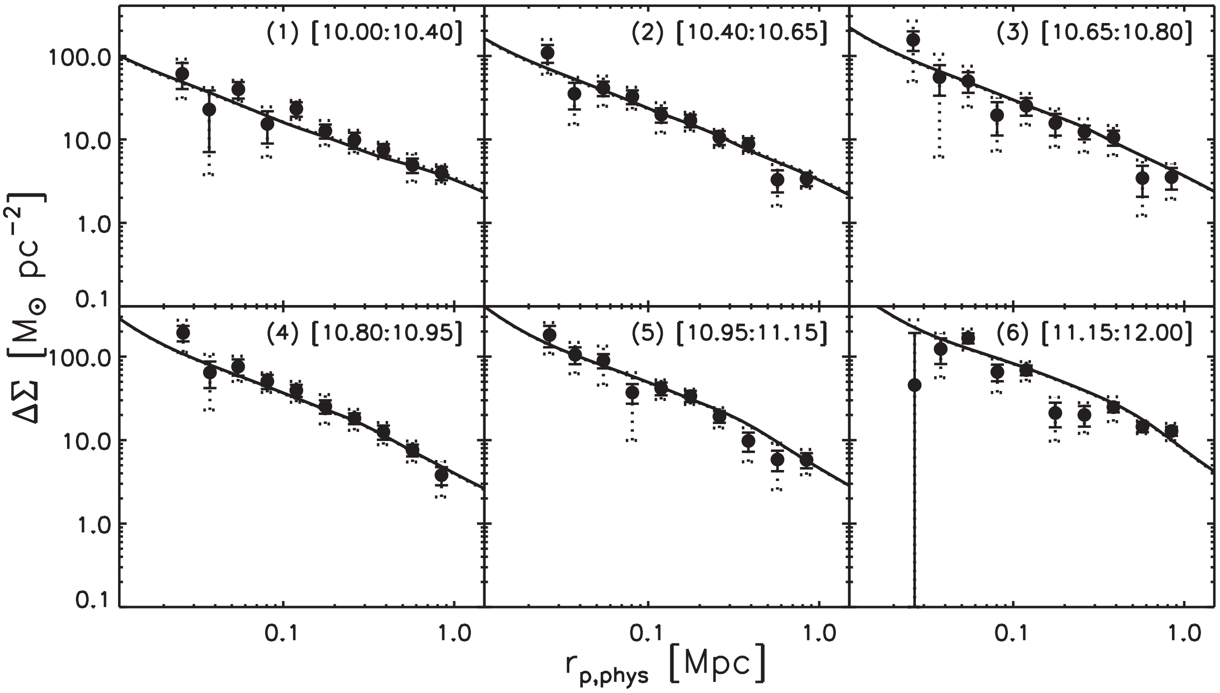

Figure 8. Galaxy–galaxy lensing signal measurements (data points with error bars) and best-fitting models (thick lines). As in Fig. 7, thick and dotted lines are for statistical and total error results, respectively. The mass ranges in the top-right corner of each panel are given in $\log(M_\star/\mathrm{M}_\odot)$.

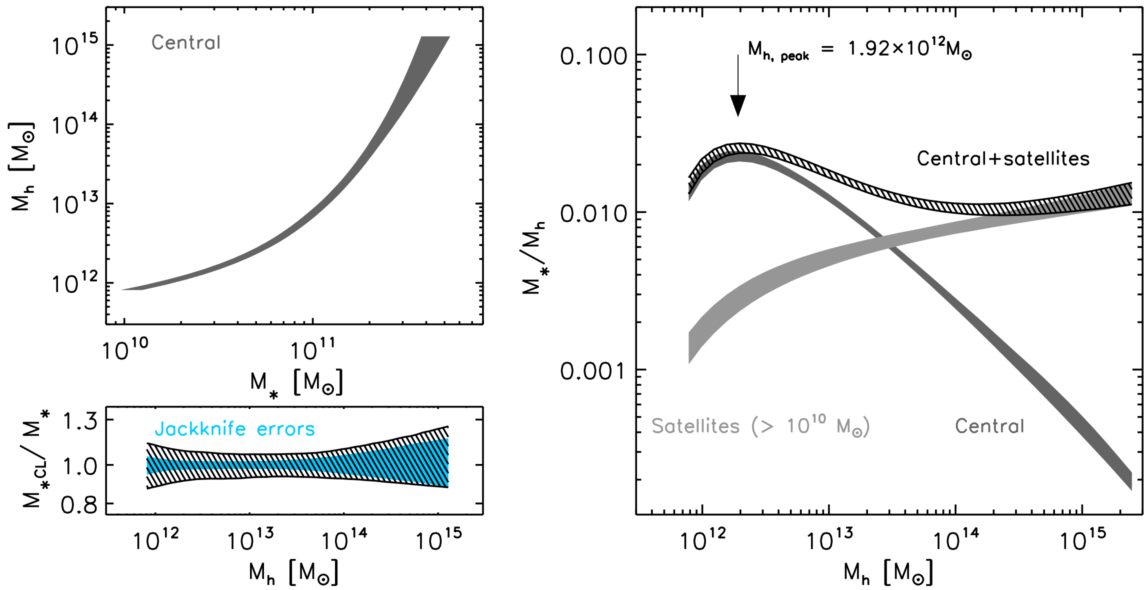

Figure 9. Best-fitting $M_\star - M_\mathrm{h}$ relationship (left) and SHMR (right). The black shaded areas represent the confidence limits from the total errors. The bottom-left panel shows the confidence limits interval as a function of halo mass in the case of statistical errors (from jackknife resampling in light blue) and total errors (in black). The SHMR is derived as function of host halo mass for the central galaxy (dark grey), the satellites (light grey) and the sum of both (black). The peak value of the central SHMR is indicated by the black arrow.

characterized the uncertainties affecting stellar mass estimates by accounting for a number of systematic errors. In particular, unlike in our systematic errors, the authors had to include uncertainties arising from the choice of the IMF and the SPS galaxy templates, necessary when combining the stellar mass functions from several works using different stellar mass measurement methods. Here, we consider their results at $z \sim 1$. A significant difference with our model resides in the assumption that satellite galaxies in larger haloes are seen as central galaxies in subhaloes. To circumvent the difficulty of accurately predicting a subhalo mass function (e.g. complications from tidal stripping), the galaxies in subhaloes at the time of interest are matched to their progenitors at the time of merging on to the central galaxy halo, under the assumption that the $M_\star - M_\mathrm{h}$ evolution at a given stellar mass is identical whether the host halo is isolated or inside a larger halo. In comparison, our model is a ‘snapshot’ of the galaxy halo occupation at a given time, where the satellite distribution is mainly constrained by galaxy clustering.

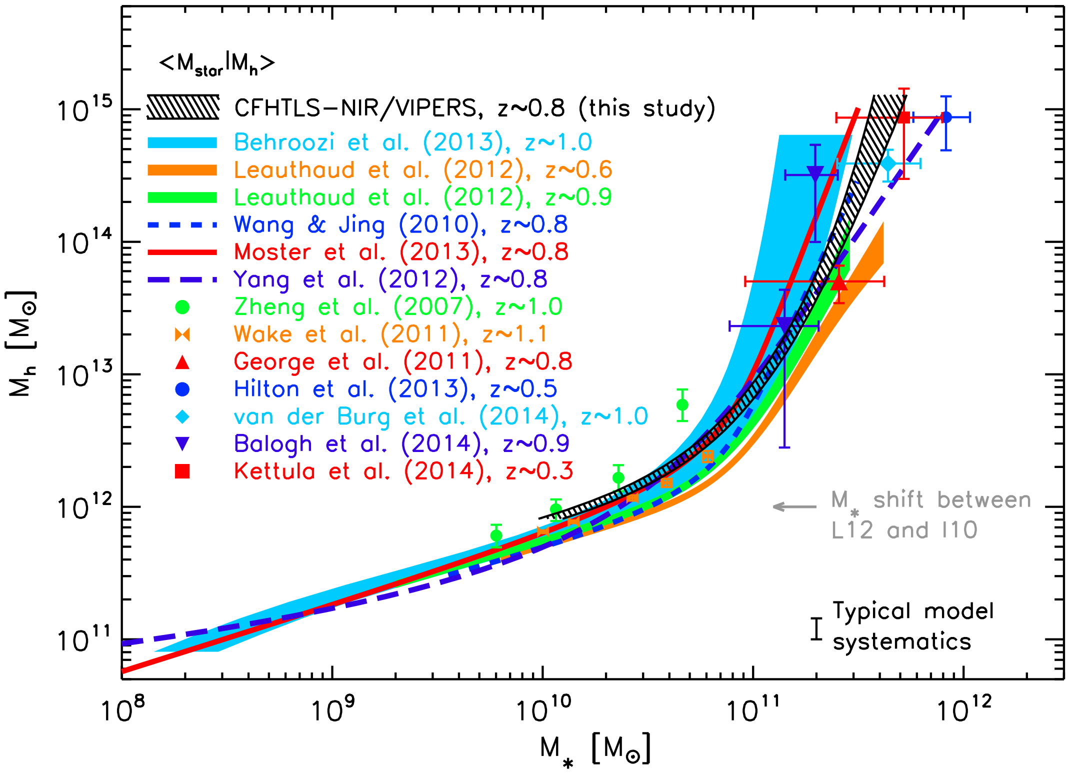

Figure 10. The best-fitting $M_\star - M_\mathrm{h}$ relationship for central galaxies, shown in the black shaded area (total-error based 68 per cent confidence limits), compared with a number of results from the literature at similar redshifts. The results shown here represent the mean stellar mass at fixed halo mass or halo-mass proxy (X-ray temperature or satellite kinematics), $\langle M_\star | M_\mathrm{h} \rangle$, but plotted M_h as function of $M_\star$ to ease the comparison with the literature. We perform appropriate halo mass conversions and IMF stellar mass corrections when required. The length of the grey arrow in the bottom-right corner shows the shift ($\sim$0.2 dex) measured from the direct comparison between stellar masses used in Leauthaud et al. (2012) and George et al. (2011), compared to those in Ilbert et al. (2010) which were estimated in a similar way to this study. The error bar on the bottom-right corner indicates the typical systematic uncertainty arising from the model.

The results from Leauthaud et al. (2012) in COSMOS are shown in brown and green at redshifts $z \sim 0.6$ and $\sim$0.9, respectively. We observe a small discrepancy which, compared to our results, is unlikely to be explained by differences in the modelling of the HOD (since the model is essentially identical), nor the sample variance as confidence limits do not overlap. A difference in stellar mass estimates on the other hand is more likely to be at the origin of the discrepancy. To check this hypothesis, we have compared the stellar mass estimates from Ilbert et al. (2010), which were measured in a similar way to this study, with those used in Leauthaud et al. (2012) with the method described in Bundy et al. (2006). We measured an offset of $\sim$0.2 dex, illustrated in Fig. 10 as the grey arrow. Part of the difference seems to be explained by the separate choice for the dust extinction law made in each study (which may typically cause a $\sim$0.14 dex offset, see Section 2.6). However, we note that in both cases the same IMF and set of SPS models were used, which leaves us without a complete understanding of the difference.

The results by Wang & Jing (2010) are shown as the blue short-dashed line. Their model is based on a HOD modelling of the stellar mass function and real-space galaxy clustering where, as in Behroozi et al. (2013b), the treatment for satellites is not based on the distribution of subhaloes in the host halo but on the $M_\star - M_\mathrm{h}$ relationship at the time of infall.

Moster, Naab & White (2013), shown as the red dot–dashed line, also used abundance matching and provided a redshift-dependent parametrization of the central $M_\star - M_\mathrm{h}$ relationship that we have calculated at $z = 0.8$. As above, the satellites are matched to their haloes at the epoch of merging. Their relation is in good agreement with ours at intermediate mass, however, it shows a steeper dependence on stellar masses at higher mass.

The green dots with error bars are from the HOD modelling results of Zheng et al. (2007), based on real-space clustering and number density measurements. Here, we show their results for DEEP2, a deep spectroscopic survey with high density $z = 1$ galaxies. Without deep NIR data, the authors have computed mean approximate stellar masses for galaxy samples selected in bins of luminosity. This source of uncertainty is not shown on the plot, however, one may expect a large scatter and potential biases due to this conversion.

The orange bow-ties with error bars represent the results[11] by Wake et al. (2011) in the NEWFIRM Medium Band Survey at redshift $z \sim 1.1$, from the combination of NIR-selected galaxy clustering and number density measurements. Their results are in good agreement with ours.

The five next results were produced using galaxy cluster samples associated with their brightest cluster galaxies (BCG). George et al. (2011) built up a catalogue of central versus satellite galaxies in COSMOS, matched to an X-ray detected group/cluster sample with robust halo masses from weak lensing (Leauthaud et al. 2009). From their catalogue, we have computed the mean of stellar mass and halo mass values for clusters in the range $0.5 < z < 1$, shown as the single red triangle (the error bars show the standard deviation

---

[11] Here, we use updated results compared to the original publication, estimated with Bruzual & Charlot (2003) templates and with rectified h-scaling (Wake, private communication)

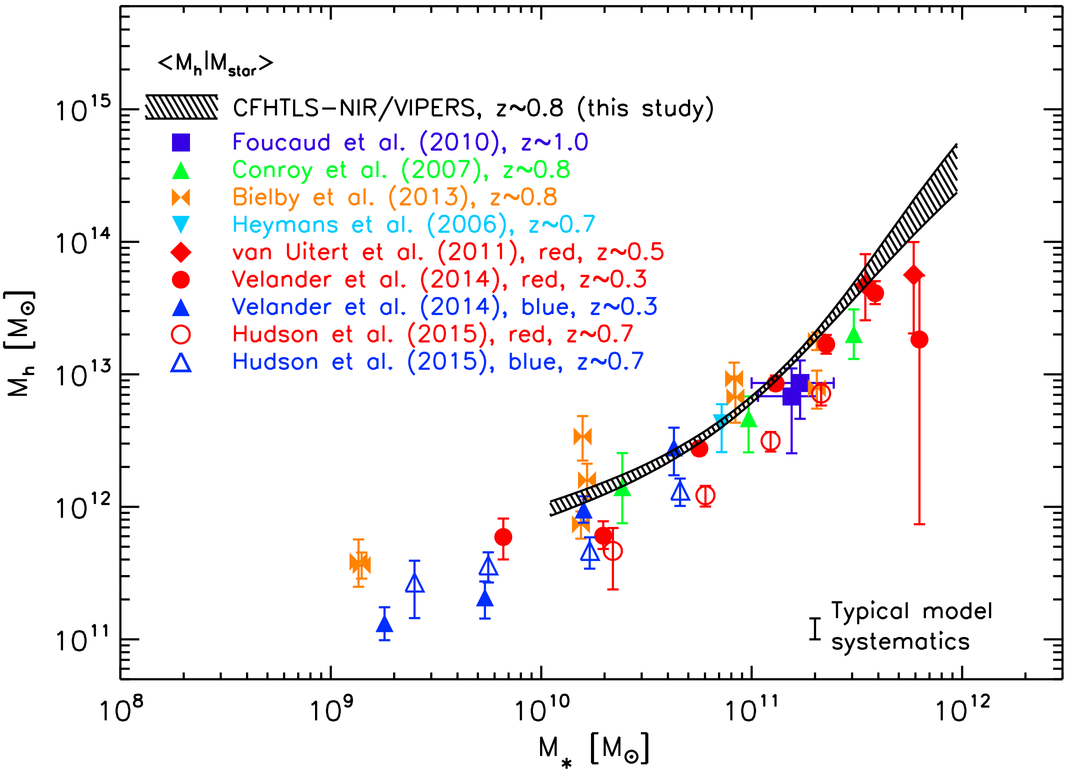

Figure 11. The best-fitting $M_\star - M_{\rm h}$ relationship for central galaxies, shown in the black shaded area (total-error based 68 per cent confidence limits), compared with a number of results from the literature at similar redshifts. Unlike in Fig. 10, the results shown here represent the mean *halo mass* at fixed stellar mass $\langle M_{\rm h}|M_\star\rangle$. We perform appropriate halo mass conversions and IMF stellar mass corrections when required. The relatively low halo masses found by Hudson et al. (2015) is linked to a different treatment of the satellite subhalo contribution to the lensing signal at small scale (see text for details).

in halo and stellar masses). As they used identical stellar masses to Leauthaud et al. (2012), we also expect a systematic difference in stellar masses compared with our estimates.

From Sunyaev–Zel'dovich detected clusters using the Atacama Cosmology Telescope, Hilton et al. (2013) presented the measurements of the galaxy properties between $0.27 < z < 1.07$. Member galaxies were identified from high-density spectroscopic observations, and stellar masses were measured from *Spitzer* IRAC1-2 mid-infrared (MIR) fluxes. Halo masses were estimated from satellite kinematics. Here, we show the mean halo mass versus mean BCG stellar mass, represented by the single blue dot with errors bars (standard deviations of both masses). Their results appear to be in good agreement with our $M_\star - M_{\rm h}$ relationship, although our constraints on such high-mass clusters are extrapolated from the few clusters more massive than $4–5 \times 10^{14}\,{\rm M}_\odot$ expected in our sample.

We show as a single light blue diamond the mean halo mass versus mean BCG stellar mass from van der Burg et al. (2014) in the GCLASS/SpARCS cluster sample at $z \sim 1$. Galaxy cluster members were identified from intensive spectroscopic observations, and halo masses were estimated from satellite kinematics. We note that stellar masses were measured from a similar combination of data, redshift range and volume size as ours, however, the methodology used to link halo mass to galaxy stellar masses was rather different. Thus, the agreement with our high-mass $M_\star - M_{\rm h}$ relationship within the sample variance is quite remarkable.

Results from Balogh et al. (2014) are shown as the downward purple triangles. Halo mass measurements were made using satellite kinematics for a sample of 11 groups/clusters in the COSMOS field. We show the mean and standard deviation of their measurements split into two halo mass bins (the 11 groups are split into 5 and 6 groups below and above $M_{\rm h} = 9 \times 10^{13}\,{\rm M}_\odot$, respectively). Although their results suffer from large sample variance, they are in broad agreement with our results and with the rest of the literature.

Finally, the single red square with error bars shows the mean of halo mass measurements from a weak lensing analysis of X-ray selected clusters in the CFHTLenS by Kettula et al. (2014), versus the mean stellar mass of associated BCGs (Mirkazemi et al. submitted). We have re-measured stellar masses of those BCGs in a consistent way to this study (with the exception of missing NIR data for most of the BCGs, which may increase the scatter in stellar mass). Despite the lower redshift range, the identical photometry and lensing catalogue makes the comparison relevant to our results, where the expected difference should arise solely from redshift evolution, although the large statistical uncertainties prevent us from drawing strong conclusions.

5.3.2 $\langle M_{\rm h}|M_\star\rangle$ results

We compare the results for $\langle M_{\rm h}|M_\star\rangle$ in Fig. 11. To express the mean halo mass at fixed stellar mass $\langle M_{\rm h}|M_\star\rangle$ from our results, we derive it from the mean stellar mass at fixed halo mass $\langle M_\star|M_{\rm h}\rangle$ using the Bayes theorem relating conditional probability distributions:

$$P(M_{\rm h}|M_\star) \propto P(M_\star|M_{\rm h}) \times P(M_{\rm h})\,. \qquad (17)$$

We can then compute $\langle M_\mathrm{h}|M_\star\rangle$ as the expectation value of $P(M_\mathrm{h}|M_\star)$:

$$\langle M_\mathrm{h}|M_\star\rangle = \frac{\int P(M_\star|M_\mathrm{h})\,P(M_\mathrm{h})M_\mathrm{h}\mathrm{d}M_\mathrm{h}}{\int P(M_\star|M_\mathrm{h})\,P(M_\mathrm{h})\mathrm{d}M_\mathrm{h}} \quad (18)$$

with

$$P(M_\star|M_\mathrm{h}) = \frac{\mathrm{d}\langle N_\mathrm{cen}(M_\mathrm{h}|M_\star)\rangle}{\mathrm{d}M_\star}\,, \quad (19)$$

the distribution of central galaxies given a halo mass, and

$$P(M_\mathrm{h}) = \frac{\mathrm{d}n}{\mathrm{d}M_\mathrm{h}}\,, \quad (20)$$

the halo mass function.

We show the results of Foucaud et al. (2010) at $z \sim 1$ from clustering measurements in the UKIDSS-UDS field as the blue squares with error bars. The UKIDSS-UDS field is a small patch of $\sim$1 deg^2 with deep NIR and optical data. They have converted their clustering amplitude measured in bins of stellar mass into halo masses, using the analytical galaxy-bias halo-mass relationship from Mo & White (1996). As they do not use any constraints from galaxy number density, their error bars are dominated by sample variance and uncertainties on the projected galaxy clustering.

Green upward triangles represent the results by Conroy et al. (2007). Halo masses were derived from satellite kinematics using spectroscopic measurements from the DEEP2 survey. Since the authors have selected their samples based on bins of stellar masses, we can compare their results with our $\langle M_\mathrm{h}|M_\star\rangle$ $M_\star{-}M_\mathrm{h}$ relationship. The agreement is found to be good.

Results from clustering measurements in the CFHTLS-DEEP/WIRDS fields by Bielby et al. (2014) are displayed by the brown bow-ties with error bars. We select all mass bin results in the range $0.5 < z < 1$. Although the total field of view is small ($\sim$2.4 deg^2), the combination of four independent fields allowed them to reduce the cosmic variance. As in Foucaud et al., they used an analytical prescription based on the large-scale clustering amplitude to estimate halo masses per bin of stellar mass, so that their results should be compared to our $\langle M_\mathrm{h}|M_\star\rangle$ results. The two points well above the other results correspond to the measurements at $z \sim 0.7$ and seem to disagree with our constraints and the rest of the literature. The authors claim to have observed an unusually high clustering signal at those redshifts, potentially explained by cosmic variance effects.

Results by Heymans et al. (2006) in the COMBO-17/GEMS field are shown as the downward light-blue triangle with error bars. Here, we have picked their unique measurement at $z > 0.5$. Halo masses were measured using weak lensing with galaxy shapes from the *Hubble Space Telescope* observations.

We show as red diamonds the results for $z \sim 0.5$ red galaxies by van Uitert et al. (2011) in the Red Sequence Cluster Survey 2, a medium-deep CFHT-MegaCam survey in three bands (*gri*) which overlaps 300 deg^2 of the SDSS. The authors have measured the galaxy–galaxy lensing signal for SDSS lens galaxies with a spectroscopic redshift using background source galaxies from the RCS2 survey. Here, the large area permits a high signal-to-noise measurement for very massive galaxies from lensing only. Their results are consistent with ours as this mass bin ($>3 \times 10^{11}$ M$_\odot$) is dominated by red galaxies.

We compare our results with those from Velander et al. (2014) at $z \sim 0.3$, shown as filled symbols (red dots and blue triangles for red and blue galaxies, respectively), and those from Hudson et al. (2015) at $z \sim 0.7$, shown as empty symbols (red dots and blue triangles for red and blue galaxies, respectively). In both studies, halo mass measurements were obtained from galaxy–galaxy lensing measured using the CFHTLenS lensing catalogue and stellar masses computed in a similar way to this study, with the exception that, in both cases, no NIR data were available at the time. This mostly affects the stellar mass estimates of Hudson et al. at $z \sim 0.7$ which, unlike Velander et al. at $z \sim 0.3$, do not benefit from the leverage of the CFHTLS z band. We expect the $M_\star{-}M_\mathrm{h}$ relationship of the full galaxy population to lie between those of the red and blue populations, however, the results from Hudson et al. lie below our results for both galaxy populations. The bias caused by the scatter in stellar mass partially explains this difference (by shifting their mean stellar mass to higher values), but not entirely: Hudson et al. account for the contribution of subhaloes around satellites occurring at small scale in the lensing signal, whereas we do not (see Section 4.5[12]). As Velander et al. also accounted for subhaloes in their lensing model, we cannot exclude that the apparent good agreement may result from a redshift evolution going in the opposite direction, and requires further investigation.

5.3.3 The total SHMR

In Fig. 12, we show the SHMR as function of halo mass compared with observations from the literature. The black shaded area represents the total SHMR as the sum of the central and satellite contributions. The central SHMR (in dashed line on the figure) is simply derived from the central $M_\star{-}M_\mathrm{h}$ relationship. The satellite SHMR (in dot–dashed line on the figure) is computed from the sum of satellite stellar masses over the halo occupation function at each halo mass, with a lower integration limit of $M_\star = 10^{10}$ M$_\odot$. The total baryon fraction compared to dark matter in the Universe is assumed to be 0.171 and represented on the figure by the grey shaded area on the top (Dunkley et al. 2009, the width of the line represents the uncertainty).

In green, we display the total SHMR from Leauthaud et al. (2012) measured at $z \sim 0.9$. The procedure to compute the total SHMR is identical to ours, i.e. the integrated stellar masses from the satellite HOD were added to the central stellar mass at each halo mass. The authors adopted a mass threshold of $10^{9.8}$ M$_\odot$, which does not change the integrated stellar mass from satellites by a large amount compared with a cut of $>10^{10}$ M$_\odot$. As shown in Fig. 10, part of the vertical shift is explained by the systematic difference in stellar mass estimates.

We show in light blue the central SHMR from Behroozi et al. (2013a). As seen in Fig. 10, the agreement with our central SHMR is good, although their peak is located at a slightly lower halo mass value.

The red triangle shows the results by George et al. (2011) in COSMOS in the redshift range $0.5 < z < 1$. The point represents the mean total stellar mass divided by the halo mass versus the halo mass, and the error bars the standard deviation in each direction. Here, we computed the total stellar mass as the sum of the central galaxy stellar mass plus the stellar masses of associated group members with $M_\star > 10^{10}$. As they used the stellar masses of Leauthaud et al. the agreement is consistently good with their results, however shifted compared to ours.

The single blue dot with error bars marks the mean and standard deviation of estimates by Hilton et al. (2013). Here, the total cluster

---

[12] This point is also investigated in detail in Appendix D of Hudson et al. (2015).

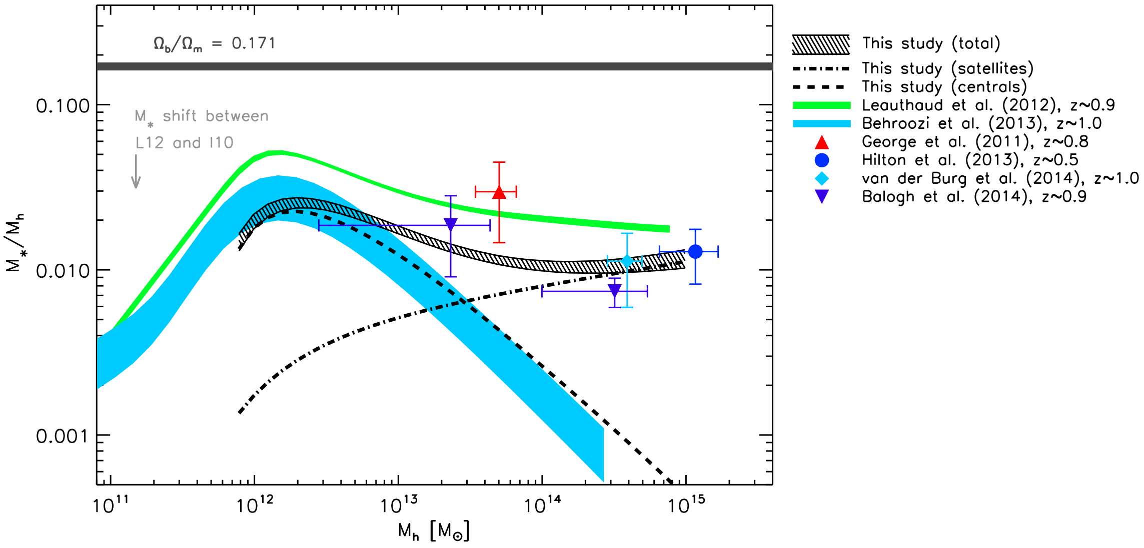

Figure 12. SHMR as function of halo mass compared with observations from the literature. Our best-fitting result for total (central plus satellites) SHMR is shown as the black shaded area. The black dashed line represents the best-fitting central relationship, whereas the dot–dashed line is for the integrated stellar-mass satellite contribution. For Behroozi, Wechsler & Conroy (2013a), only the central SHMR was published and we display it here for comparison with our central SHMR and as an illustration of typical stellar mass systematics. The length of the grey arrow represents the shift to apply to Leauthaud et al. (2012) and George et al. (2011) to reconcile their results with ours, based on the stellar mass comparison with Ilbert et al. (2010).

stellar mass is measured from the background-subtracted sum of galaxy IRAC fluxes within R_{500} from the BCG. Based on the stellar mass completeness computed by Ilbert et al. (2010), an IRAC AB magnitude cut of 24 gives a complete passive galaxy sample down to $M_\star = 10^9\,\mathrm{M}_\odot$ at $z \sim 0.5$. With an IRAC completeness AB magnitude limit of 22.6, it is therefore safe to assume that Hilton et al. are complete above $10^{10}\,\mathrm{M}_\odot$ at $z \sim 0.5$, which matches our sample. We then conclude that their measurements are in good agreement with our results.

Results from van der Burg et al. (2014) are shown as the single light-blue diamond, representing the mean SHMR versus halo mass with its standard deviation. Total stellar masses are computed as the sum of the BCG stellar mass and the stellar mass from galaxy members spectroscopically identified and corrected for TSR. The authors have checked that for $>10^{10}\,\mathrm{M}_\odot$ galaxies, which contribute the most to the total SHMR (see their Fig. 2), the spectroscopic success rate reaches 90 per cent. We note that the median stellar mass completeness $\sim 10^{10.16}\,\mathrm{M}_\odot$ is slightly higher than ours (limited by their K_s-band data), however the contribution of satellites compared to a mass limit of $10^{10}\,\mathrm{M}_\odot$ will not significantly change the total SHMR and their measurements can be fairly compared to our results, and we observe an excellent agreement. Interestingly, the authors conclude that when comparing with the literature, no redshift evolution in the total SHMR at high mass is found below $z \sim 1$ and the comparison with our results ($z \sim 0.8$) and those from Hilton et al. ($z \sim 0.5$) confirm their findings.

The two purple downward triangles represent the results from Balogh et al. (2014) in the GEEC2 survey in COSMOS. Here, we show the mean and standard deviation of the SHMR versus halo mass in two halo mass bins. Galaxy members are identified from the spectroscopic redshift when available or using the PDF-weighted photometric redshift computed from the 30-band COSMOS photometric catalogue (Ilbert et al. 2009). The spectroscopic (photometric) sample is complete for group members with $M_\star > 10^{10.3}\,\mathrm{M}_\odot$ ($M_\star > 10^9\,\mathrm{M}_\odot$). Again, since most of the contribution to the total SHMR originates from $M_\star > 10^{10}\,\mathrm{M}_\odot$ galaxies, the comparison with our results is fair. We note a slightly lower value at high mass, and good agreement within the error bars at the group-scale halo mass.

The value of the central SHMR peak may also be compared to that of Coupon et al. (2012) computed from a clustering and galaxy number density analysis of the CFHTLS-Wide. In their study, the authors have measured the evolution of the SHMR peak as function of redshift and have found a lower value compared to ours ($1.1 \times 10^{12}\,\mathrm{M}_\odot$ at redshift $z \sim 0.7$). The difference may not be fully explained by cosmic variance, first because our field significantly overlaps with the full CFHTLS and secondly because the difference is larger than our error bars. In fact, due to their selection in the optical ($i < 22.5$), the SHMR peak above $z = 0.6$ is much less constrained than for our $K_s < 22$ sample, and their peak location suffers from higher uncertainties than in this study, not properly accounted for in their published error bars.

In Fig. 13, we compare our results with a number of semi-analytic predictions from the Millennium simulation (Springel, Frenk & White 2006). In brief, semi-analytic models are anchored to the dark matter halo merger trees provided by *N*-body simulations, in which empirical recipes of physical processes drive the evolution of galaxies. The fine-tuning of those different processes aim at reproducing the observed galaxy statistical properties. In each case, to derive the total SHMR we compute the sum of the central galaxy stellar mass and the integrated stellar masses of satellites with $M_\star > 10^{10}\,\mathrm{M}_\odot$ to match our sample mass completeness limit. The central SHMR is represented as a dashed line and the shaded area represents the total SHMR with 15 and 85 per cent percentiles. All quantities were computed at redshift $z = 0.8$. The model of Bower et al. (2006) is shown in red (top left), the model of De Lucia & Blaizot (2007) in orange (top right) and the model of Guo et al. (2011) – a modified version of the latter – in green (bottom right). In both De Lucia & Blaizot and Guo et al. models, the contribution from satellites to the total SHMR is significantly below the observations. Despite a different treatment of satellite galaxies and the efficiency of stellar feedback in the latter model, compared to the former, those

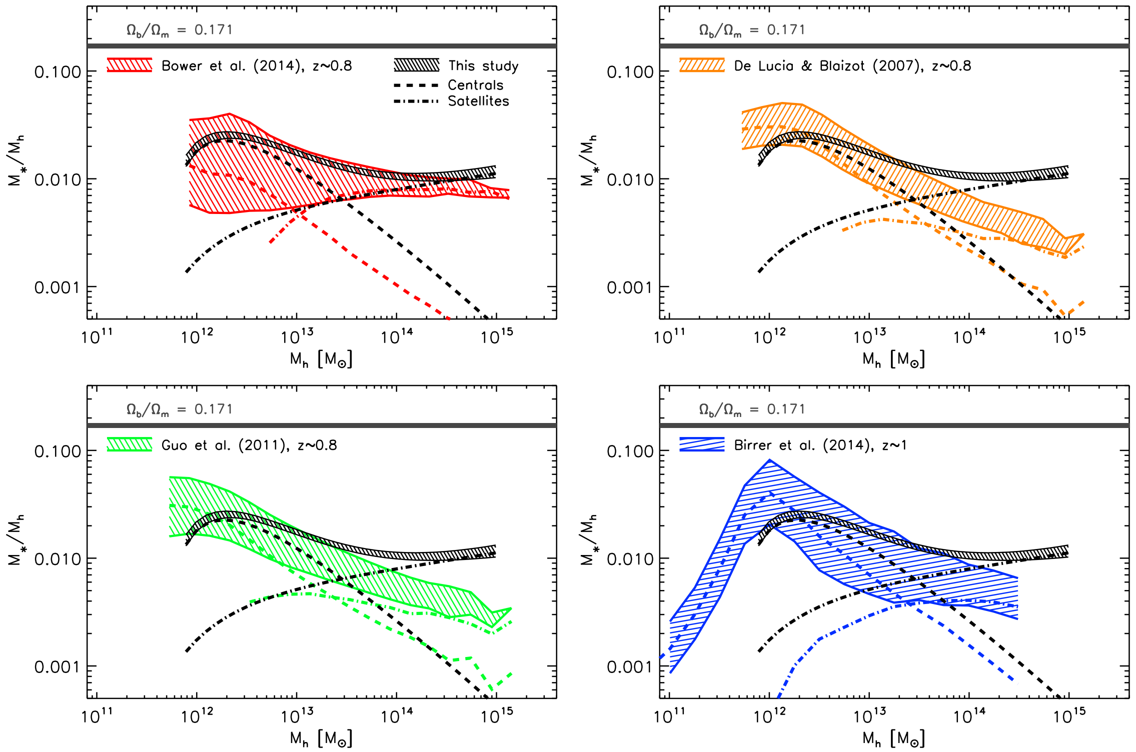

Figure 13. SHMR as function of halo mass compared with simulations from the literature. We compare our total, central and satellite SHMR results with three studies based on semi-analytic models applied to the Millennium simulation (top-row and bottom-left panels) and one study (bottom-right panel) based on the 'gas-regulator' analytical model. For each model, we also display the corresponding central (dashed line) and satellite (dot–dashed line) SHMR.

changes do not show up here. The discrepancy with our results could not arise from a limitation caused by the simulation resolution, as we imposed a cut of $M_\star > 10^{10}$ $M_\odot$ to match our observations. The model of Bower et al. better reproduces the observed satellite SHMR, however it underestimates the central SHMR and features a significant scatter in the $M_\star - M_h$ relationship.

We also show the results from the analytical model proposed by Birrer et al. (2014) in blue (bottom right). Their model is an application of the gas-regulator model (Lilly et al. 2013), in which the star formation efficiency is driven by the amount of available gas in the reservoir. In its simplest form, the model describes the inflows and outflows of the gas in the reservoir by two adjustable parameters: a star formation efficiency ϵ, and a mass-loading factor λ that represents the outflows, proportional to the SFR. We show their SHMR at $z = 1$ from the model 'C'.

6 DISCUSSION AND CONCLUSIONS

Using a unique combination of deep optical/NIR data and large area, we have combined galaxy clustering, lensing and galaxy abundance, to put constraints on the galaxy occupation function in the range $0.5 < z < 1$ and to link galaxy properties to dark matter halo masses. Our main result is an accurate measurement of the central galaxy $M_\star - M_h$ relationship at $z \sim 0.8$ ranging from halo masses at the peak of the SHMR up to the galaxy cluster mass regime. We also provide separate measurements of the SHMR for central and satellite galaxies.

We have shown that the statistical errors (computed using a jackknife estimator) were smaller than systematic errors in the stellar mass measurements caused by uncertainties in the assumed cosmology, dust modelling and photometric calibration. Due to the relatively small amount of statistical uncertainties, the low- to intermediate-mass regime of the stellar mass function is most affected by systematic errors: a factor of ~8 was found between statistical errors and total errors, increasing the error bars of parameters controlling the shape of the $M_\star - M_h$ relationship by approximately the same amount (see Table 4). Conversely, clustering and lensing measurements feature relatively higher statistical uncertainties and only a factor of ~2 increase in error of the HOD parameters describing the satellite population is observed compared to statistical errors. By probing such a large volume, nearly 0.1 Gpc^3, this study brings unprecedented constraints on the $M_\star - M_h$ relationship from statistical methods in the cluster mass regime at those redshifts. As shown in Fig. 10, our results make the link between statistical methods based on HOD applied to deep, small-volume surveys, with direct measurements of massive clusters from large-scale surveys.

For central galaxies, we have shown that when properly accounting for halo mass definition, choice of the IMF and the scatter between $M_\star$ and M_h, there is general agreement among results from the literature. We find that stellar mass estimates are the main source of uncertainty, as reflected by the light-blue shaded area from Behroozi et al. (2013b) in Fig. 10, or the stellar mass shift measured between Leauthaud et al. (2012) and Ilbert et al. (2010). We stress, however, that if stellar mass differences may induce a global shift (for instance caused by a separate choice for the IMF), it may also translate into a mass-dependent shift in the more general case (e.g. between two sets of SPS models): hence applying a constant shift may not necessarily reconcile two measurements.

In Fig. 11, stellar mass systematics do not seem to explain all of the observed differences with some results from the literature for which the stellar mass was measured in a similar way to this study. To measure the impact of some of the assumptions made in our model, we have compiled a list of potential systematics propagated through the halo mass and satellite normalization best-fitting parameters. We quote an estimate of 50 per cent error in M_1 and 20 per cent error in B_{sat}, respectively.

For satellite galaxies, the combination of lensing and clustering in this work represents a significant improvement over studies using only the stellar mass function. In Fig. 12, we have shown the measured total SHMR as function of halo mass, compared with a number of results from observations and simulations in the literature. Starting from group-size haloes up to the most massive clusters, we find that the total SHMR is gradually dominated by the contribution from satellites.

Clearly, most SAMs tend to underestimate the total amount of stellar mass produced in medium- to high-mass satellites ($10^{10} < M_\star/\mathrm{M}_\odot < 10^{11}$) at $z \sim 1$ compared to observations. This would suggest that, in SAMs, the bulk of star formation occurs in low-mass galaxies, but is quenched or suppressed at higher mass. Possible explanations for this include either a too strong quenching of haloes in the mass regime $10^{10} < M_\star/\mathrm{M}_\odot < 10^{11}$ (e.g. the work by Henriques et al. 2012, who argue that the gas could be later reincorporated into the haloes), or that low-mass subhaloes are too numerous and would 'catch' the gas in detriment of high-mass subhaloes. It is interesting to link this feature to the overabundance of low-mass galaxies found in numerical simulations compared to observations (see e.g. Guo et al. 2011; Weinmann et al. 2012; De Lucia, Muzzin & Weinmann 2014). In this context, Schive, Chiueh & Broadhurst (2014) recently proposed that cold dark matter could behave as a coherent wave and have shown using *N*-body simulations that this would suppress a large amount of small-mass haloes.

Finally, we can summarize our findings as follows:

(i) the HOD model accurately reproduces the four observables within the statistical error bars in all mass bins over three orders of magnitudes in halo mass and two orders of magnitudes in stellar mass;

(ii) our $M_\star - M_{\mathrm{h}}$ relationship shows generally good agreement with the literature measurements at $z \sim 0.8$ and we have shown that, when modelling differences are properly accounted for, we are able to make a fair comparison of a number of results derived using independent techniques;

(iii) the systematic errors affecting our measurements were propagated through the whole fitting process. For the parameters describing the $M_\star - M_{\mathrm{h}}$ relationship, we find that including systematic errors leads to a factor of 8 increase in error bars, and for the parameters describing the satellite HOD a factor of 2 increase in error bars, compared to statistical error bars;

(iv) the sum of systematic errors from the halo model and our model assumptions may be as high (but likely overestimated) as 50 per cent in halo mass and 20 per cent in the satellite number normalization;

(v) the central galaxy SHMR peaks at $M_{\mathrm{h}} = 1.9 \times 10^{12}\,\mathrm{M}_\odot$, a value slightly larger than the clustering results from the full CFHTLS from Coupon et al. (2012),

(vi) the total (central plus satellites) SHMR is dominated by the satellite contribution in the most massive haloes, in apparent contradiction with SAMs in the Millennium simulation.

We have demonstrated the power of associating a large and deep area with a combination of independent observables to constrain the galaxy–halo relationship with unprecedented accuracy up to $z = 1$. The potential of these data will undoubtedly allow us to extend this analysis to galaxies split by type in future work.

Additionally, studying the evolution in redshift of the SHMR above $z = 1$ is one of the greatest challenge in the near future. If abundance matching already probes the central galaxy–halo relationship up to high redshift, clustering and lensing are necessary to put constraints on the satellite HOD and break some of the degeneracies. Large-scale clustering measurements require wide-field imaging, whereas high-redshift lensing techniques are yet to be improved, but on-going projects such as Hyper Suprime Cam (HSC), Dark Energy Survey (DES) or COSMOS/SPLASH (Spitzer Large Area Survey with Hyper-Suprime-Cam), which will increase by orders of magnitude the currently available data, represent the ideal data sets to address those issues.

ACKNOWLEDGEMENTS

This work is primarily based on observations obtained with WIRCam, a joint project of CFHT, Taiwan, Korea, Canada, France, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France and the University of Hawaii. The WIRCAM images have been collected during several semesters and from different programs. We thank G. Morrison, J. Willis and K. Thanjavur for leading some of these programmes as PIs, and a special thanks to the canadian agency who always highly ranked these proposals. We thank the Terapix team for the reduction of all the WIRCAM images and the preparation of the catalogues matching the T0007 CFHTLS data release. The CFHTLenS project is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the CFHT. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. We thank the CFHT staff for successfully conducting the CFHTLS observations and in particular J.-C. Cuillandre and E. Magnier for the continuous improvement of the instrument calibration and the Elixir detrended data that we used. We also thank TERAPIX for the quality assessment and validation of individual exposures during the CFHTLS data acquisition period, and E. Bertin for developing some of the software used in this study. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program, and to HPC specialist O. Toader. Part of the numerical computations were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth.

We thank M. Cousin and S. de la Torre for useful discussions related to this work. We thank S. Birrer and D. Wake for providing us with their measurements. JC acknowledges the support from the laboratoire de Marseille during his stay in 2014 September. The early stages of the CFHTLenS project was made possible thanks to the support of the European Commission's Marie Curie Research Training Network DUEL (MRTN-CT-2006-036133) which directly supported six members of the CFHTLenS team (LF, HH, PS, BR, CB, MV) between 2007 and 2011 in addition to providing travel support and expenses for team meetings. BR +CH and HHo+ acknowledge support from the European Research Council under EC FP7 grant numbers 240672 (BR), 240185 (CH) and 279396 (HHo). HHi is supported by the DFG Emmy Noether grant Hi 1495/2-1. AF acknowledges support by INAF through VIPERS grants PRIN 2008 and PRIN 2010. MJH acknowledges support from the Â Natural Sciences and Engineering Research Council of Canada (NSERC).

Author Contributions: All authors contributed to the development and writing of this paper. The authorship list reflects the lead authors of this paper (JC, SA, LVW, TM, OI, EVU) followed by two alphabetical groups. The first alphabetical group includes key contributors to the science analysis and interpretation in this paper, the founding core team and those whose long-term significant effort produced the final CFHTLenS data product, and the VIPERS collaboration leaders. The second group covers authors in alphabetical order who made a significant contribution to either the projects, this paper, or both. The CFHTLenS collaboration was co-led by CH and LVW.

REFERENCES

Ahn C. P. et al., 2014, ApJS, 211, 17

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540

Arnouts S. et al., 2013, A&A, 558, A67

Balogh M. L. et al., 2014, MNRAS, 443, 2679

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379

Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJ, 770, 57

Behroozi P. S., Marchesini D., Wechsler R. H., Muzzin A., Papovich C., Stefanon M., 2013b, ApJ, 777, L10

Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

Berlind A. A. et al., 2003, ApJ, 593, 1

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bielby R. et al., 2012, A&A, 545, A23

Bielby R. M. et al., 2014, A&A, 568, A24

Birrer S., Lilly S., Amara A., Paranjape A., Refregier A., 2014, ApJ, 793, 12

Boulade O. et al., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. SPIE, Bellingham, p. 72

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Brainerd T. G., Blandford R. D., Smail I., 1996, ApJ, 466, 623

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Budzynski J. M., Koposov S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, MNRAS, 423, 104

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559

Bundy K. et al., 2006, ApJ, 651, 120

Cacciato M., van den B., Frank C., More S., Li R., Mo H. J., Yang X., 2009, MNRAS, 394, 929

Cacciato M., van Uitert E., Hoekstra H., 2014, MNRAS, 437, 377

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Chabrier G., 2003, PASP, 115, 763

Charlot S., Fall S. M., 2000, ApJ, 539, 718

Coil A. L. et al., 2011, ApJ, 741, 8

Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201

Conroy C. et al., 2007, ApJ, 654, 153

Conseil S., Vibert D., Amouts S., Milliard B., Zamojski M., Liebaria A., Guillaume M., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 107

Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

Coupon J. et al., 2009, A&A, 500, 981

Coupon J. et al., 2012, A&A, 542, A5

Coupon J., Broadhurst T., Umetsu K., 2013, ApJ, 772, 65

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839

Davidzon I. et al., 2013, A&A, 558, A23

de la Torre S. et al., 2013, A&A, 557, A54

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

De Lucia G., Muzzin A., Weinmann S., 2014, New Astron Rev., 62, 1

Dunkley J. et al., 2009, ApJ, 701, 1804

Erben T. et al., 2013, MNRAS, 433, 2545

Faber S. M. et al., 2007, ApJ, 665, 265

Foucaud S., Conselice C. J., Hartley W. G., Lane K. P., Bamford S. P., Almaini O., Bundy K., 2010, MNRAS, 406, 147

Fritz A. et al., 2014, A&A, 563, A92

Garilli B. et al., 2008, A&A, 486, 683

Garilli B. et al., 2014, A&A, 562, A23

Gelman A., Rubin D., 1992, Stat. Sci., 7, 457

George M. R. et al., 2011, ApJ, 742, 125

Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111

Guo Q. et al., 2011, MNRAS, 413, 101

Guzzo L. et al., 2014, A&A, 566, A108

Henriques B. M. B., White S. D. M., Lemson G., Thomas P. A., Guo Q., Marleau G.-D., Overzier R. A., 2012, MNRAS, 421, 2904

Heymans C. et al., 2006, MNRAS, 371, L60

Heymans C. et al., 2012, MNRAS, 427, 146

Hildebrandt H. et al., 2012, MNRAS, 421, 2355

Hilton M. et al., 2013, MNRAS, 435, 3469

Hinshaw G. et al., 2009, ApJS, 180, 225

Hoaglin D. C., Mosteller F., Tukey J. W., 1983, Understanding Robust and Exploratory Data Analysis. Wiley, New York

Hoekstra H., Yee H. K. C., Gladders M. D., 2004, in Ryder S.D., Pisano D.J., Walker M.A., Freeman K.C., eds, Proc. IAU Symp. 220, Dark Matter in Galaxies. Astron. Soc. Pac., San Francisco, p. 439

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Hudelot P. et al., 2012, VizieR Online Data Catalog, 2317, 0

Hudson M. J., Gwyn S. D. J., Dahle H., Kaiser N., 1998, ApJ, 503, 531

Hudson M. J. et al., 2015, MNRAS, 447, 298

Ilbert O. et al., 2006, A&A, 457, 841

Ilbert O. et al., 2009, ApJ, 690, 1236

Ilbert O. et al., 2010, ApJ, 709, 644

Ilbert O. et al., 2013, A&A, 556, A55

Kettula K., S. G., van Uitert E., Hoekstra H., A. F., 2014, MNRAS

Kitching T. D., Miller L., Heymans C. E., van Waerbeke L., Heavens A. F., 2008, MNRAS, 390, 149

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35

Kroupa P., 2001, MNRAS, 322, 231

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lawrence A. et al., 2007, MNRAS, 379, 1599

Le Fèvre O. et al., 2005, A&A, 439, 845

Le Fèvre O. et al., 2014, preprint (arXiv:e-prints)

Leauthaud A. et al., 2009, ApJ, 709, 97

Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, ApJ, 738, 45

Leauthaud A. et al., 2012, ApJ, 744, 159

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119

Mandelbaum R. et al., 2005a, MNRAS, 361, 1287

Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005b, MNRAS, 362, 1451

Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715

Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544

Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765

Marinoni C., Hudson M. J., 2002, ApJ, 569, 101

Martin D. C. et al., 2005, ApJ, 619, L1

Martinez-Manso J., Gonzalez A. H., Ashby M. L. N., Stanford S. A., Brodwin M., Holder G. P., Stern D., 2015, MNRAS, 446, 169

Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, MNRAS, 382, 315

Miller L. et al., 2013, MNRAS, 429, 2858

Mirkazemi M. et al., submitted

Miyatake et al., 2013, preprint (arXiv:1311.1480)

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210

More S., Miyatake H., Mandelbaum R., Takada M., Spergel D., Brownstein J., Schneider D. P., 2014, preprint (arXiv:1407.1856)
Morrissey P. et al., 2005, ApJ, 619, L7
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, ApJ, 731, 113
Moster B. P., Naab T., White S. D. M., 2012, MNRAS, 428, 3121
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Muñoz-Cuartas J. C., Macciò A. V., Gottlöber S., Dutton A. A., 2010, MNRAS, 411, 584
Muzzin A., Yee H. K. C., Hall P. B., Ellingson E., Lin H., 2007, ApJ, 659, 1106
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Norberg P., Baugh C. M., Gaztanaga E., Croton D. J., 2009, MNRAS, 396, 19
Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, ApJ, 769, L35
Planck Collaboration XVI 2014, A&A, 571, 16
Polletta M. et al., 2007, ApJ, 663, 81
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical Recipes in C++: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Puget P. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 978
Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, MNRAS, 337, 1282
Salpeter E. E., 1955, ApJ, 121, 161
Schive H.-Y., Chiueh T., Broadhurst T., 2014, Nat. Phys., 10, 496
Scoville N. et al., 2007, ApJS, 172, 1
Sheldon E. S. et al., 2004, AJ, 127, 2544
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Szalay A. S., Connolly A. J., Szokoly G. P., 1999, AJ, 117, 68
Takada M., Jain B., 2003, MNRAS, 344, 857
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, ApJ, 729, 127
Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
van der Burg R. F. J., Muzzin A., Hoekstra H., Wilson G., Lidman C., Yee H. K. C., 2014, A&A, 561, A79
van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14
Velander M. et al., 2014, MNRAS, 437, 2111
Wake D. A. et al., 2011, ApJ, 728, 46
Wang L., Jing Y. P., 2010, MNRAS, 402, 1796
Weinberg N. N., Kamionkowski M., 2002, MNRAS, 337, 1269
Weinmann S. M., Pasquali A., Oppenheimer B. D., Finlator K., Mendel J. T., Crain R. A., Macciò A. V., 2012, MNRAS, 426, 2797
Wraith D., Kilbinger M., Benabed K., Cappé O., Cardoso J.-F., Fort G., Prunet S., Robert C., 2009, Phys. Rev. D, 80, 023507
Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057
Yoo J., Tinker J. L., Weinberg D. H., Zheng Z., Katz N., Davé R., 2006, ApJ, 652, 26
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I. et al., 2011, ApJ, 736, 59
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

APPENDIX A: COMPLETENESS OF THE SAMPLES

In this section, we use the CFHTLS-Deep/WIRDS combined data to test our samples' mass completeness. The CFHTLS-Deep/WIRDS data are over 2 mag deeper in all bands compared to our CFHTLS-Wide/WIRCam data and with accurate photometric redshift and stellar mass estimates computed in a similar fashion to this study. Fig. A1 shows the galaxy distribution in WIRDS as function of stellar mass and redshift corresponding to our selection $K_s < 22$ for the photometric sample (top) and $i < 22.5$ for the spectroscopic sample (bottom).

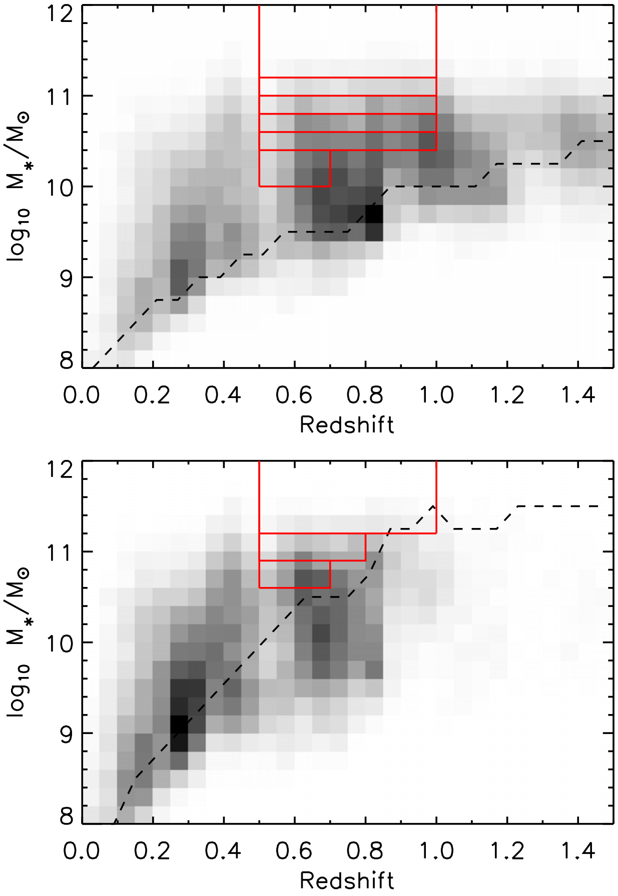

Figure A1. Galaxy distribution as function of stellar mass and redshift in WIRDS. Stellar mass 90 per cent completeness limits of $K_s < 22$ (top) and $i < 22.5$ (bottom) selected samples are represented as the dashed black line and the sample selection as the thick red line.

The density fluctuations seen as function of redshift are due to cosmic variance (the field of view is smaller than 1 deg^2), but we do not expect any significant impact on our completeness assessments. In both panels, we represent the 90 per cent completeness limits as dashed lines, and our samples' selection as red boxes. In the case of the photometric sample, a conservative $z < 0.7$ cut is adopted in the lower mass sample to prevent missing red galaxies caused by the optical incompleteness at the CFHTLS-Wide depth. Overall, these verifications show that all of our samples are complete in mass.

APPENDIX B: DETAILS ON THE DERIVATION OF THE OBSERVABLES

Here, we provide detailed calculations of the four observables used in this study and derived from the HOD model described in Section 4. For the dark matter halo profile and the distribution of satellites, we assume a (Navarro et al. 1997, NFW) profile with the theoretical mass–concentration relation from equation (16) of

Takada & Jain (2003) with $c_0 = 11$ and $\beta = 0.13$, featuring the redshift dependence $(1+z)^{-1}$ (Bullock et al. 2001). All dark matter quantities are derived at the mean redshift of the galaxy sample, computed from the expectation value of the sum of redshift PDFs. All quantities are computed in comoving units ('co'). The clustering and galaxy–galaxy lensing are then converted into physical units ('phys') to match the measurements.

B1 Stellar mass function

The stellar mass function is the integrated HOD over the halo mass function:

$$\phi_{\mathrm{SMF}}\left(M_\star^{t_1}, M_\star^{t_2}\right) = \int_0^\infty \left\langle N_{\mathrm{tot}}\left(M_{\mathrm{h}}|M_\star{}^{t1}, M_\star{}^{t2}\right)\right\rangle \frac{\mathrm{d}n}{\mathrm{d}M_{\mathrm{h}}}\mathrm{d}M_{\mathrm{h}}. \tag{B1}$$

B2 Galaxy clustering

We describe galaxy clustering using the two-point correlation function, as the sum of the one-halo and two-halo terms:

$$\xi_{\mathrm{gg}}(r_{\mathrm{co}}) = 1 + \xi_{\mathrm{gg},1}(r_{\mathrm{co}}) + \xi_{\mathrm{gg},2}(r_{\mathrm{co}})\,. \tag{B2}$$

The one-halo term, $\xi_{\mathrm{gg},1}(r_{\mathrm{co}})$, expresses the relative contribution of galaxy pairs within the halo $\langle N_{\mathrm{tot}}(M_{\mathrm{h}})(N_{\mathrm{tot}}(M_{\mathrm{h}})-1)\rangle/2$ and can be decomposed, assuming Poisson statistics for the satellites, into two terms:

$$\begin{aligned}\langle N_{\mathrm{cen}}N_{\mathrm{sat}}\rangle(M_{\mathrm{h}}) &= \langle N_{\mathrm{cen}}(M_{\mathrm{h}})\rangle\langle N_{\mathrm{sat}}(M_{\mathrm{h}})\rangle;\\ \langle N_{\mathrm{sat}}(N_{\mathrm{sat}}-1)\rangle(M_{\mathrm{h}})/2 &= \langle N_{\mathrm{sat}}(M_{\mathrm{h}})\rangle^2/2\,.\end{aligned} \tag{B3}$$

The correlation function for central-satellite pairs is given by

$$1 + \xi_{\mathrm{cs}}(r_{\mathrm{co}}, z) = \int_{M_{\mathrm{vir}}(r)}^{\infty} \mathrm{d}M_{\mathrm{h}}\, n(M_{\mathrm{h}}, z)\frac{\langle N_{\mathrm{cen}}\rangle\langle N_{\mathrm{sat}}\rangle}{n_{\mathrm{gal}}^2/2}\rho_{\mathrm{h}}(r_{\mathrm{co}}|M_{\mathrm{h}}), \tag{B4}$$

where we assume that the distribution of central-satellite pairs simply follows that of the dark matter halo profile. The lower integration limit $M_{\mathrm{vir}}(r_{\mathrm{co}})$ accounts for the fact that no halo with a virial mass corresponding to r_{co} would contribute to the correlation function.

For the satellite contribution ξ_{ss}, the distribution of satellite pairs is the convolution of the dark matter halo profile with itself, computed here in Fourier space. The satellite power spectrum is

$$P_{\mathrm{ss}}(k) = \int_{M_{\mathrm{low}}}^{M_{\mathrm{high}}} \mathrm{d}M_{\mathrm{h}}\, n(M_{\mathrm{h}})\frac{\langle N_{\mathrm{sat}}(M_{\mathrm{h}})\rangle^2}{n_{\mathrm{gal}}^2}|u_h(k|M_{\mathrm{h}})|^2\,, \tag{B5}$$

where $u_{\mathrm{h}}(k|M_{\mathrm{h}})$ is the Fourier transform of the dark-matter halo profile $\rho_{\mathrm{h}}(r_{\mathrm{co}}|M_{\mathrm{h}})$. The correlation function ξ_{ss} is then obtained via a Fourier transform.

The one-halo correlation function is the sum of the two contributions,

$$\xi_{\mathrm{gg},1}(r_{\mathrm{co}}) = 1 + \xi_{\mathrm{cs}}(r_{\mathrm{co}}) + \xi_{\mathrm{ss}}(r_{\mathrm{co}})\,. \tag{B6}$$

The two-halo term is computed from the galaxy power spectrum:

$$P_2(k, r_{\mathrm{co}}) = P_{\mathrm{m}}(k) \times \left[\int_{M_{\mathrm{low}}}^{M_{\mathrm{lim}}(r_{\mathrm{co}})} \mathrm{d}M_{\mathrm{h}} n(M_{\mathrm{h}})\frac{\langle N_{\mathrm{tot}}\rangle}{n'_{\mathrm{gal}}(r_{\mathrm{co}})}b_{\mathrm{h}}(M_{\mathrm{h}}, r_{\mathrm{co}})|u_{\mathrm{h}}(k|M_{\mathrm{h}})|\right]^2, \tag{B7}$$

where

$$n'_{\mathrm{gal}}(r_{\mathrm{co}}) = \int_{M_{\mathrm{low}}}^{M_{\mathrm{lim}}(r_{\mathrm{co}})} n(M_{\mathrm{h}})\langle N_{\mathrm{tot}}\rangle\,\mathrm{d}M_{\mathrm{h}}\,. \tag{B8}$$

The upper integration limit $M_{\mathrm{lim}}(r_{\mathrm{co}})$ accounts for halo exclusion as detailed in Coupon et al. (2012), and references therein.

Finally, the two-halo term $\xi_{\mathrm{gg},2}$ of the galaxy autocorrelation function is the Fourier transform of equation (B7) renormalized to the total number of galaxy pairs:

$$1 + \xi_{\mathrm{gg},2}(r_{\mathrm{co}}) = \left[\frac{n'_{\mathrm{gal}}(r_{\mathrm{co}})}{n_{\mathrm{gal}}}\right][1 + \xi_{\mathrm{gg},2}(r_{\mathrm{co}})]\,. \tag{B9}$$

The projected clustering $w(\theta)$ is derived from the projection of ξ_{gg} on to the estimated redshift distribution from the sum of PDFs, assuming the Limber approximation (see details in Coupon et al. 2012).

The real-space clustering $w_{\mathrm{p}}(r_{\mathrm{p,\,co}})$ is derived from the projection of the 3D correlation function along the line of sight:

$$w_{\mathrm{p}}(r_{\mathrm{p,co}}) = 2\int_{r_{\mathrm{p,co}}}^{\infty} r_{\mathrm{co}}\,\mathrm{d}r_{\mathrm{co}}\xi_{\mathrm{gg}}(r_{\mathrm{co}})\,(r_{\mathrm{co}}^2 - r_{\mathrm{p,co}}^2)^{-1/2}\,, \tag{B10}$$

converted into physical units as

$$w_{\mathrm{p,phys}} = w_{\mathrm{p,co}}/(1+z)\,. \tag{B11}$$

B3 Galaxy–galaxy lensing

The galaxy–galaxy lensing estimator measures the excess surface density of the projected dark matter halo profile:

$$\Delta\Sigma_{\mathrm{co}}(r_{\mathrm{p,co}}) = \overline{\Sigma}_{\mathrm{co}}(<r_{\mathrm{p,co}}) - \overline{\Sigma}_{\mathrm{co}}(r_{\mathrm{p,co}})\,, \tag{B12}$$

where $\overline{\Sigma}_{\mathrm{co}}(<r_{\mathrm{p,co}})$ is the projected mean surface density within the comoving radius $r_{\mathrm{p,\,co}}$ and $\overline{\Sigma}_{\mathrm{co}}(r_{\mathrm{p,co}})$ the mean surface density at the radius $r_{\mathrm{p,co}}$.

To compute the analytical projected dark matter density Σ, we write

$$\begin{aligned}\Sigma_{\mathrm{co}}(r_{\mathrm{p,co}}) &= \int \rho\left(\sqrt{r_{\mathrm{p,co}}^2 + \pi_{\mathrm{co}}^2}\right)\mathrm{d}\pi_{\mathrm{co}}\\ &= \overline{\rho}\int\left[1 + \xi_{\mathrm{gm}}\left(\sqrt{r_{\mathrm{p,co}}^2 + \pi_{\mathrm{co}}^2}\right)\right]\mathrm{d}\pi_{\mathrm{co}},\end{aligned} \tag{B13}$$

where $r_{\mathrm{p,co}}$ is the transverse comoving distance, π_{co} the line-of-sight comoving distance, $\overline{\rho}$ the mean density of the Universe, so that $\Delta\Sigma_{\mathrm{co}}(r_{\mathrm{p,\,co}})$ is related to the galaxy-dark matter cross-correlation function ξ_{gm} through

$$\begin{aligned}\Delta\Sigma_{\mathrm{co}}(r_{\mathrm{p,co}}) &= \overline{\Sigma}_{\mathrm{co}}(<r_{\mathrm{p,co}}) - \overline{\Sigma}_{\mathrm{co}}(r_{\mathrm{p,co}})\\ &= \overline{\rho}\left[\frac{4}{r_{\mathrm{p,co}}^2}\int_0^{r_{\mathrm{p,co}}}\int_0^{\pi_{\mathrm{max}}} r'_{\mathrm{p,co}}\xi_{\mathrm{gm}}\left(\sqrt{r'^2_{\mathrm{p,co}} + \pi_{\mathrm{co}}^2}\right)\mathrm{d}\pi_{\mathrm{co}}\mathrm{d}r'_{\mathrm{p,co}}\right.\\ &\qquad\left.-2\int_0^{\pi_{\mathrm{max}}}\xi_{\mathrm{gm}}\left(\sqrt{r_{\mathrm{p,co}}^2 + \pi_{\mathrm{co}}^2}\right)\mathrm{d}\pi_{\mathrm{co}}\right].\end{aligned} \tag{B14}$$

The integration along the line of sight is performed up to the scale $\pi_{\mathrm{max}} = 80$ Mpc.

The excess surface density in physical units writes

$$\overline{\Delta\Sigma}_{\mathrm{phys}} = \overline{\Delta\Sigma}_{\mathrm{co}} \times (1 + z_{\mathrm{L}})^2\,, \tag{B15}$$

where z_{L} is the redshift of the lens galaxy.

As for ξ_{gg}, ξ_{gm} can be written as the sum of the one- and two-halo terms:

$$\xi_{\mathrm{gm}}(r) = 1 + \xi_{\mathrm{gm},1}(r) + \xi_{\mathrm{gm},2}(r)\,. \tag{B16}$$

$\xi_{\mathrm{gm},1}(r)$ is itself decomposed into a contribution from the cross-correlation of the central galaxy-dark matter and from that of the

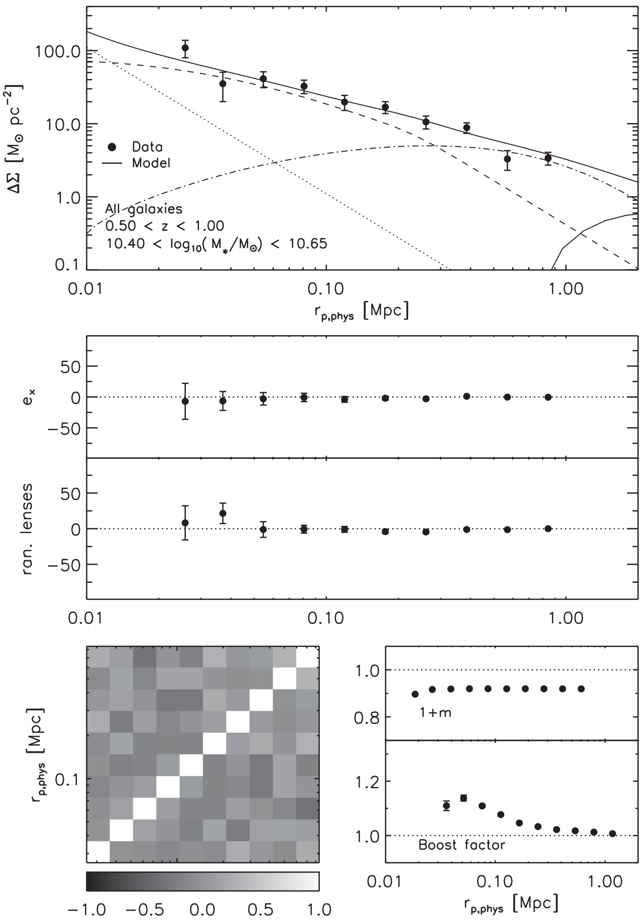

Figure C1. Galaxy–galaxy lensing measurements and systematics checks for the sample $10.40 < \log(M_\star/{\rm M}_\odot) < 10.65$. In the top panel, we show the data (dots with error bars) and the model (thick line) split into the stellar term in dotted line, the central term in dashed line, the satellite term in dot–dashed line and the two-halo term in black solid line at bottom-right corner. The lower panels show the systematic tests (rotated-shape signal and random lens positions), calibration factor (multiplicative bias correction and boost factor) and the lower-left corner the correlation coefficients of the correlation matrix from the jackknife estimate.

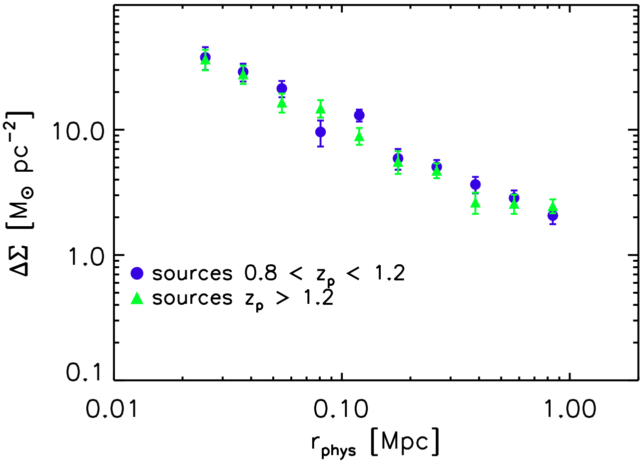

Figure C2. Galaxy–galaxy lensing measurements separating the background sample into $0.8 < z_{\rm p} < 1.2$ sources (purple dots) and $z_{\rm p} > 1.2$ sources (green triangles), keeping the same lens galaxy foreground sample (low-redshift galaxies with spectroscopic redshifts).

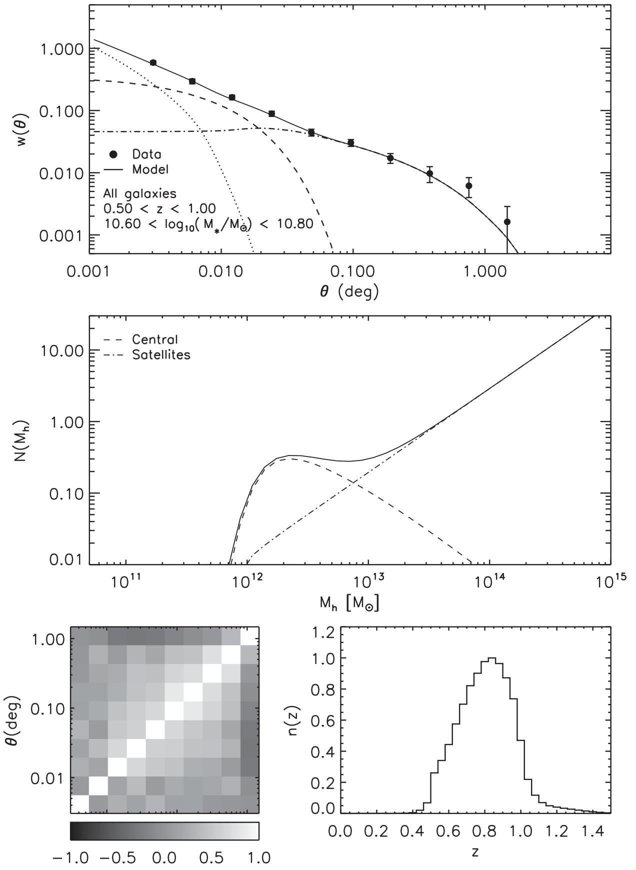

Figure C3. $w(\theta)$ measurements and the corresponding HOD function for the sample $10.60 < \log(M_\star/{\rm M}_\odot) < 10.80$. In the top panel, we show the data points with error bars and the best-fitting model: the dotted line represents the central-satellite cross-correlation, the dashed line the satellite–satellite autocorrelation, and the dot–dashed line the central–central autocorrelation (or 2-halo term). The middle panel displays the corresponding HOD, the dashed line shows the central galaxy HOD and the dot–dashed line the satellites' HOD. The lower-right panel shows the corresponding redshift distribution constructed from the sum of individual PDFs. The lower-left panel shows the correlation coefficients of the covariance matrix from the jackknife estimate.

satellite-dark matter, both assuming an NFW profile. We write the former as

$$1+\xi_{\rm gm,cen}(r,z) = \int_{M_{\rm vir}(r)}^{M_{\rm high}} {\rm d}M_{\rm h}\, n(M_{\rm h},z)\frac{\langle N_{\rm cen}\rangle}{n_{\rm gal}}\rho_{\rm h}(r|M_{\rm h})\frac{M_{\rm h}}{\bar{\rho}} \qquad \text{(B17)}$$

and the latter $\xi_{\rm gm,sat}$ from the Fourier transform of its power spectrum

$$P_{\rm gm,ss}(k) = \int_{M_{\rm low}}^{M_{\rm high}} {\rm d}M_{\rm h}\, n(M_{\rm h})\frac{\langle N_{\rm sat}(M_{\rm h})\rangle}{n_{\rm gal}}\frac{M_{\rm h}}{\bar{\rho}}|u_{\rm h}(k|M_{\rm h})|^2. \qquad \text{(B18)}$$

Finally, we compute the two-halo term $\xi_{\rm gm,2}(r)$ from the Fourier transform of the galaxy–dark matter cross-correlation power spectrum:

$$P_{\rm gm,2}(k,r) = P_{\rm m}(k) \times \int_{M_{\rm low}}^{M_{\rm lim}(r)} {\rm d}M_{\rm h} n(M_{\rm h})\frac{\langle N_{\rm tot}(M_{\rm h})\rangle}{n'_{\rm gal}(r)} b_{\rm h}(M_{\rm h},r)|u_{\rm h}(k|M_{\rm h})|, \qquad \text{(B19)}$$

with a similar treatment of halo exclusion to that of the galaxy power spectrum.

APPENDIX C: SYSTEMATICS CHECKS ON LENSING AND CLUSTERING

We have performed systematics checks for the lensing and clustering measurements. In Fig. C1, we detail the galaxy–galaxy lensing measurement for the sample $10.40 < \log(M_*/\mathrm{M}_{\odot}) < 10.65$ as an example. The top panel shows the data (dots with error bars) and best-fitting model (thick line) with the different components of the model the central galaxy term, the satellite term and the two-halo term. The lower panels show a number of systematics checks. The '$e_{\times}$' panel shows the signal measured after rotating the ellipticities by 45° and the 'ran. lenses' panel shows the signal measured by randomizing the lenses positions, both consistent with zero. The '1+m' panel shows the multiplicative bias correction applied to the galaxy–galaxy lensing measurement, estimated after replacing the ellipticities by the multiplicative calibration factor $1 + m$. The 'boost factor' was estimated from randomizing the background source positions and measuring the ratio of the number of real sources over random objects as a function of distance from the lenses, and applied to the galaxy–galaxy lensing measurement. The covariance matrix from the jackknife estimate is shown in the

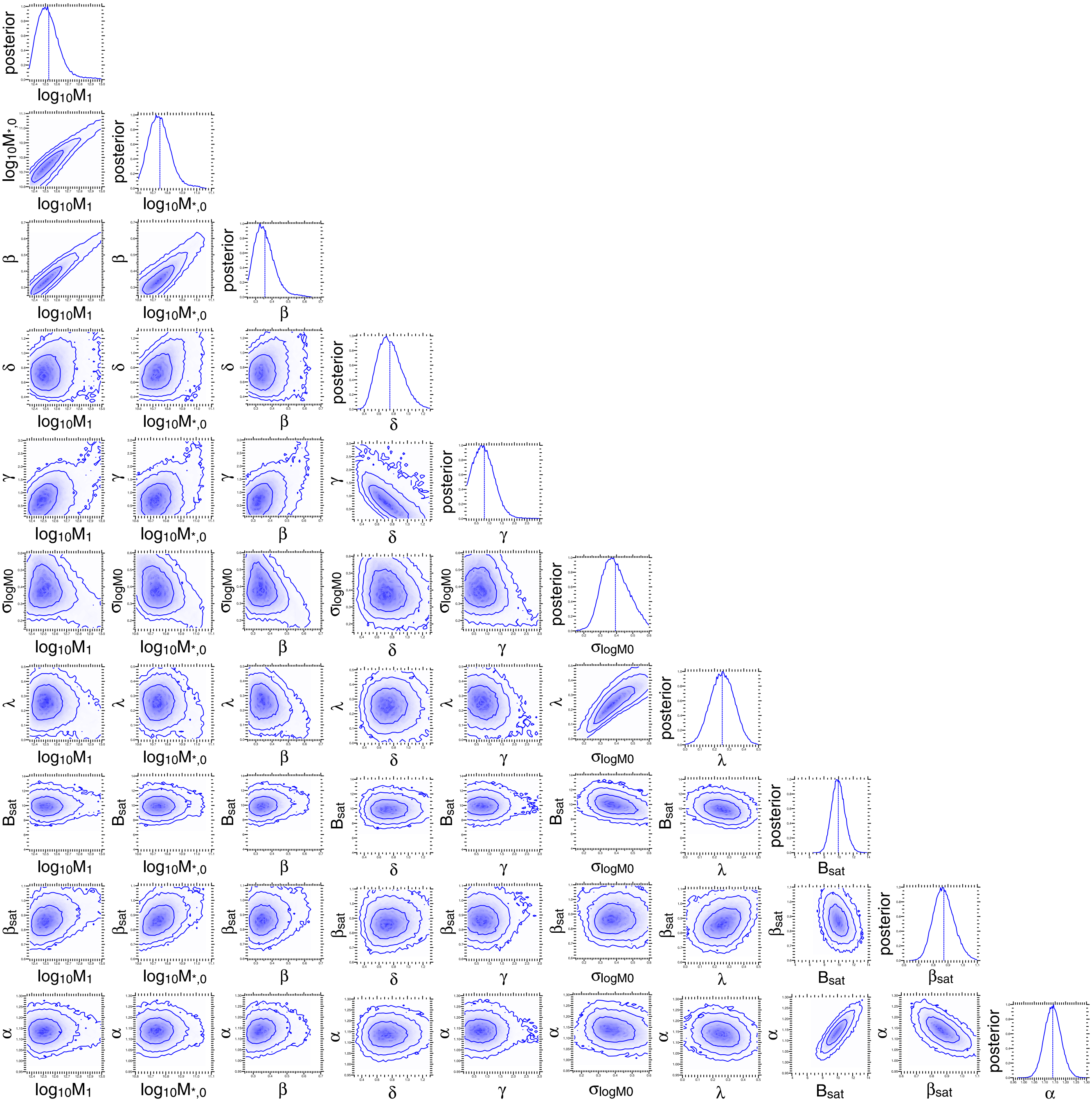

Figure D1. 1D (diagonal) and 2D likelihood distributions of best-fitting HOD parameters in the case of total errors. The 2D contours represent the 68.3, 95.5 and 99.7 per cent confidence limits. We used flat priors within the ranges shown on the figure for all parameters.

left-bottom corner of the figure. The relatively small off-diagonal values show the low correlation between data points. We repeated identical tests for all mass bins. In all cases, systematics are found to be consistent with zero.

In Fig. C2, we test the impact of including high-redshift sources beyond $z > 1.2$. To do so, we select an arbitrary sample of low-redshift lens galaxies with a spectroscopic redshift and we measured the galaxy–galaxy lensing signal using all sources with $0.8 < z_{\rm p} < 1.2$ (purple dots in the figure) and all sources with $z_{\rm p} > 1.2$ (green triangles in the figure). We see no significant difference between the two signals, meaning that the photometric redshifts and shape measurements in our catalogue are robust enough beyond $z_{\rm p} > 1.2$.

In Fig. C3, we show the projected clustering in the mass bin $10.60 < \log(M_\star/\,{\rm M}_\odot) < 10.80$. The top panel shows the data points with error bars and the best-fitting model, with the different components of the model: the one-halo term split into the central-satellite and satellite–satellite terms and the two-halo term. In the middle panel, we show the corresponding HOD, as a dashed line for the central contribution and as a dot–dashed line for the satellites' contribution.

APPENDIX D: 2D CONTOURS

We show in Fig. D1, the likelihood distributions of the best-fitting HOD parameters. Here, the results are shown for the MCMC run done with total (statistical plus systematic) errors.

[1] *Astronomical Observatory of the University of Geneva, ch. d'Ecogia 16, CH-1290 Versoix, Switzerland*
[2] *Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*
[3] *Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1,Canada*
[4] *Argelander Institute for Astronomy, University of Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[5] *INAF Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, via Bassini 15, I-20133 Milano, Italy*
[6] *INAF Osservatorio Astronomico di Brera, via Brera 28, 20122 Milano, via E. Bianchi 46, I-23807 Merate, Italy*
[7] *3 Dipartimento di Fisica, Università di Milano-Bicocca, P.zza della Scienza 3, I-20126 Milano, Italy*
[8] *The Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[9] *Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*
[10] *CEA Saclay, Service d'Astrophysique (SAp), Orme des Merisiers, Bât 709, F-91191 Gif-sur-Yvette, France*
[11] *Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*
[12] *Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[13] *CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[14] *University of Oxford, Department of Physics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[15] *IFAE, Campus UAB, E-08193 Bellaterra, Spain*
[16] *Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy*
[17] *INFN, Sezione di Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy*
[18] *INAF Osservatorio Astronomico di Roma, via Frascati 33, I-00040 Monte Porzio Catone (RM), Italy*
[19] *Dipartimento di Fisica e Astronomia - Università di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[20] *INAF Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy*
[21] *INAF Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[22] *Shanghai Key Lab for Astrophysics, Shanghai Normal University, 100 Guilin Road, 200234, Shanghai, China*
[23] *Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[24] *Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo, ON N2L 1Y5, Canada*
[25] *Kavli Institute for the Physics and Mathematics of the Universe, Todai Institutes for Advanced Study, the University of Tokyo, Kashiwa 277-8583, Japan*
[26] *INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[27] *Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

This paper has been typeset from a TEX/LATEX file prepared by the author.